

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LANGAGE ET NORMES : UNE APPROCHE PRAGMATIQUE

THÈSE  
PRÉSENTÉE  
COMME EXIGENCE PARTIELLE  
DU DOCTORAT EN PHILOSOPHIE

PAR  
YAËL SEBBAN

AVRIL 2014

VOLUME II

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## CHAPITRE IV

### DAVIDSON : SENS ET PORTÉE D'UNE NORMATIVITÉ CONSTITUTIVE DE LA SIGNIFICATION

#### 4.1. Introduction

Parmi les diverses approches quant aux normes de signification entrevues jusqu'ici, deux positions semblent bien incarner, les deux branches ou extrémités du dilemme, tel que McDowell le décrit, à savoir : d'un côté, une conception naturaliste et réductionniste de la signification et de l'autre, une conception platoniste. Chacune soulève des difficultés insurmontables et en appelle ainsi à l'adoption d'une position intermédiaire. Par position intermédiaire ou position du juste milieu, nous entendons une position qui maintiendrait l'idée kantienne d'une rationalité forte tout en veillant à ne pas la confondre avec la thèse naturaliste mais sans retenir l'idée kantienne d'un dualisme métaphysique<sup>538</sup>. Il s'agit ainsi d'examiner la thèse d'une normativité forte ou de normes de signification objectivement correctes en intégrant une dimension sociale. Un exemple d'une position intermédiaire est, d'une certaine manière, l'approche de McDowell s'inspirant ainsi de Wittgenstein et cherchant à rendre compte d'une distinction normative forte en termes d'une application correcte ou incorrecte des termes au moyen des notions sociales de coutume, de pratique et de

---

<sup>538</sup> Ce dualisme revêt, chez Kant, bien des formes : dualisme de la sensibilité et de la raison; dualisme du sensible et de l'intelligible entre le phénomène objet de l'expérience et le noumène objet de l'aperception transcendantale c'est-à-dire que les objets, en tant que matière de notre sensibilité, ne sont autre chose qu'une pure diversité ou de simples phénomènes, etc. Mais ce sont toujours là des formes du même dualisme : du dualisme radical qui caractérise le monde donné. Kant, 1944, *Analytique Transcendantale*, Livre II, Chap. III, p. 224, souligne cette distinction entre phénomènes et noumènes, comme suit : « [...] la division des objets en phénomènes et noumènes, par suite celle du monde en monde des sens et monde de l'entendement, en ce sens que la différence ne porte pas ici simplement sur la forme logique de la connaissance obscure ou distincte d'une seule et même chose, mais sur la manière dont les objets peuvent être données originellement à notre connaissance [...] quand les sens nous représentent quelque chose simplement tel qu'il apparaît, il faut cependant que ce quelque chose soit aussi une chose en soi et un objet d'une intuition non sensible, c'est-à-dire de l'entendement, c'est-à-dire qu'il doit y avoir une connaissance possible où ne se rencontre aucune sensibilité et qui a seule une réalité absolument objective, en ce sens que par elle les objets sont représentés *tels qu'il sont*, alors qu'au contraire, dans l'usage empirique de notre entendement, les choses ne nous sont connues que *comme elles apparaissent*. ».

seconde nature. Cette position intermédiaire comporte cependant ses propres problèmes : il ne va aucunement de soi que cette conception sociale d'une application correcte et objective des normes de signification évite l'approche dualiste de la métaphysique de Kant (platonisme rampant) et ne se réduise pas à une approche sociale de la normativité en termes de régularités (naturalisme brut) ou encore ne tombe pas dans l'écueil d'une circularité explicative. Nous tenterons donc de voir ici si la position de Davidson, que nous qualifions également de position du juste-milieu, est susceptible de résister à ces critiques.

En nous tournant à présent du côté de Davidson, nous laisserons néanmoins jusqu'aux prochains chapitres les questions relatives à une faible socialité. Davidson, tout comme McDowell, apporte sa contribution tant aux questions portant sur le sens et la portée que sur le fondement des normes de signification (§5.2). Nous qualifions ainsi Davidson de néo-Kantien en raison de sa conception du sens et de la portée de la contrainte normative de l'action linguistique. Davidson soutient que l'usage du langage est régi selon une contrainte normative, ne résultant donc pas simplement d'une action causale et souligne, de ce fait, que la contrainte portant sur l'interprétation et l'emploi du langage est plus que simplement pragmatique. En d'autres termes, il insiste fortement sur la distinction entre ce qui est et ce qui doit être. Par ailleurs, il défend la thèse selon laquelle cette notion d'une normativité forte concerne tant notre usage du vocabulaire logique que celui du vocabulaire non logique. Ce chapitre situe sa position sur la question du sens qu'il accorde aux normes de signification, à savoir que si le langage obéit bien à une contrainte normative forte, nous sommes amenés à nous interroger sur le sens d'une normativité forte et sur la portée significative d'une dimension constitutive de certaines normes de signification. Nous montrerons ainsi dans les premières considérations du prochain chapitre les conséquences d'une telle position (§5.3) quant à la question de la portée (§ 5.4) et du fondement des normes de signification (§5.5), alors que les chapitres VI, VII et VIII reviendront sur les problèmes entourant les approches sociales faibles des normes de signification jusqu'à cette question de leur fondement ; nous en viendrons alors à examiner une approche sociale faible (chap. VII) qui tentera de répondre à ces problèmes. Cependant, vu que Davidson s'inspire de Quine et que Quine est un naturaliste, notre classification comme néo-kantien est un peu litigieuse. C'est la raison pour laquelle une partie de ce chapitre est consacrée à justifier cette classification. En effet, bien que Davidson soit fortement influencé par Quine (§4.4.1), leur point majeur de divergence

(§4.4.2) se situe au niveau du naturalisme. Son monisme anomal confirme ainsi son anti-naturalisme en théorie de l'action et en philosophie de l'esprit. Nous nous intéressons ici à sa philosophie du langage et nous montrerons qu'à partir de là émerge un anti-naturalisme fort. Le principal point de divergence qui sépare alors Davidson de Quine, est l'accent mis sur le rôle constitutif de la charité au sens où cette présomption de rationalité doit être constitutive du comportement linguistique des êtres humains (§4.5 et §4.7). Son approche se définit donc clairement comme néo-Kantienne (par opposition au naturalisme causal) même s'il est indéniable que sa position partage certaines affinités avec Quine<sup>539</sup>. Le défi pour Davidson (§4.6) est d'arriver à concilier, tout en restant cohérent, d'un côté, son héritage quinién menant à un naturalisme éliminativiste ou à un normativisme pragmatique avec, d'un autre côté, sa reconnaissance et sa volonté de maintenir une normativité rationnelle constitutive d'inspiration Kantienne. Nous analyserons ainsi d'abord les développements de sa théorie

---

<sup>539</sup> Notre interprétation néo-Kantienne de Davidson se situe dans le prolongement de la lecture de Evnine en soulignant l'engagement de Davidson à l'égard d'un « idéal constitutif de rationalité » dans l'attribution de la croyance et de la signification, voir Evnine, 1991, et s'oppose ainsi à l'approche de Rorty, celui-ci voyant en Davidson un naturaliste post-quinién. Rorty, 1994b, définit ainsi la position de Davidson comme « a thoroughgoing anti-Kantian naturalism [...] I see the work of Donald Davidson as the best contemporary expression of the central strand of pragmatist thought, the naturalizing, Darwinizing strand ». Notre approche est cependant différente : en s'inspirant d'une distinction tirée de Gibson, 1988, p. 56, nous pensons que les réponses au travail de Quine se divisent en deux volets : une réponse radicale et une réponse conservatrice. L'option radicale, mise en valeur par Rorty, soutient que la rationalité normative devrait être éliminée et que le discours normatif doit être expliqué en termes darwiniens et évolutionnistes. Rorty avance l'idée selon laquelle l'usage du langage, la constitution de la croyance et de la signification ainsi que la formation des désirs, pourraient être parfaitement expliquées par le biais de notions relatives à la sélection naturelle. Rorty, 1994b, p.990-991, l'exprime comme suit : « Quine, Davidson and Putnam have...engaged in a project of naturalization. They attempt to naturalize semantics by getting rid of the notion of 'sameness of meaning', thereby enabling us to see linguistic ability as continuous with other human abilities. This naturalization culminates in Davidson's claim that we should 'erase the boundary between knowing a language and knowing our way around in the world generally'. Davidson suggests that we think of linguistic competence as a kind of know-how...This know-how is not interestingly different from the ability of people who work with animals to predict the animals' reactions to stimuli, and to adjust their own behaviour accordingly. This process of adjustment is not the sort of thing which has transcendental presuppositions... Davidsonians... see attainment of mutual understanding as like the mutual adjustment of two oak trees growing side by side, competing for light and space, and eventually compromising by bushing out on opposite sides... There is no more room for a moment of unconditionality than there is any other process of accommodation. ». L'option conservatrice, quant à elle, à laquelle nous pensons que Davidson souscrit, insiste sur le point de vue kantien selon lequel l'action humaine (en incluant l'usage du langage) est gouvernée par des normes de rationalité. Elle défend donc le caractère non éliminable d'une dimension fortement normative et non son remplacement en faveur de substituts naturalistes.

(§4.3) en partant de la considération de deux influences majeures : celles de Tarski (§4.2) et de Quine (§4.4).

#### 4.2. La théorie de la vérité de Tarski

Puisque quelque chose peut être vrai sans qu'on le sache, deux questions concernant la vérité devraient être distinguées : la question métaphysique et la question épistémologique.

La question métaphysique cherche à répondre à la question suivante : en quoi cela consiste-t-il pour un énoncé d'être vrai ? La question épistémologique, quant à elle, tente d'apporter un éclairage à la question suivante : en quoi cela consiste-t-il pour nous de tenir un énoncé pour vrai ? Tarski s'intéresse uniquement à la question métaphysique et propose, pour ce faire, d'attribuer une définition extensionnelle au prédicat « est vrai » pour un langage donné<sup>540</sup>. Une définition extensionnelle d'un terme fournit une liste de tous les objets auxquels le prédicat s'applique. Par exemple, une définition extensionnelle de « est rouge » définit le terme rouge en énumérant tous les objets de couleur rouge – c'est-à-dire les boîtes aux lettres anglaises, les camions de pompiers, etc. Une définition intensionnelle s'apparente, quant à elle, à la définition du dictionnaire et vise à remplacer un concept par un autre, ce dernier devant être plus transparent que le terme qu'on cherche à définir. Une définition intensionnelle de « est rouge », au lieu d'expliquer le terme rouge en énumérant tous les objets de couleur rouge, explique le concept de rouge à l'aide du concept de couleur, et le concept de couleur, lui-même, au moyen des notions de longueurs d'onde d'absorption et de réflexion des rayons lumineux etc. Les définitions extensionnelles ne substituent pas un concept à un autre concept (qui lui est synonyme); elles définissent un prédicat en faisant appel à tous les objets auxquels il s'applique. La définition extensionnelle de la vérité de Tarski définira alors la vérité en termes d'objets auxquels s'applique le prédicat « est vrai », objets que sont, selon Tarski, les phrases. Par conséquent, une définition extensionnelle de la vérité pour un langage donné fournit une méthode effective visant à déterminer ce que tout énoncé signifie (c'est-à-dire donne les conditions sous lesquelles il est vrai). Cette théorie de la vérité<sup>541</sup> a ainsi pour conséquence un ensemble d'axiomes qui énoncent les conditions de

---

<sup>540</sup> Tarski, 1944.

<sup>541</sup> La définition de la vérité ne fournit pas au sens strict une liste de phrases spécifiant les conditions de vérité de toutes les phrases d'un langage donné. Ce qu'il y a d'essentiel chez Tarski, c'est l'utilisation d'une méthode récursive pour définir la vérité; ce qui revient à définir la vérité pour un

vérité respectives de chaque phrase du langage. Ces axiomes sont de deux types : certains axiomes donnent les conditions sous lesquelles une suite d'objets satisfait une phrase complexe sur la base de conditions de satisfaction des phrases plus simples; les autres axiomes donnent les conditions de satisfaction des prédicats primitifs, et spécifient s'il y a lieu les dénnotations des noms propres primitifs.

Une telle définition ne vise pas à établir les significations d'expressions telles que le concept de « vérité » ou le prédicat « est vrai » par le biais d'autres expressions qui leur sont synonymes mais à fournir une liste de phrases auxquelles s'applique le prédicat « est vrai » dans le langage. En préconisant une définition extensionnelle sur une définition intensionnelle, Tarski cherche à éviter le danger de la régression à l'infini propre aux définitions intensionnelles : en effet, les définitions intensionnelles en substituant au concept à définir un autre concept exigent que celui-ci soit défini à son tour par un autre concept et ainsi de suite.

Tarski aborde la question métaphysique de la vérité en donnant une définition extensionnelle du prédicat « est vrai » pour tous les énoncés ou phrases d'une langue particulière. Autrement dit, pour Tarski les objets d'une théorie de la vérité sont les phrases d'un langage particulier – les véritables porteurs de vérité (*truth-bearers*) sont les phrases ou énoncés et notre définition porte sur un langage spécifique. Ces énoncés ou phrases dont il est question ici, Tarski insiste d'ailleurs là-dessus, sont des énoncés interprétés<sup>542</sup>, c'est-à-dire simplement doués de sens, d'un langage dont la structure est exactement spécifiée. Tarski est donc clair sur le fait que le problème de définir la vérité pour un langage n'a de sens que pour un langage interprété<sup>543</sup>. Pour caractériser sa position, Tarski renvoie au « formalisme

---

domaine infini (les phrases possibles du langage) sur la base d'un nombre fini d'axiomes. Pour plus de précisions, voir n. 546 ci-après. La définition de Tarski a ainsi la forme suivante: pour tout  $s$ ,  $s$  est vrai si et seulement si... (Une clause qui définit récursivement la vérité en termes de la notion de satisfaction pour un nombre fini de relations de satisfaction obtenues entre les expressions et (séquences) d'objets).

<sup>542</sup> Quand nous parlons ici d'énoncé interprété, ce n'est pas au sens moderne, de modèle-théorique, d'énoncé formel auquel a été donné une « interprétation » ensembliste, mais un sens plus relâché d'énoncé doué de sens. En termes modernes, on pourrait définir un langage interprété simplement comme un couple  $(L, M)$  où  $L$  est un langage formel et  $M$  une  $L$ -structure ; mais la notion de vérité dans une structure d'interprétation n'apparaîtra que plus tard, dans Tarski et Vaught, 1957.

<sup>543</sup> Cf. Tarski, 1983, p.166 : « It remains perhaps to add that we are not interested here in 'formal' languages and sciences in one special sense of the word 'formal', namely sciences to the signs and expressions of which no meaning is attached. For such sciences the problem here discussed has no

intuitionniste » de Lesniewski, expression sans doute malheureuse par laquelle il décrivait sa propre position théorique de prendre comme constituants des théories des énoncés formalisés et interprétés. Lesniewski écrivait :

« Having no predilection for 'various mathematical games' that consist in writing out according to one or another conventional one rule various more or less picturesque formulae which need not be meaningful or even – as some of the 'mathematical gamers' might prefer – which should necessarily be meaningless, I would not have taken the trouble to systematize and to often check quite scrupulously the directives of my system, had I not imputed to its theses a certain specific and completely determined sense, in virtue of which its axioms, definitions and final directives [...] have for me an irresistible intuitive validity.<sup>544</sup> ».

Un même énoncé, considéré du point de vue syntaxique, pouvant avoir des significations différentes dans différents langages, la définition de la vérité des énoncés sera donnée relativement à un langage interprété. Par conséquent, une définition dans le style de Tarski<sup>545</sup> est une définition de la vérité dans  $L$  (où  $L$  est un langage donné dans lequel est formulée la théorie) spécifiant l'extension de « est vrai-en- $L$  » pour toutes les phrases de  $L$ . Il ne semble donc pas approprié de parler d'une définition de la vérité dans le style de Tarski puisque chaque définition de la vérité dans le style de Tarski est établie relativement à un langage donné et il existe autant de définitions différentes de la vérité dans le style de Tarski qu'il y a de langages différents. Tarski est en effet bien conscient que la même forme verbale peut être, ou pourrait être, suivant les langages auxquels elle appartient, ici vraie, là fausse, ailleurs dénuée de sens, bref comme la vérité dépend du sens, le prédicat linguistique de vérité (« vrai » appliqué à des énoncés) doit être relativisé à un langage. Cependant, dire que les définitions dans le style de Tarski sont construites relativement à un langage donné ne veut pas dire que, pour Tarski, la vérité soit relative c'est-à-dire que ce qui est vrai change d'un langage à l'autre et correspond à ce que nous pensons être vrai. Tarski ne relativise pas la vérité au sens d'une vérité vraie-pour-moi, vraie-pour-la langue française, vraie-pour-la

---

relevance, it is not even meaningful. We shall always ascribe quite concrete and, for us, intelligible meanings to the signs which occur in the languages we shall consider. ».

<sup>544</sup> Lesniewski, 1929, p. 487; cité dans Wolenski, 2009, p.49.

<sup>545</sup> Nous utilisons l'expression « dans le style de Tarski » car les exemples que nous employons sont des définitions tirées des langues naturelles et Tarski pense que, les contraintes qui pèsent sur un usage correct, et à plus forte raison sur une éventuelle définition du prédicat de vérité, sont telles qu'elles obligeraient à un nettoyage et à une mise en ordre du langage ordinaire à un degré tel qu'il perdrait, justement, son caractère de langage naturel.

langue allemande, etc. Il cherche au contraire à aboutir à une conception claire et non ambiguë de la notion de vérité : il est crucial de noter que ce projet ne vise pas une définition simplement stipulative d'un nouveau concept sous un ancien vocable mais une définition qui capture la signification du terme « vrai », et plus spécialement celle qu'il a dans la conception classique de la vérité comme correspondance avec la réalité et que Tarski appellera, dans la version qu'il en donne, la conception sémantique de la vérité. Cette conception classique de la vérité, Tarski en identifie une expression particulièrement claire chez Aristote :

En effet, dire « l'étant n'est pas » ou « le non-étant est » est faux, par contre dire « l'étant est » ou « le non-étant n'est pas » est vrai<sup>546</sup>.

C'est cette notion, et celle-là seulement, qu'il s'agit de définir rigoureusement. Lorsque l'on se propose de définir une expression ayant déjà une signification, il faut se donner un critère de succès de cette définition, un critère qui nous garantisse, s'il est satisfait, que la définition saisit bien la notion visée. Pour parvenir à ce critère, Tarski part de l'analyse conceptuelle de la notion classique de vérité, la notion en jeu dans la définition aristotélicienne, qu'il a apprise de Kotarbinski<sup>547</sup>. Cette notion de vérité traduit ainsi l'idée selon laquelle dire d'un énoncé qu'il est vrai dépend de la manière dont le monde est, et ce qu'est le monde est indépendant de ce que nous pensons, assertons, croyons, etc. Affirmer que les définitions dans le style de Tarski sont relatives à un langage ne revient alors pas à dire que la vérité est elle-même relative – que ce qui est vrai dans un langage donné peut ne pas être vrai dans un autre langage – mais vise seulement à souligner que ce qu'exprime une vérité est relative – que ce qu'exprime la vérité dans un langage donné (au moyen d'une phrase particulière) ne correspond pas nécessairement à ce qu'exprime la même vérité dans un autre langage. Dans la mesure où, pour Tarski, les phrases sont les véritables porteurs de vérité (ce sont seulement les phrases d'un langage particulier qui sont vraies), et où les phrases diffèrent d'un langage à l'autre, la vérité doit être définie différemment relativement au langage donné. C'est dans ce sens qu'il faut entendre dans les définitions de Tarski l'expression « relativement à un langage » (« language relative »).

---

<sup>546</sup> Aristote, 1989, 1011b26; trad. Cassin/Narcy ; cette traduction est préférable à celle de Tricot qui est citée dans Tarski 1972, p.162, n. 2 : « Dire de l'Être qu'il n'est pas, ou du Non-Être qu'il est, c'est le faux ; dire de l'Être qu'il est, et du Non-Être qu'il n'est pas, c'est le vrai. ».

<sup>547</sup> Sur l'état des recherches sur la vérité dans l'École Polonaise avant Tarski, voir Wolenski, 2009.

Tarski<sup>548</sup> soutient explicitement (ce qui n'est pas le cas de Davidson qui cherche précisément à élargir la méthode de Tarski aux langues naturelles) qu'il s'agit avant tout, selon lui, d'élaborer une définition de la vérité opératoire pour les langues formelles. Les langues formelles sont des langues dont les référents des termes singuliers et les extensions des prédicats sont déterminés par *stipulation*<sup>549</sup> : la grammaire du langage doit ainsi être exactement spécifiée, c'est-à-dire la classe des formes syntaxiques qui ont une signification dans le langage en question, en particulier la classe des termes, prédicats, formules et des énoncés. Le référent d'un terme n'est alors pas sujet à changement dans la mesure où ce à quoi il fait référence n'est pas découvert mais créé par nous<sup>550</sup> : ceci exclut donc les langages contenant des termes indexicaux ou des déictiques et plus généralement des expressions dont la signification varie avec le contexte. Pour ces langages formalisés, les définitions de la vérité dans le style de Tarski sont directement données<sup>551</sup> relativement à un langage interprété

---

<sup>548</sup> Tarski, 1983, p. 267, écrit : « In my opinion, the considerations of §1 prove emphatically that the concept of truth (as well as other semantical concepts) when applied to colloquial language in conjunction with the normal laws of logic leads inevitably to confusions and contradictions. Whoever wishes, in spite of all difficulties, to pursue the semantics of colloquial language with the help of exact methods will be driven first to undertake the thankless task of a reform of this language. He will find it necessary to define its structure to overcome the ambiguity of the terms which occur in it, and finally to split the language into a series of languages of greater and greater extent, each of which stands in the same relation to the next in which a formalized language stands to its metalanguage. It may, however, be doubted whether the language of everyday life, after being 'rationalized' in this way, would still preserve its naturalness and whether it would not rather take on the characteristic features of the formalized languages. ».

<sup>549</sup> Dans un langage formel, la vérité est toujours fonction des critères à satisfaire, lesquels diffèrent d'un système à un autre. Ladrière, 1967, p. 313, relève cette relativité en ces termes : « les expressions qui figurent dans un système formel n'ont d'autre sens que celui qui résulte des possibilités opératoires stipulées dans les règles de maniement. ».

<sup>550</sup> Les langues naturelles ne fonctionnent pas ainsi. On pensait auparavant que le référent de « Vénus » était une étoile alors qu'on lui associe maintenant une planète, les extensions de « est une planète » et « est une étoile » ont changé et celles-ci peuvent continuer à varier au grès de l'évolution de notre connaissance empirique. Les langues naturelles diffèrent des langues formelles dans la mesure où la signification n'est pas, dans les langues naturelles, à l'abri d'éventuelles révisions. Ce qui est précisément le point soulevé par Quine dans son approche holistique de la relation entre signification et expérience menant à son rejet de la distinction de l'analytique et du synthétique, à savoir : qu'il n'existe pas de réelle distinction entre les phrases exprimant la signification des termes et celles rendant compte de l'état des choses dans le monde. Toutes deux ont le même statut : les significations ne sont pas déterminées de manière stipulative une fois pour toutes mais évoluent au gré des fluctuations de l'information empirique.

<sup>551</sup> Vu l'évidente simplicité des langages que nous aborderons, nous n'aurons pas besoin de faire appel à l'ingéniosité technique de la théorie de Tarski pour définir la vérité relativement à ces langages-ci. Nous nous concentrons moins sur la performance formelle et technique de la théorie de Tarski que sur sa portée philosophique. Cependant pour ne pas donner l'impression de survoler la théorie de Tarski,

– puisque les référents des termes singuliers et les extensions des prédicats sont supposés préétablis, la vérité étant définie comme la propriété qu'une phrase a, en vertu d'un état de choses qui lui correspond dans lequel le référent d'un terme singulier dans la phrase a la propriété qui lui est attribuée par le prédicat de la phrase. Autrement dit, pour Tarski, la définition de la vérité pour un langage particulier résulte de la détermination des référents de tous les termes singuliers et des extensions de tous les prédicats du langage. Une fois qu'est établi l'objet qui est dénoté par un terme singulier et les objets auxquels s'applique un

---

nous ferons quelques remarques sur ses réalisations techniques. La contribution majeure de Tarski est d'avoir établie une théorie compositionnelle de la vérité pour les langages aussi formellement complexes que la logique des prédicats quantifiée, autrement dit la logique des prédicats quantifiée est pour Tarski ce que les tables de vérité sont pour la logique propositionnelle. À partir d'un nombre fini d'axiomes et d'un nombre fini de règles s'appliquant aux connecteurs vérifonctionnels (établis par la logique propositionnelle), la théorie de Tarski définit ainsi la vérité pour un nombre infini de phrases, en incluant également les phrases faisant intervenir les termes de généralisation universelle et existentielle tels que « tous » et « quelque ». En d'autres termes, sa théorie montre comment la vérité des phrases est construite à partir des significations des parties composantes ou sub-phrastiques : noms propres, prédicats et quantificateurs. Dans la mesure où les unités syntaxiques plus petites que la phrase ne peuvent pas être vraies ou fausses (puisque seules les phrases sont les véritables porteurs de vérité), et que dans un langage intégrant des quantificateurs et des variables, les phrases sont formées par la combinaison d'expressions (phrases ouvertes et quantificateurs), aucune de ces parties composantes ne peut être porteuse de valeurs de vérité (puisque ce ne sont pas des phrases). Tarski se heurte au problème de montrer comment la valeur de vérité d'une phrase complexe est fonction de la « valeur de vérité » de ses parties composantes alors que ces mêmes parties composantes n'ont pas de véritables valeurs de vérité. Tarski doit, par conséquent, trouver une propriété sémantique appropriée à ces objets sub-phrastiques-ci telle que celle-ci soit comprise comme une contribution régulière et uniforme aux conditions de vérité des phrases. Il introduira alors la notion *auxiliaire* de *satisfaction*. Cette notion de satisfaction (et indirectement celle de vérité) est définie comme une relation entre une phrase ouverte et les objets ou séquences d'objets qui la rendent vraie (la référence d'un prédicat est ainsi définie comme la classe de ces entités qui la satisfont). Ainsi, une phrase du langage-objet est vraie si et seulement si elle est satisfaite par toute séquence d'objets sur lesquels portent les variables de quantification du langage-objet. En cherchant à expliquer la vérité en termes de satisfaction et en appliquant celle-ci aux termes singuliers, prédicats, phrases ouvertes et quantificateurs, Tarski peut à présent proposer une conception de la vérité en termes de satisfaction pour toutes les phrases. La notion de satisfaction est donc une généralisation de celle de vérité qui définit récursivement l'ensemble des énoncés d'un langage donné. La définition récursive permet alors de définir un prédicat dont l'extension est infinie en termes d'un nombre fini d'axiomes et de règles procédurales. Par conséquent, une définition de la vérité dans le style de Tarski peut générer un nombre infini de théorèmes à partir d'une axiomatisation finie. Ces notions de satisfaction et de récursivité visent ainsi à faire le lien entre théorie de la signification et théorie de la vérité. Elles répondent respectivement aux deux critères centraux d'adéquation que doit satisfaire une théorie de la vérité (ou de la signification) relativement à un langage donné : le critère de compositionnalité et le critère d'une application infinie. Une théorie de la signification, devant représenter comment une capacité infinie procède par des moyens finis, doit ainsi être récursive pour répondre à l'exigence de compositionnalité. Ce sont ces éléments que Davidson reprendra et exploitera pour construire sa théorie compositionnelle de la signification pour le langage naturel.

prédicat, on peut donner une définition de vérité pour une phrase (et ainsi de suite pour toutes les phrases auxquelles s'appliquent tous les termes singuliers et prédicats du langage). Par exemple, prenons un langage très simple  $L_1$ , composé de deux énoncés de base « Bob est américain » et « La neige est blanche »<sup>552</sup>. Un tel langage a deux noms et deux prédicats. Supposons que les référents des noms et que les extensions des prédicats sont stipulés comme suit :

- « Bob » réfère à Bob,
- « Neige » réfère à neige,
- « est américain » s'applique à tous les sujets de nationalité américaine,
- « est blanche » s'applique à tous les objets de couleur blanche.

Une définition de la vérité dans le style de Tarski pour ce langage peut être maintenant formulée comme suit :

(s) [s est vrai-en- $L_1$

ssi ( $s$  = « Bob est américain » et Bob est américain)

ou ( $s$  = « La neige est blanche » et la neige est blanche)].

Afin d'examiner la définition de Tarski, nous devons au préalable distinguer entre le langage-objet et le métalangage. Si nous appelons *langage-objet* le langage auquel appartient l'énoncé dont la vérité est définie, et *métalangage* le langage dans lequel la vérité est définie, les explications de ce en quoi consiste la vérité pour un énoncé donné du langage-objet sont fournies par les instances du schéma-T, si on peut en déduire tous les énoncés de forme :

Schéma (T)  $X$  est vrai-en- $L_1$  si, et seulement si,  $p$

Où «  $X$  » est mis pour le nom d'un énoncé donné du langage-objet, et la lettre «  $p$  » pour une traduction de cet énoncé dans le métalangage<sup>553</sup> dans lequel la définition est formulée (et par rapport auquel  $L_1$  est le langage-objet). « Tous » les énoncés, cela veut dire : un pour chaque énoncé du langage-objet  $L_1$ . Une phrase-T est donc l'une des phrases qui définit la vérité pour un énoncé donné du langage-objet et « 'Bob est américain' est vrai-en- $L_1$  ssi Bob est américain » est l'une des deux phrases-T de ce langage. Les phrases-T sont des phrases du

<sup>552</sup> Nous serons amenés à faire référence à un autre langage artificiel que nous nommerons  $L_2$  et pour ce faire, nous l'introduisons ici. À l'image de  $L_1$ , nous considérons le langage  $L_2$  comme un langage simple, composé de deux énoncés de base, à savoir : « La chaise est noire » et « le gazon est vert ».

<sup>553</sup> Remarquons que les instances de ce schéma, les équivalences-T individuelles, sont des énoncés du (méta) langage dans lequel est définie la vérité, tandis que le schéma-T lui-même n'en fait pas forcément partie.

métalangage composées d'une description structurale (noms descriptifs-structuraux de tous les énoncés du langage-objet étudié), c'est-à-dire d'une liste non interprétée de symboles, faisant donc appel à un objet purement syntaxique : par exemple, le nom structural de l'énoncé « Bob est américain » est composé de la lettre b, puis de o et de b, d'un espace, de e, etc. Mais la manière la plus simple de désigner cet énoncé consiste à le mettre entre guillemets : « Bob est américain », les guillemets créant conventionnellement le nom d'un objet linguistique. Ceci est suivi par un prédicat du métalangage « est vrai si et seulement si » et ensuite d'une traduction de la phrase du langage-objet dans le métalangage, c'est-à-dire d'une phrase du métalangage qui donne la signification de la liste syntaxique située à gauche. Une théorie adéquate doit donc avoir pour théorèmes des phrases-T qui font apparaître à droite de l'équivalence « les bonnes traductions » des énoncés sur lesquels nous avons une description structurale à gauche. Les phrases-T ne peuvent ainsi construire une définition extensionnelle du prédicat de vérité qu'à partir de phrases déjà interprétées ; autrement dit, elles définissent la vérité en présupposant la signification.

Néanmoins, il ne faudrait pas confondre les phrases-T (et le schéma qu'elles instancient, la convention T) avec le schéma décitationnel avec lequel elles semblent partager une structure similaire. Tarski admettait certes, pour différentes raisons, que l'usage des guillemets du côté gauche de l'équivalence, permettait de formuler un nom de l'énoncé mis entre guillemets. L'énoncé lui-même est donc cité, ou mentionné, plutôt qu'utilisé. Quine a transformé cette construction des guillemets en analyse de l'adjectif « vrai » : puisque citer l'énoncé pour dire qu'il est vrai, revient à affirmer cet énoncé lui-même, la fonction du prédicat de vérité est simplement décitationnelle : il annule la référence linguistique opérée du côté gauche des équivalences. D'où la thèse célèbre de Quine<sup>554</sup>, à savoir : « le prédicat de vérité est un instrument de décitation ». Pourtant, la définition de la vérité de Tarski ne devrait pas être comprise comme une conception décitationnelle de la vérité<sup>555</sup>. Le schéma décitationnel est donc purement syntaxique et implique seulement de citer et de déciter ; la signification d'une phrase citée ne joue aucun rôle dans les transformations réalisées par la

<sup>554</sup> Quine, 1975, p.24-25.

<sup>555</sup> Davidson, 1990a, p.285, souligne la différence entre la convention T de Tarski et les conceptions décitationnelles de la vérité. Rivenc, 2004, fait remarquer qu'il est certainement abusif d'imputer ce déflationnisme à Tarski, comme le fait par exemple Soames, 1999, qui parle de la conception sémantique de la vérité de Tarski comme de la « plus fameuse et influente version du déflationnisme ».

décitation, et la signification de « est vrai » n'est pas donnée au moyen de cette transformation. Une transformation syntaxique est donc exprimée par ce procédé de décitation. On pourrait penser qu'il est donc inutile; en fait, il a cependant son utilité, une fonction purement logique, celle de nous permettre de généraliser sur une classe infinie (ou non connue en détail) d'énoncés mais, il n'explique cependant pas la signification de la vérité. Ces deux aspects ont été baptisés par Gupta, respectivement, *thèse de décitation*, et *thèse de généralisation*<sup>556</sup>. Si, encouragé par l'évidente trivialité d'un énoncé comme:

« La neige est blanche » est vrai si et seulement si la neige est blanche,

on poursuit (faussant à présent compagnie à Quine) en ajoutant qu'un tel énoncé est analytiquement vrai, définitionnel du mot « vrai », « vrai en vertu du sens de 'vrai' », ou encore « épuise » la signification [*meaning*] du concept de vérité, on a tous les ingrédients d'une forme de déflationnisme<sup>557</sup>. Or, il semble que l'impression correcte sur quoi repose la plausibilité de cette forme de déflationnisme – le déflationnisme linguistique – réfute l'hypothèse qui la fonde, à savoir l'interprétation purement citationnelle des guillemets. Les équivalences T sont incontestablement triviales, et on doit pouvoir justifier l'idée qu'elles sont analytiques. Mais cela ne veut nullement dire qu'elles sont vraies « en vertu du sens du mot 'vrai' seulement », ou analytiques du concept de vérité<sup>558</sup>. Elles sont vraies « en vertu du sens » des mots qu'elles contiennent, et non seulement en vertu du sens du mot « vrai ». Pour s'en assurer, il suffit de passer des équivalences homophoniques couramment prises comme exemples, à la vérité dite *hétérophonique*:

« Schnee ist weiss » est vrai (en allemand) si et seulement si la neige est blanche,

qui, à supposer que cette forme soit attestée en français, n'est certainement pas analytique, vrai en vertu du sens de « vrai », etc. Il est possible que ce genre d'énoncé délivre une information substantielle sur la sémantique de la langue allemande (c'est le genre de « théorie

<sup>556</sup> Gupta, 1993.

<sup>557</sup> Quine, 1975 : le terme « déflationniste » a une portée plus large que le terme « décitationnel » : une conception décitationnelle de la vérité est un genre parmi d'autres de théorie déflationniste. Si on propose une théorie décitationnelle de la vérité alors on propose aussi une théorie déflationniste, cependant toutes les théories déflationnistes de la vérité ne sont pas nécessairement décitationnelles.

<sup>558</sup> Gupta, 1993, fait remarquer qu'il y a là un saut injustifié, justement du sens commun au déflationnisme.

de la vérité » que propose Davidson). Mais ce fait montre *a contrario* que dans la version homophonique<sup>559</sup>:

« La neige est blanche » est vrai si et seulement si la neige est blanche,

la signification de l'énoncé entre guillemets est active, saisie par tout locuteur compétent du français, et que « vrai en vertu du sens » doit être compris comme « vrai en vertu du sens de tous les mots figurant dans l'équivalence ». Mais c'est dire que l'expression, guillemets compris :

« La neige est blanche »

n'est pas qu'un nom propre de l'expression mise entre guillemets, au sens où cette dernière ne figurerait que de manière accidentelle, logiquement non pertinente, dans l'expression qui la mentionne. Là encore, il suffit d'opposer la citation *hétéronyme*:

L'énoncé le plus souvent pris comme exemple par Tarski est vrai si et seulement si la neige est blanche,

à la citation utilisée plus haut. Dans ce dernier cas, l'expression est mentionnée *et* utilisée à la fois. L'analyticité des équivalences T (quand elles sont analytiques!) provient du fait que les énoncés ne sont pas simplement cités, mais utilisés, quoique de manière déviante. De l'oubli de ce point, provient l'idée que ces équivalences sont vraies seulement en vertu du sens de « vrai », ou définitionnelles de la vérité. Mais ce même point justifie la conclusion suivante: si la mise entre guillemets n'est pas purement citationnelle, il n'y a aucune raison de penser que le prédicat de vérité est simplement un instrument de décitation. La thèse de décitation reposerait donc sur une mauvaise analyse de la citation.

Ainsi, pour un partisan du caractère purement décitationnel de la vérité, la vérité n'est absolument pas une propriété substantielle et le terme « vrai » n'est pas un authentique prédicat, c'est-à-dire qu'il n'exprime pas une authentique propriété des phrases ou des propositions, mais est seulement un mot logique, qui a les propriétés de décitation. Cependant, pour Tarski, la vérité est bien une propriété substantielle et les phrases-T sont bien appelées à donner une définition extensionnelle et adéquate du prédicat de vérité. Tarski restreint ainsi ses définitions de la vérité aux langages formels; il pense qu'une définition extensionnelle rigoureuse ne peut pas être donnée pour un langage naturel et que la notion

---

<sup>559</sup> Une règle de traduction homophonique établit la signification de la phrase de gauche dans le même langage que la phrase de droite.

ordinaire de vérité n'est pas définissable. L'universalité du langage ordinaire le rendrait inconsistant. Tarski écrit à cet effet :

« If we are to maintain this universality of everyday language in connexion with semantical investigations, we must, to be consistent, admit into the language, in addition to its sentences and other expressions, also the names of these sentences and expressions, and sentences containing these names, as well as such semantic expressions as 'true sentence', 'name', 'denote', etc. But it is presumably just this universality of everyday language which is the primary source of all semantical antinomies, like the antinomies of the liar or of heterological words. These antinomies seem to provide a proof that every language which is universal in the above sense, and for which the normal laws of logic hold, must be inconsistent.<sup>560</sup>».

On proteste souvent contre cette affirmation en faisant remarquer que ce sont à proprement parler des théories, plutôt que des langages, qui peuvent être dits consistants ou inconsistants. Tarski lui-même est revenu par la suite sur cette affirmation. L'une des raisons de son doute est que le langage ordinaire n'a pas une structure logique nettement spécifiée<sup>561</sup>; mais la raison la plus profonde réside dans le lien entre inconsistance et clôture sémantique : il n'y a peut-être pas de sens à se demander si oui ou non le langage est sémantiquement clos<sup>562</sup>. Autrement dit, les langages naturels tendent à effacer la distinction entre langage-objet et métalangage dans la mesure où ils contiennent eux-mêmes (c.-à-d. les langages-objet) des expressions appartenant au métalangage, celles-ci rendant possible la construction de phrases comme :

(1) c n'est pas un énoncé vrai.

Abrégeons la description « l'énoncé de cette page numéroté (1) » par la lettre « c » ; en vertu du sens ainsi conféré à la lettre « c », il est vrai que :

<sup>560</sup> Tarski, 1935, 1983, p. 164-165, Tarski a émis par la suite un diagnostic plus nuancé.

<sup>561</sup> Tarski, 1944, p.349 est clair sur ce point et corrige Tarski, 1935, 1983, qui attribuait le problème au caractère universel du langage ordinaire (Cf. Tarski, 1983, p.160 et Rouilhan, 1984, 1998 pour une description de ce changement) : «... Our everyday language is certainly not one with an exactly specified structure. We do not know precisely which expressions are sentences, and we know even to a smaller degree which sentences are to be taken as assertible. ».

<sup>562</sup> La notion de clôture sémantique provient d'une certaine reconstruction du paradoxe du menteur que Tarski attribue à Lukasiewicz. Cf. Tarski, 1944, §8, p.348 : « (I) We have implicitly assumed that the language in which the antinomy is constructed contains, in addition to its expressions, also the names of these expressions, as well as semantic terms such as the term 'true' referring to sentences of this language; we have also assumed that all sentences which determine the adequate usage of this term can be asserted in the language. A language with these properties will be called 'semantically closed.' ». Il faut insister sur ce point : la présence de toutes les instances du schéma d'équivalence (les « énoncés-T » de Davidson) est plus un réquisit d'adéquation qu'un fait du langage ordinaire.

(2) *c* est identique à « *c* n'est pas un énoncé vrai ».

et supposons maintenant que nous puissions donner en français une définition adéquate de vrai (-en-français). À partir de cette définition, une instance de (T) obtenue en substituant à « *p* » l'énoncé (1) est l'énoncé (noté (b) par Tarski<sup>563</sup>) :

« *c* n'est pas un énoncé vrai » est vrai si et seulement si *c* n'est pas un énoncé vrai.

Et une instance de (T), obtenue toujours en substituant à « *p* » l'énoncé (1), et à « *X* » le nom « *c* » de l'énoncé (1), ce qui paraît permis au vu de (2), donne finalement :

(3) *c* est un énoncé vrai si et seulement si *c* n'est pas un énoncé vrai.

qui est une contradiction patente. De quelles hypothèses la dérivation de cette antinomie dépend-elle crucialement ? Tarski en identifie trois :

1. avoir travaillé avec un langage sémantiquement clos, par quoi Tarski entend un langage possédant
  - a) un nom de chacune de ses expressions,
  - b) un prédicat de vérité pour ses énoncés,
  - c) et tel que toutes les équivalences-T qui déterminent l'usage adéquat de « vrai » y sont assertables.
2. avoir utilisé les lois de la logique classique
3. avoir fait une hypothèse empirique relativement à la dénotation d'une certaine expression, nommément «  $\lambda$  ».

Tarski remarque que l'hypothèse empirique selon laquelle *c* dénote l'énoncé « *c* n'est pas vrai » est en fait éliminable, c'est-à-dire qu'il est possible de reconduire le paradoxe sans elle. La preuve repose également sur l'usage des lois de la logique classique. Mais renoncer à ces lois est un prix que Tarski n'est pas prêt à payer.

Parce que, comme le note Tarski, l'hypothèse (3) *Ici l'hypothèse empirique est que c dénote l'énoncé « c n'est pas vrai »*, est en fait éliminable, il n'y a pas de raison de s'y attarder. Par ailleurs, Tarski n'est pas prêt à abandonner les lois de la logique classique, ce qui élimine (2). Reste donc (1): le langage sémantiquement clos. Tarski conclut:

« Accordingly, we decide *not to use any language which is semantically closed* in the sense given<sup>564</sup>. ».

<sup>563</sup> Tarski, 1983, p.158: « The source of this contradiction is easily revealed: in order to construct the assertion (b) we have substituted for the symbol 'p' in the scheme (T) an expression which itself contains the term 'true sentence'. ».

<sup>564</sup> Tarski, 1944, p.349.

Les langages (classiques) sémantiquement étant inconsistants, un langage (classique) ne peut pas contenir un prédicat adéquat de vérité pour lui-même, et par conséquent, *a fortiori*, il n'est pas possible de définir adéquatement la vérité pour un langage *L* dans *L* lui-même, puisqu'alors *L* devrait contenir les moyens de décrire ses expressions, un prédicat de vérité, et permettre de dériver tous les énoncés qui « déterminent l'usage adéquat du terme 'vrai' » c'est-à-dire toutes les équivalences-T, autrement dit être sémantiquement clos. Par conséquent, il est en fait nécessaire, pour que la définition de la vérité pour *L* soit possible, de conduire cette définition dans un métalangage distinct du langage-objet.

Nous avons vu que l'instanciation du schéma-T par l'énoncé du Menteur donnerait lieu à une contradiction. Et que par conséquent un langage ne pouvait pas même contenir de prédicat de vérité adéquat pour lui-même. Et l'inconsistance des langages sémantiquement clos nous force à les rejeter en tant qu'objets de l'entreprise sémantique. En ce point, on a envie de protester : pourquoi ne pas explorer une autre voie et maintenir le premier aspect de la clôture sémantique, si naturel, quitte à traiter autrement les équivalences-T (les limiter, les interpréter autrement)? C'est qu'ici intervient une seconde considération, qui touche à la nature du critère d'adéquation matérielle de toute définition du prédicat de vérité.

Tarski ne veut pas perdre l'intuition liée au concept classique de vérité, selon laquelle un énoncé est vrai si et seulement si l'état de chose qu'il décrit est réalisé. Mais c'est précisément le contenu des équivalences-T. Donc, toute théorie adéquate de la vérité doit permettre de retrouver les équivalences-T relatives aux énoncés de ce langage, cela pour régler l'usage du prédicat de vérité; et si de plus une définition est concevable, elle doit avoir pour conséquences toutes les équivalences analogues, où figurera, à la place du mot « vrai », le prédicat défini. On reconnaît la Convention T, que Tarski présente ainsi:

« Now at last we are able to put into a precise form the conditions under which we will consider the usage and the definition of the term « *true* » as adequate from the material point of view: we wish to use the term « *true* » in such a way that all equivalences of the form (T) can be asserted, and we shall call a definition of truth « *adequate* » if all these equivalences follow from it.<sup>565</sup> ».

Après avoir montré que les équivalences-T constituent une explication complète de ce en quoi consiste la vérité de chaque énoncé pris individuellement, Tarski est donc en mesure de formuler le critère d'adéquation matérielle de tout usage de « vrai ». C'est à la

---

<sup>565</sup> Ibid., §4.

Convention-T<sup>566</sup> qu'il revient de le formuler : celle-ci précise ainsi que pour appliquer correctement le prédicat « est vrai » à un énoncé du langage donné, l'énoncé doit exprimer qu'une certaine expression linguistique a une certaine propriété, la vérité, si et seulement si un certain état de chose est réalisé. La Convention-T pose donc deux conditions d'adéquation sur l'application du prédicat « est vrai » à un énoncé, la première concerne la signification des termes qui composent l'énoncé, la seconde porte sur la manière dont le monde est :

« What convention T, and the trite sentences it declares to be true, like « 'Grass is green' spoken by an English speaker, is true if and only if grass is green », reveal is that the truth of an utterance depends on just two things: what the words as spoken mean, and how the world is arranged<sup>567</sup>. ».

Par conséquent, si la phrase-T répond à ces deux exigences-ci, alors elle répond bien à la condition d'adéquation établie par la Convention-T et fournit ainsi une définition adéquate de la vérité pour la phrase du langage-objet. La Convention-T généralise ensuite cette exigence afin d'imposer le critère d'adéquation de la définition de la vérité à tous les énoncés du langage : elle soutient qu'une théorie ou qu'une définition de la vérité pour un langage donné est adéquate à partir du moment où elle spécifie pour tous les énoncés du langage qu'une certaine expression linguistique a une certaine propriété, la vérité, si un certain état de chose est réalisé. Autrement dit, une théorie adéquate de la vérité implique une phrase (c.-à-d. une phrase-T) qui instancie cette forme (c.-à-d. la convention T) pour toutes les phrases du langage-objet. Les phrases-T telles que « 'Bob est américain' est vraie-en-L<sub>1</sub> si et seulement si Bob est américain », « 'La neige est blanche' est vraie-en-L<sub>1</sub> si et seulement si la neige est blanche », « 'Cette chaise est noire' est vraie-en-L<sub>2</sub> si et seulement si cette chaise est noire », « 'Le gazon est vert' est vraie-en-L<sub>2</sub> si et seulement si le gazon est vert », etc., satisfont les deux exigences ci-dessus et expliquent donc, au moyen de la Convention-T, ce en quoi consiste la vérité pour ces énoncés-ci du langage-objet, c'est-à-dire que ces phrases-T définissent la vérité de chaque énoncé pris individuellement pour ces langages-objet particuliers. Il est facile de voir qu'elles instancient ainsi la forme générale : *s* est vraie-en-*L* ssi *p*, et la condition d'adéquation matérielle énonce qu'une définition adéquate de la vérité pour un langage doit imposer à tous les énoncés du langage donné la forme générale mise en valeur par ces instances du schéma-T ou équivalences-T individuelles. Davidson pose ainsi

<sup>566</sup> Tarski, 1983, p. 186-187.

<sup>567</sup> Davidson, 2001b, 1983, p.139.

la question suivante : les phrases-T nous disent-elles en quoi consiste la vérité pour ces phrases-ci du langage-objet, et s'il s'avère qu'elles définissent ce qu'est la vérité, qu'est-ce qui est présupposé par ces phrases-T pour qu'elles puissent arriver à formuler une définition adéquate de la vérité ?

Davidson souligne que les phrases-T (ou encore les équivalences-T) arrivent à une définition adéquate de la vérité seulement sur la base de certaines hypothèses<sup>568</sup>. Les phrases-T rendent bien compte du sens usuel du mot « vrai », ce qui peut être tenu pour une « définition partielle de la vérité ». Mais qu'il y ait une telle traduction, et que nous la connaissions, suppose évidemment que le langage-objet soit déjà compris par nous. Donc il semble que la Convention-T suppose, pour pouvoir fonctionner, comme critère d'adéquation, que nous comprenions le langage-objet antérieurement à l'élaboration de la théorie<sup>569</sup>. Les phrases-T se basent ainsi sur deux hypothèses pour fournir une définition adéquate de la vérité :

1. Dire que « Bob est américain » signifie que Bob est américain et dire que « La neige est blanche » signifie que la neige est blanche, c'est-à-dire que ces énoncés signifient la même chose que ce que ces phrases-ci expriment en français.
2. Dire « est vrai » doit être appliqué lorsque l'état de choses décrit ou exprimé par l'énoncé est réalisé.

La Convention-T s'appuie donc sur deux hypothèses – la signification des phrases qui définissent la vérité pour un langage-objet particulier et une compréhension préthéorique (c.-à-d. antérieure à la définition de Tarski) de la vérité comme nous le mentionnons plus haut. Afin de démontrer que cette Convention-T repose bien sur ces deux hypothèses-ci, supposons que nous n'avons pas formulé de telles hypothèses et posons-nous la question suivante : qu'est-ce que le prédicat « vrai-en-L » cherche à définir et définit-il la vérité pour le langage en question ?

Si la Convention T ne partait pas de ces hypothèses alors elle serait équivalente au schéma dénotationnel. Or la question est de savoir si l'action de simplement citer et déciter pourrait définir la vérité pour les phrases situées du côté gauche de la phrase-T ? Si nous

<sup>568</sup> Davidson, 1990a, p.285-288. Le point que soulève Davidson dans ces pages est également mentionné par d'autres comme : Blackburn, 1984b, p.267 et Putnam, 1994, p.316-320.

<sup>569</sup> Tarski restreint, comme nous l'avons vu, ses définitions de la vérité aux langages formels, c'est-à-dire aux langages dont la signification des termes et des prédicats est fixée par stipulation.

renonçons à supposer au préalable la signification des phrases et si nous n'avons aucune idée de la corrélation à établir entre les phrases du langage-objet et celles du métalangage, alors le côté gauche se réduit juste à n'être qu'une liste de symboles dénuée de sens; « La neige est blanche » n'aurait pas plus de sens pour nous que cette suite de symboles « @#\$^&\* » et construire une définition de la vérité au sens de Tarski pour cette liste de symboles serait une tentative désespérée. Nous ne saurions effectivement comment procéder ; tant que nous ne savons pas ce que ces expressions signifient, nous ne pouvons déterminer leurs conditions de vérité :

« Tarski showed how to define a truth predicate for each of a number of well-behaved languages, but his definitions do not, of course, tell us what these predicates have in common. Put a little differently: he defined various predicates of the form « s is true<sub>L</sub> », each applicable to a single language, but he failed to define a predicate of the form « s is true in L » for variable « L » [...] The central difficulty [...] is due simply to the fact that Tarski's definitions give us no idea how to apply the concept to a new case, whether the new case is a new language or a word newly added to a language<sup>570</sup>. ».

Supposons que contrairement aux apparences, les phrases de L<sub>1</sub> ne sont pas des phrases formulées en Français mais en Franglais où le Franglais est une langue dans laquelle « est américain » et « est blanc » possèdent les mêmes extensions que les prédicats français « est allemand » et « est vert ». En formant le prédicat « vrai-en-Franglais » au moyen de la procédure de citation et de décitation, nous n'aurons pas défini la vérité pour le Franglais. Nous devons au préalable connaître les références et extensions de ces expressions-ci en Franglais avant de donner une définition extensionnelle adéquate de la vérité pour le Franglais. Cependant, même si nous connaissons la signification des phrases du langage-objet, nous ne savons toujours pas si le schéma de citation et de décitation appréhende bien la notion de vérité à moins que nous disposions déjà d'une compréhension de la vérité relativement à ce schéma. Autrement dit, tant que nous n'avons pas de définition d'une notion précise de vérité, nous ne pouvons absolument pas déterminer si le prédicat « est vrai-en-L », résultant de l'opération de citation et de décitation, est conforme à la vérité. Supposons ainsi que nous connaissons l'extension des termes en Franglais, et que nous suivons le schéma de citation et de décitation pour définir un prédicat particulier, « vrai-en-Franglais ». Les phrases-T en Franglais sont alors formulées comme suit : « 'La neige est

---

<sup>570</sup> Davidson, 1990a, p. 285, p.287.

blanche' est vraie-en-Franglais si et seulement si la neige est blanche » et « 'Bob est américain' est vraie-en-Franglais si et seulement si Bob est américain. ». Si nous ne disposons d'aucune compréhension préalable d'une définition de la vérité (au sens d'une définition adéquate), et que nous considérons que la vérité est juste la propriété que nous définissons extensionnellement, alors cette procédure définit la vérité. Cependant, ceux d'entre nous qui comprennent ce que signifie la vérité, diront que les phrases-T ne définissent pas la vérité pour le Franglais, ou si c'est la manière dont « vrai-en-Franglais » est défini alors celui-ci ne définit rien qui ressemble à une notion adéquate de vérité. C'est parce que nous comprenons au préalable cette notion de vérité et que nous savons que le prédicat « vrai-en-Franglais » ne correspond pas à cette définition de la notion de vérité, que nous pouvons affirmer qu'en formant des phrases-T telles qu'énoncées ci-dessus, nous n'avons pas donné une définition de la vérité pour le Franglais.

Par conséquent même si nous connaissons l'extension des termes en Franglais, nous sommes en mesure de dire que, le prédicat « vrai-en-Franglais » ne rend pas adéquatement compte de la notion de vérité, seulement si nous savons qu'une phrase a une certaine propriété, la vérité, si un certain état de choses est réalisé. C'est seulement parce que nous savons que les prédicats « vrai-en- $L_1$  », « vrai-en- $L_2$  » sont sensés rendre compte de cette notion préthéorique de vérité que nous comprenons déjà, que nous pouvons rejeter les définitions ci-dessus de « vrai-en-Franglais » visant à définir un usage adéquat cohérent du prédicat de vérité. Par conséquent, la conception extensionnelle de Tarski de vrai-en- $L$  pour un langage donné  $L$  n'explique pas en quoi la vérité consiste mais la présuppose. En effet, une fois que nous savons ce que la phrase signifie et en quoi consiste la vérité, nous connaissons alors ce que vrai-en- $L$  veut dire, et si vrai-en- $L$  est une définition de la vérité. Si  $L_1$  est un segment du Français et non du Franglais, alors les phrases-T ci-dessus définissent la vérité pour  $L_1$ , et ceci seulement parce que « vrai-en- $L_1$  » s'accorde avec la notion préthéorique de vérité (et nous comprenons ce point une fois que nous comprenons aussi la signification des termes, c'est-à-dire une fois que nous identifions qu'il s'agit du Français et non du Franglais). L'ordre explicatif s'énonce donc comme suit : c'est parce que nous comprenons ce qu'est la vérité que nous pouvons définir vrai-en- $L$  pour tout langage-objet  $L$ , et non parce que nous pouvons définir vrai-en- $L$  pour un langage-objet particulier donné  $L$ , que nous arrivons à démontrer un usage cohérent de cette notion de vérité. En d'autres

termes, la Convention-T définit la vérité pour un langage donné *L* en présupposant au préalable la signification des termes du langage et la signification de la vérité. La propriété du prédicat vrai-en-*L* de Tarski dépend de la vérité et non l'inverse. Ce résultat souligne les limites des définitions de Tarski. Son projet est avant tout d'expliquer la vérité par le prédicat vrai-en-*L* c'est-à-dire de donner une définition de la vérité au moyen de l'extension du prédicat vrai-en-*L* pour tout langage *L*. Cependant les définitions extensionnelles ne définissent pas notre notion préthéorique, translinguistique de vérité mais se contentent de la présupposer.

Néanmoins, aucun de ces points ne constitue une objection pour Tarski dans la mesure où il est bien conscient des limites de ses définitions extensionnelles; il reconnaît explicitement qu'elles devraient être restreintes aux langages particuliers et qu'elles s'appliquent seulement aux langages formels (et qu'elles présupposent donc la signification des phrases pour lequel ce prédicat est défini). Par conséquent, Davidson n'insiste pas sur les considérations ci-dessus et ne s'attarde pas sur les objections adressées aux définitions de Tarski comme certains (notamment Blackburn et Putnam) le préconisent<sup>571</sup>. Néanmoins, il pense que nous ne devons pas en rester aux limites que Tarski impose à sa théorie, et donc que Tarski ne saisit pas tout ce qui peut être dit de la vérité. Ainsi, l'idée selon laquelle les définitions de Tarski ne nous disent rien sur la notion de vérité translinguistique peut être envisagée de deux façons :

« As showing that he did not capture essential aspects of the concept of truth, or as showing that the concept of truth is not as deep and interesting as many have thought<sup>572</sup>. ».

Davidson tire la première conclusion, les déflationnistes la seconde. Autrement dit, bien que la Convention T doive être distinguée du schéma déflationnel, les déflationnistes s'approprient néanmoins le travail de Tarski. Ses définitions excluent l'élément intensionnel du concept de vérité. La vérité n'a pas de signification profonde, ce qui peut être repris par les

<sup>571</sup> En effet, Davidson, 1990a, p.285-286, est surpris que certains aient considéré ces points comme des objections adressées à Tarski : « Tarski made this [the limitations of his account] thunderously clear from the start by proving that no such single predicate could be defined in a consistent language, given his assumptions concerning truth predicates [...] Considering how evident it is that Tarski did not give a general definition of truth, and the fact that perhaps his most important result was that it could not be done along lines that would satisfy him, it is remarkable how much effort some critics have put into the attempt to persuade us that Tarski failed to provide such a definition. ».

<sup>572</sup> Ibid., p.288.

déflationnistes pour montrer que la vérité n'a pas de nature autre que celle circonscrite pas sa fonction décitationnelle. Les partisans du caractère purement décitationnel de la vérité font ainsi valoir qu'« en disant que l'énoncé est vrai, nous disons que la neige est blanche<sup>573</sup> ».

Davidson ne défend pas, quant à lui, une conception déflationniste de la vérité, bien que certains (comme Rorty) aient affirmé le contraire<sup>574</sup>. Selon Davidson, la vérité est non seulement une notion explicative mais surtout *la* notion primitive explicative en sémantique, à savoir que : « truth is beautifully transparent compared to belief and coherence, and I take it as primitive<sup>575</sup> ». Rorty qui se définit lui-même comme un déflationniste, confond cette dimension primitive explicative avec l'absence de toute dimension explicative<sup>576</sup>. Il est vrai que Davidson n'explique pas la vérité en faisant appel à d'autres notions voisines. Mais, si Davidson ne cherche pas à définir le concept de vérité, ce n'est pas parce qu'il pense que la vérité est un concept non explicatif (comme Rorty le laisse entendre), mais bien parce qu'il affirme que la vérité, en tant que concept primitif explicatif, ne peut pas être expliquée en faisant intervenir d'autres notions (comme celle de signification, de référence, de correspondance, etc.). En effet, concevoir la vérité comme un concept primitif implique que, si ces notions-là sont douées d'un contenu quelconque, alors elles doivent être expliquées en ayant précisément recours au concept de vérité. Par conséquent, en définissant la vérité comme une notion primitive explicative, Davidson s'écarte non seulement des déflationnistes, tels que Quine et Rorty, mais exprime également son désaccord à l'endroit d'une théorie de la vérité au sens de Tarski. Davidson pense d'une part, contrairement à Tarski, qu'il faut cerner ou au moins reconnaître une notion translinguistique de vérité c'est-à-dire, un élément commun à toutes les définitions de la vérité et, d'autre part, que la vérité précède, dans l'ordre explicatif, la signification. Cependant, bien qu'il la défende, Davidson ne nous donne pas une analyse approfondie de ce qu'il entend par cette notion

<sup>573</sup> Quine, 1975, 1992a. Quine pense que les phrases-T de Tarski illustrent ce caractère décitationnel du prédicat de vérité, ce qui lui permet de faire de Tarski l'un des pères de la conception décitationnelle (quoique non tout à fait « évanescence ») de la vérité.

<sup>574</sup> Rorty, 1986, p.345. Davidson, 1990a, p.287-288, formule explicitement son refus à défendre une position déflationniste.

<sup>575</sup> Davidson, 2001b, 1983, p. 308.

<sup>576</sup> Williams, Mi., 1999, p.548, n.2, note cette confusion chez Rorty. Plutôt que d'affirmer que Davidson est un déflationniste, Williams soutient qu'il *devrait* en être un, ou qu'il *pourrait*, au vu du rôle que celui-ci assigne à la vérité dans la construction de théories adéquates de la signification, en être un, sans néanmoins vouloir porter atteinte au rejet explicitement formulé par celui-ci du déflationnisme.

translinguistique de vérité; certainement en raison de sa thèse en faveur d'une dimension primitive et explicative de la vérité, à savoir que : la vérité n'est pas analysable ou explicable par le biais d'autres notions. Il s'agit du concept sémantique le plus primitif, ce qui implique que vouloir en donner une définition est impossible. Davidson souligne plutôt l'importance de la connexion de la vérité à la notion de signification, et sa priorité au cœur d'une théorie satisfaisante de la signification pour les langues naturelles. Il pense ainsi que ces caractéristiques-ci de la vérité – telles que sa connexion à la notion de signification et son rôle dans la détermination d'une théorie adéquate de la signification d'une autre langue donnée (ce qu'il appelle le « contenu empirique » d'une théorie de la vérité) – sont laissées pour compte dans les définitions de Tarski. En effet, Davidson ne cherche pas à définir le concept de vérité, mais bien à le caractériser dans le contexte empirique de son usage. Il n'est donc pas question, pour lui, de donner une explication de la vérité des phrases dans l'absolu, mais de fournir aussi et surtout à la théorie de la vérité une assise empirique, en explicitant les conditions empiriques dans lesquelles une phrase quelconque est compréhensible pour un interprète lorsqu'elle est énoncée par un locuteur. Davidson impose donc à une théorie de la signification certaines conditions empiriques : une théorie de la signification doit pouvoir être testable, c'est-à-dire se prêter à des attributions vérifiables de significations aux locuteurs d'une langue donnée. Les phrases-T servent alors à donner la signification des phrases d'une langue donnée, et sont également dotées d'un contenu empirique; autrement dit, construire une définition de la vérité pour un langage donné sous la forme d'une théorie-T au sens de Tarski permet de donner de manière satisfaisante les conditions de vérité et de ce fait les conditions de signification des phrases d'une langue et d'une communauté données. Pourtant, pouvons-nous, pour autant affirmer qu'une théorie de la vérité satisfaisant strictement la convention T est suffisante pour permettre de comprendre un langage et l'interpréter, c'est-à-dire est-elle réellement une théorie de la signification pour les langues naturelles ? Nous souhaitons donc nous pencher à présent sur ces problèmes affectant la manière dont une théorie de la vérité (une théorie-T) sur le modèle de celle de Tarski peut servir de théorie de la signification pour un langage et notamment d'une théorie de la signification empiriquement testable.

#### 4.3. La théorie de la signification de Davidson

Davidson<sup>577</sup> pose la question de savoir quelle forme devrait prendre une théorie de la signification pour les langues naturelles<sup>578</sup>. Il tente d'apporter une réponse aux deux questions suivantes : d'une part, en quoi consiste la signification des phrases d'un locuteur dont la langue nous est totalement étrangère, ou plus précisément, l'idiolecte d'un locuteur particulier et d'autre part, comment peut-on arriver à élaborer une théorie satisfaisante de la signification de l'idiolecte en question ?

Davidson cherche à partir d'un concept plus primitif, à partir duquel la signification pourra être déterminée. Ainsi, en donnant à la vérité un rôle central dans la construction d'une théorie de l'interprétation, il renoue avec la conception vériconditionnelle<sup>579</sup> de la signification : les conditions sous lesquelles la phrase d'un idiolecte est vraie donnent la signification de cette phrase (et non la signification isolée du concept de signification<sup>580</sup>). La question est alors de définir les conditions d'adéquation que devrait remplir une théorie vériconditionnelle afin d'être adéquate. Une première condition d'adéquation requise pour la

<sup>577</sup> Davidson, 1993b, 1967.

<sup>578</sup> Williams, Mi., 1999, p. 552-554, distingue, dans l'œuvre de Davidson, entre un sens large et un sens étroit de l'expression « théorie de la signification ». Dans un sens large, une théorie de la signification est une analyse du concept de signification. Dans un sens étroit, une théorie de la signification spécifie la signification de toutes les phrases d'une langue naturelle en termes de conditions de vérité. Williams soutient que les théories de la signification de Davidson devraient être comprises dans un sens étroit et avance ainsi que celles-ci ne doivent pas expliquer la signification elle-même en termes de conditions de vérité, mais plutôt spécifier ou exprimer la signification des phrases particulières en termes de conditions de vérité des phrases. En outre, puisque la vérité, telle qu'elle est envisagée dans un sens étroit de la théorie de la signification, a seulement un rôle expressif et non explicatif, Williams poursuit en affirmant que Davidson défend une conception déflationniste de la vérité. Nous emploierons, dans ce paragraphe ainsi que dans les paragraphes suivants, l'expression de la « théorie de la signification » dans le sens étroit décrit ici par Williams.

<sup>579</sup> Le point de départ d'une conception vériconditionnelle de la sémantique est la remarque relativement triviale suivante : si dans une situation possible, par exemple si  $\alpha$  est vraie et  $\beta$  est fausse, alors il faut en déduire que les phrases  $\alpha$  et  $\beta$  doivent avoir des significations différentes. Cette remarque conduit inévitablement à l'idée que les conditions de vérité constituent une partie fondamentale de ce que l'on appelle la signification d'une phrase.

<sup>580</sup> Davidson restreint sa théorie de la signification étroitement construite aux locuteurs particuliers et à des moments de temps particuliers de leurs énonciations. La signification *via* les conditions de vérité n'est donc pas une propriété des phrases mais des actes de discours particuliers de locuteurs particuliers, c'est une propriété qui s'applique à « des triplets ordonnés de phrases, de locuteurs et de temps » (1993b, 1967, p.65). Davidson mentionne d'ailleurs que la particularité de toutes les phrases d'une langue naturelle « varient en valeur de vérité selon le temps où elles sont émises, le locuteur et même, peut-être, le public auxquels elles s'adressent » (1993b, 1970, p.97). Il sera nécessaire dès lors, pour tout interprète voulant comprendre quelles phrases un individu tient pour vraies, de spécifier l'environnement dans lequel prend place ces phrases (temps, lieu, et sujet qui énonce ces assertions).

construction d'une théorie de la signification est le principe sémantique de *compositionnalité*<sup>581</sup> : comprendre une expression linguistique, c'est être capable de discerner sa structure, c'est-à-dire la manière dont sa signification est déterminée par celle de ses parties composantes<sup>582</sup>. Davidson, en suivant Frege, ajoute qu'un locuteur compétent d'un langage *L*, c'est-à-dire un locuteur en mesure d'interpréter les phrases de son langage et de leur assigner une signification, est normalement capable de comprendre, à partir d'un ensemble fini d'expressions, un ensemble infini d'autres expressions<sup>583</sup>. Cette condition suggère que le langage a, à la fois d'un point de vue syntaxique et sémantique, une structure compositionnelle, à savoir que : les phrases sont construites systématiquement à partir des mots et leur signification est fonction de la signification de leurs parties composantes. En ce sens, il n'y a pas d'autre choix, pour une théorie sémantique, que de procéder, pour employer l'expression de Blackburn<sup>584</sup>, « de bas en haut » (*bottom up*), en décomposant les phrases en prédicats, en noms et en autres catégories grammaticales, et en montrant comment des objets sont dénotés par des noms et satisfont les prédicats, rendant ainsi vraies les phrases dans lesquelles ils figurent. Une théorie de la vérité qui ne révélerait pas une telle structure compositionnelle ne serait pas une théorie de la vérité. Mais, une chose est la structure de la théorie elle-même, et ce que cette théorie explique au moyen de notions comme celle de prédicat, de nom ou de satisfaction, et autre chose est l'application de cette théorie. Davidson ne nie donc pas que nous ayons besoin d'une caractérisation syntaxique des expressions. Mais, il soutient que nous ne pouvons pas donner de fonction sémantique à celles-ci indépendamment du rôle qu'elles jouent dans des phrases.

Selon Davidson, une telle théorie doit alors a) rendre compte de la signification de toute phrase du langage considéré comme construite à partir d'éléments empruntés à un stock *fini*,

---

<sup>581</sup> Davidson, 1993b, 1967, p.48; Frege, 1971, p.107-110. À l'origine, il s'agit d'un principe frégeén en vertu duquel le sens et la référence ou la dénotation des expressions complexes sont fonction du sens et de la référence des expressions plus simples. Ainsi, on ne peut pas, par exemple, se prononcer sur la vérité ou la fausseté d'une expression « Ulysse fut déposé sur le sol d'Ithaque dans un profond sommeil », car le terme « Ulysse » est dépourvu de dénotation; c'est-à-dire ne désigne pas d'objet. En fait, on ne peut pas dire d'un objet qui n'existe pas qu'il a été ou non déposé sur le sol d'Ithaque dans un profond sommeil. Par conséquent, une proposition dont l'une des composantes est dépourvue de dénotation ne peut avoir de dénotation.

<sup>582</sup> Davidson, 1993b, 1977b, p.316 : « [...] la vérité de toute phrase d'un langage *L* particulier dépend de sa structure et de ses constituants. ».

<sup>583</sup> Davidson, 1993b, 1967, p.41.

<sup>584</sup> Blackburn, 1984a, p.273.

qui entrent dans l'énoncé de ses conditions de vérité. Elle doit b) nous permettre de *décider*, pour toute phrase arbitrairement choisie, quelle est sa signification. Et c) les énoncés des conditions de vérité des phrases impliquées par la théorie doivent être axés sur les *mêmes* concepts que ceux des phrases dont elle est la théorie<sup>585</sup>. Pour ce faire, nous avons besoin d'une théorie de la signification qui nous permette de discerner la manière dont la signification des phrases est déterminée par celle de ses parties composantes, de telle sorte qu'à partir de l'application d'un nombre fini d'axiomes donnant la signification de ses parties composantes à un nombre fini de règles donnant les effets sémantiques résultant de la combinaison de ces parties respectives, nous puissions générer un nombre infini de théorèmes et rendre compte de la signification d'un nombre infini de phrases. C'est ainsi la structure de la théorie de la signification que proposera Davidson, celle-ci répondant à la condition d'adéquation selon laquelle une théorie de la signification doit être compositionnelle.

Par conséquent, comprendre une expression, c'est d'après le principe de compositionnalité, comprendre les expressions plus simples dont elle est composée. Il y a ainsi deux lectures possibles de ce principe. Nous pouvons le lire comme impliquant que la signification d'une phrase est déterminée par celle de ses parties composantes c'est-à-dire de ses constituants et de leurs modes de construction. En ce sens, nous soutenons une version atomiste du principe. Mais, nous pouvons aussi le lire comme impliquant que la signification des parties composantes d'une phrase est déterminée elle-même par la signification de la phrase entière. En ce sens, nous soutenons une version holiste du principe de compositionnalité. Quel que soit le sens du principe contextuel de Frege, selon lequel « ce n'est que dans le contexte d'une phrase que les mots ont une signification<sup>586</sup> », c'est dans ce sens holistique que Davidson le comprend<sup>587</sup>. En d'autres termes, ce principe équivaut apparemment à une certaine forme limitée de holisme linguistique (faible), qu'on pourrait

<sup>585</sup> Davidson, 1993b, 1970.

<sup>586</sup> Frege 1971, p.48: le principe frégeén est lui-même sujet à un certain nombre d'interprétations. Seymour, 2000, n. 6. Même si Frege endosse le principe de contextualité selon lequel les mots n'exercent leur fonction de signification que dans le contexte d'une proposition, il demeure quand même un partisan de la conception atomiste. Le principe de contextualité constitue tout au plus une contrainte sur le type de signification qu'un mot isolé peut être en mesure de signifier. Il faut que sa signification contribue à la détermination de la signification des énoncés dans lesquels il se trouve. Or, cela s'accorde parfaitement avec l'idée que le mot a, lorsque pris isolément, une signification déterminée se situant dans une réalité objective, fixe, transculturelle et intemporelle. Cf. Dummett, 1982.

<sup>587</sup> Davidson, 1993b, 1967, p.48.

aussi appeler le « holisme de la phrase », puisqu'il revient à dire que la phrase complète est l'unité de base de la signification. Il semble, par ailleurs, qu'en rejetant la théorie du langage comme « jeu de cubes<sup>588</sup> » ou comme « théorie des blocs empilés<sup>589</sup> », Davidson s'engage à ce holisme de la phrase. Les principes de contextualité et de compositionnalité paraissent ainsi s'opposer : l'un avance que la phrase est l'unité de la signification et l'autre que la signification d'une phrase dépend exclusivement de la signification de ses parties composantes, ne dépendant d'*aucune* autre expression. La théorie de Davidson s'efforce pourtant de réconcilier ce conflit en étant, d'un point de vue formel, compositionnelle mais d'un point de vue explicatif, holistique. Autrement dit, comprendre un langage, c'est d'abord comprendre l'ensemble des phrases du langage, et ensuite l'ensemble des éléments qui les composent; une phrase n'ayant jamais de signification indépendamment du contexte de l'ensemble des autres phrases d'un langage, et les expressions qui les composent n'acquièrent de sens que dans le contexte de toutes les phrases où elles figurent<sup>590</sup>. Il faut donc noter que le holisme de la phrase en lui-même n'est pas incompatible avec la théorie du jeu de cubes, puisqu'il n'est pas contradictoire de soutenir à la fois que la signification d'une expression simple est déterminée par la signification des phrases où elle apparaît et qu'elle est (indépendamment) déterminée (en plus) par quelque chose d'autre (par exemple par certaines relations causales avec le monde).

Il est également possible, comme le souligne Laurier<sup>591</sup>, que le holisme de la phrase, joint au principe de compositionnalité, mène à une certaine forme extrême de holisme linguistique, selon lequel la signification de n'importe quelle expression, qu'elle soit simple ou complexe, dépend de la signification de toutes les autres expressions de la langue donnée<sup>592</sup>. Prenons ainsi, en suivant Laurier, l'exemple de la relation « dépendre de » (ou « être déterminé par ») et supposons qu'elle soit transitive, et qu'étant données deux

<sup>588</sup> Laurier, 1994, p.149.

<sup>589</sup> Davidson, 1993b, 1967, p.24, 1993b, 1977b, p. 318-321, p.325 : la méthode atomiste et moléculaire, qui consiste à expliquer progressivement le sens d'une phrase à partir de la référence à ses parties composantes, « de bas en haut », ou, selon l'image de Davidson lui-même, par « empilage de blocs » (*building block theory*) successifs est donc vouée à l'échec.

<sup>590</sup> Davidson, 1993b, 1967, p.48: « Si la signification des phrases dépend de leur structure, et si nous comprenons la signification de chaque élément de la structure que par abstraction à partir de la totalité des phrases dans lesquelles il figure, alors nous ne pouvons donner la signification d'une phrase quelconque (ou d'un mot) qu'en donnant la signification de toutes les phrases (et mots) du langage.»

<sup>591</sup> Laurier, 1994, p.150

<sup>592</sup> Ibid., p. 150.

expressions quelconques, il y a au moins une phrase qui les contient toutes les deux. Choisissons une phrase au hasard : par compositionnalité, sa signification dépend de la signification de ses constituants. Prenons maintenant l'un de ses constituants : en vertu du holisme de la phrase, sa signification dépend de la signification de toutes les phrases où il apparaît; et la signification de chacune d'entre elles, à son tour, dépend de la signification de ses constituants, et ainsi de suite à travers la langue dans sa totalité. Nous voyons ainsi que la signification de n'importe quelle expression donnée doit dépendre de celle de toutes les autres expressions, car toutes les autres expressions apparaissent dans au moins une des phrases contenant l'expression donnée. Il semblerait alors qu'une source d'au moins une forme de holisme linguistique attribuable à Davidson se trouve dans son rejet de la théorie des cubes ou des blocs empilés, c'est-à-dire de l'idée que la signification des expressions simples peut être déterminée indépendamment de la signification des phrases dans lesquelles elles apparaissent. Aussi, rejette-t-il cette approche parce qu'il soutient que c'est seulement au niveau des phrases qu'une théorie du langage est confrontée aux données non linguistiques (c'est-à-dire aux données empiriques plausiblement accessibles à un interprète radical) capables de la confirmer. Et, il soutient cela, non seulement en raison de la longue histoire des échecs de la théorie de jeu de cubes à laquelle il fait allusion<sup>593</sup>, mais parce qu'une théorie du langage doit relier la langue à l'usage, et que les mots ne sont (presque) jamais utilisés autrement qu'en parties de phrases. Il apparaîtrait donc que l'obligation de relier la théorie du langage à des données empiriques impose d'emblée le rejet de la théorie du jeu de cubes et le holisme qui vient avec ce rejet (étant donné le principe de compositionnalité).

Par conséquent, si comprendre le sens d'une expression, c'est comprendre le rôle qu'elle joue et la contribution qu'elle apporte au sens d'une phrase, et si comprendre une phrase, c'est comprendre le rôle qu'elle joue au sein d'un ensemble d'autres phrases, de quelle nature peuvent être ces rôles ou ces contributions ? Il faut donc préciser de quelle sorte de composition ou de structure il s'agit. La suggestion la plus évidente est que les phrases ont une structure grammaticale, en vertu de laquelle elles sont articulées. Une théorie de la signification devra donc nous permettre de discerner la structure grammaticale des phrases, et la récurrence de cette structure au sein de l'ensemble des phrases. Mais cela ne suffit pas à rendre compte des relations que les phrases, en vertu de l'hypothèse holistique, ont entre

---

<sup>593</sup> Davidson, 1993b, 1977b, p.318-321, p.325.

elles. De quelles sortes de relations peut-il s'agir ? Davidson suggère que ces relations ont, avant tout, des relations logiques ou inférentielles : de certaines phrases on peut en inférer d'autres, et ces relations logiques ont un rôle important à jouer dans la signification des phrases. Mais, ces relations logiques ne peuvent elles-mêmes exister qu'en vertu du caractère des éléments qui composent ces phrases. Cela suggère que la structure compositionnelle devra aussi avoir certaines propriétés logiques, et qu'elle s'identifiera, au moins jusqu'à un certain point, avec la structure logique ou la forme logique discernable dans ces phrases. Cette structure compositionnelle repose ainsi sur une axiomatisation finie : Davidson propose que la théorie soit pourvue d'un nombre fini d'axiomes, et qu'elle soit récursivement construite à partir d'un ensemble de notions sémantiques primitives. Ainsi, elle peut être apprise (condition (a) ci-dessus<sup>594</sup>). Quant à sa structure holistique, celle-ci est mise en évidence par le fait que les axiomes, exprimant la référence des termes singuliers du langage et donnant l'extension des prédicats, ne sont pas définis indépendamment et antérieurement aux déterminations des conditions de vérité des phrases dans lesquelles ils apparaissent. De plus, puisque le point où le langage peut rencontrer la réalité ne peut pas être situé au niveau des expressions simples ou des mots qui composent les phrases, mais seulement au niveau des phrases elles-mêmes, la première valeur sémantique que doit tester une théorie de la signification est la vérité, et non la référence. Une théorie de la référence ne peut avoir qu'un rôle dérivé par rapport à une théorie de la vérité pour un langage. La vérité est une notion explicative primitive, les conditions de vérité des phrases sont d'abord déterminées, et ensuite les références des termes singuliers au sein des phrases sont des déterminations fonctionnelles des conditions de vérité – les références des termes sont une fonction de l'attribution des conditions de vérité. La vérité n'est donc ni construite à partir des termes de références ni expliquée en termes de référence, c'est, au contraire, la référence qui est construite à partir de la vérité. De cette manière, bien que la théorie de la signification de Davidson soit, d'un point de vue formel, compositionnelle, elle demeure holistique (et non à l'image d'une théorie des blocs empilés) dans son ordre d'explication sémantique<sup>595</sup>.

<sup>594</sup> Cf. condition a) rendre compte de la signification de toute phrase du langage considéré comme construite à partir d'éléments emprunté à un stock fini, qui entrent dans l'énoncé de ses conditions de vérité.

<sup>595</sup> La différence entre Davidson et Tarski, mentionnée précédemment, est mis en exergue ici. Pour Tarski, la vérité peut seulement être définie une fois que les attributions référentielles des noms et des prédicats ont été finalisées; pour Davidson, les secondes sont une fonction de l'attribution des

La théorie de la signification de Davidson répond donc, comme nous venons de le voir, à deux contraintes formelles : celle de compositionnalité et de contextualité. Nous cherchons maintenant à déterminer la forme d'une théorie capable d'assigner une signification à toutes les phrases du langage, c'est-à-dire d'une théorie capable d'identifier la propriété sémantique explicative fondamentale des phrases et donc des mots. Depuis la distinction formulée par Frege<sup>596</sup> entre le *Sinn* et la *Bedeutung* d'une expression – assimilée à celle entre l'intension (son sens) et l'extension (sa référence ou sa dénotation) de cette expression – nous disposons de deux notions sémantiques, des notions intensionnelles et extensionnelles; Davidson doit donc choisir entre employer, comme notion primitive explicative, une notion intensionnelle ou extensionnelle. De plus, étant donné son engagement à l'égard des principes de compositionnalité et de contextualité, la notion sémantique primitive devra nous donner la signification des phrases d'un langage en montrant comment est déterminée leur structure compositionnelle. Par conséquent, seule une analyse qui établirait comment la signification des phrases est déterminée compositionnellement à partir des éléments qui affectent leurs conditions de vérité peut effectivement donner leur signification. Pour être compositionnelle, l'extension d'une phrase doit être une fonction des extensions de ses parties composantes, ce qui implique, que si l'extension de ses parties demeure identique, alors l'extension de la phrase l'est aussi. Nous pouvons ainsi identifier l'extension d'une phrase comme suit : il s'agit de la propriété qu'une phrase a d'être vraie dans certaines conditions, propriété qui demeure inchangée sous la substitution de termes co-extensionnels ou coréférentiels. Par conséquent, la valeur de vérité d'une phrase est son extension, et la définir comme une notion

---

conditions de vérité aux phrases. Comme nous le verrons plus loin dans la discussion relative au principe de charité, cette différence est une conséquence du holisme généralisé de Davidson et de la thèse selon laquelle l'interprétation est un projet obéissant à des contraintes normatives, à savoir que l'interprétation est régie par des normes constitutives de maximisation de la vérité et de rationalité telles que nous ne pouvons pas interpréter les autres locuteurs en leur attribuant des croyances irrationnelles ou/et erronées, selon nos propres normes. Dans la mesure où les phrases peuvent être vraies (ou fausses) et sont reliées logiquement et inférentiellement les unes aux autres, les conditions de vérité des phrases et les relations rationnelles les régissant sont, selon Davidson, les notions sémantiques primitives dont la référence en est une notion dérivée. La référence n'est pas une relation allant du langage vers le monde, ce n'est pas le point où le langage rencontre le monde et qui se traduit ensuite par une relation causale entre un objet dans le monde et la production d'un signe ou d'un son, mais c'est plutôt un *posit* (une entité postulée) théorique, une construction du processus général interprétatif lui-même. Une implication problématique de cette conception de la référence est examinée au §4.7.

<sup>596</sup> Frege, 1971, p.102.

sémantique extensionnelle équivaut à dire qu'elle permet la substitution de termes co-extensionnels, *salva veritate*<sup>597</sup> (c'est-à-dire sans modifier la valeur de vérité). La signification d'une phrase est, dans un contexte intensionnel, la pensée qu'elle exprime, et afin d'être compositionnelle, la pensée qu'elle exprime doit être construite à partir du contenu des autres expressions ou parties composantes de la phrase de telle sorte que deux phrases expriment la même pensée si quiconque qui les comprend ne peut pas (rationnellement) avoir, à leur égard, des attitudes cognitives différentes, c'est-à-dire croire que l'une est fautive et que l'autre est vraie. Autrement dit, si nous considérons les propositions suivantes :

(i) Hespérus est une planète.

(ii) Phosphorus une planète.

Il importe également de savoir que l'expression « Hespérus » est généralement considérée comme synonyme de « l'Étoile du soir » alors que « Phosphorus » est considérée comme synonyme de « l'Étoile du matin ». Puisque Hespérus est identique à Phosphorus, ces deux propositions devraient avoir un même contenu conceptuel, et les termes « Hespérus » et « Phosphorus » être substituables l'un à l'autre dans n'importe quel énoncé, *salva veritate*.

Or, il se pose de toute évidence un problème. Si ces deux propositions avaient un même contenu conceptuel, alors elles devraient dire strictement la même chose, et utiliser des expressions comme « étoile du matin » ou « étoile du soir » ne serait que l'affaire de choisir arbitrairement un de ces deux signes. Pourtant, il semble bien que ces propositions ne disent pas la même chose, puisqu'une des parties de l'expression complexe a un sens différent : le soir n'est pas le matin et comme le fait remarquer Philippe de Rouilhan, « le fait que l'étoile du matin soit illuminée par le soleil est une chose, le fait qu'il en soit de même pour l'étoile du soir en est une autre, les deux énoncés ne sont pas logiquement, c'est-à-dire analytiquement équivalents<sup>598</sup> ». Par conséquent, il devient difficile de considérer que les deux

<sup>597</sup> Quine, 1975, p. 110 affirme à ce sujet « tout langage qui a une grammaire canonique est *extensionnel*, c'est-à-dire les prédicats coextensifs y sont interchangeable *salva veritate*. Il n'y a pas dans ce langage d'énoncé qui puisse devenir faux quand une occurrence d'un prédicat y est remplacée par un prédicat coextensif. » (En théorie des ensembles, deux ensembles sont identiques lorsqu'ils ont exactement les mêmes éléments, c'est-à-dire la même *extension*; ils sont alors substituables *salva veritate* à chacune de leurs occurrences dans les énoncés.) C'est ce langage purement référentiel de la théorie des ensembles qui est pour Quine la norme suprême du jugement philosophique. Enfin, loi d'extensionnalité et principe de substituabilité des termes s'impliquent réciproquement: dans la grammaire canonique, il est nécessaire que deux termes coextensifs soient substituables, et que deux termes substituables soient coextensifs.

<sup>598</sup> Rouilhan, 1988, p. 29.

noms propres aient une identité de contenu, tandis que les énoncés dans lesquels ils apparaissent, présenteraient – eux – un contenu conceptuel différent, car cela violerait la définition leibnizienne de l'identité<sup>599</sup>, et il serait illégitime de procéder à la substitution des termes.

Selon ce critère, des énoncés comme « Hespérus est une planète » et « Phosphorus est une planète » ont des significations (intensionnelles) différentes, ils expriment des pensées différentes, parce que nous pouvons les comprendre et pourtant, si nous ignorons que « Hespérus » et « Phosphorus » font référence au même objet (la planète Vénus), nous sommes alors enclins à considérer qu'une phrase est vraie et que l'autre est fausse. C'est donc à partir d'une réflexion de ce type que Frege va procéder à une scission du contenu conceptuel, le scinder en deux éléments distincts : le sens et la référence. À la lumière de cette distinction, il devient alors possible de considérer les propositions comme signifiant quelque chose de différent, et ce, tout en montrant qu'elles sont du point de vue de leur valeur de vérité équivalentes, c'est-à-dire identiques. « Hespérus » et « Phosphorus » sont bien des expressions équivalentes en cela qu'elles font référence au même objet, à savoir Vénus. Elles désignent la même chose. L'identité est donc celle de leur référence. Mais, si leur référence est identique, comment peut-il alors se faire que les propositions disent quelque chose de différent ? La réponse est en fait simple : c'est parce qu'elles ont un sens différent. « Hespérus » a un sens différent de « Phosphorus », et ces deux sens différents déterminent une même référence. En ce sens, la signification n'est pas une notion extensionnelle, elle n'autorise pas la substitution de termes coextensionnels (ou coréférentiels), *salva veritate*, ce

---

<sup>599</sup> Selon une conception anti-frégéenne, la notion d'identité peut être définie à partir de la *loi de Leibniz* :

Deux choses sont identiques si et seulement si elles ont toutes leurs propriétés en commun.

$x=y$  si et seulement si  $\forall F(Fx \equiv Fy)$

La loi de Leibniz peut être exploitée pour définir l'identité au moyen de la notion apparemment différente d'indiscernabilité. Plus précisément, selon cette conception, le principe de l'*indiscernabilité des identiques* est une condition nécessaire de l'identité (« seulement si »), alors que le principe de l'*identité des indiscernables* est une condition suffisante de l'identité (« si »).

*Principe de l'indiscernabilité des identiques* : si deux choses sont identiques, elles ont toutes leurs propriétés en commun.

*Principe de l'identité des indiscernables* : si deux choses ont toutes leurs propriétés en commun, elles sont identiques.

De la loi de Leibniz (conjonction de ces deux principes), on peut déduire les propriétés formelles de l'identité.

qui revient à dire qu'elle appartient à une classe différente de notions sémantiques, à savoir celle des notions intensionnelles.

Par conséquent, décider d'une notion sémantique sur laquelle fonder une théorie de la signification se résume ici à choisir entre la signification d'une phrase (la pensée qu'elle exprime) ou sa valeur de vérité. Davidson analyse et rejette une théorie qui adopterait comme concept explicatif primitif la notion intensionnelle de signification, c'est-à-dire une théorie composée de théorèmes intensionnels de la forme : *s* signifie que *p* (ce que Dummett a nommé les phrases-*M*). Deux raisons majeures motivent sa position : la première est de nature quinienne, la seconde porte sur la contrainte de compositionnalité.

La première raison est une conséquence du holisme de Quine, à laquelle Davidson adhère. Elle met en évidence le problème émanant de l'emploi des significations au sens d'entités intensionnelles, à savoir : l'absence de critère d'identité satisfaisant nous permettant d'identifier les significations au sens d'entités intensionnelles, ce qui se traduit par la disparition ultérieure de la notion de signification<sup>600</sup> comme notion primitive au centre de l'explication de toute théorie sémantique. Revenons donc sur la portée de l'apport de Quine dans la construction de la théorie de la signification et de l'interprétation de Davidson.

Quine rejette ainsi les significations si l'on entend par « significations » des entités idéales existant entre le monde et le langage. Conformément à la norme de l'engagement ontologique, l'univers de Quine est un univers où il est vrai que les entités que sont les noms sont significatives, mais il est faux, pour Quine, qu'existent des entités qui soient des significations; c'est une hypostase dont il fait l'économie. Quel est l'argument, auquel Quine se contente ici de faire allusion, sur lequel il se base pour rejeter les significations? Et comment Quine conçoit-il le caractère significatif du langage? Encore une fois, Quine a été très clair sur les réponses qu'il donne à ces deux questions. La raison pour laquelle Quine rejette les significations de son ontologie se trouve dans son fameux principe: « pas d'entité sans identité ». S'il faut concevoir, en plus des énoncés, des « propositions » ou des « significations », ou des « entités idéales » auxquelles ces énoncés renvoient, alors on devrait pouvoir disposer d'un critère d'identité afin de pouvoir trouver des significations identiques entre elles, substituables grâce à des énoncés synonymes. Or il est très difficile de

---

<sup>600</sup> Nous verrons que dans le cadre de la théorie de l'interprétation radicale de Davidson (1993b, 1973a; 1993b, 1974a; 1993b, 1976), il paraît douteux qu'on puisse éliminer le concept de signification au profit des seuls concepts extensionnels de vérité et de référence.

décider si tel énoncé a une signification parfaitement identique avec tel autre, ou si leurs significations diffèrent. Sans critère d'identité, nous ne pouvons pas savoir ce qu'est une chose et par conséquent, si cette chose existe. Autrement dit, sans critère d'identité, nous ne pouvons pas garantir que nous ne formulons pas des attributions contradictoires de la même chose et que nous respectons le principe de non-contradiction. Quine pose alors la question suivante : à quel critère d'identité répondent les significations au sens d'entités intensionnelles ?

Vient avec la contrainte de « pas d'entité sans identité » l'équivalence entre savoir ce qu'est une chose et spécifier ce que l'on entend par être identique avec telle autre chose. Appliquée au cas de la signification, la question de savoir s'il existe des choses telles que les significations, revient à se demander s'il existe bien un critère capable de déterminer lorsque deux expressions ont la même signification ou sont synonymes :

« Defining the notion of meaning for sentences may properly be said to consist in specifying the circumstances in which sentences have the same meaning. We can settle for no less, because there is no entity without identity, no meaning without sameness of meaning<sup>601</sup>. ».

En un mot, significations et similitude de signification posent un seul et même problème<sup>602</sup>.

En reliant les deux questions, à savoir: celle portant sur la détermination de la signification d'une phrase et celle relative à la relation de synonymie des phrases entre elles, Quine tente de définir ou d'identifier la signification au moyen d'une relation d'équivalence, la relation de « a la même signification que ». Les relations d'équivalence produisent des classes ou des ensembles d'équivalence, et dans la mesure où ces ensembles ont un critère d'identité (extensionnel), Quine pense arriver à établir un critère d'identité extensionnel ou un ensemble théorique de signification. Par exemple, avoir un QI de 120 peut être défini théoriquement ou extensionnellement en termes de la relation d'équivalence « avoir le même QI que », dans laquelle avoir un QI de 120 veut dire appartenir à l'ensemble de toutes les choses ayant le même QI, à savoir le QI de 120. Cette propriété est donc définie extensionnellement en termes de choses appartenant à une classe d'équivalence particulière. Quine tente de procéder de la même manière avec la notion d'« identité de signification » en définissant la

---

<sup>601</sup> Quine, 1995, p.75-76.

<sup>602</sup> Quine, 1993, p.84.

signification d'une phrase (et à cet effet, les expressions sub-phrastiques) en termes de phrases appartenant à une classe d'équivalence :

« [O]nce we defined sameness of meaning for sentences we can define the meaning of a sentence...as the class of all sentences alike it in meaning<sup>603</sup>. ».

Si nous pouvions mettre au point une relation de similitude acceptable entre significations, la distance serait courte vers une définition acceptable des significations. Car [...] nous pourrions alors définir la signification d'une expression comme la classe de toutes les expressions qui lui sont semblables par la signification<sup>604</sup>.

Cette tentative visant à définir l'identité de signification (et donc la signification elle-même) de manière extensionnelle au moyen d'autres phrases appartenant à la même classe d'équivalence, déréifie la signification : une phrase est significative, non pas en admettant des entités abstraites comme significations, mais en s'intégrant au sein d'une activité pratique dans laquelle la phrase ne prend sens qu'en étant corrélée à d'autres phrases. Quine et Davidson préconisent ainsi, tous deux, une approche déréifiée de la signification, appréhendant la signification par l'usage c'est-à-dire par le biais des pratiques de traduction, d'interprétation, de substitution. La question qui se pose alors est celle de savoir si ces pratiques-ci peuvent véritablement définir la signification en évitant toute indétermination et s'il est possible d'envisager un critère d'identité de signification en termes de substituabilité d'énoncés appartenant à une classe particulière d'équivalence.

Quine<sup>605</sup> illustre cette difficulté à l'aide de l'exemple suivant: si « célibataires » et « non mariés » étaient parfaitement synonymes et substituables, *salva veritate*, dans toutes les expressions où ils figurent, il semblerait que l'on soit en droit de conclure à l'identité de signification de deux expressions. Mais la différence entre les énoncés qui suivent est immédiatement perceptible:

- (1) Nécessairement tous les célibataires et seulement les célibataires sont célibataires.
- (2) Nécessairement tous les célibataires et seulement les célibataires sont non mariés.

Même si nous affirmons que (2) suit de la définition de « célibataire » par « non marié » et que les deux énoncés indiquent la même chose, il est malgré tout impossible de soutenir qu'ils expriment rigoureusement la même chose. Il suffit de remarquer que dans (1) la signification

<sup>603</sup> Quine, 1995, p.76.

<sup>604</sup> Quine, 1993, p. 84.

<sup>605</sup> Quine, 2003, p. 58-64.

de « célibataires » est indifférente à la vérité de l'énoncé, alors que dans (2) la signification qu'on accorde à ce même terme conditionne la valeur de vérité, pour se persuader qu'il est vain d'admettre dans notre univers des *entités* comme les significations; il est plus juste de considérer que le sens dérive de l'usage<sup>606</sup>. Lorsque Quine, « congédie les propositions<sup>607</sup> », il entend mettre un terme à cette mythologie chosiste de la signification en démontrant qu'il est à la fois plus rationnel et plus simple de dire que des énoncés sont, ou ne sont pas, significatifs.

La doctrine des propositions semble vaine, à tout prendre, même si nous imaginons résolu le problème de leur individuation. Car la solution de ce problème consisterait en une définition appropriée de l'équivalence des énoncés; et alors pourquoi ne pas parler d'énoncés et d'équivalence tout simplement, en laissant tomber les propositions ? L'affaire n'est ni plus ni moins que ceci: les propositions sont les ombres portées des énoncés, pour reprendre une comparaison de Wittgenstein. En mettant les choses au mieux, elles ne peuvent rien nous donner que ne nous donnent déjà les énoncés. Si elles promettent davantage, c'est seulement que, sans critique, nous avons admis pour elles une individuation qui ne correspond, entre énoncés, à aucune équivalence que nous sachions définir. Ces ombres ont encouragé à penser n'importe quoi<sup>608</sup>.

Sur cette question, Quine n'a jamais varié. « Ce que je rejette, c'est l'objectification rigide du sens » écrit-il dans *Quiddités*, puis à la fin de l'article consacré au mot « Signification », il enfonce le clou:

---

<sup>606</sup> En jouant sur les mots, Quine montre la précarité de la synonymie: « bachelor » signifie en anglais « célibataire », mais aussi « licencié » ou, en un sens ancien, « bachelier ». On peut également remarquer que l'on ne sait plus si l'on doit accorder un sens univoque à l'adverbe « nécessairement », car la nécessité de la vérité de (1) est triviale, à la différence de la vérité de (2) qui est fondée sur la connaissance de la synonymie de « célibataire » et « non marié ». On pourrait considérer l'énoncé (2) comme un énoncé analytique, mais l'argumentation de Quine vise précisément à rejeter cette notion d'analyticité. L'explication la plus claire et la plus concise qu'il a donnée de ce rejet est la suivante: « Mon rejet de la notion d'analyticité veut juste dire mon refus de tracer une frontière entre ce qui contribue à la pure compréhension des phrases d'une langue et tout ce dont la communauté tombe d'accord. Je doute qu'on soit capable d'opérer une distinction objective entre la signification et une information indirecte que posséderaient tous les locuteurs d'une communauté. ». (Quine, 1977b, p. 100-101.)

<sup>607</sup> Quine, 1975, Chap.1.

<sup>608</sup> Ibid., p. 21.

A la fin de l'entrée consacrée aux IDEES, je soutiens qu'il n'y a pas de place en science pour les idées et, à l'entrée CONNAISSANCE, je soutiens de même qu'il n'y a pas de place pour la connaissance dans la théorie de la connaissance. Or que fais-je ici, sinon soutenir qu'il n'y a pas de place pour les significations (telles qu'on les conçoit d'ordinaire) dans la théorie de la signification? La signification et l'usage y ont leur place, mais pas les significations. Mais je ne veux pas qu'on m'accuse de déprécier la science ou la connaissance, et je ne voudrais pas non plus que l'on m'accuse de déprécier la sémantique ou la théorie de la signification. Débarrassée de ce qui l'entrave, elle ne s'en portera que mieux<sup>609</sup>.

La signification que Quine donne au rejet des significations permet de comprendre l'humour de ces déclarations qui ne sont paradoxales qu'en apparence. La signification n'est rien d'autre pour Quine que l'usage, c'est la raison pour laquelle il pense préférable de l'expliquer « en termes de comportement », en donnant bien sûr à ce terme le sens le plus large possible, car alors même la signification ou l'usage des symboles utilisés en mathématiques est affaire de comportement. Mais si l'on veut ici faire l'économie des explications de type behavioriste et donner malgré tout une théorie de la signification, ce que Quine écrit de façon aphoristique dans « Deux dogmes de l'empirisme » pourrait suffire:

Les choses avaient des essences pour Aristote, mais seules les formes linguistiques ont des significations. La signification est ce que l'essence devient lorsqu'elle divorce de l'objet de la référence et se marie au mot<sup>610</sup>.

Quine ne veut sans doute rien dire d'autre que ceci: il ne peut pas y avoir d'essences pour une théorie de la signification. Si l'on admet que les mots ont pour référence des objets qui sont des essences (le mot « homme » faisant par exemple référence à l'espèce humaine), les significations des mots ne sont pas séparables des mots eux-mêmes et de leur usage (le mot « homme » est différemment significatif selon les contextes). Davidson semble ainsi suivre jusqu'à un certain point, qu'il nous faudra préciser, la trame d'analyse de Quine.

Venons-en au second problème que posent ces théorèmes intensionnels-ci, à savoir qu'ils ne permettent pas la substitution de termes coextensionnels *salva veritate*, ce qui signifie qu'une théorie composée de phrases-*M* risque de ne pas pouvoir satisfaire la condition d'adéquation de compositionnalité. En logique du premier ordre, nous disposons d'une théorie de la compositionnalité des valeurs de vérité c'est-à-dire d'une théorie qui décrit la manière dont la valeur de vérité d'une phrase ou d'un énoncé est une fonction de ses

<sup>609</sup> Quine, 1992, p. 219-220.

<sup>610</sup> Quine, 2003, p. 22.

parties composantes. Les tables de vérité de la logique propositionnelle rendent compte de la vérifonctionnalité des phrases construites au moyen de connecteurs propositionnels tels que « et », « ou », « non », « Si...alors », et « si et seulement si ». Étant donné que la valeur de vérité d'une phrase est son extension, la logique propositionnelle nous dit comment l'extension (la valeur de vérité) d'un énoncé complexe est une fonction de l'extension (la valeur de vérité) de ses énoncés élémentaires. En effet, la logique propositionnelle rend bien compte de la nature compositionnelle de la notion sémantique extensionnelle de valeur de vérité des phrases complexes. Le travail de Frege et de Tarski en logique des prédicats étend ce résultat de la logique propositionnelle afin de montrer que la valeur de vérité d'une phrase (non complexe) dépend de la notion extensionnelle appropriée à ces parties sub-phrastiques. En tant que tel, la logique de premier ordre (logique quantifiée, logique des prédicats « polyadiques<sup>611</sup>») nous donne une théorie compositionnelle de la notion sémantique extensionnelle de vérité. Le problème dans l'usage du concept de signification comme notion sémantique fondamentale est qu'il ne rend pas compte de la structure compositionnelle et extensionnelle des expressions linguistiques. Vu que d'une part, seule une conception fonctionnelle de la vérité conforme aux réquisits usuels de l'extensionnalité (substitution *salva veritate* des termes coréférentiels, etc.) peut rendre compte de la structure compositionnelle des notions extensionnelles et d'autre part, que nous disposons déjà d'un système formel en vigueur (logique de premier ordre) qui en explique les rouages, toute théorie employant l'expression « signifie que » dans ses théorèmes éprouvera de la difficulté à expliquer la nature compositionnelle de la signification et ne pourra donc pas exploiter dans ce type d'explication les ressources de la logique du premier ordre. Le problème est alors de trouver une notion appropriée susceptible d'asseoir notre théorie de la signification, mais qui, contrairement au concept de signification, soit une notion nous donnant accès aux ressources de la logique des prédicats du premier ordre afin de nous permettre d'expliquer la compositionnalité des significations. La solution de Davidson est de substituer à la notion logiquement obscure de signification celle de vérité logiquement claire c'est-à-dire de remplacer l'opérateur intensionnel « signifie que » par l'opérateur extensionnel « est vrai si et seulement si ».

---

<sup>611</sup> Kenny, 1963.

Sautons finalement le pas audacieusement, et essayons de traiter la position occupée par « p » extensionnellement : pour réaliser cela, écarterez l'obscur « signifie que », donnez à la phrase qui remplace « p » un connecteur propositionnel approprié, et donnez à la description que remplace « p » son propre prédicat. Le résultat plausible est :

(T) *s* est *T* si et seulement si *p*

Ce que nous demandons à une théorie de la signification pour un langage *L* est que sans faire appel à aucune notion sémantique (supplémentaire) elle place suffisamment de restrictions sur le prédicat « est *T* » pour impliquer toutes les phrases obtenues à partir du schéma *T* quand « *s* » est remplacée par une description structurale d'une phrase de *L* et « *p* » par cette phrase.

Deux prédicats satisfaisant cette condition ont la même extension, en sorte que si le métalangage est suffisamment riche, rien ne s'oppose à ce que nous donnions « est *T* ». Mais qu'on le définisse explicitement ou qu'on le caractérise récursivement, il est clair que les phrases auxquelles le prédicat « est *T* » s'applique seront justement les phrases vraies de *L*, parce que la condition que j'ai imposée aux théories satisfaisantes de la signification est dans son essence la Convention *T* de Tarski qui teste l'adéquation d'une définition formelle de la vérité<sup>612</sup>.

L'avantage majeur de ce changement est que les théorèmes de sa théorie de la signification auront la même forme que la convention *T*, et Tarski a ainsi montré comment rendre compte de manière compositionnelle de théorèmes de ce genre. En effet, la condition formelle principale d'adéquation d'une théorie de la signification est précisément la Convention *T*, qui est pour Tarski, la « condition d'adéquation formelle » d'une théorie de la vérité et une théorie de la signification doit devenir, en un sens qu'il faudra préciser, une théorie de la vérité (ce que nous appelons une théorie-*T*). Néanmoins, ce passage énonce mais ne montre pas qu'une théorie de la signification doit prendre la forme d'une théorie de la vérité au sens de Tarski. Davidson ne montre pas qu'une théorie de la signification doit être extensionnelle. Il n'établit pas que le connecteur « si et seulement si » est le connecteur extensionnel approprié pour prendre la place de « est *T* »<sup>613</sup>. Il se sert d'une définition tarskienne de la vérité qui, chez Tarski, ne fait rien de plus que de définir la vérité pour un langage donné<sup>614</sup>. Une définition tarskienne de la vérité n'est pas à proprement parler une théorie de la signification puisqu'elle cherche seulement à définir la vérité relativement à un langage. Elle fournit dans le meilleur des cas ce que l'on appelle une « inférence vers la meilleure explication », qui nous suggère de traiter les réquisits de Tarski pour une définition de la vérité comme ceux qui permettent de satisfaire le mieux les conditions constitutives

<sup>612</sup> Davidson, 1993b, 1967, p.49-50.

<sup>613</sup> Comme le remarque Platts, 1980, p.55-56.

<sup>614</sup> Tarski, 1944, 1972.

(de compositionnalité, de vérifonctionnalité, etc.) sous lesquelles le prédicat de vérité peut être attribué aux énoncés du langage.

Par conséquent, le fait de se servir d'une théorie tarskienne a pour effet de pousser Davidson à privilégier le calcul des prédicats du premier ordre afin de construire une théorie compositionnelle de la signification pour les langues naturelles. En effet, puisque la sémantique tarskienne est d'abord et avant tout un outil adapté à la définition des conditions de vérité d'un langage quantificationnel extensionnel du premier ordre, le programme de Davidson va de pair avec la possibilité d'extensionnaliser plusieurs phrases du langage qui semblent à première vue être intensionnelles, et de représenter leur forme logique à l'aide d'un langage quantificationnel du premier ordre. Davidson ne prétend pas que toutes les phrases soient réductibles à des phrases extensionnelles, mais il cherche néanmoins à démontrer l'efficacité du calcul des prédicats du premier ordre pour la représentation des langues naturelles. C'est dans cette optique qu'il faut se placer pour apprécier la théorie démonstrative de la citation, la théorie parataxique du discours indirect, et celle des attitudes propositionnelles, l'analyse des énoncés singuliers de cause et celle des énoncés d'action. Ces théories ont pour effet d'extensionnaliser des énoncés qui ont toutes les apparences d'énoncés intensionnels.

L'exemple le plus frappant d'une telle stratégie est donné par la théorie parataxique du discours indirect en vertu de laquelle les énoncés du discours indirect qui sont en tant que tels un exemple parfait d'énoncés intensionnels sont analysés comme des couples d'énoncés extensionnels. Ainsi, un énoncé comme « Galilée dit que la Terre tourne » s'analyse en « La Terre tourne. Galilée dit cela ». Le deuxième énoncé contient une expression démonstrative qui désigne l'énonciation du premier énoncé. Les deux énoncés sont parfaitement extensionnels et peuvent être représentés à l'aide du calcul des prédicats du premier ordre<sup>615</sup>. Davidson a appliqué ce même genre d'analyse à plusieurs autres sortes d'énoncés du langage, ce qui montre la puissance du calcul des prédicats du premier ordre pour la représentation des langues naturelles.

Mais une question se pose quant à savoir si Davidson peut s'approprier totalement le travail de Tarski, étant donné la nature différente de leurs projets respectifs. En effet, soutenir qu'une théorie de la signification peut gagner à être mise sous la forme d'une théorie-T au

---

<sup>615</sup> Davidson, 1993b, 1968.

sens de Tarski, et soutenir qu'une théorie de la signification est, ou s'identifie purement et simplement avec une théorie-T au sens de Tarski, sont deux choses différentes. Davidson ne soutient pas la seconde thèse, mais seulement la première : dans certaines conditions, une théorie-T pour  $L$  peut jouer le rôle d'une théorie de la signification pour  $L$ .

Il y a un cas où cette exigence est clairement remplie : une théorie de la signification pour un langage  $L$  est bien une théorie-T, et sous des conditions tout à fait spécifiées, lorsque  $L$  est un langage formel. C'est précisément ce que Tarski a cherché à établir. En posant le problème pour le cas où  $L$  est une langue naturelle, Davidson change sa nature, mais du même coup sa proposition paraît infondée, et non pas simplement « audacieuse », si elle revient simplement à transposer aux langues naturelles les conditions de Tarski pour la construction d'une sémantique pour les langues formelles. Une des hypothèses des définitions de la vérité de Tarski est qu'elles puissent définir la vérité des énoncés du langage seulement une fois la signification des expressions sub-phrastiques et donc des énoncés du langage fixée. C'est ainsi que Tarski entend entreprendre de donner une définition du terme « phrase vraie » dans un langage, une définition qui soit « matériellement adéquate » et « formellement correcte ». Une définition de la vérité est formulée dans un métalangage, pour un langage-objet. Elle ne peut donc être formulée que pour un langage, dont la structure formelle soit déjà spécifiée, c'est-à-dire pour les langues formelles : pour les langues formelles, les attributions de référence et d'extension sont fixées par stipulation. Il est alors nécessaire de distinguer la convention T du schéma dénotationnel dans la mesure où celle-ci n'est pas strictement syntaxique mais construite à partir des concepts de signification ou de traduction. En effet, la Convention T est chez Tarski une condition d'adéquation matérielle d'une définition de la vérité. Elle requiert qu'une définition de « phrase vraie » (pour  $L$ ) ait comme conséquences des phrases de la forme :

(T)  $S$  est vrai (dans  $L$ ) si et seulement si  $p$

qui sont déduites des axiomes de la théorie et qui seront donc les théorèmes d'une théorie de la vérité. Appelons désormais celles-ci des « phrases-T ». Elles sont empiriquement testables. Ceci répond à la condition (b)<sup>616</sup>. Selon Tarski, la Convention T implique que la phrase «  $p$  » du métalangage située à droite du biconditionnel soit une traduction de la phrase «  $s$  » du

---

<sup>616</sup> Nous évoquions plus haut cette condition comme suit: (b) nous permettre de *décider*, pour toute phrase arbitrairement choisie, quelle sa signification.

langage-objet (la langue étudiée) située à gauche. Ce type d'équivalence rendrait ainsi compte du sens usuel du mot « vrai ». Mais qu'il y ait une telle traduction, et que nous la connaissions, suppose évidemment que le langage-objet soit déjà compris par nous. Il semble alors que la Convention T suppose, pour pouvoir fonctionner comme critère d'adéquation, que nous comprenions le langage-objet antérieurement à l'élaboration de la théorie. Par conséquent, la convention T de Tarski considère la signification ou la traduction comme des notions primitives servant à expliquer le concept de vérité. Or, Davidson ne requiert rien de tel. Présupposer la notion de traduction, ou celle d'identité de signification, ce serait présupposer la notion de signification, et violer la condition d'immanence<sup>617</sup>. Davidson ne requiert ainsi qu'une chose : que les phrases de gauche et celles de droite aient la même valeur de vérité, en sorte que le biconditionnel correspondant soit vrai :

Dans l'œuvre de Tarski, les phrases-T sont tenues pour vraies parce que le côté droit du biconditionnel est supposé être une traduction de la phrase dont les conditions de vérité sont établies...Ce que je propose est de renverser l'ordre d'explication : en supposant la traduction donnée, Tarski était capable de définir la vérité; l'idée est de traiter la vérité comme primitive et d'extraire une analyse de la traduction ou de l'interprétation... Il n'y a pas de difficulté à reformuler la Convention T sans faire appel au concept de traduction : une théorie acceptable de la vérité doit impliquer, pour chaque phrase *s* du langage objet, une phrase de la forme : *s* est vraie si et seulement si *s* l'est<sup>618</sup>.

Par conséquent, afin de ne pas présupposer ce dont il est question, la théorie de la signification de Davidson ne peut pas être définie à l'aide de théories, ou phrases-T au sens de Tarski. Il semble alors que la Convention T ne puisse plus être un critère d'adéquation, puisque, antérieurement à la théorie, nous n'avons aucune idée de la corrélation à établir entre les énoncés du langage-objet et les énoncés de notre langage (le métalangage). Le point de la Convention de Tarski était de vérifier que la théorie nous donne comme conséquences les bonnes équivalences-T, que nous connaissons déjà. Mais, si c'est à la théorie elle-même qu'il revient de nous apprendre ces équivalences, on ne voit plus bien à quoi on pourrait la

<sup>617</sup> Laurier, 1983 : une théorie de la signification ayant pour structure celle d'une théorie tarskienne de la vérité doit respecter la condition d'immanence, selon laquelle la phrase-T ne doit pas contenir de concepts sémantiques, hormis celui de vérité, qui ne sont pas déjà présents dans la phrase de la langue objet, sous peine de pétition de principe étant donné l'interdéfinissabilité des concepts sémantiques; il serait effectivement malvenu d'utiliser le concept de signification dans une théorie qui vise à expliquer et à rendre compte de la signification.

<sup>618</sup> Davidson, 1993b, 1973a, p.134, p.199, 1993b, 1974a, p.150 et 1993b, 1977a, p.204.

confronter pour tester son adéquation. Bref, dans le cas où la théorie de la vérité sert à l'interprétation d'une langue étrangère, la Convention T a-t-elle encore un sens<sup>619</sup> ?

Pour dramatiser la question, voici l'opinion de Kripke : quand le langage-objet n'est pas déjà compris, y a-t-il place pour la Convention T ? La réponse de Kripke est non :

La Convention T de Tarski ne s'applique que quand le langage-objet en question est un langage que nous comprenons déjà, peut-être par quelque procédure informelle de traduction dans le langage naturel (c'était le cas pour les langages quantificationnels à l'époque où écrivait Tarski) ou par quelque autre définition de la vérité<sup>620</sup>.

La réaction de Davidson face à ce problème a été, non pas d'abandonner l'invocation de la Convention T, mais de proposer de la « lire autrement ». La Convention T n'exigera plus que la théorie soit confrontée avec les équivalences-T, mais demandera que la théorie soit assujettie à des contraintes telles qu'on soit raisonnablement justifié à penser que les énoncés qu'elle produit sont de quelque manière des « traductions » des énoncés du langage-objet :

Puisque Tarski voulait définir la vérité, et opérait avec des langages artificiels où la stipulation pouvait remplacer l'illumination, il pouvait prendre le concept de traduction pour acquis. Mais c'est justement ce qu'on ne peut assumer dans l'interprétation radicale. Aussi j'ai proposé à la place certaines contraintes empiriques sur l'acceptation d'une théorie de la vérité qui puisse être formulée sans faire appel à des concepts de signification, traduction, ou synonymie, quoique non sans une certaine compréhension de la notion de vérité. En argumentant, j'ai tenté de montrer que si les contraintes sont satisfaites par une théorie, alors les énoncés-T qui découlent de cette théorie auront en fait des traductions de S pour remplacer « p »<sup>621</sup>.

Davidson a alors proposé l'image suivante de sa relation à Tarski : il y aurait, de Tarski à lui, comme un renversement de point de départ. Sa stratégie vise, au contraire, à inverser l'ordre d'explication suggéré par Tarski. On retrouverait alors entre la théorie de Tarski et la sienne, une sorte d'échange ou de chassé-croisé des « concepts de base » : là où l'un présupposerait

<sup>619</sup> Davidson a été sensible au genre d'inquiétude que nous venons d'évoquer : la méthodologie de Tarski a-t-elle encore sa place dans la situation d'« interprétation radicale »? En témoigne entre autres ce texte de Davidson, 1993b, 1973b, p.73 : « Dans tous les cas qui comptent [c.-à-d. quand nous construisons la théorie d'une langue étrangère dans notre langage], nous ne pouvons assumer que le langage décrit et le langage de description coïncident. De fait, nous ne pouvons utiliser un critère formel de traduction sans pétition de principe concernant l'application empirique. Mais alors, que devient la Convention T ? Comment peut-on reconnaître un énoncé-T, sans parler de sa vérité ? ».

<sup>620</sup> Evans-McDowell, 1976, p.409.

<sup>621</sup> Davidson, 1993b, 1979a, p.172.

la traduction et la signification pour définir la vérité, l'autre prendrait la vérité<sup>622</sup> pour base pour accéder à la signification. Cependant, à lire rétrospectivement le passage crucial de « Vérité et Signification », nous constatons que Davidson a cru que la théorie pourrait à la fois donner le sens (les conditions de vérité) des énoncés du langage-objet, et en même temps définir un prédicat de vérité :

Le résultat plausible est :

(T) *s* est *T* si et seulement si *p*

[...]

Deux prédicats satisfaisant cette condition ont la même extension, en sorte que si le métalangage est suffisamment riche, rien ne s'oppose à ce que nous donnions « est *T* ». Mais qu'on le définisse *explicitement* ou qu'on le *caractérise récursivement*, il est clair que les phrases auxquelles le prédicat « est *T* » s'applique seront justement les phrases vraies de *L*, parce que la condition que j'ai imposée aux théories satisfaisantes de la signification est dans son essence la Convention T de Tarski qui teste *l'adéquation d'une définition formelle de la vérité*<sup>623</sup>.

Inutile d'épiloguer sur la confusion de ce texte puisque Davidson lui-même l'a reconnue ultérieurement :

« My confusion on this point is most apparent in « Truth and Meaning » [...] My mistake was to think that we could *both* take a Tarski truth definition as telling us all we need to know about *and* use the definition to describe an actual language.<sup>624</sup>».

Les théories de Tarski expliquent la vérité seulement en présupposant la signification ; par conséquent, utiliser la théorie de la vérité de Tarski pour expliquer la signification revient toujours à un raisonnement circulaire : puisque la vérité présuppose la signification (selon la théorie de Tarski) et ne peut être définie indépendamment de celle-ci, employer la vérité pour expliquer la signification équivaut encore à présupposer la signification. Mais, dans la mesure où le projet de Tarski était différent à savoir, définir un prédicat métalinguistique approprié aux langues formelles, ce point ne lui posait pas problème. En effet, pour Tarski, l'adéquation des phrases-T est jugée selon leur conformité aux relations de traduction entre les phrases du langage-objet et celles de la métalangue. Tarski soutenait que les définitions de la vérité

<sup>622</sup> Davidson, 1993b, 1979b, p.134, parle ainsi de « prendre la vérité pour base et en extraire de quoi rendre compte de la traduction ou de l'interprétation »; Davidson, 1990a, n'a pas cessé de s'expliquer avec Tarski; dans cet article, Davidson propose ainsi de conserver la « structure » de la théorie de Tarski tout en affirmant qu'il faut doter la notion de vérité d'un « contenu » qu'elle n'a pas chez Tarski.

<sup>623</sup> Davidson, 1993b, 1967, p.49-50 : nous soulignons.

<sup>624</sup> Davidson, 1990a, p.286, n.20.

devaient se restreindre aux langues formelles et reconnaissait donc ouvertement que c'est, seulement une fois que la signification des phrases située du côté gauche du biconditionnel est spécifiée, c'est-à-dire seulement une fois que la phrase située du côté droit du biconditionnel est bien une traduction de la phrase située du côté gauche du biconditionnel, que les phrases-T peuvent définir la vérité des phrases du langage-objet. Ces relations de traduction ne sont pas problématiques, puisqu'elles sont purement stipulatives. Pour tester une théorie de la vérité, il n'y a donc qu'à examiner ses théorèmes et déterminer si la phrase utilisée à la droite du biconditionnel est la bonne traduction de la phrase mentionnée à gauche. Ainsi, la théorie n'a besoin d'aucun test empirique, puisque les phrases-T sont des conséquences logiques de définitions stipulatives. La définition de la vérité est adéquate si elle entraîne toutes les équivalences de la forme de (T)<sup>625</sup>.

Davidson modifie sensiblement l'approche de Tarski, de façon à pouvoir l'appliquer aux langues naturelles. Tarski présuppose la connaissance d'un manuel de traduction du langage-objet dans la métalangue pour définir la vérité; selon l'approche de Davidson, le test proposé par Tarski pour juger de la vérité des phrases-T n'est pas approprié, puisque nous ne pouvons présupposer aucune connaissance de la signification des phrases du langage-objet. Le projet de Davidson est alors de déterminer les conditions d'application correcte d'une théorie de la signification pour un autre idiolecte et suppose, à cet effet, que la théorie soit empiriquement testable ou repose sur une base empirique. Ainsi, plutôt que de présupposer la signification, il soutient que la base empirique implique que l'interprète soit capable d'identifier, dans les attitudes et comportements de l'interprété, ceux qui peuvent être tenus avec vraisemblance pour des manifestations de la croyance en la vérité d'une énonciation (nous pouvons repérer le « tenir pour vrai »). Autrement dit, une hypothèse centrale de Davidson précise que nous pouvons identifier qu'une phrase est tenue pour vraie sans connaître le contenu de cette phrase (et Davidson soutient qu'à partir d'une phrase tenue pour vraie, nous pouvons ensuite définir son contenu), mais que pour être capable d'identifier en quoi consiste cette attitude de « tenir pour vrai », nous avons besoin au préalable de comprendre ce qu'est la vérité. Davidson n'explique pourtant pas ce qu'il entend par cette notion de vérité préalablement comprise, mais la considère néanmoins comme une notion primitive c'est-à-dire comme une notion préthéorique de vérité. Par conséquent, une fois que

---

<sup>625</sup> (T) est défini comme suit : *S* est vrai (dans *L*) si et seulement si *p*.

cette notion préthéorique de vérité est indépendamment présupposée, et que Davidson y adhère explicitement – en soutenant que la vérité est un concept non défini et primitif de la théorie de la signification – alors une théorie vériconditionnelle de la signification sur le modèle de Tarski, qu'il est possible d'évaluer empiriquement, peut être construite. La conjecture à la fois originale et controversée de Davidson est finalement qu'une théorie tarskienne de la vérité possède la forme appropriée pour une théorie de la signification.

Selon Davidson, pour qu'une théorie de la vérité au sens de Tarski puisse aboutir à une théorie de la signification pour une langue naturelle, celle-ci doit répondre à trois conditions :

1. Il doit être raisonnable, au sens d'une théorie de la vérité de Tarski, de caractériser, pour une langue naturelle, quelque chose comme la vérité sans pour autant que cette caractérisation prenne la forme d'une définition de la vérité.
2. Il doit être possible d'affirmer qu'une telle théorie est correcte sur la base des données empiriques dont on peut supposer qu'elles sont accessibles à l'interprète radical, c'est-à-dire à un interprète ne disposant d'aucune connaissance préalable de la langue à interpréter.
3. Si la théorie est tenue pour vraie, c'est-à-dire si elle répond de manière satisfaisante à 2, alors elle doit nous donner une interprétation du langage-objet c'est-à-dire pouvoir interpréter les énonciations des locuteurs du langage<sup>626</sup>.

Nous nous concentrerons presque exclusivement sur la seconde exigence, mais avant de nous y intéresser, nous exposerons brièvement les problèmes émanant des deux autres conditions. Le premier problème porte sur la possibilité même de fournir une théorie de la signification pour l'ensemble du langage naturel en ayant seulement recours aux ressources standards de la logique du premier ordre – c'est-à-dire aux phrases construites à partir de termes singuliers, de prédicats, de quantificateurs et de connecteurs propositionnels vérifonctionnels.

Le problème est que le langage naturel contient des locutions et des constructions qui résistent particulièrement à ce type de traitement – notamment, les modificateurs adverbiaux et le discours indirect<sup>627</sup>. Sous cette condition, le travail de Davidson est consacré, à part entière, à l'élaboration d'une analyse sémantique extensionnelle de ces cas problématiques. La troisième condition concerne la question plus fondamentale de savoir si nous pouvons

---

<sup>626</sup> Davidson, 1993b, 1973a, p.196.

<sup>627</sup> Davidson, 1993a, Essai 6.

adéquatement rendre compte de la signification au moyen d'une notion proprement extensionnelle de la vérité, à savoir si une définition de la vérité au sens de Tarski pour un langage donné suffit bien à produire une théorie de la signification<sup>628</sup>. Nous avons exposé plus haut les avantages d'une telle approche : elle évite la réification des significations et elle contribue à nous « donner la signification » en montrant comment est déterminée leur structure compositionnelle. Pourtant, un problème ressort inévitablement de cette approche, à savoir le fait que la signification est une notion plus riche que la vérité. Ce ne sont pas toutes les phrases vraies qui ont la même signification. Cependant, puisque la valeur de vérité d'une phrase s'identifie à son extension alors toutes les phrases vraies ont la même extension, et si donner l'extension (la valeur de vérité) d'une phrase revient à déterminer sa signification alors, toutes les phrases vraies (ou fausses) auront la même signification. Ce qui est manifestement faux, car, même si « la neige est blanche » et « le gazon est vert » sont toutes deux des phrases vraies, elles n'ont pas pour autant la même signification. C'est précisément ce genre de considération qui a conduit Frege à établir la distinction entre le sens et la référence, à distinguer la signification (ou intension) de l'extension des phrases<sup>629</sup>. Par ailleurs, la signification semble être une notion plus concrète non seulement que celle de valeur de vérité mais aussi que celle de condition de vérité. En effet, connaître sous quelles conditions une phrase est vraie revient à en connaître moins sur le contenu de la phrase que si nous déterminons ce qu'elle signifie, dans la mesure où deux phrases peuvent bien être vraies et satisfaire les mêmes conditions de vérité et pourtant ne pas être synonymes. C'est

<sup>628</sup> Foster, Fodor et Lepore le poussent à reconnaître ce problème. Cf. Foster, 1976; Fodor et Lepore, 1992, Chap.3.

<sup>629</sup> Nous avons d'ailleurs précédemment mentionné que la réponse de Quine 1977a, Chap.6; 1977b, est différente de Frege sur ce point. Selon Quine, la signification au sens d'entités intensionnelles n'a pas d'utilité explicative; tout ce que cette notion de signification est supposée expliquer – à savoir la différence de signification entre deux phrases vraies, etc. – peut l'être en faisant seulement appel aux données empiriques relatives au contexte d'énonciation. En tant que telle, cette conception de la signification devrait être éliminée. On retrouve ici une touche d'inspiration quinienne dans la réponse de Davidson même si celui-ci entend dissocier son objection de celle de Quine. Tout ce que nous pouvons dire de la signification (au sens étroit) est recueilli par les théories de la vérité et découle des conditions requises pour qu'une théorie soit empiriquement testable. Autrement, tout autre sens du concept de signification peut se trouver indéterminé du point de vue de la théorie de la signification (au sens de la théorie Tarskienne de la vérité); il ne s'agit pas, comme le souligne Davidson, 1993b, 1967, p. 46, d'un signe de faiblesse de la théorie mais plutôt d'une confusion autour de cette notion de signification : « Mon objection aux significations dans une théorie de la signification n'est pas qu'elles soient abstraites ou que leurs conditions d'identité soient obscures, mais qu'elles n'aient pas d'usage démontré. ».

notamment le cas de « Il y a un lapin » et de « Il y a une partie non détachée de lapin », ou encore, comme nous l'avons vu de « Hespérus est une planète » et « Phosphorus est une planète », qui sont des phrases qui, non seulement ont la même valeur de vérité, mais donc chaque paire est vraie (ou fausse) sous les mêmes conditions, bien que les premiers et seconds membres de chaque paire ne signifient pas la même chose. Il semble donc que la signification soit une notion plus nuancée que celle de valeur de vérité et de condition de vérité, et que la stratégie de Davidson visant à expliquer la signification en termes de l'une ou de l'autre de ces notions soit vouée à l'échec. Une autre manière de faire valoir ce même point est de supposer que la stratégie de Davidson est correcte et que la propriété sémantique, requise pour une théorie de la signification et pour chaque phrase du langage donné, s'apparente à son extension (c'est-à-dire à sa valeur de vérité). Nous devrions alors être capables de substituer, au sein de la phrase donnant la signification, une phrase correspondante dotée de la même valeur de vérité sans affecter la vérité de la phrase du métalangage qui donne la signification. Puisque « La neige est blanche » et « le gazon est vert » ont la même valeur de vérité, alors on devrait pouvoir substituer à une phrase du métalangage exprimant la signification de « La neige est blanche » telle que « 'La neige est blanche' signifie que la neige est blanche », la phrase « le gazon est vert » sans affecter la valeur de vérité de la phrase qui donne la signification. Cependant, ce n'est pas le résultat escompté, dans la mesure où de :

(1) « 'La neige est blanche' signifie que la neige est blanche ».

et de

(2) « La neige est blanche » et « le gazon est vert » ont la même valeur de vérité

on ne peut inférer que :

(3) « 'La neige est blanche' signifie que le gazon est vert »

sans formuler un énoncé faux. En effet, concevoir que « 'La neige est blanche' signifie que l'herbe est verte » revient à comprendre de manière erronée la signification des termes en français. Le problème tient à ceci que, comme le dit Davidson, si nous cherchons « à batailler avec la logique du terme apparemment non extensionnel 'signifie que' nous rencontrerons des problèmes aussi difficiles que ceux que notre théorie cherche à résoudre<sup>630</sup> ». C'est ici qu'intervient la condition d'immanence, qui nous prescrit de formuler une théorie de la

---

<sup>630</sup> Davidson, 1993b, 1967, p.48.

signification sans recourir explicitement au concept de signification et par conséquent de substituer un autre prédicat, un prédicat sémantique au « signifie que », sans pour autant renoncer à l'idée que la phrase de droite du métalangage établisse la signification de la phrase de gauche du langage-objet :

La théorie aura fait son travail si elle fournit, pour chaque phrase du langage étudié, une phrase correspondante (remplaçant « p ») qui, d'une certaine manière qui doit être clarifiée, « donne la signification » de  $s$ <sup>631</sup>.

Par conséquent, l'approche extensionnelle de Davidson requiert seulement que nous recueillions un biconditionnel vrai ou que nous pouvons raisonnablement tenir pour tel; et la vérité d'un biconditionnel n'exige rien de plus que les phrases de gauche et celles de droite aient la même valeur de vérité, ce qui n'a rien à voir, évidemment, avec leurs significations. Comme le fait remarquer Bennett<sup>632</sup>, si de «  $X$  signifie que  $p$  » on peut inférer «  $x$  est vrai si et seulement si  $p$  », la réciproque ne vaut pas; le biconditionnel suivant est vrai :

« La gravité varie selon la loi inverse du carré » est vrai si et seulement si la vitesse de la lumière est finie.

on n'en conclura pas que les énoncés ont même signification<sup>633</sup>. Tout ce que disent les énoncés-T en tant qu'équivalences matérielles vraies, c'est que l'énoncé cité à gauche et l'énoncé utilisé à droite ont même valeur de vérité : en conclure sans plus de précaution que l'énoncé de droite « donne les conditions de vérité » de l'énoncé mentionné à gauche est trompeur; à tout le moins, cela procède d'un usage déviant de l'expression « condition de vérité », comme Foster l'a fait remarquer :

---

<sup>631</sup> Ibid., p.49.

<sup>632</sup> Bennett, 1985.

<sup>633</sup> Id. Davidson a ainsi soutenu que du côté de la confirmation empirique de la théorie, rien de plus n'était demandé que des biconditionnels vrais : « En fait, tout ce dont on a besoin est la capacité de reconnaître quand les biconditionnels requis sont vrais » (1993b, 1967, p .61).

« I have said that these theorems, T-sentences as Davidson calls them, state the truth conditions of the sentences they refer to. But this is so only in a special sense. As ordinarily understood, the truth conditions of a sentence are those conditions necessary and sufficient for its truth: two sentences are said to have the same truth conditions just in case they would, with their meaning held constant, have the same truth value in all possible circumstances [...] But this is not the sense in which a T-sentence states truth conditions. A T-sentence does not say that such and such a structural type *would be* true (meaning held constant) in all and only those circumstances in which it *was* the case that..., but merely that, things being as they are, this structural type *is* true if and only if... It suffices for the truth of a T-sentence that what fills the blank has the same truth value as what it structurally designates, and the only sense in which a T-sentence states truth conditions is a sense whereby two sentences have the same truth *conditions* if and only if they have the same truth *value*<sup>634</sup>. ».

Davidson a reconnu le problème posé par l'usage non critique de l'expression « conditions de vérité », du moins à partir de « l'interprétation radicale » :

Mais, à la réflexion il est clair qu'un énoncé-T ne donne pas la signification de l'énoncé qu'il concerne : les énoncés-T fixent la valeur de vérité relativement à certaines conditions, mais ne disent pas que l'énoncé du langage-objet est vrai parce que ces conditions ont lieu. Pourtant si les valeurs de vérité étaient la seule chose qui compte, l'énoncé-T pour la « neige est blanche » pourrait aussi bien dire que ce dernier est vrai si et seulement si l'herbe est verte ou  $2 + 2 = 4$ , que dire qu'il est vrai si et seulement si la neige est blanche<sup>635</sup>.

Le bénéfice attendu est donc celui-ci : bien que les énoncés-T ne soient que des biconditionnels matériels, les contraintes doivent aboutir à ce que les énoncés cités à gauche soient justement ceux qui assignent les bonnes conditions de vérité<sup>636</sup>. Par conséquent, bien que les notions de valeurs de vérité et de conditions de vérité puissent paraître, aux yeux de Davidson, plus primitives que la notion de signification, il se peut également que cette qualité soit leur talon d'Achille<sup>637</sup>.

<sup>634</sup> Foster, 1976, p.11

<sup>635</sup> Davidson, 1993b, 1979b, p.138; voir aussi, p.134.

<sup>636</sup> Ibid., p.134 : « L'espoir est qu'en soumettant la théorie dans sa totalité à des contraintes formelles et empiriques appropriées, les énoncés-T individuels serviront en fait à donner les interprétations. ».

<sup>637</sup> Davidson tente de rendre compte de cette troisième condition en maintenant d'une part, une théorie vériconditionnelle qui soit compatible avec la logique du premier ordre et en soutenant d'autre part, que la contrainte holistique portant sur l'attribution de phrases-T et la référence des termes singuliers réduira de manière significative l'incidence des théorèmes déviants. Il s'agit ici de mettre en évidence l'idée selon laquelle la signification et l'interprétation d'une phrase comme « La neige est blanche » est reliée à d'autres phrases comme « Les nuages sont blancs », « La neige est froide », etc., de telle sorte qu'une phrase-T comme « 'La neige est blanche' est vraie si et seulement si le gazon est vert », dont les axiomes sont « 'neige' réfère à gazon » et « 'est blanche' s'applique à tous et seulement aux objets

La troisième condition implique donc une question métathéorique, à savoir : quelle forme devrait prendre une théorie de la signification (au sens étroit) qui spécifierait la signification de toutes les phrases d'une langue naturelle ? Une théorie de la vérité est-elle suffisante pour construire une théorie de la signification ? La seconde condition, quant à elle, soulève davantage un problème épistémologique qui peut se formuler comme suit : en supposant qu'une théorie de la vérité est une théorie de la signification, nous est-il alors possible de connaître empiriquement que nous disposons bien d'une théorie correcte de la vérité (signification) pour un langage donné ? La seconde exigence mentionne que, si une théorie de la vérité au sens de Tarski vaut comme théorie de la signification, alors nous devons être capable de donner une interprétation du langage-objet à l'aide des phrases-T, c'est-à-dire de pouvoir évaluer empiriquement la vérité de ces phrases-T. Davidson aborde ainsi cette question d'ordre épistémologique *via* la théorie de l'interprétation radicale; il pense ainsi qu'il s'agit du test le plus approprié pour juger de l'adéquation épistémologique d'une théorie de la signification. Si une telle théorie satisfait aux contraintes que nous avons spécifiées, alors elle ne doit disposer *a priori* d'aucune idée de ce que peut être une théorie de la signification ; nous devons donc pouvoir identifier la signification des phrases d'une langue donnée, sans présupposer une définition explicite de cette notion. Afin de satisfaire cette condition, c'est-à-dire éviter de présupposer une connaissance préalable de la signification des phrases du langage-objet (comme par exemple formuler des hypothèses homophoniques) au cours de la procédure d'évaluation de la théorie de la signification eu égard à un langage donné (notamment dans le cas où le langage objet est contenu dans le métalangage et où la procédure interprétative consiste à inscrire à droite l'énoncé cité à gauche pour construire l'équivalence-T demandée<sup>638</sup>), Davidson avance que nous devons avant tout examiner les conditions requises afin de déterminer empiriquement si une théorie de la signification s'applique correctement à une langue qui demeure totalement inconnue à l'interprète<sup>639</sup>. Par

---

de couleur verte », donnerait lieu à des phrases-T manifestement erronées ou incohérentes avec l'ensemble des phrases-T déjà construites. Cf. Davidson, 1993b, 1976.

<sup>638</sup> C'est ce que Davidson appelle « un critère formel de traduction » (1993b, 1973b) ou encore « un test syntaxique » (1993b, 1974a), C'est cette règle qu'illustrent les exemples canoniques, selon lesquels la théorie doit avoir pour conséquence, entre autres, l'énoncé suivant : « La neige est blanche » est vraie (en français) si et seulement si la neige est blanche.

<sup>639</sup> Nous ne voulons pas dire qu'interpréter un locuteur français soit différent, d'un point de vue qualitatif, d'interpréter un locuteur parlant une langue radicalement étrangère. Au contraire, l'interprétation radicale, tout comme la traduction radicale, commence *chez soi*. L'idée est davantage

exemple, la phrase du langage objet « La neige est blanche » sera interprétée à l'aide de la phrase-T comme suit : « 'La neige est blanche' est vraie (en français) si et seulement si la neige est blanche ». Nous comprenons la phrase-T située à droite du biconditionnel parce que nous comprenons le français et que nous comprenons le langage dans lequel elle est formulée. Mais, qu'est-ce qui justifie un tel présupposé ? S'il s'agit d'une phrase formulée dans un second langage qui nous est inconnu, alors la phrase-T « 'La neige est blanche' est vraie si et seulement si le gazon est vert » est correcte. Ce que nous tenons pour correct dépend de la langue à laquelle nous supposons que la phrase est rattachée, c'est-à-dire dépend des hypothèses que nous formulons quant à la signification de la phrase du langage-objet. Le problème épistémologique soulevé par la seconde condition met davantage l'accent sur la nécessité de justifier plutôt que de simplement présupposer, à partir des axiomes donnant les conditions sous lesquelles les phrases-T sont correctes, comment déterminer la signification des phrases du langage-objet (identifier les relations de significations qui sont exprimées) sans supposer d'emblée une identité de langue entre le langage objet et le métalangage. La théorie de l'interprétation radicale tente donc de garantir ce réquisit minimal, en répondant à la seconde condition à laquelle la théorie de la vérité au sens de Tarski doit répondre. Elle doit ainsi être empirique non seulement quand le langage-objet est contenu dans le métalangage et qu'elle implique des phrases-T qui donnent les conditions de vérité de chaque phrase et que celles-ci peuvent être vérifiées, mais aussi quand elle porte sur le langage naturel. Ainsi, la traduction du langage-objet dans le métalangage doit être vérifiée, et tout critère de traduction doit répondre au « préjugé de l'application empirique<sup>640</sup> ». Cette vérification consiste en la détermination de l'extension du prédicat « est vrai », c'est-à-dire à établir les phrases-T qui sont vraies, à partir des attitudes intentionnelles que l'interprète

---

de souligner que le modèle de l'interprétation radicale révèle plus nettement la nature de leurs caractéristiques communes. Davidson, 1993b, 1973a, p. 187-188, l'exprime ainsi : « Le problème de l'interprétation se pose tout autant pour notre langue que pour une langue étrangère : il se présente pour les locuteurs d'un même langage sous la forme de la question de savoir comment on peut déterminer que ce langage est le même. Les locuteurs d'un même langage peuvent s'appuyer sur l'hypothèse qu'ils doivent interpréter les mêmes expressions de la même manière, mais cela ne nous dit pas en quoi cette hypothèse est justifiée. Toute compréhension du langage de quelqu'un d'autre implique une interprétation radicale. Mais les présupposés passeront moins inaperçus si nous nous concentrons sur les cas où l'interprétation est le plus évidemment requise : l'interprétation dans un certain idiome de ce qui se dit dans un autre. ».

<sup>640</sup> Davidson, 1993b, 1973b, p.115.

manifeste à l'égard des énoncés du locuteur. Ce faisant, nous donnons les conditions de vérité des phrases, et par conséquent, leur signification.

Davidson examinera donc cette seconde condition en exploitant la relation entre les concepts de vérité et de signification. L'objection envisagée antérieurement à ce sujet touche aux problèmes qui se posent quand nous expliquons la notion de signification exclusivement en termes de conditions de vérité. Cependant, il ne s'agit pas de dire que les concepts de vérité et de signification n'ont aucun lien entre eux mais seulement que la signification ne peut pas uniquement être appréhendée en termes de conditions de vérité<sup>641</sup>. Il est donc vain de chercher à « réduire » les significations dans un langage naturel aux conditions de vérité. Mais, cela n'implique pas pour autant qu'on doive les traiter comme des notions primitives d'une sémantique, c'est-à-dire comme des entités intensionnelles, directement applicables aux phrases et aux expressions. Elles apparaîtront peu à peu, au terme d'une démarche holistique qui aura donné les conditions d'ensemble de la vérité dans un langage naturel<sup>642</sup>. Néanmoins, si nous voulons que nos théories soient encore testables, la forme d'une théorie de la vérité devra être maintenue : pas de sémantique sans conditions de vérité et pas de théorie de la signification sans sémantique. L'idée directrice d'une sémantique vériconditionnelle repose donc sur cette interdépendance de la signification et de la vérité. Celle-ci consiste ainsi à reconnaître qu'en comprenant notre langue, nous comprenons (au moins) la relation entre nos phrases et le monde : nous savons alors sous quelles conditions nous tenons une certaine phrase pour vraie. C'est ainsi que le concept de vérité est reliée au concept de signification : le concept de vérité est ce qui affecte les relations langage-monde de notre propre langage. Davidson suggère que la seule manière de progresser de manière significative dans la compréhension d'une langue étrangère est d'exploiter la compréhension que nous avons déjà de notre propre langue c'est-à-dire la manière dont celle-ci interagit avec le monde. Dans une situation d'interprétation radicale, nous avons deux

---

<sup>641</sup> La signification dans les langues naturelles n'est pas indépendante de la manière dont nous attribuons certaines attitudes propositionnelles, comme la croyance et le désir, aux locuteurs d'un langage, dans la mesure où parler un langage, c'est manifester une certaine forme de comportement intentionnel. Interpréter un langage n'est pas alors différent d'interpréter des actions, qu'on rationalise avec des attitudes des agents. La signification repose sur des croyances, tout comme l'assignation des croyances suppose qu'on connaisse le contenu des assertions, donc leur signification. Une théorie de l'interprétation devra nous dire comment on peut construire croyances et significations ensemble. Une théorie de la vérité n'en sera qu'un composant, même s'il reste essentiel.

<sup>642</sup> Davidson, 1993a ; cf. Engel, 1981.

locuteurs (un locuteur indigène et l'interprète) parlant différentes langues ou idiolectes et ne se comprenant donc pas mais nous partageons néanmoins un monde commun. Nous connaissons les conditions de vérité des phrases que nous formulons; par conséquent, si les locuteurs indigènes émettent des énonciations dans une situation qui correspond aux conditions de vérité de l'une de nos phrases, alors nous pouvons avancer l'hypothèse interprétative selon laquelle la signification de leurs énonciations est similaire aux nôtres sous ces mêmes conditions. Nous pouvons ainsi comprendre la significations de leurs phrases en partant du fait que nous parlons du même monde<sup>643</sup> que nous avons en commun et que nous communiquons tous deux en nous basant sur la connaissance de nos relations respectives du langage au monde. Autrement dit, nous pouvons en principe isoler, dans le contexte mondain de l'émission verbale, des traits de la situation globale qui rendent vraie telle de nos énonciations, énonciation que nous allons corréler avec l'énonciation du locuteur comme ayant au moins la même valeur de vérité que cette dernière, et possiblement les mêmes conditions de vérité; au bout du compte, un énoncé-T fondé sur l'observation de l'usage est produit, qui constitue la base empirique de la théorie<sup>644</sup>. C'est du moins la stratégie de base de l'interprétation radicale de Davidson; cependant la question est loin d'être aussi évidente. En tant qu'interprètes, nous ne pouvons pas directement inférer qu'à chaque fois qu'un locuteur indigène émet des sons dans des situations où nous aurions dit, qu'« il y a un lapin » que ce qu'il veut dire au moyen de ces sons est similaire à ce que nous signifions par la phrase, « Il y a un lapin ». Ce point puise sa source dans le holisme de Quine, ou dans la thèse de l'interdépendance de la signification et de la croyance, ce qui nous conduit du côté de l'autre influence majeure qui marque le travail de Davidson c.-à-d., Quine. Davidson adopte la démarche et quelques thèmes centraux de Quine; cependant il émet tout de même certaines objections à l'égard de son maître. Dans les deux prochains paragraphes, nous chercherons ainsi à préciser et clarifier leurs points de convergence et de divergence.

<sup>643</sup> La signification est donc définie, selon Davidson, dans un sens minimalement social: sa détermination exige toujours au moins la présence d'autrui – la relation entre un interprète et un interprété. À partir de la manière dont l'interprète comprend sa propre langue en interagissant avec le monde, et de l'hypothèse d'un monde commun affectant tant l'interprète que l'interprété, l'interprète tente de construire les relations langage-monde émanant des propos de l'interprété (et *vice versa*). La dimension sociale de la théorie de l'interprétation de Davidson sera discuté au §5.6.2.

<sup>644</sup> Davidson, 1993b, 1979b, p.135-139; la même analyse est présente dans Davidson 1990a, p.298-300: le concept de vérité est au fond une sorte de concept d'observation applicable à des faits linguistiques publiques.

#### 4.4. Quine et Davidson

##### 4.4.1. Similarités entre Quine et Davidson

La similarité centrale entre Quine et Davidson est leur rejet des significations « réifiées<sup>645</sup>», c'est-à-dire de toute notion fondationnelle de la signification. De plus, tous deux défendent un antiréalisme<sup>646</sup> en théorie de la signification, à savoir que le rôle d'une théorie systématique de la signification est avant tout de nous fournir une certaine analyse de ce que connaissent les locuteurs quand ils utilisent leur langage. La signification n'a donc aucune réalité en dehors des données empiriques sur lesquelles devraient reposer la traduction ou l'interprétation et des contraintes d'adéquation requises par la traduction et l'interprétation. Par conséquent, si, comme Quine l'affirme, les données empiriques et les contraintes d'adéquation régissant la traduction sont insuffisantes pour déterminer un schème de traduction unique de la langue de l'indigène dans celle du linguiste, alors la traduction est indéterminée. De plus, puisque la signification n'a de réalité autre que ce l'on pourrait abstraire de la situation de traduction radicale, elle est indéterminée. Davidson est tout autant en désaccord avec Quine quant aux données empiriques sur lesquelles reposent toute traduction ou interprétation que sur les contraintes les gouvernant, mais il défend, tout comme Quine, une position anti-réaliste en théorie de la signification. Selon lui, la signification n'a d'autre réalité que ce que l'on qu'on pourrait tirer des principes normatifs et holistiques de l'interprétation elle-même; son approche globale de la signification se rapproche donc de celle de Quine.

Cet antiréalisme sur la question centrale de la théorie de la signification requiert alors d'envisager la signification *via* les contraintes qui régissent son usage. Cependant, Quine et Davidson soutiennent, par la suite, qu'il est nécessaire de recourir à une conception radicalement externaliste de ces pratiques. L'image d'un traducteur ou d'un interprète radical, ne connaissant absolument pas la langue de ceux qu'il a à interpréter et ne disposant d'aucune donnée, si ce ne sont les énonciations et comportements des locuteurs, est la pièce maîtresse de leurs approches respectives. En effet, les seules ressources empiriques accessibles pour le

<sup>645</sup> Williams, Mi., 1999, p. 553: « Davidson's split-level approach to meaning reflects his acceptance of Quine's skepticism about reified 'meanings.' For Quine, since there are no meanings – in the head or anywhere else – meaning is just whatever translation preserves. [...] For Davidson, meaning is just whatever the practice of interpretation reveals. ».

<sup>646</sup> L'antiréalisme est une thèse sémantique qui affirme que la nature de la signification n'est pas indépendante de la connaissance que nous en avons.

traducteur ou l'interprète radical sont son expérience du comportement verbal des locuteurs indigènes, sa propre expérience de l'environnement physique (Quine) ou de l'environnement physique en tant que tel (Davidson) au sein duquel les locuteurs indigènes produisent des sons et des mouvements ainsi que la compréhension de sa propre langue. Cependant, cette situation de départ commune à Quine et Davidson d'un observateur (radicalement) externe ne provient pas de leur position antiréaliste en théorie de la signification. Wittgenstein et McDowell abordent tous deux la signification *via* la capacité qu'ont les locuteurs, quand ils énoncent certaines phrases, de connaître leur signification, et de manifester cette capacité de manière publique (et dans ce sens, ils ont une position anti-réaliste en théorie de la signification), et évitent ainsi l'approche mettant en scène un observateur extérieur en privilégiant davantage le point de vue d'un locuteur-acteur participant de l'intérieur aux « jeux de langage<sup>647</sup> ». Wittgenstein et McDowell pensent que l'approche externaliste de Quine et de Davidson ne rend pas adéquatement compte de la signification linguistique dans la mesure où comprendre la signification requiert une immersion au sein de la variété de ses usages<sup>648</sup>, réels ou possibles, c'est-à-dire d'acquérir cet arrière-plan au terme duquel un langage fait sens. Cet arrière-plan incarne ainsi cette habileté acquise par l'usage nous permettant de juger de la similarité et de la différence, c'est-à-dire de l'acquisition d'une capacité à formuler des jugements de similarité normative, celle-ci ne pouvant ainsi être réalisée qu'une fois que nous avons été initiés aux pratiques linguistiques et sociales pertinentes. Les jugements de similarité normative sont ainsi des jugements de similitude qui peuvent s'avérer tant corrects qu'erronés; ce sont des jugements de similitude sur lesquels tout ce que nous considérons comme similaire n'est pas considéré, *ipso facto*, comme tel. De tels jugements portent sur la question de savoir ce que signifie agir ou employer un concept de la même manière, sur ce que cela implique pour une expérience d'être identique à une autre expérience, et ainsi de suite. Dans la mesure où ces jugements de similitude sont tous des jugements sur lesquels nous pouvons nous tromper, ce sont des jugements de similarité

<sup>647</sup> Cet important contraste entre l'approche de la signification *via* le point de vue d'un observateur extérieur commune à Quine et à Davidson et celle *via* le point de vue d'un locuteur-acteur participant de l'intérieur aux jeux de langage mise en avant par l'approche de Wittgenstein-McDowell aura des implications sur la conception de la signification, que celle-ci soit perçue comme exclusivement publique ou comme sociale. Ce point sera discuté dans un prochain chapitre.

<sup>648</sup> Wittgenstein, 2005, §43 : « Pour une large classe de cas – bien que pas tous – dans lesquels nous employons le mot 'signification', il peut être défini ainsi : la signification d'un mot est son usage dans le langage. ».

*normative*. L'application correcte des jugements de similarité normative ne peut pas, en tant que telle, être associée aux hypothèses (ou interprétations) auxquelles nous arrivons en ce qui concerne l'activité de suivre une règle de la même manière, ou à ce que nous considérons comme étant des similarités observables. C'est précisément ce que montre l'argument autour de l'activité de suivre une règle de Wittgenstein : un problème de similarité normative apparaît lorsque nos observations sont identiques, ou lorsque nos interprétations ou hypothèses identifient correctement la relation appropriée de similarité dans le cours de l'action ou de l'application d'un concept. Cependant, en limitant les seules ressources accessibles pour une perspective externaliste aux similarités dans la formation des observations et des hypothèses (ou interprétations) entre l'interprète ou le traducteur et le locuteur indigène, le problème d'une similarité normative souligne précisément que ces ressources ou ces régularités entre les causes et structures comportementales ne peuvent jamais être suffisantes pour identifier les objets de référence. Wittgenstein et McDowell concluent ainsi que l'approche préconisant le point de vue d'un observateur extérieur, qui développe des hypothèses et des interprétations sur des similarités normatives, ne peut pas rendre compte de la dimension normative de ces mêmes jugements de similitude et doit donc être abandonnée au profit d'une approche privilégiant le point de vue du locuteur-acteur participant de l'intérieur à la pratique linguistique.

Par conséquent, si une conception antiréaliste en théorie de la signification n'exige pas de faire appel à une approche de la traduction ou de l'interprétation radicale qui préconiserait le point de vue d'un observateur extérieur, et que cette approche rencontre des difficultés avec le problème d'une similarité normative, quelle est alors la raison de la maintenir ? Davidson, mais non Quine, dont la conception a souvent été considérée comme apparentée à un nihilisme de la signification, tente de proposer une théorie adéquate de la signification pour les langues naturelles c'est-à-dire une théorie capable d'assigner une signification à toutes les phrases d'un langage donné. Davidson entend donc déterminer ce que serait une théorie de la signification, afin d'éclairer en retour la notion de signification. Une condition est ainsi requise afin d'espérer élucider cette notion et éviter toute pétition : que nous n'ayons au préalable aucune idée de ce qu'est, pour une phrase d'un langage, une bonne attribution de signification. Wittgenstein et McDowell affirment que le problème d'une similarité normative concerne une approche externaliste radicale, et ne peut donc pas être résolu par

cette approche. Quine et Davidson, en adoptant cette approche, n'ignorent certes pas ce problème. Celui-ci est d'ailleurs clairement mis en exergue par le processus d'interprétation ou de traduction radicale : dans une situation où nous avons comme sujet un locuteur radicalement étranger, parlant une langue que nous ne connaissons pas, il est alors fort probable que celui-ci ne partage pas le même espace normatif d'observations et d'applications de concepts que nous. Quine et Davidson ne pensent néanmoins pas qu'il soit nécessaire de rejeter une telle perspective, dans la mesure où tous deux sont convaincus qu'écarter une telle perspective reviendrait, d'une certaine manière, à énoncer une pétition de principe en tentant de justifier si les jugements de similarité normative des locuteurs indigènes sont appropriés au regard des observations et applications des concepts du traducteur ou de l'interprète. Wittgenstein et McDowell considèrent que ce problème d'une similarité normative exige de renoncer à la perspective exclusivement publique qui considère le point de vue d'un observateur extérieur; autrement, ce problème demeure aporétique. Mais, comme nous l'avons vu à la fin du chapitre précédent, l'autre extrémité de l'anti-réductionnisme, visant à rejeter la conception normative préconisant la perspective d'un observateur extérieur, est d'un point de vue explicatif circulaire : adopter l'approche du locuteur participant de l'intérieur « aux jeux de langage » revient alors à adopter une conception de la normativité de la signification en termes d'immersion dans nos habitudes et pratiques linguistiques et sociales. Quine et Davidson, en abordant la signification à partir de la perspective d'un observateur extérieur, tentent d'éviter l'écueil d'une circularité explicative. Par conséquent, leur approche externaliste tente de répondre au problème d'une circularité explicative. Ils reconnaissent certes que le problème d'une similarité normative demeure, dans une certaine mesure, sans réponse. Tous deux partagent la même conclusion, à savoir que l'indétermination affecte inévitablement la traduction ou l'interprétation; la perspective externe d'un observateur radical met bien en évidence ce problème. Par ailleurs, d'autres similarités entre Quine et Davidson s'étendent au-delà de ces considérations méthodologiques-ci et touchent à des thèses philosophiques plus substantielles, telles que les conséquences de leur engagement en faveur du holisme : le rejet de la distinction de l'analytique et du synthétique et la thèse de l'indétermination de la signification.

Le holisme de Quine est épistémologique, c'est-à-dire qu'un énoncé ne peut jamais être confronté à l'expérience isolément, parce que seuls des ensembles d'énoncés sont vérifiés par

rapport à la totalité de l'expérience possible. Si une expérience ne réussit pas à confirmer un énoncé donné, un énoncé observationnel ou non observationnel, cela ne signifie pas nécessairement que l'énoncé est infirmé par l'expérience; le fait qu'il le soit ou qu'il ne le soit pas dépendra des autres énoncés tenus pour vrais. Quine exprime ce point en affirmant que les conséquences empiriques d'un énoncé sont étroitement liées à l'information collatérale ou stockée, c'est-à-dire aux autres énoncés tenus pour vrais ou aux autres croyances envers lesquelles on est engagées<sup>649</sup>. De plus, étant donné que la signification d'un énoncé est la différence que cet énoncé fait par rapport à la totalité de l'expérience possible (Quine défend une position vérificationniste ou empiriste de la signification), la signification (c'est-à-dire les conséquences empiriques) d'un énoncé varie selon l'information collatérale disponible. Si cette information collatérale n'est pas déterminée et est sujet à changer, alors il en est de même de la signification et des conséquences empiriques qui lui sont rattachées, et *vice versa*. Chez Quine, le holisme épistémologique est bien l'une des prémisses conduisant à l'indétermination de la traduction : parce que l'on ne peut vérifier les phrases une à une, leur signification n'est jamais unique. Davidson défend précisément ce même point qu'exprime l'idée formulée par Quine d'une « inséparabilité de la signification et de l'information collatérale » en avançant la thèse d'une « interdépendance de la signification et de la croyance ». Cependant, nous ne voulons pas dire pour autant que Davidson ne propose pas une perspective nouvelle eu égard au holisme de Quine : il étend la portée du réseau holistique d'implications, à l'origine de la détermination de la signification et du contenu de la croyance d'un énoncé, en incluant les désirs et les actions de l'agent linguistique. Ce point sera discuté dans les lignes qui suivent.

Toutefois, avant d'en venir à ce point, une précision semble s'imposer ici : selon Davidson, le holisme n'est pas la thèse selon laquelle les significations n'existent pas, mais la thèse selon laquelle la nature des faits qui constituent la croyance et la signification n'est pas telle que leur individuation puisse s'effectuer en examinant des croyances et des phrases isolées. Alors qu'il revêt chez Quine une forme négative, le holisme de Davidson prend une

---

<sup>649</sup> Quine, 1977a, §9 : « Même dans les phrases occasionnelles types comme 'lapin' et 'Gavagai', la similitude de significations-stimulus n'est pas une candidate sans défaut à la relation de synonymie. Par exemple, l'acquiescement ou le refus de l'informateur à 'Gavagai' peut dépendre d'une information annexe (*collateral information*) acquise antérieurement et qui vient s'ajouter au stimulus excitateur actuel. ».

forme positive : le fait que pour reconnaître un élément il soit nécessaire de recourir à d'autres éléments est la garantie de la découverte d'une structure, et non la menace de la destruction du sens. L'idée qu'il existe une structure commune de la vérité et de la signification nous conduit donc aux antipodes du scepticisme ou au nihilisme de la signification qu'on a parfois prêté à Quine<sup>650</sup>. Davidson considère que les langages humains sont essentiellement inter-traduisibles<sup>651</sup>. Cette thèse découle, comme nous le verrons, non seulement de l'extension considérable qu'il donne au principe de charité, mais aussi de sa conception de l'interprétation : interpréter et communiquer est possible, non seulement parce qu'un fonds de vérités et de croyances est nécessairement commun à l'interprète et à celui ou ceux qu'il interprète, mais aussi parce que l'objectif de l'interprétation est l'accord ou la compréhension. Ses précisions apportées, venons-en maintenant aux conséquences du holisme sur la thèse de l'indétermination de la signification.

Davidson admet également les conséquences que Quine tire du holisme: le rejet de la distinction de l'analytique et du synthétique<sup>652</sup> (*a priori / a posteriori*) et l'indétermination de la signification :

« Dummett and Quine...disagree about holism, the claim that the truth of our sentences must be tested together rather than one by one. And they disagree also, and consequently, about whether there is a useful distinction between analytic and synthetic sentences, and about whether a satisfactory theory of meaning can allow the sort of indeterminacy Quine argues for. (On all of these points, I am Quine faithful student)<sup>653</sup>. ».

La conception holistique de la relation justificative entre les phrases et les expériences affirme que les énoncés individuels ne sont pas confrontés à l'expérience de manière individuelle et qu'aucun énoncé, observationnel ou non observationnel, n'est confirmé s'il est seulement mis en relation avec l'expérience de manière isolée, mais seulement lorsque

<sup>650</sup> Nous ne voulons pas dire ici que l'indétermination de la traduction conduit à ce scepticisme et à ce nihilisme. La distinction entre le holisme quinién et le holisme davidsonien est bien analysée par Heal, 1989, Chap. 5. Nous soutiendrons que le holisme davidsonien s'oppose en tout cas au scepticisme quant à la signification qu'on a tirée de certains arguments wittgensteiniens.

<sup>651</sup> Davidson, 1993b, 1973b, p. 117.

<sup>652</sup> Nous reviendrons plus en détail sur ce point dans le prochain chapitre.

<sup>653</sup> Davidson, 2001b, 1983, p.144: Notons que si l'on en juge à partir de l'emploi du terme « consequently » dans la seconde phrase, que Davidson souscrit à la structure argumentative que nous attribuons à Quine selon laquelle le holisme épistémologique est la prémisse à partir de laquelle nous arrivons à tirer les conclusions du rejet de la distinction de l'analytique et du synthétique ainsi que de l'indétermination de la signification, et non l'inverse.

associé à d'autres énoncés auxquels il est lié. Une vérité analytique (*a priori*) répond à la définition d'un énoncé dont la vérité serait en principe confirmée, peu importe l'expérience : un énoncé analytique serait simplement un énoncé qui serait confirmé « quoi qu'il arrive<sup>654</sup> ». Le holisme ne permet donc pas d'établir qu'une catégorie particulière d'énoncés possède de telles propriétés épistémiques et conclut alors à l'abandon de la distinction entre énoncés analytiques et énoncés synthétiques dans la mesure où aucun énoncé n'appartient à la première classe.

Une autre conséquence qui découle également du holisme – c'est-à-dire la relation indissociable de la signification et de l'information collatérale ou stockée (dans les termes de Quine), ou de l'interdépendance de la signification et de la croyance (dans les termes de Davidson) – est celle de l'indétermination de la traduction et de l'interprétation<sup>655</sup>. Davidson formule ce problème d'indétermination dérivé du holisme, comme suit :

L'interdépendance de la croyance et de la signification est ainsi évidente : un locuteur juge une phrase vraie en raison de ce que signifie la phrase (dans son langage), et en raison de ce qu'il croit. Sachant qu'il juge la phrase vraie, et connaissant la signification, nous pouvons inférer sa croyance; avec suffisamment d'information sur ses croyances, nous pourrions peut-être inférer la signification. Mais l'interprétation radicale reposerait sur des données qui ne présupposent pas la connaissance de significations ou une connaissance détaillée des croyances<sup>656</sup>.

Alors que pour Quine, les données accessibles au traducteur radical concernent les stimulations causant l'assentiment ou le dissentiment à des phrases, ce sont, pour Davidson, les circonstances observables provoquant l'attitude de tenir certaines phrases pour vraies, qui constituent les données rendant possible l'interprétation. « Tenir une certaine phrase pour vraie » est ainsi une attitude qui, selon Davidson, est accessible à l'interprète antérieurement

<sup>654</sup> Selon la théorie vérificationniste, un énoncé analytique est un énoncé qu'aucune donnée empirique ne peut falsifier, et, par conséquent, pour la vérification duquel aucune donnée n'est requise. La théorie vérificationniste est surtout invoquée pour proposer une méthode permettant de délimiter la classe des énoncés pourvus d'une signification qui s'appuie sur la manière dont les énoncés synthétiques sont confrontés à l'expérience, contrairement aux énoncés de la métaphysique, qu'aucune donnée ne peut confirmer ni falsifier.

<sup>655</sup> Notons que d'une part, le holisme tire sa plausibilité de l'examen des liens qu'ont, de fait, les phrases de notre langue avec les données empiriques et que, d'autre part, le holisme est, pour Quine la source principale d'une telle indétermination. Quine développe deux arguments sur la thèse de l'indétermination : l'indétermination de la signification (des phrases) et l'inscrutabilité de la référence (des termes singuliers). Nous nous intéressons ici davantage à la première qui résulte du holisme (et du caractère indissociable de la signification et de l'information collatérale).

<sup>656</sup> Davidson, 1993b, 1973a, p. 200.

à la connaissance de la signification d'une phrase. Il pourrait sembler que la notion de tenir-pour-vrai ne soit qu'une transposition de la notion quinienne d'assentiment à une phrase provoquée par des stimulations. Mais, si Davidson s'accorde avec Quine sur le fait que les données empiriques de base de la traduction et de l'interprétation concernent l'usage de phrases complètes, car seules celles-ci peuvent être utilisées de manière autonome par les locuteurs<sup>657</sup>, il insiste sur le fait que l'attitude de tenir une certaine phrase pour vraie constitue un état intentionnel<sup>658</sup>, puisqu'elle est une forme de croyance. L'attitude de tenir une phrase pour vraie résulte de deux facteurs : ce que l'agent croit et ce que la phrase signifie. Ce qui est observable, le fait que l'agent tient une certaine phrase pour vraie, est donc le produit de deux facteurs qui, eux, ne le sont pas, la croyance et la signification. Davidson soutient donc que les phrases qu'un locuteur tient pour vraies dépendent de ce qu'il croit, et que l'on ne peut supposer que le contenu de ses croyances nous soit accessible indépendamment et antérieurement à l'assignation d'une signification à ses phrases. Or, les données empiriques et psychologiques du comportement linguistique sous-déterminent radicalement la signification des phrases qui sont tenues pour vraies par les locuteurs.

Tentons donc d'illustrer cette notion d'interdépendance de la signification et de la croyance, en partant de l'idée de Quine selon laquelle l'interprétation ou la traduction radicale commence « chez nous<sup>659</sup> » à travers l'exemple suivant d'un locuteur s'exprimant dans notre idiolecte du français. Supposons que ce locuteur tient pour vrai la phrase « La terre est ronde » et que l'interprète sache qu'il croit que la terre est bien ronde, dans ce cas, l'interprète peut alors interpréter cette phrase « La terre est ronde » pour vouloir dire que la terre est ronde. Cependant, si le locuteur donne son assentiment à cette phrase mais que l'interprète sait que celui-ci ne croit pas que la terre est ronde, alors il se trouvera dans l'impossibilité d'interpréter ainsi l'énonciation de l'interprété. La phrase exprimera une proposition différente, autrement dit, si l'interprété croit que la terre est plate et affirme

<sup>657</sup> Quine, 1977a, §3, §5, 1993, p. 8-9 et Davidson, 1993b, p.133, 220, 1990a, p. 300.

<sup>658</sup> C'est, certainement là l'aspect essentiel de cette différence entre Quine et Davidson, puisque Quine s'en tient « aux manifestations de surface » comme l'assentiment et le dissentiment, qui ne relèvent pas du mental.

<sup>659</sup> Quine, 1977b, p.59 : « quand on y réfléchit bien, la traduction radicale commence *chez soi* » et il ajoute un peu plus loin : « Notre règle de traduction domestique est la règle d'homophonie qui applique chaque suite de phonèmes sur elle-même. » « Pourtant, continue-t-il, nous tempérons notre règle d'homophonie pour la concilier avec ce que Neil Wilson a appelé 'le principe de charité'. Nous continuons çà et là, de manière de rendre moins absurde son message. ».

pourtant que la « La terre est ronde », alors la locution « est ronde » dans son idiolecte pourrait bien vouloir dire plat ou encore « La terre » pourrait bien signifier ballon ou n'importe quoi d'autre ; toutes les possibilités sont ainsi envisageables. Ce ne peut donc pas être le cas que « La Terre » signifie la terre et « est ronde » signifie est ronde si ces faits « collatéraux », relatifs à ses croyances et au fait qu'il tienne la phrase « La terre est ronde » pour vraie, s'avèrent vrais.

Davidson en conclut donc que ce qu'une phrase signifie dépend de ce qu'un agent croit et qu'il n'est pas possible de connaître la langue qu'un agent parle et d'interpréter les phrases qu'il utiliserait sans individuer précisément les croyances qu'il entretiendrait pour exprimer celles-ci. Mais, le problème est que l'attitude de tenir une phrase pour vraie peut être isolée par l'interprète sans que celui-ci sache quelle croyance elle exprime, et sans qu'il sache non plus quelle est sa signification. En ce sens, tenir-pour-vrai  $p$  n'est pas la même chose que juger que  $p$ , ou qu'accepter que  $p$ , qui supposent que le sujet connaisse le contenu ou la signification de  $p$ . Supposons par exemple que je lise dans un ouvrage scientifique que « la sérotonine est dérivée du tryptophane ». Je peux savoir que cette phrase est vraie sans en connaître la signification, ni sans savoir quelle vérité elle exprime. L'attitude de tenir une phrase pour vraie sous-détermine alors radicalement la croyance. Un interprète ne peut inférer la croyance d'un interprété à partir des phrases que celui-ci tient pour vraies à moins de présupposer la signification de la phrase à interpréter : l'interprète peut, dans ces conditions, conclure que l'interprété croit que la terre est ronde à partir du moment où celui-ci tient cette phrase pour vraie, « La terre est ronde » si l'interprète juge que cette phrase-ci signifie bien que la terre est ronde. Autrement dit, soit une certaine phrase tenue pour vraie : si l'interprète ne comprend pas la signification de la phrase, il ne peut identifier la croyance que l'interprété exprime par celle-ci, et si l'interprète connaît au préalable quelle est la croyance exprimée par cette phrase, alors il peut inférer sa signification. Nous devons donc désamorcer cette impasse que constitue l'interdépendance de la croyance et de la signification à laquelle se heurte l'interprète radical si nous voulons établir une théorie de l'interprétation. En outre, en raison de l'antiréalisme de Davidson en ce qui a trait à la nature même d'une théorie de la signification et des contenus de pensée (ceux-ci ne sont pas réels mais attribués ou interprétés), cette impasse ou indétermination n'est pas seulement épistémologique mais ontologique : l'épistémologie de la signification (ce que nous savons de la signification par le

biais de l'interprétation radicale) détermine l'ontologie de la signification (le fait qu'il existe bien un fait décisif, *fact of the matter*, permettant d'établir la signification d'une expression d'un langage quelconque). Ce n'est pas juste que l'interprète n'est pas capable de connaître la signification de la phrase de l'interprété puisque, la phrase a bien une signification déterminée dans la mesure où, tout ce que l'on peut connaître de la signification d'une phrase s'arrête à ce que l'interprète radical peut nous révéler : l'interprète doit nécessairement chercher à projeter le contenu de ses pensées sur celles de l'interprété, et de ce fait, simuler chez autrui ce qu'il pense chez lui. Par conséquent, s'il est, malgré tout, incapable de déterminer la signification de la phrase, le problème revêt alors la forme d'une thèse ontologique sur l'indétermination de la signification elle-même.

Davidson pense ainsi que nous devons, comme Quine le suggérait déjà, utiliser le principe de charité si nous voulons surmonter cette impasse et comprendre comment les données dont dispose l'interprète peuvent être adéquatement employées pour construire une théorie de l'interprétation. Le principe de charité s'avère donc nécessaire en raison de l'interdépendance de la croyance et de la signification et nous incite à adopter la stratégie esquissée ci-dessus selon laquelle l'interprète ne peut faire autrement que supposer ce que l'interprété est susceptible de croire, pouvant ainsi, sur cette base, utiliser ce que l'interprété tient pour vrai pour identifier ce que celui-ci veut dire au moyen d'une énonciation. En appliquant cette stratégie à l'exemple précédent, nous avons : l'interprète suppose que l'interprété croit ce qu'il (projection de l'interprète) fait ou dit, c'est-à-dire qu'il croit que *la terre est ronde*; si l'interprété ne tient pas la phrase « La terre est ronde » pour vraie, nous avons la preuve empirique (*evidence*) que « La terre » ne signifie pas *la terre* ou que « est rond » ne signifie pas *est rond*. Le terme « la terre » veut peut-être dire *cette table* ou « est rond » signifie peut-être *est rouge*. Ces possibilités-ci d'interprétation seront réduites en considérant un plus grand nombre d'éléments provenant du réseau holistique des énonciations de l'interprété. Cependant, il est difficile de passer de plusieurs possibilités à une seule possibilité, il demeurera toujours une indétermination résiduelle inéliminable. Mais, en introduisant le principe de charité, une différence majeure entre Quine et Davidson se laisse entrevoir : celle de déterminer si la nature de son emploi est pragmatique et optionnelle, ou constitutive et nécessaire. Cette question a, à son tour, de profondes implications quant à

leurs conceptions respectives sur la nature de l'interprétation, et de ce fait sur leurs conceptions du langage et de la signification.

#### 4.4.2. Divergences entre Quine et Davidson

Avant d'aborder la différence cruciale, pour notre propos, entre Davidson et Quine, relative à la nature du principe de charité au sein de leurs conceptions de la traduction ou de l'interprétation, nous nous pencherons d'abord sur d'autres points de divergence pour lesquels Davidson a pris fait et cause :

Le terme d'« interprétation radicale » est destiné à suggérer la parenté étroite qu'il a avec le terme quinién de « traduction radicale ». Toutefois, parenté ne veut pas identité, et « interprétation » mis à la place de « traduction » indique l'une des différences : un plus grand accent mis sur ce qui est explicitement sémantique dans le premier cas<sup>660</sup>.

Les lecteurs qui mesurent à quel point cette analyse est parallèle à celle que fait Quine de la traduction radicale au chapitre 2 de *Le mot et la chose* noteront aussi les différences : [...] la notion de signification-stimulus ne joue aucun rôle dans ma méthode, mais elle est remplacée par celle de référence aux caractéristiques objectives du monde qui changent en liaison avec les changements d'attitude à l'égard de la vérité des phrases<sup>661</sup>.

« I now think it is essential, in doing radical interpretation, to include the desires of the speakers from the start, so that the springs of action and intention, namely belief and desire, are related to meaning<sup>662</sup>. ».

« [O]n my present plan, intention and intentional action won't directly explain meaning. Rather meaning, belief and desire will be treated as fully co-ordinate elements in an understanding of action... The aim is a theory for the interpretation of a speaker's word, a theory that also provides a basis for attributing beliefs and desires to the speaker<sup>663</sup>. ».

« Despite the limitations that have been identified... in Tarski's work on truth, a number of philosophers[...] have endorsed that work as embracing all of the truth's essential features. These philosophers include [...] Quine [...] I do not belong on this list, however. The basic argument, which was intended to reveal Tarski as a deflationist, can be taken in two ways: as showing that he did not capture essential aspects of the concept of truth, or as showing that the concept of truth is not as deep and interesting as many have thought. I think we must take it in the first of these two ways<sup>664</sup>. ».

<sup>660</sup> Davidson, 1993b, 1973a, p.188, n.1.

<sup>661</sup> Ibid., p. 203, n.16.

<sup>662</sup> Davidson, 2001b, 1983, p148, n.6.

<sup>663</sup> Davidson, 1980, p.2.

<sup>664</sup> Davidson, 1990a, p.287-288.

Davidson souligne quatre différences avec Quine :

1. La différence dans le contenu sémantique d'un manuel de traduction quinién et d'une théorie de la signification au sens de Davidson.
2. La différence sur les données empiriques sur lesquelles se fonde la traduction pour établir un manuel de traduction par opposition à celles sur lesquelles s'appuie une théorie de la signification et de l'interprétation de Davidson.
3. L'extension de la portée de l'interprétation et donc des implications holistiques dans la théorie de l'interprétation de Davidson par opposition à l'étroitesse du cadre empirique et des présupposés mis en œuvre (béhaviorisme et physicalisme) dans la situation de traduction radicale de Quine.
4. La différence sur la portée de la signification du concept de vérité.

1. Davidson note certes une parenté étroite entre le concept d'interprétation radicale et celui de traduction radicale mais il insiste pour souligner que « parenté » ne signifie pas « identité » entre les deux projets. Le point de divergence semble être que l'interprétation, à la différence de la traduction, met davantage l'accent sur ce qui est explicitement sémantique. Selon Davidson, l'objet d'une théorie de la signification est de déterminer la signification des énoncés du locuteur alors qu'un manuel de traduction ne fait qu'établir une corrélation systématique entre les phrases d'une langue *L* et celles d'une autre langue *L'*; mais étant donné qu'il est possible que nous ne connaissons pas la langue dans laquelle le langage-objet (à traduire) est traduit, il ne nous est alors pas nécessaire de connaître la signification de la phrase dans le langage-objet même si le manuel de traduction lui a attribuée une traduction. Davidson souligne ce point en affirmant que la traduction<sup>665</sup> est une « notion syntaxique » (qui ne rend pas compte de la signification des phrases traduites), alors que l'interprétation emploie « la notion sémantique de vérité » et vise, par conséquent, à établir une théorie de la signification<sup>666</sup> (une théorie de la signification devra reposer sur une théorie de la vérité, ou des conditions de vérité). Par exemple, un manuel de traduction du français vers l'anglais peut assigner une signification au moyen du théorème suivant : « la neige est blanche » a la même signification que la phrase allemande, « Schnee ist weiss ». On pourrait être à même de

<sup>665</sup> Quine entreprend plutôt d'analyser les conditions de la formulation de manuels de traduction d'un langage dans un autre et laisse de côté sa tentative de définition de la signification.

<sup>666</sup> Davidson, 2001b, 1983, p.149.

voir que le théorème est vrai en consultant un expert en linguistique ou un dictionnaire français-allemand; néanmoins, on ne peut ainsi pas savoir ce que la phrase en français signifie à moins de savoir ce que la phrase en allemand veut dire. Ce théorème de traduction nous permet de connaître la signification de la phrase en français seulement si nous connaissons au préalable la signification de la phrase en allemand; l'emploi d'un tel manuel de traduction ne rend donc pas compte de la signification de la phrase en français, mais la présuppose. Par conséquent, nous ne comprenons pas la phrase du langage-objet en lui assignant un théorème de traduction à moins que nous comprenions déjà la proposition exprimée par la phrase du langage-objet (telle que celle-ci est exprimée dans la phrase traduite). La difficulté tient finalement au fait que la possession d'un manuel de traduction d'une langue dans une autre ne nous permettrait pas de savoir ce que les phrases de *L* signifient si nous ne savions pas ce que signifient celles de *L'*. En d'autres termes, quelqu'un qui saurait que :

(1) « Kiler ne radi kako treba » signifie (traduit) en serbe ce que « The radiator does not work » signifie en anglais.<sup>667</sup>

n'aurait aucune idée de ce que signifie la phrase en serbe en question s'il ne savait pas ce que signifie la phrase anglaise. On peut ainsi envisager ici que Davidson ait quelque peu exagéré la différence entre la traduction et l'interprétation. Davidson semble ainsi affirmer que les théorèmes de la traduction assignent une signification aux phrases seulement en supposant la traduction dans une langue connue, mais nous notons pourtant que l'interprétation procède également de la même manière : nous établissons, sur la base de ce qu'énonce un locuteur étranger dans une occasion donnée, ce que ses mots signifient en corrélant ses significations d'expressions linguistiques avec nos significations linguistiques dans notre propre langue, c'est-à-dire en associant cette phrase étrangère à une phrase dont nous comprenons au préalable la signification. Rappelons que Davidson exige d'une théorie de la signification que sa connaissance puisse suffire pour interpréter et comprendre le langage. Il n'y a rien à objecter, de ce point de vue, à une théorie qui nous donnerait des spécifications du type :

(3) « Kiler ne radi kako treba » signifie (en serbe) que le radiateur me marche pas.

La différence entre (1) et (2) est que la phrase française qui nous donne la signification de la phrase serbe mentionnée n'est pas elle-même mentionnée; elle est utilisée. Nous la

---

<sup>667</sup> Engel, 1994, p. 15.

comprenons parce que nous comprenons le français, et parce que la relation entre la phrase française et la phrase serbe est une relation de traduction ou de synonymie, mais ce n'est pas pour autant une théorie de la traduction. Dans (1) au contraire, la phrase anglaise qui traduit la phrase serbe est mentionnée. Cela montre qu'une théorie de la signification pour  $L$  formulée dans un second langage n'est pas une théorie de la traduction. Elle implique que nous comprenons le langage dans lequel elle est formulée. Il serait absurde de supposer qu'une théorie de la signification ne puisse pas être articulée dans un langage quelconque, et que ce langage ne doive pas être compris par ceux qui connaissent cette théorie. Il ne s'agit donc pas seulement d'établir une relation entre des expressions linguistiques et leur signification, mais entre ces dernières et les locuteurs qui les utilisent. Ainsi, interpréter ce qu'un locuteur signifie en une occasion donnée implique une connaissance de ce que ses mots signifient, ou pourraient signifier, en d'autres occasions, et par conséquent de ce que les expressions de son langage signifient. En ce sens, la connaissance d'une théorie de la signification pour le langage d'un locuteur fait partie du savoir requis pour l'interprétation de ses énonciations. Selon Davidson, toute compréhension du langage suppose une capacité à interpréter. En d'autres termes, si l'interprète d'un langage a besoin d'une compétence linguistique, cette compétence pourra être décrite comme une compétence de l'interprète. Ici encore, la question posée n'est pas celle des connaissances effectives que nous mettons en jeu quand nous interprétons le discours de locuteurs étrangers, mais la question normative de savoir ce que nous devrions connaître pour interpréter. De plus, le problème n'est pas de chercher à éviter tout recours à une notion comme celle de traduction, mais de spécifier les règles précises qui nous permettraient d'obtenir une relation qui ait le même effet que la notion de traduction.

Mais, cela ne nous dit pas pourquoi il y a un problème dans l'interprétation du langage. Supposons que nous ayons une théorie de la signification pour une langue  $L$  que nous comprenons déjà. En ce cas, il nous suffira d'examiner les phrases produites par une théorie de la signification, et de voir si elles sont vraies. Il n'y aura pas de problème parce que nous pourrions vérifier si les phrases de la théorie de la signification (notre métalangage) sont vraies quand celles de  $L$  (le langage-objet) le sont, puisque la théorie de la signification sera homophonique. Mais supposons que nous ayons affaire à une langue  $L$  que nous ne comprenons pas, dont les phrases ne sont pas interprétées. Dans ce cas, il ne suffit pas que les

phrases de la théorie de la signification et celles de *L* soient homophoniques. Il faut encore que chaque phrase de *L* signifie la même chose que son homologue dans une théorie de la signification. Or c'est précisément ce que nous ne savons pas dans ce cas. Si nous voulons savoir si une théorie particulière s'applique à une langue *L* particulière, nous ne pouvons pas présupposer que les phrases de notre métalangage signifient la même chose que celles de la langue *L*, sous peine de faire une pétition de principe. C'est donc bien relativement à ce type de situation que Davidson appelle « interprétation radicale », que le problème de l'interprétation se pose : il s'agit de se placer dans la situation où nous ne connaissons pas le langage de ceux que nous avons à interpréter, et où nous ne disposons pas d'autres données que les énonciations et le comportement des locuteurs.

Néanmoins, ce problème de l'interprétation n'aboutit pas nécessairement à une pétition de principe, notamment en raison de la thèse antiréaliste de la signification à laquelle adhère Davidson (et Quine également) : une fois qu'on abandonne l'idée que la signification est une donnée objective qui existerait indépendamment de la connaissance qu'en ont les locuteurs, on soutient qu'il ne peut y avoir de contenus de significations sans manuels de traduction et sans théories interprétatives. Si ces considérations impliquent de traduire et d'interpréter dans des langues connues, rien n'est alors introduit ou supposé de manière illégitime dans de telles procédures; il n'y a pas, relativement aux données observables, de *fact of the matter*, permettant d'établir la signification d'une phrase d'une langue quelconque. Par conséquent, étant donné la similarité mise en évidence dans le paragraphe précédent entre leurs conceptions de la signification dans son ensemble, la différence entrevue ici entre traduction (Quine) et interprétation (Davidson) est peut-être d'une portée moindre que Davidson ne l'envisage.

2. Une différence plus significative entre Quine et Davidson concerne leurs amorces respectives de la traduction ou de l'interprétation, c'est-à-dire les données empiriques sur lesquelles s'appuient une théorie interprétative ou une situation de traduction radicale. Tous deux s'accordent pour poser comme point de départ du processus de traduction ou d'interprétation l'examen des phrases occasionnelles. Quine distingue ainsi dans l'ensemble des phrases occasionnelles<sup>668</sup> (« voici un lapin », « Voici un célibataire ») le sous-ensemble des phrases observationnelles (« voici un lapin ») et des phrases non observationnelles, mais

---

<sup>668</sup> Quine, 1977a, §12.

ce sont, les phrases occasionnelles observationnelles qui bénéficient d'un statut épistémologique particulier dans la mesure où elles sont en corrélation régulière avec les simulations et sont directement liées aux stimulations sensorielles qui provoquent l'assentiment ou le dissentiment du locuteur étranger. Les énoncés que Quine appelle « énoncés d'observation » dans « *Grades of Theoreticity*<sup>669</sup> » sont ceux qu'il appelle explicitement « énoncés occasionnels observationnels ». Les énoncés occasionnels observationnels sont les énoncés qui décrivent « un événement observable contemporain de l'élocution verbale<sup>670</sup> ». Un énoncé est observationnel pour un locuteur si celui-ci est disposé à donner le même verdict à cet énoncé dans n'importe quelle situation où le même ensemble de terminaisons nerveuses est déclenché (autrement dit, si la phrase est occasionnelle pour ce locuteur). Une phrase est alors observationnelle pour une communauté de locuteurs si elle est telle pour chacun d'entre eux, et si chacun donnerait le même verdict à celle-ci lorsque placé dans les mêmes circonstances. Notons que Quine attribue aux phrases occasionnelles observationnelles<sup>671</sup> un contenu de signification empirique autonome, et à elles seules. Les phrases occasionnelles non observationnelles, suscitent, quant à elles, également l'assentiment de l'informateur uniquement lorsqu'elles sont prononcées après la simulation, mais contrairement aux phrases occasionnelles observables, elles ne sont pas associées à un trait physique observable. La phrase « voici un célibataire » suscitée par l'approche d'un homme ayant cet état civil en est un exemple. Dans le cas de la phrase « Voici un lapin », les stimulations présentent à l'ethnologue comme aux indigènes, une configuration anatomique commune qui permet à notre linguiste de généraliser, par induction, à partir de quelques exemples. Rien de tel dans le cas de la phrase « Voici un célibataire » où il ne peut être question de repérer un trait observable commun :

---

<sup>669</sup> Quine, 1970, p.4.

<sup>670</sup> Gochet, 1978, p. 54.

<sup>671</sup> Davidson rejette la distinction introduite par Quine entre les énoncés d'observation et les autres. Ce qu'il rejette, c'est l'idée que certaines phrases, c'est-à-dire les énoncés occasionnels observationnelles, aient un statut épistémologique particulier, qui serait tel que leur vérité serait justifiée par les stimulations sensorielles qui provoquent leur acceptation. Mais, Davidson, 2001b, 1983, p.143, maintient tout de même qu'une distinction en termes de verdicts donnés dans des circonstances observables est possible entre les phrases occasionnelles et les autres.

Une méthode inductive analogue entreprise à propos d'une phrase [...] du type de « voici un célibataire » nous aurait enlisés dès les premières étapes. Les échantillons de stimulations appartenant à la signification-stimulus de cette phrase [...] ne présenteraient aucun trait commun qui invitât à risquer une conjecture pour les cas futurs<sup>672</sup>».

Les phrases observationnelles s'opposent ainsi aux énoncés dits « stables » ou aux phrases « perdurables » telles que « La terre est ronde », « L'herbe est verte », ou «  $2 + 2 = 4$  », parmi lesquels on trouve les phrases éternelles qui sont « des phrases qui demeurent toujours vraies ou toujours fausses, indépendamment des circonstances qui entourent leur production écrite ou parlée<sup>673</sup> » et les phrases non éternelles telles que « il pleut à Boston le 14 juillet 1970 » où « il pleut » est traité comme intemporel. L'assentiment ou le dissentiment du locuteur ne dépend donc pas du contexte ou des circonstances observables. Le locuteur étranger peut manifester son assentiment à une infinité de vérités dans une situation donnée sans qu'il soit possible pour le traducteur radical de déterminer, parmi les stimulations à la suite desquelles le locuteur accepte ou refuse une phrase, celles qui provoquent son verdict, autrement dit celles qui sont causalement responsables de ce verdict : seules ces stimulations peuvent servir d'indices concernant la signification de la phrase. En effet, un locuteur donné peut être disposé à accepter durant toute la journée une phrase concernant le fait que le soleil est levé, et ce, peu importe la stimulation, pourvu que certaines stimulations pertinentes aient à un moment donné déclenché cette disposition. Une façon de distinguer les stimulations pertinentes de celles qui ne le sont pas est de fournir l'une à la suite de l'autre deux stimulations telles que, suite à la première, le locuteur refuse la phrase, tandis que lorsqu'il est soumis à la seconde, il l'accepte. Ces considérations permettent de définir de manière plus précise la notion de signification-stimulus. En effet, pour Quine, la traduction de ces phrases occasionnelles-ci repose sur la notion de « stimulation stimulus » (signification stimulus), affirmative ou négative d'une phrase donnée, c'est-à-dire la classe de toutes les stimulations qui provoquent l'assentiment ou le dissentiment de l'indigène. Ainsi, une stimulation  $\alpha$  appartient à la signification-stimulus affirmative d'une phrase  $S$  pour un locuteur donné si, et seulement si, il existe une stimulation  $\alpha'$  telle que si le locuteur est soumis successivement à  $\alpha'$  et à  $\alpha$ , et qu'on lui demande après chaque stimulation s'il accepte  $S$ , alors il donne son refus dans le premier cas, et son assentiment dans le second. La signification-stimulus

---

<sup>672</sup> Quine, 1977a, p. 46.

<sup>673</sup> Quine, 1975, p. 26.

négative est définie de la même façon, en interchangeant « refus » et « assentiment ». La signification-stimulus d'une phrase est alors la paire ordonnée des significations-stimuli affirmatives et négatives de cette phrase. Selon Quine, la signification-stimulus est la donnée objective à partir de laquelle le traducteur radical peut produire son manuel de traduction. Mais, bien que de telles stimulations soient des données purement physiques (définies en termes d'excitation de terminaisons nerveuses, c'est-à-dire de patrons, de structures d'irradiation ou d'irritations chromatique de l'œil<sup>674</sup>), ce sont des événements en tant qu'elles sont des descriptions de stimulations sensorielles qui font réagir le locuteur, et non des événements ou des descriptions du monde lui-même. Quine défend ainsi non seulement une position anti-réaliste eu égard aux prémisses régissant de l'intérieur la théorie de la signification qu'une thèse anti-réaliste quant à la nature de la signification.

Cependant, pour Davidson, ces données ne peuvent pas être purement comportementales. Elles ne sont pas, en particulier, les conditions de stimulations sensorielles qui, selon Quine, « déclenchent » l'assentiment ou le dissentiment face à une phrase, et qui constituent les raisons empiriques qu'un locuteur a de donner son assentiment ou son dissentiment. En fait, Davidson rejette l'empirisme quinié qui suppose que soient associées à chaque type de phrase des données sensorielles proximales qui la vérifient. Selon lui, les significations des phrases du langage d'un individu doivent être rattachées aux objets et événements distaux qui figurent dans l'environnement commun à l'interprété et à son interprète<sup>675</sup>. L'interprète suppose que ce sont les objets saillants dans cet environnement sur lesquels portent les phrases de l'interprété, sans recourir à l'idée que des stimulations sensorielles confirment ces phrases. Ce sont donc les objets du monde extérieur, tels qu'ils sont donnés intersubjectivement, qui servent à l'interprète à établir la référence des phrases. L'interprète va supposer que l'interprété entre dans une relation causale avec ces objets, mais

---

<sup>674</sup> Voir Quine, 1993, p.2. Dans le *Mot et la chose*, Quine identifie les stimulations à des patrons ou structures d'irritation. Par exemple, une stimulation visuelle est une structure d'irradiation chromatique de l'œil (1977a, p. 27-28, p. 81). Les stimulations sont donc quelque chose d'extérieur au sujet : une stimulation visuelle peut en principe être captée par une plaque photographique. Dans ses écrits ultérieurs toutefois (1977b, p. 157-160), les stimulations nerveuses sont explicitement caractérisées en termes de déclenchement de terminaisons nerveuses appartenant au sujet. Voir Follesdal, 1975, 1995 pour certaines remarques dans le même sens.

<sup>675</sup> Les stimuli proximaux, dans la terminologie behavioriste, sont ceux qui affectent causalement les récepteurs sensoriels de l'individu, les stimuli distaux sont ceux qui partent des objets et événements de l'environnement au sujet.

il n'y a aucune raison de postuler des intermédiaires entre le sujet et ces objets, comme des stimulations sensorielles. Cela veut dire que Davidson rejette la notion quinienne de signification-stimulus<sup>676</sup>. C'est pourquoi, il est trompeur de parler ici de « données » si cela doit suggérer que Davidson souscrit à une théorie vérificationniste de la signification. Par ailleurs, l'accent mis par Davidson sur le concept central de vérité (et de conditions de vérité) comme notion sémantique fondamentale plutôt que sur celui de signification-stimulus (et de conditions de vérification) contribue à lier étroitement la signification aux objets saillants du monde, et non aux expériences individuelles. Par conséquent, alors qu'il est un anti-réaliste en ce qui concerne la nature de la signification, il ne semble pas défendre un point de vue anti-réaliste en ce qui a trait aux prémisses régissant de l'intérieur la théorie de la signification : la notion d'expérience ou de vérification ne joue aucun rôle et la signification d'une phrase est donnée par ses conditions de vérité. Davidson s'écarte donc de Quine en ce qu'il rejette la notion de signification-stimulus et l'idée même de stimulations sensorielles comme intermédiaires entre le langage et le monde et conçoit le langage comme directement en contact avec les objets et événements du monde.

3. L'interprétation radicale est une entreprise de plus large envergure que la traduction radicale : il ne s'agit pas simplement d'établir une corrélation entre les phrases du locuteur indigène et le langage du traducteur, il est plutôt question de rendre adéquatement compte des conditions de vérité de chaque phrase émise par le locuteur indigène et d'identifier le contenu de ses croyances et de ses désirs afin d'expliquer ses actions. Nous n'interprétons donc pas seulement les contenus ou les significations des expressions linguistiques, mais aussi des contenus de pensée, en attribuant à autrui et à nous-mêmes des attitudes propositionnelles, telles que des croyances, des désirs, des souhaits ou des espoirs. Davidson pose ainsi le problème de l'interprétation dans le cadre d'une théorie de la pensée et de l'action : il cherche à arriver à une « théorie unifiée de la signification et de l'action<sup>677</sup> », et l'interprétation est bien une activité totalement unifiée dans la mesure où à son holisme du langage ou de la signification<sup>678</sup>, selon lequel la signification des expressions simples et des phrases d'une langue dépend de la signification d'autres, voire de toutes les phrases et expressions de cette

---

<sup>676</sup> Voir Davidson, 1993b, 1979c.

<sup>677</sup> Davidson, 1980, 1985, 1990a.

<sup>678</sup> Engel, 1994, p.264.

langue, s'ajoute ce que l'on peut appeler son holisme de l'interprétation, venant renforcer sa thèse d'une interdépendance des significations, croyances, désirs et actions des individus en affirmant que l'on ne peut interpréter les unes sans interpréter les autres, et que l'on ne peut interpréter un contenu sémantique ou psychologique sans le rattacher à d'autres contenus. Une phrase n'a ainsi jamais de signification indépendamment du contexte de l'ensemble des autres phrases d'un langage, et les expressions qui les composent n'acquièrent de sens que dans le contexte de toutes les phrases où elles figurent. Rappelons donc ce que nous avons appelé le holisme du langage de Davidson :

Si la signification des phrases dépend de leur structure, et si nous ne comprenons la signification de chaque élément de la structure que par abstraction à partir de la totalité des phrases dans lesquelles il figure, alors nous ne pouvons donner la signification d'une phrase quelconque (ou d'un mot) qu'en donnant la signification de toutes les phrases (et mots) du langage. Frege disait que ce n'est que dans le contexte d'une phrase qu'un mot a un sens; dans la même veine, il aurait pu ajouter que ce n'est que dans le contexte du langage qu'une phrase (et par conséquent un mot) a une signification<sup>679</sup>.

À cette thèse du holisme du langage se joint donc l'idée que les significations sont étroitement liées aux croyances, désirs et actions du locuteur. Ce holisme de l'interprétation est alors étroitement associé au principe de l'interdépendance des croyances et des significations et il conduit Davidson à considérer comme indissociable le holisme sémantique (portant sur le contenu des significations) et le holisme du mental (portant sur le contenu des croyances). En effet, le problème de l'interprétation linguistique, comme nous l'avons vu, et celui de l'interprétation des pensées et des actions sont plus que similaires. En effet, la plupart du temps, nous ne savons pas ce qu'un agent fait, quelles actions il accomplit, ni quelles croyances et désirs il a, si nous ne sommes pas en mesure d'interpréter ce qu'il dit, et si nous ne savons pas ce qu'il fait, désire et croit. Le holisme de Davidson implique finalement d'étendre l'idée de Quine d'interdépendance des croyances et des significations aux croyances, significations, désirs et actions. Davidson insiste régulièrement sur ce point : un état mental n'est pas une croyance si nous ne pouvons pas le relier à d'autres croyances, désirs et actions. Mais, relier une croyance à d'autres croyances et à d'autres états mentaux, c'est découvrir une structure rationnelle minimale dans la créature, un ensemble quelconque de raisons qu'elle a de croire ce que nous lui attribuons, ou d'agir de telle façon. Une théorie

---

<sup>679</sup> Davidson, 1993b, 1967, p. 48.

de l'interprétation peut donc être représentée comme un triangle dont les trois sommets sont le langage, la pensée et l'action; on ne peut déterminer l'un des facteurs sans déterminer les autres. L'ensemble du projet de Davidson s'inscrit dans cette triple perspective, et c'est pourquoi ce qu'il vise n'est pas seulement une théorie du langage, mais « une théorie généralisée du langage et de l'action » qui met en jeu ces trois dimensions. La théorie de la signification linguistique de Davidson est ainsi tout aussi bien une théorie générale du contenu mental qu'une théorie de l'action linguistique et de l'action non linguistique. Nous en dirons plus là-dessus dans la section suivante.

4. Comme nous l'avons mentionné au §4.2, Davidson prend ses distances à l'égard d'une conception déflationniste de la vérité<sup>680</sup>, ce qui contraste avec l'attitude dédaigneusement déflationniste de Quine quant à la vérité, celui-ci considérant qu'il n'y a rien de plus dans les notions de vérité et de signification que des platitudes et ôtant toute portée significative et explicative à la notion de vérité. La vérité est, selon Quine, une notion expressive et non explicative : la vérité n'est pas une propriété substantielle que toutes les phrases vraies ont en commun en vertu du fait qu'elles sont vraies. Pour une phrase donnée, le prédicat de vérité est contenu dans son usage expressif comme un « instrument de décitation » pour une ascendance ou montée sémantique<sup>681</sup> (« semantic ascent »), c'est-à-dire n'est pleinement clair que lorsqu'est possible la décitation, à savoir la suppression des guillemets qui l'entourent et donc le passage d'un énoncé du statut de simple citation à celui d'énoncé affirmé sérieusement<sup>682</sup>. Nous pouvons, par exemple, traiter la théorie correspondantiste de la vérité de façon déflationniste, « décitationnelle » pour être précis : au lieu de dire « la neige est blanche », si et seulement si c'est un fait que la neige est blanche, nous effaçons le « c'est un fait que » et décrétons que : « la neige est blanche » si et seulement si la neige est blanche. Ainsi, attribuer la vérité à cette phrase, c'est attribuer la blancheur à la neige. L'attribution de vérité ne fait donc qu'effacer les marques de citations, c'est dire que le prédicat de vérité est superflu lorsqu'on l'attribue à une phrase donnée. Quine décrit ainsi ce qu'il reste de la conception de la vérité comme correspondance une fois

<sup>680</sup> Davidson, 1990a, 1996, p. 263-278.

<sup>681</sup> Quine, 1977a, p.371.

<sup>682</sup> Cf. Quine, 1993, p.116-119. Le terme « affirmation » est utilisé par Austin, 1994, p.96 : « Une phrase est faite *de* mots, une affirmation est faite *en* mots. [...] Nous parlons de *mon* affirmation, mais de *la* phrase française (si une phrase est de moi, je l'ai construite, mais je ne construis pas d'affirmations). ».

cette dernière débarrassée de la phraséologie philosophiquement suspecte dans laquelle elle est ordinairement plongée, comme suit :

Qu'est-ce qui du côté des énoncés vrais est supposé correspondre à quoi du côté de la réalité ? Si nous cherchons une correspondance mot à mot, nous nous retrouvons à créer dans la réalité un supplément d'objets abstraits fabriqués aux fins de la correspondance [...] un énoncé est vrai si et seulement s'il rapporte un fait. Mais là encore nous avons fabriqué une substance pour une doctrine vide. Le monde est plein de choses, reliées de façon variées, mais que sont, en plus de tout cela, les faits ? Ils sont projetés à partir des énoncés vrais pour les fins de la correspondance.

Mais réfléchissons à cette manœuvre un moment. La vérité de « la neige est blanche » est due, nous dit-on, au fait que la neige est blanche. L'énoncé vrai « la neige est blanche » correspond au fait que la neige est blanche. L'énoncé « la neige est blanche » est vrai si, et seulement si, c'est un fait que la neige est blanche. [...] La combinaison « c'est un fait que » est vide et peut être éliminée. « C'est un fait que la neige est blanche » se réduit à « la neige est blanche ». Notre explication de la vérité en termes de faits se réduit maintenant à ceci : « la neige est blanche » est vraie si, et seulement si, la neige est blanche<sup>683</sup>.

Citer l'énoncé pour dire qu'il est vrai, revient ainsi à affirmer cet énoncé lui-même, la fonction du prédicat est simplement décitationnelle : il annule la référence linguistique, ou la montée sémantique, opérée du côté gauche des équivalences. On pourrait penser qu'il est donc inutile ; mais, il est utile dans le cas de phrases qui ne sont pas des données ; il est aussi utile pour universaliser une proposition. Quel est donc le rôle exact du prédicat de vérité dans l'expression de ces généralités ? Si le besoin de généralité nous conduit à une montée sémantique, le rôle du prédicat de vérité est de « nous ramener sur terre », pour reprendre l'expression de Rivenc<sup>684</sup>. On le voit sur les énoncés du type

« La neige est blanche » si et seulement si la neige est blanche

où l'attribution de la vérité à un énoncé cité a un effet décitationnel, c'est-à-dire que l'attribution de la vérité à l'énoncé cité est « équivalente » à l'affirmation de cet énoncé lui-même. Ce cas particulier d'attribution de la vérité à un énoncé cité n'est qu'une illustration, particulièrement claire, du fait que le rôle plus général de la vérité est « l'annulation de la référence au langage » dans notre discours sur le monde<sup>685</sup>. Et c'est dans les contextes d'expression de la généralité comme ceux que nous mentionnerons ci-dessous que cet effet rend le prédicat de vérité indispensable :

<sup>683</sup> Quine, 1992, p. 213.

<sup>684</sup> Rivenc, 2008.

<sup>685</sup> Quine, 1975, p. 24.

Nous pouvons affirmer un énoncé singulier simplement en l'énonçant, sans nous aider de guillemets ni d'un prédicat de vérité; mais si nous voulons affirmer une collection infinie d'énoncés que nous ne pouvons délimiter qu'en parlant de ses membres, nous sommes bien contents de trouver le prédicat de vérité. Il nous le faut pour rétablir l'effet qu'a la référence à des objets, lorsque, en vue de quelques généralisations, nous avons recours à la montée sémantique<sup>686</sup>.

L'apport de Quine est donc d'avoir identifié l'utilité des attributions de vérité dans l'expression de certaines généralisations. Il y a ainsi, selon Quine, des circonstances dans lesquelles « bien qu'ayant en tête une réalité d'ordre non linguistique, nous devons procéder indirectement et parler des énoncés<sup>687</sup> ». Pourquoi cette « montée sémantique » est-elle nécessaire ? Elle répond, nous dit Quine, au besoin d'exprimer certaines généralisations :

Nous pouvons généraliser sur « Tom est mortel », « Richard est mortel », etc., sans parler de vérité ou d'énoncés. Nous pouvons dire : « Tous les hommes sont mortels ». [...] Quand par ailleurs nous voulons généraliser sur « Tom est mortel ou Tom n'est pas mortel », « La neige est blanche ou la neige est non blanche », et ainsi de suite, nous nous élevons jusqu'à parler de vérité et d'énoncés en disant : « Tout énoncé de la forme 'p ou non p' est vrai », ou « Toute disjonction d'un énoncé avec sa disjonction est vraie ». Ce qui nous contraint à cette montée sémantique, ce n'est pas que « Tom est mortel ou Tom n'est pas mortel » porterait de quelque façon sur des énoncés, tandis que « Tom est mortel » et « Tom est Tom » porterait sur Tom. Ces trois énoncés portent tous sur Tom<sup>688</sup>.

Qu'est-ce, alors, qui nous contraint à cette montée ? La généralisation sur « Tous les hommes sont mortels » peut s'écrire :

« x est mortel pour tous les hommes x » - c'est-à-dire tous les objets x d'une espèce qui est telle que « Tom » est le nom de l'un d'entre eux. Or que serait l'interprétation analogue de la généralisation de « Tom est mortel ou Tom est non mortel » ? Elle s'écrirait « p ou non p pour tous les objets p d'une espèce qui est telle que des énoncés en sont des noms ». Mais les énoncés ne sont pas des noms, et cette interprétation est incohérente, car elle emploie « p » à la fois dans des places qui appellent des membres de phrase et dans une place qui appelle un substantif. Donc, pour parvenir à l'assertion générale que nous cherchons, nous montons d'une marche et nous parlons des énoncés : « Tout énoncé de la forme 'p ou non p' est vrai<sup>689</sup> »

Par conséquent, dans certains cas, lorsque par exemple, nous voulons généraliser une propriété, et non pas un nom, nous sommes obligés de recourir à une ascendance sémantique,

<sup>686</sup> Ibid., p. 25.

<sup>687</sup> Ibid., p. 23.

<sup>688</sup> Ibid., p. 23.

<sup>689</sup> Ibid., p. 23.

en universalisant sur un domaine dans lequel il y a des objets : le domaine du langage; on généralise alors sur des objets sémantiques. Le prédicat de vérité est alors, selon Quine, un intermédiaire entre le monde et les mots. Ce qui est vrai, c'est la phrase, mais sa vérité consiste dans le fait que le monde est comme la phrase dit qu'il est. La phrase doit donc dire comment est le monde pour qu'on lui attribue le prédicat de vérité. On comprend bien ici l'usage du prédicat de vérité à la phrase comme permettant l'ascendance sémantique : on applique le prédicat de vérité à la phrase qui dit l'état du monde, plutôt qu'à l'état du monde. On parle alors de ce qui parle du monde plutôt que du monde lui-même, mais cela n'est qu'une procédure sémantique particulière qui permet encore de parler du monde, de façon « indirecte ».

De plus, Quine, en se basant sur sa conception éliminativiste de la signification et son approche déflationniste de la vérité, soutient qu'il n'existe pas de lien explicatif fort entre la vérité et la signification. Quine donne les exemples suivants dans l'article « Vérité » de *Quiddités* :

On ne peut ainsi se passer de [l'adjectif « vrai »] lorsque nous disons que certaines (ou toutes les) phrases ayant telle ou telle forme spécifiée sont ou ne sont pas vraies, ou que l'assertion de quelqu'un - qu'on est dans l'impossibilité de citer - était ou n'était pas vraie, ou que les lois contre la diffamation ne s'appliquent pas aux assertions vraies, ou lorsqu'on jure de dire la vérité, toute la vérité et rien que la vérité<sup>690</sup>.

À l'opposé, pour Davidson, la notion de vérité est une notion explicative centrale à laquelle toute théorie de la signification doit avoir recours. La vérité explique la signification : connaître la signification d'une phrase, c'est connaître les conditions dans lesquelles cette phrase est vraie c'est-à-dire identifier les états de choses qui doivent être réalisés dans le monde pour que cette phrase soit vraie. Nous entrevoyons donc ici le lien entre d'une part, une première différence sur la conception de la vérité et d'autre part, une seconde différence eu égard aux données empiriques sur lesquelles s'appuient une théorie interprétative ou un manuel de traduction pour déterminer le contenu. Étant donné que selon Davidson, la vérité joue un rôle explicatif central dans la détermination de la signification, on comprend qu'il doive lier étroitement la détermination du contenu des phrases aux états de chose ou objets du monde dans la mesure où ce sont ces objets saillants qui permettent à l'interprète d'établir la référence des phrases. De la même manière, il n'est pas surprenant de

---

<sup>690</sup> Davidson, 1992a, p. 214.

voir Quine lier la signification des phrases aux expériences du locuteur qui les énoncent et non aux conditions données intersubjectivement dans le monde. Pour Quine, la vérité, comme le contenu (compris en termes de conséquences expérientielles publiquement observables) est indéterminée en vertu de principes holistiques, et ne peut qu'avoir un statut non explicatif et de ce fait, déflationniste. De plus, en raison du statut déflationniste conféré à la vérité, les états de chose, événements ou objets du monde extérieur ne jouent aucun rôle dans la détermination du contenu.

Michael Williams<sup>691</sup> soutient qu'il s'agit là d'une autre différence entre Quine et Davidson qui devrait être considérée comme plus apparente que réelle. Williams, en distinguant entre deux sens distincts de la théorie de la signification, en vient à affirmer que les théories de la signification de Davidson doivent être comprises plutôt dans un sens étroit et déflationniste qui spécifie et exprime la signification des phrases particulières en termes de leurs conditions de vérité que dans un sens plus large qui vise à expliquer le concept de signification par le biais d'autres concepts voisins comme celui de vérité. La vérité possède davantage un rôle expressif qu'un rôle explicatif, la signification des phrases est exprimée en donnant ses conditions de vérité (l'ensemble des phrases-T). La notion de vérité a donc dans ces théories-ci de la signification une fonction proprement expressive, tout comme le préconisent les déflationnistes<sup>692</sup>. Williams<sup>693</sup> en conclut alors qu'il n'y a rien de tel dans la forme d'une théorie davidsonienne<sup>694</sup> qui s'apparenterait à une définition de la vérité dans le style de Tarski et qui imposerait à la vérité d'avoir un sens autre que déflationniste. Davidson ne s'intéresse d'ailleurs pas juste à la forme que devrait prendre une théorie de la vérité, mais également aux contraintes spécifiques permettant de déterminer les conditions suivant

<sup>691</sup> Williams, Mi., 1999, p. 546: « To be sure, Davidson's views diverge from Quine's in important respects. But, it is far from obvious that the differences mandate a wholly distinct attitude to truth. In my view, they do not. »

<sup>692</sup> Ibid., p. 556-557.

<sup>693</sup> Williams, Mi., note: « Davidson argues [...] we cannot both treat a Tarski-style truth definition as an empirical theory of meaning for a natural language and as telling us about truth, because we cannot assume an ability to recognize what T-sentences to count as correct [...] But for Davidson, the main problem is to explain the empirical basis for a Tarski-style theory's being taken to give correct truth conditions for the sentences of some target language. » (Ibid., p. 555).

<sup>694</sup> Davidson, 1993b, p. xiv: «... alors que Tarski analysait le concept de vérité en faisant appel (sous le nom de Convention-T) au concept de signification (au moyen de l'identité de signification ou de traduction), j'ai en tête l'idée inverse. Je considère la vérité comme le concept primitif central, et j'espère qu'en examinant en détail la structure de la vérité, nous arriverons à une théorie satisfaisante de la signification. ».

lesquelles de telles théories sont correctes pour un langage quelconque; autrement dit, il s'intéresse à la manière dont on peut confirmer une interprétation du langage sur une base empirique suffisante. Par conséquent, pour que Williams développe adéquatement son argument, il doit démontrer qu'aucune de ces contraintes ne justifie l'emploi d'une notion de vérité plus riche que celle d'une notion de vérité déflationniste. Le principe de charité est ainsi la contrainte majeure portant sur le contenu empirique d'une théorie; Williams va donc chercher à montrer que le principe de charité est compatible avec la notion d'une vérité déflationniste<sup>695</sup>. Le principe de charité, comme nous le préciserons par la suite, revêt différentes formulations, une de ces formulations en termes de vérité prescrit aux interprètes de supposer que ceux qu'ils interprètent tiennent en grande partie des croyances vraies. Le principe de charité implique également que l'interprète forge ses hypothèses à partir de ce qu'il tient pour vrai (ses croyances) afin de repérer les phrases que l'interprété tient pour vraies (leurs croyances) et comprendre la signification des phrases de l'interprété. On peut alors dire que le principe de charité prescrit d'attribuer à l'interprété non pas ce qu'il est, en général, et pour tout individu possible, rationnel et correct de croire, mais ce qu'il est correct de croire, *pour lui interprète*. Par conséquent, puisque la vérité telle qu'elle est décrite dans le principe de charité est toujours une vérité sur laquelle on projette notre propre univers doxastique, la réelle contrainte sous-jacente au processus sur l'interprétation demeure l'accord<sup>696</sup>; autrement dit, ceux que nous interprétons doivent entretenir le maximum de croyances similaires aux nôtres. Par conséquent, la charité nous demande d'attribuer au locuteur indigène le maximum de croyances possibles que nous tenons, nous-mêmes, pour vraies. Cette maxime se lit ainsi : n'attribuez pas, en grande partie, au locuteur étranger la croyance que *p* à moins que *p*, la croyance que *q* à moins que *q*, et ainsi de suite. Cette maxime, comme le fait remarquer Michaël Williams, emploie le prédicat « vrai » comme un instrument de généralisation, tel que le préconise une approche déflationniste. Soutenir que l'accord repose sur un présupposé charitable c'est-à-dire que nous devons nous mettre

---

<sup>695</sup> Williams, Mi., 1999, p. 560-562.

<sup>696</sup> Davidson, 1993b, 1973a, p. 203 : « ce qu'on réalise en assignant des conditions de vérité à des phrases étrangères qui donnent raison aux locuteurs dont c'est la langue lorsqu'on peut supposer que cela est possible, en fonction naturellement de notre propre conception de ce que veut dire avoir raison. Ce qui justifie la procédure est le fait que le désaccord comme l'accord ne sont intelligibles que sur fond d'un accord massif. ».

d'accord sur le maximum de vérités possibles, ne va absolument pas à l'encontre d'un point de vue déflationniste sur la vérité.

Par conséquent, la procédure proposée par Davidson se base, comme on le voit bien sur les idées suivantes. Tout d'abord, une théorie de la signification repose, et est constituée en partie, par une théorie de la vérité. Elle n'est pas seulement cela, car Davidson admet, comme Quine, que plusieurs théories de la vérité équivalentes pour un langage *L* rendraient aussi bien compte des données, et qu'en ce sens l'interprétation restera indéterminée. Ensuite, comprendre un langage, c'est l'interpréter. Plus exactement, la compréhension qu'a le locuteur du langage *L* est fournie par la théorie que l'interprète extérieur lui assigne. Enfin, cette théorie de l'interprétation n'est possible que si le locuteur et l'interprète sont supposés partager, et avoir en commun, un grand nombre, voire une majorité de croyances qu'ils tiennent pour vraies. C'est précisément le sens du principe de charité, qui n'est pas une simple maxime heuristique de l'interprétation, mais une précondition sans laquelle celle-ci ne peut avoir lieu. La vérité n'est donc pas, du point de vue du principe de charité, indépendante de l'accord et celle-ci peut très bien être comprise en termes déflationnistes, c'est-à-dire comme possédant seulement une utilité expressive. Par conséquent, Williams conclut que les théories de la signification de Davidson ne requièrent du point de vue de leur forme et des contraintes gouvernant une application correcte des concepts rien de plus qu'une vérité déflationniste. Cette quatrième différence entre Quine et Davidson paraît donc quelque peu exagérée.

En ayant ainsi dégagé les différences entre Davidson et Quine sur lesquelles Davidson ne manque pas d'insister (et dont certaines sont discutables), nous allons, à présent, nous intéresser à la différence la plus significative de leurs deux théories.

#### 4.5. Le principe de charité de Davidson

Quine et Davidson soutiennent donc tous deux que le holisme mène à une impasse quant à la détermination de la signification des énonciations de l'interprété ou du locuteur indigène, et proposent de surmonter cette impasse en employant le principe de charité. Par conséquent, le principe de charité nous permet de répondre à la seconde condition épistémologique d'adéquation qui s'applique à la théorie de la signification : il nous fournit une manière d'amorcer l'interprétation, ou en ayant déjà adopté un schéma interprétatif, une manière de

tester empiriquement une théorie de la signification afin d'examiner si notre interprétation (attribution de phrases-T) des énonciations d'autrui est correcte. En tant que principe régissant la traduction ou l'interprétation, le principe de charité incite le traducteur ou l'interprète à être charitable en assignant des significations aux énonciations d'autrui, c'est-à-dire en exigeant de l'interprète qu'il attribue, le plus possible, des croyances vraies et/ou rationnelles (logiquement cohérentes) à ses locuteurs. La différence majeure entre Quine et Davidson provient de la nature du principe de charité c'est-à-dire de la force de l'obligation avec laquelle il est prescrit d'interpréter en étant charitable. Cette différence dans leur compréhension respective de la nature du principe de charité indique une différence profonde dans leurs conceptions de l'activité linguistique : la différence entre d'un côté, une conception naturaliste faiblement normative et de l'autre, une conception néo-Kantienne fortement normative. C'est ainsi dans la compréhension de la portée significative du rôle du principe de charité au sein du processus interprétatif que Davidson manifeste sa divergence par rapport au naturalisme béhavioriste et au pragmatisme déflationniste de Quine et son engagement à l'égard d'une pratique linguistique fortement normative.

Nous devons donc maintenant approfondir et étoffer le contenu du principe de charité. Jusqu'à présent, nous sommes restés à dire que le principe de charité impose d'interpréter charitablement c'est-à-dire, que l'interprète doit attribuer à l'interprété le maximum de croyances et de significations cohérentes qu'ils tiennent tous deux pour vraies. Examinons donc les trois formulations non équivalentes que propose Davidson du principe de charité :

1. Supposons que l'interprète tienne les énonciations du locuteur qu'il interprète comme vraies c'est-à-dire, que lorsqu'un locuteur tient une phrase pour vraie, celle-ci est vraie. [Formulation de l'hypothèse ou de la contrainte de vérité]
2. Supposons que l'interprète et l'interprété partagent et aient en commun un grand nombre, voire une majorité de croyances. [Formulation de l'hypothèse ou de la contrainte de l'accord]
3. Supposons que l'interprété soit minimalement rationnel dans ses croyances et significations (selon les normes de l'interprète) [Formulation de l'hypothèse ou de la contrainte de rationalité]<sup>697</sup>

---

<sup>697</sup> Dans Davidson, 2001b, 1983, p.211. En ce sens, le principe de charité est un principe de *cohérence logique* ou de *rationalité des croyances*. Nous pouvons ainsi le formuler comme suit : (PCR)

Ces trois formulations ci-dessus appartiennent à deux classes distinctes : les deux premières se classent sous ce que nous appelons la version « large » du principe de charité, alors que la troisième appartient à ce que nous nommons la version « étroite » du principe de charité. La troisième formulation du principe de charité a un contenu étroit : elle ne prescrit pas à l'interprète d'interpréter autrui de telle sorte à ce que celui-ci tienne des croyances dotés d'un contenu spécifique. C'est plutôt une injonction à interpréter autrui afin d'éviter toute imputation d'erreurs à l'interprété par violation de principes logiques; autrement dit, elle prescrit de préserver les lois de la logique usuelle telles que « interprétez toujours de manière à ce que ceux que vous interprétez ne croient pas à la fois  $p$  et  $\neg p$  », ou « si ceux que vous interprétez croient que si  $p$  alors  $q$ , et  $p$ , alors interprétez-les toujours comme croyant  $q$  » et ainsi de suite. Les deux premières formulations, quant à elles, visent au contraire à attribuer à autrui un contenu spécifique. Elles impliquent des injonctions telles que « s'il pleut aux alentours, interprétez autrui comme tenant la croyance qu'il pleut », « si, ce locuteur étranger est amené à croire, sur la base de sa perception, dans des conditions observationnelles normales et sous une bonne visibilité, qu'il voit un lapin, alors interprétez-le comme tenant la croyance qu'il y a un lapin », « Interprétez donc toujours de manière à ce que, ceux que vous interprétez croient ce que nous considérons comme vérités empiriques évidentes telles que la terre est ronde, la neige est blanche, etc. ». Le principe de charité se divise donc en deux hypothèses distinctes : interprétez toujours de manière à ce que, ceux que vous interprétez aient des croyances logiques et cohérentes dans un sens formel (étroit) et de manière à ce qu'ils croient évidents (pour nous) nos vérités empiriques (dans un sens large).

Pour Davidson, construire une théorie de l'interprétation est bien une activité obéissant à une contrainte normative. Il est important ici de clarifier le sens à donner à cette contrainte normative. Tout d'abord, l'interprétation est normative dans la mesure où, le principe de charité sur lequel elle repose, se compose de principes normatifs, de principes incarnant des normes de rationalité telles que, des lois logiques et des principes épistémiques tels que, faire

---

*interprétez toujours de manière à rendre les croyances de ceux que vous interprétez cohérentes et non contradictoires.* Mais Davidson ne limite pas le principe de charité à cette maxime de rationalité : il l'étend à l'ensemble de la procédure interprétative (Davidson, 1993b, 1973a, p.203). Dans ce cas, le principe de charité prend la forme d'un principe de *correspondance* ou de *véridicité des croyances* prescrivant à l'interprète de « maximiser l'accord » et de « minimiser le désaccord ». Nous pouvons l'énoncé ainsi : (PCV) *interprétez toujours de manière à maximiser l'accord entre vos croyances et celles de ceux que vous interprétez.*

confiance à nos sens dans des conditions normales, et croire fermement aux vérités empiriques établies (par nos canons de la logique). Le principe de charité suppose donc que les normes que nous acceptons pour justifier nos propres croyances conditionnent toute interprétation. Tout ce que nous tenons comme correct, et nos critères mêmes de correction, doivent intervenir. Davidson veut ainsi souligner qu'il n'y a pas de différence précise entre ce que croit un interprète et ce qu'il croit que les autres devraient croire. Il projette aussi bien sa propre psychologie et ses croyances que les raisons qu'il peut avoir pour soutenir ses croyances. La distinction entre le sens normatif du principe, qui nous prescrit d'idéaliser la véridicité et la rationalité des sujets, et son sens psychologique, est destinée à s'estomper, si nous envisageons le principe de charité comme supposé guider l'interprète dans les premières étapes de son interprétation, et non pas prescrire une assignation définitive de croyances vraies et rationnelles. Le premier élément d'une contrainte normative semble avoir d'abord un statut pragmatique : le principe de charité s'apparente alors à une règle à suivre et conçoit avant tout le processus d'interprétation comme un processus dynamique, au cours duquel l'interprète peut réviser ses attributions ou hypothèses initiales, que la suite confirmera ou infirmera.

Pourtant, le principe de charité n'est pas seulement une règle méthodologique, que nous pourrions décider de suivre ou de ne pas suivre, une maxime utile dont on pourrait comparer les avantages et les inconvénients par rapport à d'autres maximes possibles. Davidson soutient ainsi que le principe de charité est un principe normatif *a priori*, la condition de possibilité minimale de toute interprétation et de toute compréhension. Le second élément d'une contrainte normative est donc sa thèse d'une normativité constitutive : l'idée selon laquelle le principe de charité (dans un sens étroit et large) est un principe constitutif de l'attribution de croyances et de significations :

Il ne faudrait *pas* croire que le conseil méthodologique qui consiste à interpréter d'une manière qui optimise l'accord repose sur un présupposé charitable concernant l'intelligence humaine qui *puisse se révéler faux*. Si nous ne parvenons pas à trouver le moyen d'interpréter les paroles et autres comportements d'un être comme révélant un ensemble de croyances pour une bonne part cohérentes et vraies selon nos propres normes, nous n'avons considéré aucune raison de considérer que cet être est rationnel, qu'il a des croyances, ou qu'il dit quoi que ce soit<sup>698</sup>.

<sup>698</sup> Davidson, 1993b, 1973a, p.203-204 : nos italiques. Nous notons ici que les trois différentes formulations du principe de charité sont présentes dans cette citation. Dans la première phrase, Davidson traite de l'interprétation comme prescrivant « d'optimiser l'accord », dans la seconde phrase,

Par conséquent, les principes normatifs font en sorte que le principe de charité n'a pas simplement une force pragmatique mais bien constitutive : ce n'est pas que nous devrions interpréter autrui en supposant que celui-ci respecte le principe de non contradiction, autrement dit qu'il veuille bien dire lapin par « gavagai », si nous voulons le comprendre plus facilement, mais c'est plutôt que nous devrions interpréter autrui de cette manière si nous devons l'interpréter *tout court* c'est-à-dire, si nous voulons le compter comme locuteur prenant part de manière significative à l'activité linguistique et être capables d'interpréter ses croyances et significations<sup>699</sup>. Concevoir la charité comme une contrainte constitutive de l'interprétation implique que la relation entre les principes normatifs qui la composent et les objets (c'est-à-dire les énonciations, les croyances, les désirs, etc.) que ceux-ci régissent soit très différente de la relation entre les principes empiriques ou lois naturelles<sup>700</sup> et les objets que ces derniers gouvernent. Nous pouvons, en effet, seulement dégager des lois empiriques une fois l'expérience réalisée : nous regroupons ensemble les objets d'une certaine forme, nous les soumettons à des expériences, et nous tirons ensuite des lois qui régissent leur comportement. Nous arrivons ainsi de cette manière aux conclusions suivantes quant à la nature des métaux, à savoir qu'ils sont malléables et ductiles<sup>701</sup>, qu'ils se dilatent sous l'action

---

il glisse vers l'idée qu'il ne nous est pas possible d'interpréter un individu quelconque si nous ne supposons pas qu'il est largement véridique et minimalement rationnel dans ses croyances.

<sup>699</sup> Puisque Davidson défend un anti-réalisme en théorie de la signification – tout ce que nous pouvons dire sur la signification, nous le connaissons à partir de la pratique interprétative – et puisqu'il soutient également que tout ce qui s'applique à la détermination de la croyance et de la signification dans le cas de la troisième personne s'applique aussi dans le cas de la première personne, son approche implique immédiatement des engagements ontiques quant à ce qu'est, en tant que tel, un contenu de croyance ou un contenu de signification et non simplement des engagements épistémologiques quant à la question de savoir sur quelles données empiriques se baser pour attribuer des significations et des croyances au locuteur. Dans ce qui suit, nous nous déplacerons tantôt entre des formulations épistémologiques du principe de charité au sens de principe normatif constitutif afin de clarifier ce que signifie attribuer des croyances ou considérer un individu quelconque (y compris soi-même) comme minimalement rationnel dans ses croyances, et des formulations ontiques cherchant à expliciter ce qu'est une croyance.

<sup>700</sup> Davidson, 2001b, 1991, p.217: « The normative and the causal properties of mental concepts are related. If we were to drop the normative aspect from psychological explanations, they would no longer serve the purposes they do. We have such a keen interest in the subject's reasons for acting and for his or her changes of belief that we are willing to settle for explanations that cannot be made to fit perfectly with the laws of physics. Physics, on the other hand, has an aim laws that are as complete and precise as we can make them; a different aim. The causal element in mental concepts helps make up for precision they lack; it is part of the concept of an intentional action that it is caused and explained by beliefs and desires; it is part of the concept of a belief or a desire that it tends to cause, and so explain, actions of certain sorts. »

<sup>701</sup> *Ibid.*, p. 216.

de la chaleur, etc. Les normes interprétatives gouvernant la signification des mots et le contenu des croyances procèdent différemment : nous n'avons pas recours à des expériences pour analyser les actions, croyances et désirs des individus afin d'établir des hypothèses inductives ou des généralisations tirées de nos expériences respectives à l'égard d'un grand nombre d'instances particulières de croyances, de désirs et d'actions. Les principes normatifs de rationalité sont des conditions d'adéquation constitutives des croyances. Ce sont des principes qui s'appliquent nécessairement aux croyances; on ne peut donc pas dire qu'il s'agit d'une croyance, si elle n'obéit pas à ces principes :

[L]a charité n'est pas une option, mais la condition [...] La charité nous est imposée; que cela nous plaise ou non, si nous voulons comprendre les autres, nous devons considérer qu'ils ont raison sur la plupart des sujets.<sup>702</sup>

Une créature ne peut donc pas être considérée comme un individu doté de croyances à moins que ses croyances et actions soient conformes à ces principes. C'est précisément parce que nous ne pouvons pas être sûrs de la présence de cette structure rationnelle chez les animaux que nous ne pouvons pas valablement leur attribuer des pensées et des croyances, ni même un langage. Par conséquent, interpréter autrui de manière charitable est une condition *a priori* de toute interprétation et de toute compréhension et non une hypothèse pragmatiquement utile ou empiriquement bien confirmée. Ces normes *a priori* de l'interprétation ne peuvent pas être mises en défaut, car elles déterminent les conditions auxquelles un esprit est pour nous objet d'une possible rencontre. Ce principe n'est pas une option, comme le précise ci-dessus Davidson, mais une condition de toute interprétation. Il n'est donc pas susceptible d'être réfuté par l'expérience.

En effet, selon Davidson, l'impératif ou le présupposé charitable visant à interpréter autrui, c'est-à-dire comme respectant (en grande partie) les lois logiques (version étroite) et comme croyant à la vérité de nos propositions empiriques (version large), revêt une autorité beaucoup plus forte que s'il s'agissait d'une force simplement pragmatique. Il s'agit donc bien d'une condition *a priori* de l'interprétation, à savoir d'une condition transcendantale de la possibilité même de toute interprétation du langage ou d'attribution de croyances. Ce ne sont donc pas des normes d'interprétation qui « peuvent se révéler fausses » mais bien des normes constitutives de l'attribution de pensées : si une créature est rationnelle, alors elle doit

---

<sup>702</sup> Davidson, 1993b, 1974b, p. 287.

suivre, d'une manière ou d'une autre, les principes logiques ou les principes bayésiens, c'est-à-dire être cohérente dans ses croyances et ses préférences, qui doivent être closes sous une certaine relation déductive, ou respecter des principes de formation de ses croyances, comme le principe d'information totale (*total evidence*) selon lequel un agent doit adopter les croyances qui correspondent à l'ensemble de l'information inductive dont il dispose<sup>703</sup>. Il n'existe donc pas d'autre alternative que celle d'une interprétation selon nos normes ou d'un échec qui relègue l'interprété dans le monde de la nature. L'interprétation étant la mesure du mental, il ne peut y avoir d'échec du principe de charité interne au monde humain. Si nous ne comprenons pas les autres à partir de nos croyances et de nos principes de rationalité normative, ce ne sont pas (pour nous) d'autres esprits. Par conséquent, cette thèse d'une normativité constitutive du principe de charité exprime bien l'engagement de Davidson pour une normativité forte : si certaines normes de signification ou croyances sont constitutives au sens où elles déterminent de manière significative nos actions, alors le sens d'une application correcte se rattachant aux contenus de signification et de croyance est également constitutif, fort ou objectif. Par exemple, dire que les normes de significations ou que les croyances doivent respecter le principe de non contradiction est une condition constitutive de leur emploi, ce qui équivaut à affirmer que nous possédons bien des normes de signification pour les termes « et » et « non », qui soient objectivement correctes et qui ne peuvent donc pas être fausses au risque de ne vouloir plus rien dire du tout. La notion de statut constitutif conféré au principe de charité ne veut donc pas juste dire que ces normes de signification sont correctes dans le sens où celui-ci faciliterait leur compréhension. Le principe de charité peut ainsi tout aussi bien se lire au moyen d'une version large que d'une version étroite; Davidson a d'ailleurs souligné, dans la citation ci-dessus, que toutes deux impliquent la thèse d'une normativité constitutive. La version large, quant à elle, implique l'interprétation de concepts ordinaires empiriques. Par conséquent, soutenir que le principe de charité est constitutif de l'attribution de signification implique, par exemple, qu'il soit objectivement correct que ces termes empiriques signifient par « gavagai », lapin et non partie non-détachée de lapin. Davidson propose donc « un argument transcendantal<sup>704</sup> » portant sur le statut constitutif du principe de charité (c'est-à-dire que ceux que nous interprétons partagent nos vérités

---

<sup>703</sup> Davidson, 1993a, Essai 2.

<sup>704</sup> Davidson, 1993b, 1973b, p.117.

empiriques et admettent que leurs croyances respectent les lois logiques) quant aux contenus de croyance et de signification, celui-ci impliquant ainsi que la signification de certains termes de notre langage appartenant au vocabulaire logique et non logique soit fortement normative. Ainsi, le fait que Davidson confère au principe de charité un statut constitutif dans ses deux versions respectives, vient appuyer son argument transcendantal en faveur d'une normativité forte eu égard aux deux domaines lexicaux. Ainsi, Davidson, en mettant l'accent sur le rôle constitutif du principe de charité<sup>705</sup>, adhère à l'idée d'une normativité forte de la signification; autrement dit, à l'idée que les phrases exprimant explicitement la signification d'expressions logiques et non logiques ne sont pas simplement que des hypothèses empiriques ou des alternatives pragmatiques. Il soutient également que les attributions de significations doivent obéir à des principes normatifs de rationalité n'ayant pas d'écho dans la théorie physique et conçoit donc l'activité linguistique comme une activité exclusivement rationnelle. En effet, selon Davidson, la pratique humaine de communication verbale est beaucoup plus riche que le phénomène purement causal et naturaliste consistant à corrélérer avec régularité les images et les sons, les stimuli et leurs réponses; parler une langue contient indubitablement un élément irréductiblement rationnel. Cet élément rationnel est ainsi, selon Davidson, une caractéristique *a priori* inhérente au comportement humain et de ce fait, communiquer au moyen du langage s'avère être une activité essentiellement régie par des principes de rationalité. Par conséquent, faire de la rationalité une contrainte constitutive portant sur la détermination du contenu signifie que la raison est une caractéristique *a priori* de la signification linguistique.

La théorie interprétative de Davidson est, comme nous l'avons mentionné précédemment, un projet de plus large envergure que la traduction radicale quinienne : elle n'implique rien de moins qu'une théorie unifiée des significations, croyances, désirs et actions. Ainsi, l'extension de la portée de l'interprétation (qui inclut maintenant les croyances, désirs et actions) s'accompagne à la fois d'une extension du réseau holistique de connexions interagissant au sein de la pratique interprétative et des principes normatifs de rationalité constitutive de l'interprétation. Le holisme quinien implique que nous ne pouvons

---

<sup>705</sup> Davidson, 1985, p.92: « It should be emphasized that these maxims of interpretation are not mere pieces of useful or friendly advice; rather they are intended to externalize and formulate (no doubt very crudely) essential aspects of the common concepts of thought, affect, reasoning and action. What could not be arrived at by these methods is not thought, talk, or action. ».

pas juste confirmer ou infirmer une phrase ou une croyance particulière à partir d'une expérience individuelle, et ce principe concerne les phrases ou croyances provenant de la théorie scientifique, telles que « le papier tournesol devient rouge au contact d'un acide et bleu au contact d'une base », ainsi que les phrases issues de la théorie de l'interprétation ou de la traduction, telles que « Gavagai veut dire qu'il y a un lapin ». Le holisme de Davidson considère, quant à lui, d'étendre la thèse d'une interdépendance de la signification (des phrases) et de la croyance, à celle d'une interdépendance de la signification-croyance-désir-action. Par conséquent, le contenu d'un désir et la portée significative d'une action ne peuvent donc pas juste être extraites des dispositions comportementales; leur contenu dépend de l'interprétation d'autres phrases, croyances, désirs et actions. Par ailleurs, la détermination holistique du contenu des phrases, croyances, désirs et actions doit respecter la contrainte constitutive de rationalité – les attributions particulières de significations, croyances, désirs et actions ne doivent pas seulement être individuées sur la base d'une détermination plus large de significations, croyances, désirs et actions (holisme), mais doivent l'être de telle manière à ce que les normes rationnelles jouent un rôle constitutif dans ces déterminations-ci (la version d'une normativité constitutive du principe de charité). Par conséquent, les principes normatifs de rationalité composant le principe de charité doivent s'étendre jusqu'à inclure non seulement nos normes contraignant de manière constitutive l'interprétation des autres locuteurs afin que ceux-ci respectent les canons de la logique, mais également celles qui déterminent de manière contraignante nos désirs et nos interprétations de leurs actions afin que ces dernières soient aussi considérées comme constitutivement rationnelles. Une contrainte rationnelle constitutive de l'attribution des désirs, pouvant faire partie d'une conception élargie du principe constitutif de charité peut prendre la forme suivante : attribuer des désirs à un individu quelconque de telle sorte à ne pas lui donner un désir qui soit en son pouvoir de satisfaire mais bien un désir qu'il ne peut pas satisfaire (à moins que vous lui attribuez d'autres croyances et désirs qui rendraient cette situation rationnelle). La contrainte rationnelle constitutive de l'interprétation de l'action de telle sorte que l'attribution de significations, de croyances, de désirs et d'actions forme un ensemble totalement coordonné et rationnellement cohérent, pourrait s'exprimer ainsi : rendre compte de l'action de *X* à poursuivre le but *Y* en termes de sa croyance *B* et de son désir *D* comme suit : *X* a désiré *D* et a cru qu'en poursuivant l'action *Y*, il pourrait obtenir *D*. Il devrait ainsi être plus clair à

présent que les deux traits majeurs de la théorie de l'interprétation de Davidson sont la rationalité normative (sa version d'une normativité constitutive du principe de charité) et son holisme (au sein duquel la portée de l'interprétation s'étend outre la question de l'interdépendance de la signification des énonciations et du contenu des croyances).

Le principe constitutif de la charité soulève alors deux questions cruciales, à savoir :

1. Qu'est-ce qui justifie le bien-fondé de son emploi, c'est-à-dire comment pouvons-nous éviter un raisonnement circulaire quant à la question d'une testabilité empirique ?
2. Comment pouvons-nous rendre compte des croyances fausses et/ou irrationnelles de l'interprété si nous devons nécessairement présumer que la plupart des croyances de l'interprété sont vraies et cohérentes en vertu de la nature de l'interprétation ?

Dans les deux prochains paragraphes, nous évaluerons les réponses qu'apporte la théorie de l'interprétation de Davidson à ces deux questions. Il semble que l'origine ultime des problèmes que rencontre la théorie interprétative de Davidson provienne de la tension entre deux sources opposées, à savoir : l'héritage quinién et la reconnaissance des intuitions kantienne. L'héritage quinién envisage l'adoption de la perspective de l'observateur extérieur de la traduction ou de l'interprétation radicale, un holisme généralisé ou global et le rejet de la distinction de l'*a priori* et de l'*a posteriori*. Cet héritage le conduit vers un naturalisme *éliminativiste* ou un normativisme pragmatique (comme Quine le souligne); pourtant, Davidson a bien conscience des difficultés propres à chacune de ces deux positions et fait valoir, pour une approche adéquate de la pratique interprétative, la nécessité de prendre en considération cette dimension rationnelle et normative inhérente à toute activité humaine. Les problèmes soulevés touchent ainsi à la cohérence interne de sa position, une position qui tente de concilier les conclusions de Quine menant au naturalisme et au pragmatisme avec une tonalité plus kantienne en faveur d'une rationalité normative forte.

#### 4.6. Le dilemme de Davidson

Le problème initial auquel l'interprète radical fait face et que le principe de charité est appelé à surmonter est l'aporie générée par le holisme, ou l'interdépendance de la signification et de la croyance. Cette impasse revient ainsi à une forme de scepticisme sémantique : en se basant sur tous les faits pertinents à l'interprétation, c'est-à-dire, selon

Davidson, sur les états des objets ou des événements dans le monde ainsi que sur les phrases tenues pour vraies par l'interprété, il semble qu'il n'existe pas de manière de déterminer la signification correcte d'une phrase d'un langage quelconque (et donc des termes la composant). La seconde condition qu'une théorie de la vérité dans le style de Tarski doit satisfaire, selon Davidson, en tant que théorie de la signification, est de faire en sorte que nous puissions identifier les phrases-T correctes pour le langage-objet, c'est-à-dire que nous soyons capables de déterminer la signification des phrases, ou une fois la théorie construite, que nous soyons à même de tester notre théorie empiriquement afin de vérifier si elle attribue des assignations correctes aux conditions de vérité des phrases du langage-objet. C'est ici qu'intervient le principe de charité; il fournit un outil pour tester empiriquement la théorie de l'interprétation. Le principe de charité prescrit ainsi à l'interprète de présumer que l'interprété tient, pour la plupart, des croyances similaires aux siennes, ou des croyances, pour la plupart, vraies ou rationnelles. Par conséquent, si ce principe nous permet de déterminer la signification des phrases, alors il semble que l'on puisse tirer de ce principe une réfutation *a priori* du scepticisme. Davidson l'applique systématiquement au processus d'interprétation. Il le conçoit comme le vecteur directeur aidant l'interprète à projeter ses propres standards de vérité sur l'ensemble des phrases tenues pour vraies par l'interprété, de telle sorte qu'il favorise au maximum la cohérence et la justesse dans les croyances qu'il lui attribue. En fait, « [l]'objectif du principe est de rendre le locuteur intelligible<sup>706</sup> » à savoir : que l'interprète parvienne à passer de la vérité à la signification en établissant une corrélation entre les phrases tenues pour vraies par l'interprété et celles qu'il tient lui-même pour vraies et en associant les circonstances dans lesquelles l'interprété donne régulièrement son assentiment aux conditions de vérité de ces phrases. Ce point soulève alors la question suivante : ne sommes-nous pas, en adoptant cette hypothèse interprétative, en train de formuler une pétition de principe ? Sur quels fondements est-il légitime d'établir cette présomption de rationalité ou de vérité ? Il semble bien qu'il y ait une impasse entre la croyance et la signification dans la mesure où il apparaît quelque peu circulaire de prétendre que si l'on connaît, au moyen du principe de charité (*via* l'hypothèse de rationalité et de vérité), les contenus de croyance que les phrases expriment alors, nous serions en mesure d'interpréter ou de comprendre les énonciations de quelqu'un. Davidson répond à ce problème en

---

<sup>706</sup> Davidson, 1993b, 1967, p. 52.

conférant au principe de charité un rôle constitutif et en attribuant à la présomption de vérité et de rationalité un statut *a priori*. Nous soutiendrons cependant que cette réponse met en avant la notion d'analyticité que Davidson n'est pourtant pas autorisé légitimement à défendre, dans la mesure où une telle position va directement à l'encontre de son engagement en faveur d'une conception holistique et de ce fait, de sa position préconisant l'abandon de la classe particulière des vérités analytiques (*a priori*)<sup>707</sup>.

Davidson soutient ainsi que l'emploi du principe de charité n'est pas circulaire dans la mesure où ce principe n'est pas une hypothèse empirique, c'est-à-dire une hypothèse « qui pourrait se révéler fausse<sup>708</sup> » mais un principe constitutif de l'interprétation intervenant dans l'explication et la compréhension des comportements linguistiques et non linguistiques des agents. Par conséquent, si de tels comportements ne respectent pas le principe de charité, alors il ne nous est plus possible de les interpréter : en effet, s'ils ne se conforment pas au principe de charité, c'est donc la preuve que nous ne pouvons pas interpréter leur langage et non que le principe de charité est erroné. Le principe de charité a donc un statut *a priori* ou transcendantal dans l'interprétation – il ne peut alors pas simplement être révoqué par l'expérience dans la mesure où il s'agit bien d'une condition de possibilité de toute activité susceptible d'interpréter les pensées, croyances, désirs, et actions d'autrui. De plus, puisque le principe de charité (dans sa version étroite) est composé de normes de rationalité, tel que le principe de non contradiction; dire qu'il est constitutif de l'attribution de significations, de croyances et de désirs revient à affirmer que la rationalité (c'est-à-dire la cohérence logique) est constitutive de l'attribution de ces mêmes entités – nous ne pouvons attribuer les significations, croyances, désirs et actions intentionnelles à autrui à moins que le contenu de ses énonciations, croyances et désirs s'articule rationnellement, ou soit cohérent, et *vice versa*<sup>709</sup>. Par conséquent, affirmer que le principe de charité est une contrainte constitutive de l'interprétation implique d'admettre, comme une condition *a priori* de l'interprétation, l'hypothèse selon laquelle ceux que nous interprétons sont rationnels.

---

<sup>707</sup> Affaiblir la thèse d'une normativité constitutive du principe de charité (comme nous le soutenons dans la prochaine section) élimine cette incohérence.

<sup>708</sup> Davidson, 1993b, 1973a, p. 203.

<sup>709</sup> Nous avons énoncé ce point en évoquant la formulation de l'hypothèse de rationalité dans le cadre d'une version étroite du principe de charité, mais nous aurions très bien pu l'exprimer au moyen de la formulation de la présomption de vérité ou d'accord dans le cadre d'une version large du principe de charité.

Les propos de Davidson autour d'un usage constitutif du principe de charité peuvent ainsi être rapprochés de la position kantienne sur deux points, à savoir : l'importance qu'il accorde à la dimension rationnelle et normative au sein de sa théorie de l'action (linguistique) humaine, et sa conception transcendantale du langage, des pensées et des actions d'un agent. Davidson pourrait ainsi être considéré comme néo-Kantien dans la mesure où il semble adopter la stratégie transcendantale de Kant et c'est précisément cette continuité avec Kant que nous examinerons ici. En effet, tout comme Kant produit un argument transcendantal visant à établir les conditions de possibilité de l'expérience, Davidson cherche à établir les conditions de possibilité de l'interprétation ou de la compréhension des autres (et de nous-mêmes). Alors que Kant tente ainsi de montrer que les formes *a priori* de la sensibilité ou intuitions pures (l'espace et le temps) et les formes *a priori* de l'entendement ou concepts purs (identité, causalité, etc.) sont des éléments constitutifs et nécessaires à l'expérience de tout objet de connaissance, Davidson s'efforce, quant à lui, de souligner que la présomption de vérité ou de rationalité (à savoir les lois logiques) est nécessaire à l'attribution de croyances et de significations aux contenus et comportements de l'interprété. Ce type d'arguments transcendantsaux établit des conclusions qui ont un statut synthétique *a priori* : ils sont sensés nous délivrer des vérités empiriques sur le monde sensible, qu'il s'agisse d'objets matériels (Kant) ou d'action linguistique (Davidson), et ne sont donc pas analytiques (dans la mesure où les propositions analytiques clarifient les connexions entre nos concepts et n'énoncent donc rien sur le monde); pourtant, ils sont *a priori* c'est-à-dire que les objets ou événements dans le monde dépendent pour leur vérité des conditions de possibilité de l'expérience tout en n'ayant de sens que par l'expérience. De telles conclusions ne sont donc pas justifiées par l'expérience (*a posteriori*), l'expérience ne pouvant pas les confirmer dans la mesure où elles sont, en premier lieu, les conditions qui rendent possible toute expérience, et sont donc, de ce fait, *a priori*. Mais, c'est seulement en les considérant comme un principe *a priori* ou transcendantal et non comme une hypothèse empirique ou heuristique, que le principe de charité évite de présupposer ce qui est précisément en question au sujet de l'interprétation. Autrement dit, nous pourrions penser que le principe de charité aurait tout d'une pétition de principe : en effet, si nous l'identifions à une hypothèse empirique, alors nous pourrions envisager la possibilité qu'il soit erroné; donc présumer dès le début que le principe de charité est nécessairement vrai reviendrait à formuler une pétition de principe. Par

exemple, nous ne pouvons pas présupposer, sans pétition de principe, qu'une proposition empirique, telle que toutes les émeraudes sont vertes, ne pourrait pas s'avérer fautive alors qu'il s'agit précisément d'une hypothèse empirique. Une hypothèse empirique doit donc, pour être tenue pour vraie, être confirmée par l'expérience et donc afin d'écarter, avant même la réalisation de l'expérience, la possibilité d'une expérience venant l'infirmar, elle suppose exactement ce que l'expérience est appelée à établir. Cependant, ce n'est pas, selon Kant, faire une pétition de principe que de dire d'une expérience qu'elle doit être située dans un cadre spatio-temporel puisqu'il ne s'agit pas ici d'une hypothèse empirique mais d'une proposition synthétique *a priori* qui ne peut se révéler fautive précisément parce que la proposition d'être situé dans l'espace et dans le temps est une condition transcendantale de l'expérience. Il en est de même pour le principe de charité : son statut constitutif et transcendantal de l'interprétation ne l'apparente pas à une pétition de principe. Davidson conçoit donc la charité sur le modèle kantien comme condition de possibilité de l'interprétation : la charité a alors le statut d'une proposition synthétique *a priori*. La procédure d'interprétation radicale est ainsi supposée être, chez Davidson, la seule position à partir de laquelle on puisse avoir accès non seulement aux contenus de signification et de pensée des locuteurs d'un langage, mais également aux contenus de toute pensée portant sur une quelconque réalité objective. L'interprétation radicale est bien la situation à partir de laquelle on peut envisager les conditions de possibilité d'une expérience possible, et elle est, en ce sens kantien, transcendantale. C'est donc bien ce que soutient la thèse d'une normativité constitutive du principe de charité, à savoir : être rationnel est constitutif de la croyance, de la signification, etc., tout comme pour Kant, être soumis à l'expérience par le biais des formes de la sensibilité de l'espace et du temps est constitutif de la connaissance des objets accessibles à l'expérience. Par conséquent, Davidson tente de démontrer au moyen d'une relation constitutive entre la rationalité et la croyance (ainsi qu'entre la rationalité et la signification, etc.) que concevoir les agents comme des agents obéissant à des contraintes de rationalité est une condition nécessaire de possibilité du processus interprétatif. L'articulation rationnelle de la signification des phrases et du contenu des croyances, désirs et actions doit être satisfaite, et de ce fait en faire une hypothèse clé (adopter la thèse constitutive du principe de charité) n'équivaut pas à formuler une pétition de principe quant à la question de la détermination de la croyance et de la signification.

Par conséquent, considérer le principe de charité comme une condition transcendantale de possibilité de l'interprétation, et lui reconnaître, selon nos canons de rationalité un statut synthétique *a priori*, semble peut-être permettre à Davidson de surmonter l'écueil d'un raisonnement circulaire quant à la question de l'interprétation radicale, mais risque de le mener à une contradiction : en effet, il est ainsi amené à admettre des propositions *a priori* ce qui va à l'encontre de son engagement pour un holisme généralisé ou global. Davidson défend un holisme global selon lequel la justification de toute phrase d'un langage est déterminée par la justification des phrases du langage en entier. Il suggère alors que cesse toute tentative de relier nos croyances particulières à une réalité externe, compte tenu qu'il est impossible de connaître les croyances une à une, et exige, en conséquence, que toute justification provienne du corps même des croyances. L'expérience pourrait ainsi modifier le statut justificatif de toute phrase, impliquant en retour l'absence de distinction entre propositions : des changements au statut justificatif de propositions considérés jusqu'ici comme *a posteriori* peuvent entraîner des changements dans la justification de propositions envisagées jusqu'alors comme *a priori*. En d'autres termes, un holisme global mène à un faillibilisme global c'est-à-dire à la thèse selon laquelle tout énoncé ou proposition peut se révéler faux. Mais des propositions *a priori* sont traditionnellement définies comme des propositions dont la justification est indépendante ou antérieure à l'expérience; l'expérience ne peut donc ni les justifier ni les réfuter. Par conséquent, une des conséquences du holisme global est de rejeter cette classe des propositions *a priori* et de soutenir la thèse selon laquelle tous les énoncés sont faillibles, c'est-à-dire qu'ils ne sont pas à l'abri d'une éventuelle réfutation ou confirmation empirique<sup>710</sup>. Cependant, en affirmant que le principe de charité est constitutif de l'attribution de significations et de croyances, Davidson en vient à soutenir que, la proposition selon laquelle les croyances et les significations sont rationnellement articulées, est une vérité (synthétique) *a priori*. Cette proposition (et les croyances particulières et normes de significations qui l'expriment) ne peut donc pas s'avérer fausse,

---

<sup>710</sup> Le holisme global de Quine (auquel Davidson adhère) implique ainsi le rejet des vérités analytiques. Nous y reviendrons par la suite au §5.4. Il serait ainsi adéquat d'en proposer une lecture épistémologique compte tenu du fait que, pour Quine, les vérités « analytiques » sont définies comme étant à l'abri de toute révision sur la base de l'expérience. Le holisme global de Quine implique donc un faillibilisme global : l'idée que tout énoncé peut s'avérer faux ou qu'aucun énoncé ne peut être à l'abri d'une éventuelle révision sur la base de l'expérience, ce qui implique alors l'élimination des vérités *a priori*, c'est-à-dire de vérités immunisés contre toute tentative de révision par l'expérience.

l'expérience ne peut pas nous conduire à la réviser (ou à les réviser), c'est-à-dire que nous ne devrions pas adopter, à son égard, d'attitude faillibiliste. Par conséquent une position comme celle de Davidson mettant en avant l'idée d'un statut *a priori* de la charité est incompatible avec un holisme global (et un faillibilisme global).

Nous pourrions penser qu'il ne s'agit pas là d'un problème majeur dans la mesure où, en affirmant qu'une proposition est *a priori*, Davidson ne ferait que suivre l'image chère à Quine d'un réseau de croyances (*web of belief*) étalant sa toile du centre à la périphérie et donc qu'affirmer que les croyances *a priori*, en se situant au centre du réseau de croyances, ne sont pas, comme le souligne Quine, à l'abri de la révision. Selon cette lecture, les croyances « *a priori* » de Davidson ont le même statut que les croyances « au centre » du réseau de croyances de Quine, et puisque celles-ci sont, selon Quine, sujettes à affronter le tribunal de l'expérience sensible (étant donné qu'aucune croyance n'est à l'abri de la révision, même les croyances « au centre » doivent s'y soumettre), il n'y aurait alors pas de contradiction à dire que certaines croyances sont *a priori* (dans ce sens) et qu'aucune croyance ne peut se soustraire au tribunal de l'expérience sensible. Une solution de ce genre semble ainsi être proposée par Davidson. Davidson affaiblit sa thèse d'une normativité constitutive du principe de charité, ce qui a pour effet d'« atténuer » le statut *a priori* des principes qui le composent; ces derniers s'apparentent ainsi aux propositions « profondément enracinées » de Quine, propositions qui ne sont pas totalement à l'abri de la révision. Cependant, il est difficile de poursuivre ce but dans la mesure où si nous avons d'abord attribué au principe de charité un statut *a priori* fortement constitutif au sein du processus interprétatif, c'était afin d'éviter l'écueil d'une pétition de principe vers laquelle autrement nous tendions en adoptant conjointement au sein d'une approche de l'interprétation radicale le principe de charité et le point de vue d'un observateur extérieur à la détermination de la signification et de la croyance.

Davidson fait donc face à un dilemme : soit, le principe de charité est constitutif de la détermination de la signification et de la croyance auquel cas son approche semble contredire sa thèse d'un faillibilisme holistique (avec son rejet des *vérités a priori*), soit le principe de charité n'est pas constitutif de la compréhension et son approche, bien que cohérente, conduit à un raisonnement circulaire. Le dilemme résulte ainsi du choc d'éléments incompatibles : l'élément néo-Kantien qui conçoit l'action linguistique comme contrainte *a priori* par des

normes de rationalité et l'élément naturaliste ou pragmatique Quinien qui retire à la rationalité tout rôle essentiel dans la description du comportement linguistique. Davidson semble alors devoir choisir entre, soit renoncer (ou affaiblir) à son engagement à l'égard d'une normativité constitutive du principe de charité dans la détermination de la croyance et de la signification et adopter, de ce fait, une position plus fortement naturaliste ou pragmatique de la charité au sens de Quine, soit revenir sur son rejet des vérités *a priori* et admettre que certaines croyances (comme celles composant le principe de charité) sont immunisées contre toute révision empirique.

#### 4.7. La thèse d'une normativité constitutive et le Faillibilisme : le problème de l'erreur et de l'irrationalité.

Nous souhaitons donc ici examiner de plus près le premier volet du dilemme : l'incohérence présumée entre une version large ou étroite du principe de charité envisagée comme principe constitutif de la détermination de la signification et de la croyance, et un faillibilisme global. Notre lecture néo-Kantienne de Davidson met l'accent sur son engagement à l'égard d'un idéal constitutif de vérité ou de rationalité au cœur de la procédure interprétative d'attribution de significations et de croyances; pourtant, celui-ci soulève aussi le problème de l'erreur ou de l'irrationalité. Ce problème revêt ainsi deux formes, chacune correspondant aux deux versions générales du principe de charité : la formulation de l'hypothèse de vérité ou d'accord (sous une version large du principe de charité) et la formulation de l'hypothèse de rationalité (sous une version étroite du principe de charité). Nous traiterons donc des deux versions du problème. Bien qu'apparemment, tout semble nous laisser croire qu'il faille abandonner la thèse d'une normativité constitutive de la charité, à l'origine de ce problème, il apparaît néanmoins que procéder ainsi n'aurait que pour effet de nous ramener au second volet du dilemme c'est-à-dire à une forme de naturalisme réductionniste. Nous soutiendrons donc qu'une réponse plus satisfaisante implique d'introduire une troisième formulation du principe de charité en termes d'une hypothèse ou d'une contrainte d'intelligibilité, constitutive de l'attribution et de la détermination de la croyance et de la signification. Cette solution repose certes bien sur la thèse d'une normativité constitutive mais sous une forme plus faible, autorisant ainsi un faillibilisme global, ce qui contribue à effacer la contradiction initiale. Nous demeurons néanmoins aux prises avec la

question de savoir si nous pouvons, à juste titre, abandonner ou proposer une version plus faible de notre engagement à l'égard d'un statut constitutif de la charité (dans les deux sens discutés jusqu'ici), et de ce fait, d'une normativité forte de la signification, sans tomber dans un naturalisme non normatif ou un pragmatisme faiblement normatif. Cette démarche pourra bien nous mener au second volet du dilemme. Cette question sera abordée dans le dernier chapitre.

Le principe de charité dans sa version large contraint l'interprète à supposer que ceux qu'il interprète partagent des croyances vraies (selon les normes de l'interprète). Le problème de l'erreur soulève la question de savoir comment l'interprète peut attribuer des croyances fausses (selon ses propres normes) à l'interprété, si cette présomption de vérité est constitutive de l'interprétation :

Par exemple, dans quelle mesure pouvons-nous dire que les anciens – ou du moins certains anciens – croyaient que la terre est plate ? *Cette terre* ? Or cette terre, la nôtre, fait partie du système solaire, lequel est pour nous identifié en partie par le fait que c'est une masse de corps célestes volumineux, froids, et solides qui tournent autour d'une étoile très grosse et très chaude. Si quelqu'un ne croit *aucune* de ces choses au sujet de la terre, est-il bien certain que c'est de la terre qu'il parle ?<sup>711</sup>

Le premier Davidson met conjointement en évidence la contrainte constitutive de la présomption de vérité avec la relation holistique entre croyances relativement à l'identification d'un fait objectif permettant de déterminer une croyance donnée. Nous ne pouvons en effet, identifier une croyance qu'à travers sa localisation dans une trame de croyances; c'est donc cette trame qui détermine ce sur quoi porte une croyance et ce à propos de quoi elle est une croyance. Ainsi, avant qu'un objet, ou une propriété d'un objet, faisant partie du monde, puisse faire partie de ce sur quoi porte une croyance (vraie ou fausse), il doit y avoir une infinité de croyances vraies rattachées à la croyance en question. Les croyances fausses, quant à elles, invalident la description de la croyance comme portant sur tel ou tel objet. Par conséquent, si ce qui détermine ici l'objet de notre croyance sont les autres

<sup>711</sup> Davidson, 1993b, 1975, p.247-248. Davidson introduira un autre principe dans l'interprétation afin de traiter de ce problème de l'erreur. Ce principe, le principe causal, est introduit dans « A coherence Theory of truth and knowledge » et dans ses articles ultérieurs. Comme « Pensée et Discours » a été publié dans sa version originale en 1975, alors que « A coherence Theory of truth and knowledge » l'a été en 1983, nous ferons référence aux diverses stratégies que Davidson adopte à l'égard de ce problème de l'erreur et de l'irrationalité en distinguant entre les stratégies respectives du premier Davidson et celles du second ou du dernier Davidson.

croyances que nous pouvons avoir, la possibilité d'identifier l'objet d'une croyance comme étant cette Terre dépend alors de notre capacité à attribuer de nombreuses autres croyances vraies à propos de corps célestes volumineux, froids et solides tournant autour d'une étoile très grosse et très chaude. La référence d'une croyance ou d'une expression est donc une construction théorique tirée des contraintes d'une vérité constitutive et de l'approche holistique globale de la théorie de l'interprétation. Ainsi, si l'objet auquel une croyance ou une expression est censée faire référence correspond à ce que la majorité des autres croyances que l'on peut avoir à son sujet ne tient pas pour vrai, alors il ne s'agit absolument pas de la référence de la croyance ou de l'expression. La thèse d'une normativité constitutive de la charité et la thèse holistique nous imposent donc d'attribuer des conditions de vérité aux phrases, et par là même des référents aux termes singuliers de telle sorte que la plupart des autres phrases et croyances que nous formons soient vraies (selon nos propres normes). Dans l'exemple ci-dessus des anciens, leur attribuer des croyances ou leur assigner comme référent à l'un de leurs termes singuliers, la terre, rendrait fausse un grand nombre de leurs croyances astronomiques. Par conséquent, notre théorie nous impose de les réinterpréter de telle sorte à ce qu'ils n'entretiennent pas, pour la plupart, des croyances fausses à propos de notre terre (« cette terre qui est la nôtre »), mais des croyances (vraies) portant sur un autre objet, « leur terre ». Dans la conception de l'interprétation du premier Davidson, en tant que celle-ci respecte la contrainte holistique et le principe constitutif de charité, la simple possibilité d'attribuer des croyances erronées à autrui s'avère problématique. Plutôt que d'attribuer aux anciens de fausses croyances quant à la terre, aux étoiles, etc., il semble que nous soyons comme forcés d'interpréter les anciens comme tenant des croyances vraies quant à leur terre, étoiles, etc. Cette théorie semble alors ouvrir la voie au relativisme : au lieu de nous autoriser à interpréter les anciens comme tenant des croyances fausses relativement à un monde commun que nous partageons dans lequel ils ont jadis vécu et au sein duquel nous continuons à vivre, cette conception de l'interprétation paraît nous encourager à conclure que nous tenons tous deux des croyances vraies, mais à propos d'objets différents<sup>712</sup>.

---

<sup>712</sup> Le relativisme est pourtant une position que Davidson rejette vigoureusement. Voir à ce sujet, Davidson 1993b, 1974b.

Dummett s'attaque ainsi à la position du premier Davidson quant à la possibilité de soutenir une thèse holistique et une conception de la vérité constitutive de l'interprétation<sup>713</sup>. Dummett soutient que le holisme généralisé à savoir, l'idée selon laquelle il n'est pas possible de déterminer un contenu de croyance ou de signification individuel indépendamment des contenus d'autres croyances ou phrases, est problématique dans la mesure où la signification et la référence ne devraient pas dépendre de croyances théoriques mystérieuses. Dans les termes de l'exemple précédent, Dummett avance qu'une théorie de la signification devrait pouvoir interpréter les anciens comme tenant des croyances fausses à propos de « cette terre qui est la nôtre ». Cependant, c'est précisément ce qu'une conception holiste du principe constitutive de charité ne peut pas facilement permettre. Dummett, en soulignant ainsi une faiblesse de la théorie de l'interprétation du premier Davidson, met en évidence le problème de l'erreur. Dummett mentionne qu'il est inacceptable qu'une théorie de l'interprétation soutienne que formuler des croyances fausses soit impossible (ce qui est le risque encouru par une théorie holistique et une conception de la vérité constitutive de la signification); les erreurs sont des données du comportement linguistique qu'une théorie adéquate de la signification devrait expliquer, et non un élément dont il faut se défaire et exclure du processus interprétatif. Nous analyserons quatre réponses à ce problème. Nous pensons que les trois premières ne sont pas satisfaisantes, alors que la quatrième l'est davantage, celle-ci impliquant d'affaiblir quelque peu la thèse d'une normativité constitutive.

1. Il semblerait ainsi que le principe de charité dans sa formulation de l'hypothèse de rationalité traite du problème de l'erreur généré par la formulation de la présomption de vérité. La formulation de l'hypothèse de rationalité du principe de charité – « supposons que le locuteur est rationnel » – impose le respect des principes logiques tels que le principe de non contradiction<sup>714</sup>. Ce principe rend possible l'attribution de croyances fausses notamment dans le cas où les croyances que nous attribuons à l'interprété ne coïncident pas, à moins

---

<sup>713</sup> Dummett, 1993.

<sup>714</sup> La quatrième réponse est une variante de la première mais elle donnera à la notion de rationalité une interprétation plus large et plus flexible. La « Rationalité » ne signifiera pas seulement cohérence logique, mais une notion plus proche du sens commun entendu sous le qualificatif de « ce qui est raisonnable » et inclura donc non seulement le principe formel logique de non contradiction, mais aussi des principes épistémiques bien fondés tels que le principe d'induction, le principe de fiabilité des sens (si les objets apparaissent à un observateur standard dans des conditions d'observation optimales comme ayant telle caractéristique alors les objets ont bien telle caractéristique), etc.

d'être incohérentes, avec son corpus de croyances. Selon cette lecture, le principe de charité est bien la présomption constitutive que ceux que nous interprétons tiennent un tissu de croyances logiquement cohérentes entre elles. Nous pouvons par conséquent leur attribuer des croyances erronées ou fausses : à l'endroit où ils tiennent des croyances incohérentes, nous en concluons qu'ils sont dans l'erreur. Cependant, nous ne pouvons pas leur attribuer d'emblée des croyances irrationnelles ou incohérentes d'une part, parce que la procédure d'interprétation radicale repose en partie sur les contraintes de rationalité et que la possibilité de l'interprétation radicale conditionne la compréhension, et d'autre part, parce que les contraintes de rationalité sont constitutives des concepts d'attitudes et que l'attribution d'attitudes conditionne la compréhension<sup>715</sup>.

Cette solution au problème soulevé par Dummett quant à la possibilité de rendre compte de l'attribution de croyances fausses s'avère insatisfaisante dans la mesure où elle ne fait que dupliquer le problème, en insistant davantage sur la présomption de rationalité que sur celle de vérité c'est-à-dire en impliquant l'idée selon laquelle tenir ou attribuer des croyances irrationnelles (au sens d'incohérence logique) est impossible. L'origine du problème réside dans le statut constitutif accordé à la charité (dans un sens étroit ou large). Or, une condition d'adéquation portant sur la théorie de l'interprétation doit expliquer la possibilité d'attribuer et de tenir des croyances fausses. En effet, en se référant à la contrainte de rationalité, le principe de charité, en tant que règle méthodologique, est supposé guider l'interprète, et non pas prescrire une assignation définitive de croyances vraies. Néanmoins, en tant que contrainte constitutive de l'interprétation, il ôte la possibilité, voire même rend impossible à l'interprété de tenir des croyances logiquement incohérentes. Autrement dit, le principe de charité n'est pas seulement une règle méthodologique, que nous pourrions décider de suivre ou de ne pas suivre mais bien un principe normatif et *a priori*, la condition de possibilité minimale de toute interprétation et de toute compréhension. Pourtant, autoriser la possibilité d'une incohérence logique est tout aussi bien une condition d'adéquation portant sur le contenu de nos significations et de nos croyances que l'est celle d'autoriser la possibilité de croyances fausses. Cette intuition néo-Kantienne semble ainsi résulter de l'idée qu'il est nécessaire d'attribuer un arrière-plan de rationalité à autrui pour interpréter adéquatement son

---

<sup>715</sup> Davidson, 1933b, 1974b, p.287 : « [...] la charité nous est imposée; que cela nous plaise ou non, si nous voulons comprendre les autres, nous devons considérer qu'ils ont raison sur la plupart des sujets. ».

comportement en termes intentionnels : nous nous distinguons en tant que *sapient creatures*<sup>716</sup> (et non comme *sentient creatures*) parce que nous sommes capables de répondre à des normes de rationalité<sup>717</sup>, à ce qui doit être, en étant le sujet d'une soumission à ces normes, et non pas seulement l'objet de forces causales, de ce qui est. Ainsi, ce qui distingue nos jugements et nos actions d'autres types de phénomènes naturels est que ce sont des phénomènes envers lesquels nous sommes responsables. Pour reprendre les termes de Sellars, toute activité humaine s'inscrit au sein de « l'espace logique des raisons<sup>718</sup> » et non seulement dans l'« espace logique des causes » (« realm of law »). Le premier Davidson tente ainsi de retrouver cette idée à travers sa version d'une normativité constitutive de la rationalité, selon laquelle la formation de la croyance et de l'agir humain et de ce fait de l'activité linguistique, sont constitutivement rationnelles. Mais, en faisant de la rationalité une propriété constitutive des pratiques humaines, la thèse d'une normativité constitutive transforme l'intuition néo-Kantienne, à savoir qu'il est *possible* que les êtres humains soient rationnels, en une intuition assurément fautive, à savoir que les êtres humains sont *nécessairement* rationnels. Cette intuition qui souligne le fait que nous nous différencions en tant que *sapient creatures* précisément parce qu'il est possible pour nous d'être rationnel, revient, au moyen de sa version d'une normativité constitutive de la rationalité, à affirmer que nous nous distinguons parce qu'il nous est impossible d'être irrationnel. Or, une théorie adéquate de l'interprétation doit certes rendre compte de l'intuition néo-kantienne, mais ne pas exclure la possibilité de l'irrationalité.

2. On pourrait néanmoins répondre à cette première critique en apportant un nouvel éclairage au problème de l'erreur et de l'irrationalité dans la mesure où, le principe de charité

---

<sup>716</sup> C'est-à-dire les créatures cognitives humaines, qui ont quelque chose de plus que des capacités sensorielles, recognitionnelles et comportementales (présentes chez la paramécie, le thermostat et le chimpanzé).

<sup>717</sup> Sellars, 1949, p. 311: « To say that man is a rational animal is to say that man is a creature not of *habits*, but of *rules* ».

<sup>718</sup> Sellars, 1956: §37, p.80: « en caractérisant comme connaissance un épisode ou un état, nous n'offrons pas une description empirique de cet épisode ou état, mais nous le situons dans l'espace logique des raisons, des justifications et des aptitudes à justifier ce que l'on affirme. ». L'espace des raisons, chez Sellars, est avant tout un espace logique : un espace de descriptions ou d'explications particulières, où les comportements des agents sont décrits et expliqués en termes de raisons, accessibles à ces agents. Le schème des raisons n'est pas réductible ou traductible dans l'espace logique des causes; les entités conceptuelles posées dans cet espace ne sont pas déterminables et compréhensibles dans un vocabulaire naturaliste.

ne prétend pas soutenir que *toutes* les croyances que nous attribuons à l'interprété doivent être vraies (ou rationnelles) mais seulement que *la plupart* ou la *majorité* d'entre elles doivent l'être<sup>719</sup>. En effet, dans le passage cité ci-dessus, Davidson met le terme « aucune » en italique : il précise que nous ne pouvons pas attribuer aux anciens des croyances au sujet de la terre s'ils ne font appel à aucune de ces autres croyances à propos des corps célestes, des étoiles, etc. Ainsi, ce qui détermine l'objet d'une croyance, ce sont les autres croyances que nous pouvons avoir. Par conséquent, les anciens doivent tenir un certain nombre de ces croyances-ci pour vraies, pas nécessairement toutes ces croyances dans leur intégralité, afin d'être considérés comme tenant des croyances au sujet de la terre. Ainsi, dans la mesure où le principe de charité requiert seulement que ceux que nous interprétons tiennent certaines de ces croyances-ci pour vraies, nous laissons entendre alors qu'ils peuvent aussi tenir des croyances fausses, et néanmoins être toujours considérés comme ayant des croyances à propos de la terre. Davidson ne veut cependant pas dire que nous devons interpréter le locuteur comme un individu capable de tenir des croyances quelles que soient les circonstances, mais plutôt que nous pouvons lui attribuer des croyances fausses seulement si lui attribuer ces croyances erronées demeure rationnel c'est-à-dire seulement si lui attribuer des erreurs ne devient pas une pratique largement ou massivement répandue. Cet argument semble ainsi pouvoir être défendu : Davidson émet la possibilité que ceux que nous interprétons puissent tenir des croyances fausses mais interdit néanmoins d'envisager l'option selon laquelle toutes ou la plupart de leurs croyances puissent se révéler fausses : l'attribution de l'erreur peut uniquement survenir sur « un arrière-plan de croyances vraies<sup>720</sup> ». Ce qui rend donc l'interprétation possible est le fait que nous puissions écarter *a priori* la possibilité d'erreur massive<sup>721</sup>. Par conséquent, une théorie holistique du principe constitutif de charité doit permettre que le locuteur, tout comme l'interprète, puisse se tromper. Néanmoins, elle ne peut pas être correcte si elle suppose qu'un individu adhère à de nombreuses croyances

<sup>719</sup> Nous traiterons de ce point spécifique seulement quand nous aborderons la formulation du principe de charité au moyen de la contrainte ou de la présomption de vérité et donc au sujet du problème de l'erreur. Un traitement similaire peut également être apporté à la formulation du principe de charité au moyen de la présomption de rationalité relative au problème de l'irrationalité.

<sup>720</sup> Davidson, 1993b, 1974b, p. 248.

<sup>721</sup> Ibid., p. 248 : « plus on a de croyances correctes, plus les erreurs que l'on peut faire sont repérables. Quand il y a trop d'erreurs, cela brouille tout simplement l'objectif. ».

fausses : il faut supposer qu'en général une phrase est vraie quand un locuteur la tient pour vraie<sup>722</sup>.

Cependant, un tel argument pourrait suggérer qu'il soit requis de compter les croyances pour arriver à dire que la plupart d'entre elles sont vraies. Davidson admet néanmoins que cette idée est problématique<sup>723</sup>. Elle soulève en effet deux problèmes : tout d'abord, en supposant que les croyances telles que la terre est ronde puissent être individuées, comment peut-on alors les compter ? Si je me lève chaque matin pendant dix ans et que je nourris la croyance, en observant le lever du soleil, que la terre est ronde, dois-je la compter comme une croyance ou comme 3650 croyances ? Et certainement rien d'aussi important que le contenu de la croyance devrait pouvoir nous aider à résoudre ce problème. Un second problème émane initialement d'une question antérieure liée à l'individuation des croyances, celle-ci découlant de la conception holistique du principe de charité. Pour compter les croyances, il n'est pas nécessaire de les individuer, l'identification et l'individuation des croyances les unes à l'égard des autres (et de ce fait, s'il s'agit d'une ou de deux croyances) devraient être basées sur leur contenu c'est-à-dire ce sur quoi elles portent. Le contenu de la croyance est ainsi le produit de l'interprétation, et l'interprétation est elle-même régie par la présomption constitutive de vérité. Ce n'est pas comme si nous pouvions d'abord identifier l'objet d'une croyance et que nous procédions ensuite à l'examen du nombre de croyances vraies s'y rapportant. Selon Davidson, ces deux processus ne doivent pas être dissociés mais considérés conjointement. L'élément faisant en sorte que la plupart de nos croyances soient tenues pour vrai est ainsi l'objet sur lequel porte ces mêmes croyances, mais sans moyen d'identifier

---

<sup>722</sup> Sur cette ligne de pensée, le principe de charité n'a pas pour but d'attribuer systématiquement à un sujet des croyances vraies ou rationnelles, et n'exclut en rien l'erreur de sa part. Il nous prescrit de minimiser l'erreur inexplicable :

Certains désaccords sont plus propres à détruire la compréhension que d'autres, et une théorie sophistiquée doit prendre ce point en considération. Le désaccord quant à des sujets théoriques peut (dans certains cas) être plus tolérable que le désaccord quant à ce qui est plus évident; le désaccord quant à la manière dont les choses nous apparaissent est moins tolérable que le désaccord quant à ce qu'elles sont en réalité; le désaccord quant à la vérité des attributions de certaines attitudes à un locuteur par ce même locuteur peut ne pas être tolérable, ou à peine tolérable. Il est impossible de simplifier les considérations pertinentes, car tout ce que nous savons où croyons quant à la manière dont nos croyances sont confirmées peut être mis à profit pour décider où la théorie peut le mieux s'accommoder avec l'existence d'erreurs, et pour décider quelles sont les erreurs qui détruisent le moins la compréhension. La méthodologie de l'interprétation n'est, de ce point de vue, rien d'autre que l'épistémologie au miroir de la signification. (Davidson, 1993b, 1975, p. 204).

<sup>723</sup> Davidson, 2001b, 1983, p. 308: « [T]here is probably no useful way to count beliefs, and so no clear meaning to the idea that most of a person's beliefs are true. ».

l'objet de ces croyances-ci indépendamment de ce qui les rend vraies, le glissement de la plupart des croyances qui sont vraies à toutes celles qui le sont constitue un réel danger. Vu que le fait de compter les croyances s'avère problématique, il semble difficile d'attribuer un sens strictement qualitatif à l'expression « la plupart des croyances sont vraies ». Peut-être est-il plus pertinent d'attribuer un sens qualitatif au terme « la plupart », pour lequel ce qui semble avoir le plus d'importance n'est pas le simple nombre de croyances, mais si certaines de ces croyances que nous tenons pour vraies sont centrales ou d'une importance majeure, de telle sorte que si nous les maintenons comme telles alors nos croyances se réfèrent aux objets que celles-ci (ces croyances centrales) tiennent pour vraies nonobstant le fait que d'autres croyances que nous tenons demeurent malgré tout fausses. Cette notion qualitative de croyances digne d'importance (« weighty ») laisse entendre que toutes les croyances n'ont pas le même poids, à savoir que certaines sont plus essentielles que d'autres lorsqu'il s'agit de déterminer la signification d'un terme par exemple : pour déterminer la signification du terme « La terre », il semble que nous devons considérer certaines croyances au sujet de la terre comme plus primordiales que d'autres. Les croyances relatives à sa forme et à sa distance au soleil pourraient avoir plus de poids que des croyances astrophysiques bien plus mystérieuses à propos de son origine et de sa taille. Cette notion de croyances d'une plus grande importance (« weighty ») quant à la terre repose sur une distinction entre les croyances théoriques importantes c'est-à-dire les croyances centrales et les croyances plus obscures. Cependant, il devrait être évident que cette distinction entre croyances, en termes de leur importance eu égard à la signification d'un terme, jette un anathème sur la théorie globalement holistique en laissant entrevoir une position en faveur de l'atomisme et de la distinction de l'analytique et du synthétique, les dogmes mêmes que le holisme global critique. Davidson doit donc, pour envisager la possibilité de générer des croyances fausses, clarifier la notion d'un arrière-plan de croyances massivement ou majoritairement vraies (ou rationnelles) afin de pouvoir rendre compte de manière significative de l'idée de croyances fausses (ou irrationnelles), mais les deux façons de procéder – la manière quantitative qui consiste à compter les croyances et l'idée qualitative de croyances classées par niveaux d'importance – contreviennent toutes deux à une conception holiste.

C'est d'ailleurs précisément l'analyse que fait Dummett du problème que pose ici le premier Davidson. Selon Dummett, une conception holiste fait obstacle à la construction

d'une théorie adéquate de la croyance et de la signification – c'est-à-dire à une théorie autorisant la possibilité de croyances fausses ou irrationnelles – et il propose de développer une image anti-holiste<sup>724</sup> globale de la signification et de l'interprétation. Dans l'argument anti-holiste, il n'est pas nécessaire de supposer que certaines phrases dépassent nos capacités de reconnaissance : d'après la conception holiste, toute condition de vérité d'une phrase du langage est vouée, en un certain sens, à échapper à nos capacités de reconnaissance, puisque sa signification ne peut être attestée que relativement aux conditions de vérité de toutes les autres. En ce sens, Dummett va jusqu'à dire que le holisme n'est même pas « une théorie de la signification » ; il en nie la possibilité même<sup>725</sup> :

---

<sup>724</sup> Nous employons le terme quelque peu maladroit « anti-holisme » pour souligner le fait que Dummett, en rejetant le holisme, ne préconise pas un retour au holisme, ce que Quine (et Davidson) a si durement critiqué. En rejetant un holisme global, Dummett ne soutient pas que la signification puisse être déterminée directement et individuellement à partir des fragments particuliers du comportement linguistique (atomisme). Il suggère qu'une position plus intermédiaire entre l'atomisme et le holisme global est possible et même souhaitable, une approche qu'il nomme le *molécularisme*. Le molécularisme est plutôt un holisme modéré que global, il cherche donc à éviter les difficultés qui affectent tant l'atomisme que le holisme global. Le molécularisme est la thèse selon laquelle les expressions d'une langue appartiennent à des domaines discursifs contenant les clauses récursives de la théorie et la signification d'une expression est, au sein de ce domaine discursif, une fonction des relations holistiques liant les autres expressions du domaine. Le langage, conçu comme un tout, est la collection de ces domaines discursifs holistiques autonomes plutôt qu'un unique réseau holistique continu de phrases interreliées (comme la conception holistique de Quine-Davidson le présente). Le molécularisme n'est pas l'atomisme, c'est-à-dire qu'il est holistique en ce qu'il fait en sorte que la signification est dépendante d'autres usages d'expressions reliés les uns aux autres; il ne s'agit néanmoins pas d'un holisme global dans la mesure où les usages interreliés ne portent pas sur toutes les expressions sous-tendant notre usage du langage tout entier, mais concernent seulement ceux qui sont assez proches pour faire partie de son domaine discursif. Par conséquent, le molécularisme de Dummett est sensible à l'interdépendance de la signification et de la croyance tout en tentant de circonscrire cette interdépendance holistique à l'intérieur de certaines limites exploitables. Sa position moléculariste espère éviter les problèmes qui touchent les deux positions antagonistes que sont l'atomisme et le holisme global. En d'autres termes, le molécularisme reconnaît l'absence de corrélation une à une entre la signification et l'expérience, ou entre le contenu propositionnel et le comportement linguistique (ce qui était le problème avec l'atomisme), alors qu'il autorise la possibilité de croyances erronées (ce qui est le problème rencontré avec le holisme global). Le molécularisme comme Dummett le reconnaît trouve des affinités avec l'image du langage chère à Wittgenstein d'une collection de jeux de langage relativement autonomes.

<sup>725</sup> Dummett, 1978, p.309: « Holism is not, in this sense, a theory of meaning: it is the denial that a theory of meaning is possible. ».

« On a holistic view, no model for the individual content of a sentence can be given: we cannot grasp the representative power of any one sentence save by a complete grasp of the linguistic propensities underlying our use of the entire language; and, when we have such a grasp of the whole, there is no way in which this can be systematised so as to give us a clear view of the contribution of any particular part of the apparatus. No sentence can be considered as saying anything on its own: the smallest unit which can be taken as saying something is the totality of sentences believed, at any given time, to be true; and of what this complex totality says no representation is possible – we are part of the mechanism, and cannot view it from outside<sup>726</sup>. ».

Dans la théorie du premier Davidson, si l'interprétation débute avec une croyance erronée de l'interprété alors, c'est la preuve qu'il s'agit d'une erreur *dans l'interprétation* (l'interprète s'est trompé en identifiant la référence ou l'extension des expressions de l'interprété) plutôt que d'une erreur de la part des interprétés. Une manière d'en rendre compte est ainsi de préciser que les référents et extensions des termes singuliers ainsi que les prédicats sont avant tout des constructions théoriques ou des entités postulées (*posits*) de la théorie interprétative – ils ne font qu'établir les principes à suivre pour que l'ensemble du processus interprétatif attribue aux interprétés des croyances, pour la plupart, vraies. Les axiomes assignant la référence ne sont néanmoins pas à l'abri d'une éventuelle révision et notamment d'une révision, non sur la base de l'observation du comportement linguistique des interprétés, mais sur celle de la contrainte constitutive du principe de charité, c'est-à-dire de l'exigence de maximisation de la vérité de l'interprétation des phrases de leur langage dans sa totalité. Par conséquent, pour le premier Davidson, toute interprétation doit passer au crible du principe de charité ; il n'y a pas lieu de considérer une classe de phrases transcendantes ou indécidables dont nous serions incapables de représenter les conditions de reconnaissance; toute preuve présumée externe doit toujours pouvoir être ré-interprétable afin de s'ajuster holistiquement à d'autres fragments du comportement linguistique obéissant ainsi à l'exigence constitutive de maximisation de la vérité de l'interprétation du langage dans son intégralité. Dummett soutient que selon le premier Davidson, la justification de notre pratique linguistique et d'une théorie de la signification, est interne à cette même théorie – autrement dit la charité entre dans la constitution de la pratique linguistique<sup>727</sup>. Ainsi, dans la mesure où la théorie interprétative telle que présentée par le premier Davidson, est globalement holistique et régie par le principe constitutif de charité, elle rend impossible l'idée d'un usage

<sup>726</sup> Ibid., p.309.

<sup>727</sup> Dummett, 1993, p. 28.

erroné des termes et donc le fait que les interprétés puissent avoir des croyances fausses. Dummett conclut que pour pouvoir rendre compte de l'erreur, la preuve<sup>728</sup> d'une théorie de la signification devrait être extérieure à la théorie. Ce point requiert d'envisager la possibilité que certaines expressions bénéficient d'un usage privilégié. Les usages privilégiés ne sont pas des produits du processus interprétatif mais sont antérieurs et extérieurs au principe constitutif de charité, au réseau globalement holistique de la théorie interprétative. Ils correspondent à des fragments particuliers du comportement linguistique et en établissant la preuve qu'un locuteur signifie au moyen d'une expression ceci et non cela, ils s'avèrent de ce fait plus importants que d'autres énoncés. Supposons que nous observons que celui que nous interprétons emploie le prédicat « est rond » lorsqu'il désigne de la main une balle de tennis mais non lorsqu'il fait référence à des ballons de forme ovale : nous sommes alors disposés, en interprétant ses propos, à formuler la maxime, « 'est rond' réfère à tous les objets et seulement aux objets de forme ronde ». Supposons également que, sur la base d'autres fragments de son comportement, nous soyons conduits à formuler l'axiome interprétatif, « 'La terre' réfère à la terre », et que, bien que son comportement linguistique appuie les deux axiomes, il ne donne pas son assentiment à la phrase, « La terre est ronde ». L'approche de Dummett considère ainsi les capacités de reconnaissance de l'interprété quant aux formes de balles comme des usages privilégiés qui s'avèrent cruciaux et centraux pour exprimer ce que celui-ci veut dire au moyen du prédicat « est rond ». Dummett ne soutient certes pas qu'une seule expérimentation soit suffisante pour déterminer ce que l'interprète signifie; il proposera de varier les nombres, les couleurs, les formes et les types de balles afin de réduire l'indétermination présente dans une instanciation unique. Il tente ainsi de démontrer que le comportement de l'interprété dans les contextes de reconnaissance de forme (là où les cas testés sont suffisamment variés) constitue un usage privilégié du prédicat « est rond », c'est-à-dire qu'il est suffisant pour établir l'interprétation de « est rond »; nous ne devons pas nécessairement le considérer au sein de l'interprétation de toutes les autres phrases dans le contexte d'un langage constitutivement contraint par l'impératif de maximiser la vérité. Selon Dummett, si un locuteur dispose bien de ces capacités de reconnaissance lui permettant de

---

<sup>728</sup> La sémantique de Dummett est vérificationniste : pour lui, connaître la proposition *p* revient à disposer d'une preuve de *p*, ou à connaître les conditions de vérité de *p*.

différencier les formes, comme son comportement l'atteste dans diverses situations, alors il veut bien dire rond par « est rond » et a donc une croyance erronée au sujet de la terre.

3. Il apparaît ainsi que les engagements du premier Davidson en faveur d'un holisme global et d'un rôle constitutif du principe de charité l'empêchent de fournir une réponse satisfaisante au problème de l'erreur et de l'irrationalité. Le second Davidson de « A Coherence Theory of truth and Knowledge » semble y être arrivé en introduisant un nouveau principe au sein du projet de l'interprétation, le *principe causal*. Étant donné le contexte épistémologique de ce travail, nous nous intéresserons aux deux objectifs qui justifient l'emploi de ce principe causal, à savoir: garantir que la plupart des croyances soient vraies (et de ce fait, il n'est pas possible que toutes les croyances puissent être fausses, ce qui est le principal argument anti-sceptique de ce texte-ci), et envisager la possibilité, même si la plupart de nos croyances doivent être vraies, que certaines puissent être fausses, ce qui est précisément le problème dont il est question ici. Davidson introduit ainsi le principe de causalité comme suit:

« [An interpreter] interprets sentences held true (which is not to be distinguished from attributing beliefs) according to the events and objects in the outside world that cause the sentence to be held true [...] we cannot in general first identify beliefs and meanings and then ask what caused them. The causality plays an indispensable role in determining the content of what we say and believe [...] we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief [...] The method applies directly, at best, only to occasion sentences – the sentences' assent to which is caused systematically by common changes in the world<sup>729</sup>. ».

La causalité, comme contrainte de l'interprétation, énonce que la signification d'une phrase ou le contenu d'une croyance est généralement causé par son environnement extérieur, et les causes sont des caractéristiques objectives provenant d'objets du monde. Ce point nous ramène à l'une des différences entre Quine et Davidson entrevue au §4.4.2 : Davidson conçoit que l'interprété entre dans une relation causale avec les objets du monde extérieur, mais il n'y a aucune raison de postuler un intermédiaire entre le sujet et le monde, comme des stimulations sensorielles. Le principe causal insiste sur ce point : le contenu de certaines propositions (et de croyances) est déterminée par leurs relations causales à l'environnement et

---

<sup>729</sup> Davidson, 2001b, 1983, p.150-151.

les causes sont des objets ou états de chose dans le monde<sup>730</sup>. Par conséquent, les croyances ne peuvent pas radicalement différer de la manière dont le monde est dans la mesure où le monde détermine le contenu des croyances (c'est la fonction anti-sceptique du principe). Une autre caractéristique de l'introduction du principe causal insiste sur la primauté, d'un point de vue méthodologique, des phrases occasionnelles au sein du projet de l'interprétation radicale. Le principe causal a donc trois caractéristiques importantes : la première souligne que le contenu de signification ou d'une croyance est déterminé par les causes extérieures, la seconde précise que les causes doivent être comprises en termes d'événements ou d'états de chose dans le monde, et la troisième, spécifie que la portée du principe causal, quant à la détermination du contenu, se restreint d'un point de vue méthodologique, à une classe de phrases particulières, à savoir aux phrases occasionnelles.

Deux lectures du principe de causalité semblent ainsi envisageables pour comprendre son rôle au sein de la théorie de l'interprétation de Davidson – la première vise à répondre au problème de l'erreur et de l'irrationalité mais ne coïncide pas avec son holisme global, son principe constitutif de charité et son rejet de la théorie causale de la référence, alors que la seconde s'accorde bien avec les engagements de Davidson mais ne réussit pas à rendre compte du problème de l'erreur et de l'irrationalité. Nous pensons que la seconde lecture est celle qui traduit le plus adéquatement la manière dont nous devrions comprendre l'emploi que fait Davidson du principe de causalité, et par conséquent, le fait que le problème de l'erreur et de l'irrationalité demeure toujours, après l'introduction du principe de causalité, en suspens.

Selon la première lecture, l'image du dernier Davidson admet, qu'alors que l'interprétation est largement holistique et régie par le principe de charité – à savoir que

---

<sup>730</sup> L'expérience se rencontre également dans les relations causales aux croyances, en fait, elle peut seulement se trouver dans les relations causales. Cependant, la raison pour laquelle le contenu d'une croyance est lié à sa cause dans le monde, plutôt qu'à l'expérience, est de garantir l'objectivité du sens. Mais Davidson ne soutient pas une version forte du réalisme quant à la vérité et à la signification : il rejette à la fois les théories causales de la référence et les théories correspondantistes de la vérité. La vérité ne peut être établie indépendamment de l'accord, elle n'est donc pas le produit d'une correspondance avec le monde réalisée au moyen d'une relation causale entre les mots et les objets. L'objectivité du sens et de la vérité est le résultat de la « triangulation » entre le monde, le locuteur (l'interprète) et l'auditeur (l'interprété), c'est-à-dire que nous déterminons conjointement avec autrui l'image que nous nous faisons du monde et identifions les causes de nos croyances et de nos énonciations respectives. La notion d'objectivité ne prend sens qu'au sein d'un accord intersubjectif; la cause des croyances et des énonciations (et donc de leurs significations selon le principe causal) est socialement constituée au sein de la situation de triangulation de l'interprétation. La dimension sociale de l'approche de Davidson sera abordée au §5.6. 2.

lorsque nous interprétons, nous ajustons l'interprétation des phrases et des croyances de telle sorte à ce qu'elles soient, pour la plupart, vraies et rationnelles – le contenu de certaines phrases, en l'occurrence des phrases occasionnelles, n'est pas déterminé ainsi. En effet, ces phrases bénéficient d'un statut épistémologique particulier – elles doivent être préalablement interprétées pour permettre au processus interprétatif de suivre son cours, et l'interprétation de cette classe particulière de phrases ne dépend pas de la thèse d'une interdépendance holistique de la croyance et de la signification, ni même de la contrainte constitutive de rationalité. Ainsi, en suivant cette lecture, l'interprétation devient un processus bipartite fondé sur une distinction entre deux classes de phrases : en premier lieu, la classe privilégiée des phrases occasionnelles dont les significations sont déterminées indépendamment et directement par les facteurs causaux externes et en second lieu, les phrases perdurables (*standing sentences*) et les croyances dont les significations sont déterminées au moyen de considérations normatives et rationnelles. Elles sont ainsi interprétées afin de préserver autant que faire se peut la contrainte de rationalité et de vérité étant donné les déterminations au préalable des significations des phrases occasionnelles par le biais du principe causal. L'attribution de l'erreur et de l'irrationalité devient possible dans la mesure où des usages et des énonciations peuvent à présent être interprétés indépendamment de la présomption constitutive de vérité ou de rationalité. Leur interprétation n'obéit pas à cette contrainte et a, en tant que nous admettons que ce sur quoi porte nos usages et nos énonciations est déterminé par ce qui les cause dans le monde, un statut privilégié. Par conséquent, s'il s'avère qu'une fois que nous avons identifié la référence d'une expression ou l'objet d'une croyance selon la contrainte causale, que les autres croyances que nous tenons à propos de telles choses sont fausses, nous ne sommes plus tenus d'ajuster la détermination initiale de l'objet d'une croyance ou de la référence d'un terme en nous efforçant de garantir leur vérité; nous pouvons plutôt admettre la possibilité de l'erreur. Ainsi, il apparaît, en revenant sur « l'exemple des anciens », que cette lecture du principe de causalité mène à une interprétation des anciens les autorisant à tenir des croyances fausses au sujet de la terre, c'est-à-dire à affirmer, par exemple, que la terre est plate. Les significations et les références des anciens sont donc liées directement à des propriétés et à des objets qui causent leurs croyances. Une fois que la causalité détermine le contenu d'une croyance, il se peut que le contenu d'une croyance et l'objet qui la cause ne coïncident pas. En d'autres termes, il se peut que la

croissance porte sur un certain objet, mais que ce qui est cru au sujet de l'objet soit faux. La cohérence des croyances, ou leur caractère holistique, tendent à réduire la possibilité de l'erreur, alors que leur propriété d'être causées par des objets extérieurs doit envisager la possibilité de l'erreur<sup>731</sup>.

Selon cette lecture, le principe de causalité représente une entorse significative à l'approche globalement holistique contrainte par le principe constitutif de charité, telle que décrite par le premier Davidson. En introduisant le principe de causalité, un certain externalisme sémantique est à prendre en considération : le contenu n'est pas seulement une fonction des relations allant de croyances en croyances et des connexions inférentielles et rationnelles obtenues entre elles; le contenu est déterminé, au moins pour certaines phrases, par des relations causales allant du langage au monde. Tous les contenus de croyances et de significations ne sont pas des constructions théoriques du principe constitutif de charité au sein d'une théorie interprétative globalement holistique; les objets ou références de certaines croyances et propositions sont établis indépendamment du principe de charité. Le principe de causalité, en tant que principe de détermination du contenu, n'est pas régi par le principe de charité, c'est-à-dire que le contenu de ces phrases et de ces croyances-ci, en étant directement déterminé par ce qui les cause objectivement dans le monde, n'obéit pas à la contrainte de rationalité ou de vérité. Il apparaît alors que, selon cette lecture, le premier Davidson préconise une théorie causale de la référence, au moins pour la référence des termes singuliers et les prédicats, survenant dans les phrases occasionnelles. Cette idée est d'ailleurs renforcée lorsqu'on constate que c'est précisément un réquisit des théories causales de la référence d'admettre la possibilité de tenir et d'attribuer des croyances erronées, et c'est ainsi dans ce contexte d'évaluation de la possibilité de l'erreur que Davidson introduit le principe de causalité. Selon ces théories causales, la référence d'un terme est déterminée par l'objet dans le monde qui la cause et en tant que telle, la référence d'un terme n'est pas dépendante des autres croyances portant sur le monde, et donc de la vérité de ces autres croyances; il est alors possible d'avoir des croyances erronées à propos d'un état de choses dans le monde tout en tenant toujours des croyances à son sujet.

---

<sup>731</sup> Cette lecture de Davidson se rapproche de la manière dont Dummett tente de rendre compte de l'erreur – certaines phrases ont un statut privilégié et leur contenu n'est pas déterminé par la contrainte constitutive (entendue comme une construction théorique) de rationalité et de vérité.

Cependant, le second Davidson, en refusant explicitement d'associer le principe de causal à une théorie causale de la référence, souligne le problème généré par la lecture ci-dessus du principe de causalité. Deux raisons motivent ce refus. Tout d'abord, la théorie causale de la référence permet que non seulement certaines de nos croyances soient fausses, mais aussi que la plupart d'entre elles le soient, et ce sont précisément ces implications radicales sceptiques qui vont à l'encontre de l'argument anti-sceptique de Davidson<sup>732</sup>. La seconde raison, plus centrale pour notre propos, souligne clairement une tension entre d'une part, la théorie causale selon laquelle les objets de nos croyances, tels qu'ils seront identifiés par un interprète, sont les causes de nos croyances, et d'autre part la thèse holistique selon laquelle les assignations de référence doivent être articulées inférentiellement et non dépendre directement de relations causales entre les mots et les objets dans le monde. Si le principe de causalité n'obéit pas, conformément à la première lecture, au principe de charité interprétative, il entre alors en conflit avec le second réquisit en faveur d'une théorie de l'interprétation privilégiant des considérations holistiques et de rationalité propres aux contenus mentaux. Si la signification déterminée par des facteurs causaux physiques est « première » dans l'ordre explicatif, alors il semble qu'il n'y ait aucune façon de garantir que l'interprétation et l'activité linguistique soient régies par des principes constitutifs de rationalité. Davidson refuse de concevoir que la causalité puisse ne pas répondre au principe de charité et il la considère alors soumise à l'autorité du principe de charité : ce qui caractérise la cause d'une croyance est déterminé par des considérations normatives de cohérence logique et de véridicité. Il s'ensuit donc que l'identification de la cause, tout comme l'identification de la référence, de la croyance et de la signification devient un artefact de l'interprétation, régi par les contraintes holistiques de rationalité et de vérité. En d'autres termes, Davidson refuse d'apparenter son principe de causalité à une théorie causale de la référence. La référence obéit bien à *une contrainte causale* mais il ne s'agit pas d'une *théorie causale de la référence* dans la mesure où l'identification de la cause d'une croyance n'est pas indépendante de la détermination holistique et rationnellement constitutive du contenu d'autres croyances auquel elle se trouve inférentiellement liée. Par conséquent, ce qui détermine l'objet de notre croyance ici, ce sont les autres croyances inférentiellement liées aux premières que nous pouvons avoir en vertu de principes logiques normatifs. La cause

---

<sup>732</sup> Davidson, 2001b, 1983, p.151, n. 7.

d'une croyance est ainsi, dans ce sens, le produit des contraintes constitutives de rationalité et de véridicité.

Ainsi, selon cette seconde lecture de la relation entre causalité et charité, le conflit potentiel entre ces deux principes, c'est-à-dire entre d'un côté, une approche néo-Kantienne et de l'autre, une approche naturaliste et causale du contenu, s'estompe. Plutôt que de concevoir la causalité et la charité comme des éléments séparés d'une théorie interprétative hiérarchique à deux niveaux appliquant d'abord le principe de causalité aux phrases observationnelles et aux prédicats dont l'interprétation doit, en premier lieu, être établie; l'appliquant ensuite aux phrases non observationnelles et aux prédicats dont l'interprétation, une fois établie, doit venir préserver les attributions de significations des phrases observationnelles bénéficiant d'un statut privilégié eu égard à l'application du principe de causalité, tel que la première lecture le mentionne; la détermination de la signification des phrases observationnelles et des prédicats *via* le principe de causalité implique déjà, depuis le début, de tenir compte du contenu d'autres phrases non observationnelles. C'est pourquoi, selon le holisme global, toutes les phrases et prédicats sont inférentiellement reliés; les phrases observationnelles et les prédicats déterminent les usages des phrases et des termes non observationnelles. Par conséquent, selon cette seconde lecture, le principe de causalité est contraint par, ou subsumé sous, le principe de charité. Le principe de causalité a seulement une primauté méthodologique : afin d'entrer dans l'interprétation radicale, nous devons d'abord considérer les phrases observationnelles et supposer qu'elles sont généralement causées par des objets du monde extérieur. Cependant, la signification de ces phrases-ci conformément au principe de causalité n'est pas déterminée ou fixée une fois pour toutes; leur signification ne peut être attestée que relativement aux relations inférentielles les liant aux autres phrases et les soumettant de ce fait aux exigences constitutives de cohérence logique et de maximisation de la vérité, celles-ci pouvant ainsi conduire à ce que la signification initiale, en tant qu'établie au moyen du principe de causalité, soit révisée. Le principe de causalité ne s'oppose donc pas, voire même semble rejoindre, selon cette lecture, le principe de charité ; il propose en quelque sorte de reformuler le principe de charité en l'appliquant aux phrases observationnelles et aux prédicats. Cependant, en faisant de l'identification de la cause une construction de l'entreprise interprétative globalement holistique et régie selon des principes constitutifs de rationalité, nous ne pouvons plus faire

appel au principe de causalité pour nous fournir une réponse satisfaisante quant au problème de l'erreur et de l'irrationalité. Par conséquent, le second Davidson est toujours aux prises avec le dilemme initial : si le principe de causalité vise à s'acquitter de la tâche de répondre au problème de l'erreur et de l'irrationalité, alors il n'est pas nécessaire qu'il se soumette au principe de charité interprétative, ce qui revient à adopter une théorie causale de la référence, au moins dans certains cas privilégiés. Mais, soutenir une thèse de ce genre, irait à l'encontre du refus de Davidson d'adhérer à une théorie causale de la référence, refus qui se justifie par son engagement néo-Kantien défendant l'idée selon lequel les contenus des croyances doivent respectés la contrainte constitutive de rationalité. Cependant, si la rationalité (c'est-à-dire la cohérence logique) ou la vérité est constitutive de toute détermination du contenu, y compris de celle établie au moyen du principe de causalité, alors le problème de l'erreur et de l'irrationalité est laissé en suspens. Nous pensons que l'emploi du principe de causalité tel que le conçoit Davidson devrait être compris de la seconde manière dans la mesure où cette lecture coïncide avec ses autres thèses centrales, ce qui nous pousse à conclure que le problème de l'erreur et de l'irrationalité n'est toujours pas résolu.

4. Une cause de ce problème est donc le principe de charité, tant dans sa formulation d'une contrainte constitutive de rationalité que dans celle d'une contrainte de vérité : en énonçant la présomption constitutive forte de vérité ou de rationalité sans laquelle nous ne pourrions même pas tenir ou attribuer des croyances fausses, il rend impossible le fait même de tenir ou d'attribuer des croyances erronées ou irrationnelles. Ce point pourrait nous laisser croire qu'une solution consisterait à renoncer à la thèse d'une normativité constitutive, mais nous avons souligné que nous serions toujours aux prises avec le second volet du dilemme. Par conséquent, plutôt que d'abandonner complètement la thèse d'une normativité constitutive, nous suggérons de modifier le contenu du principe de charité (qui est considéré comme constitutif), ce qui aura pour effet d'affaiblir la thèse d'une normativité constitutive, et d'autoriser la possibilité de l'erreur et de l'irrationalité. Cette compréhension plus charitable du principe de charité est entendue sous le concept de « principe d'humanité » ou sous une « version améliorée du principe de charité<sup>733</sup> » (*the improved principle of charity*);

---

<sup>733</sup> Le terme « le principe de charité » est dû à Grandy, 1973. Lewis, 1974, envisage un principe identique auquel il fait référence sous l'expression de « version améliorée du principe de charité ». Lewis distingue entre une version brute et une version améliorée du principe de charité comme suit,

certains passages laissent entendre que Davidson veut dire par le terme de charité une idée plus proche de celle d'humanité<sup>734</sup>. Le principe d'humanité est plus faible que chacune des formulations du principe de charité envisagées jusqu'ici. Le principe de charité, comme nous l'avons mentionné, se présente sous deux formes : une version large plus forte s'énonçant sous la forme d'une formulation d'une présomption de l'accord ou de la vérité et une version étroite plus faible s'exprimant sous la forme d'une formulation d'une présomption de rationalité (au sens de cohérence logique). La formulation d'une présomption ou d'une contrainte forte de vérité postule que nous *devons* (il s'agit d'une exigence constitutive *a priori*) interpréter autrui de telle sorte à ce que ses croyances soient, pour la plupart, vraies. Nous en revenons donc ici au problème de l'erreur. La formulation plus faible mais toujours trop forte de la contrainte de rationalité nous impose d'interpréter autrui de telle sorte à ce que ses croyances soient, pour la plupart, rationnelles, au sens où la rationalité est strictement

---

p.336 : « A crude version of the principle of charity might just require that, so far as other constraints allow, the beliefs and desires ascribed to Karl [the interpretee]...should be the same as our own beliefs and desires...But it would be more charitable to make allowances for the likelihood that Karl's circumstances – his life history of evidence and training... – may have led him understandably into error. We should at least forbear from ascribing to Karl those of our beliefs and desires which...he has been given no reason to share. We should even ascribe to him those errors which we think we would have made, or should have made, if our evidence and training had been like this. Perhaps an improved principle of charity would require Karl's beliefs and ours to be related as follows: there must exist some common inductive method *M* which would lead to approximately our present system of beliefs ascribed to Karl...if given Karl's life history of evidence. ».

<sup>734</sup> Davidson, 1974, adhère à l'idée de Lewis d'une version améliorée du principe de charité, ou affirme qu'il entend la charité précisément dans le sens de Lewis tout au long de ses « Replies to Lewis and Quine », p.346 : « The improved principle of charity says why assume Karl is right about everything...Surely we can do better by allowing for explicable error. I completely agree. The improved principle of charity, insofar as it says there are cases where you can make exceptions right from the beginning, is what I espouse ». Voir également Davidson, 1993b, 1974b, p.287 : « Nous parvenons à une première approximation de ce que sera une théorie achevée en assignant aux phrases d'un locuteur des conditions de vérité qui ne sont effectivement réalisées (à notre sens) que quand le locuteur tient ces phrases pour vraies. La politique à suivre consiste à procéder de cette façon aussi longtemps que possible, moyennant des conditions de simplicité, des indications sur les effets du conditionnement social, et bien entendu notre connaissance, commune ou scientifique, de l'erreur explicable. La méthode n'est pas destinée à éliminer le désaccord, pas plus qu'elle ne peut le faire; son but est de rendre possible le désaccord doué de sens, et cela dépend entièrement d'un fondement – de *quelque* fondement dans l'accord. »; et Davidson, 2001b, 1983, p.318 : « The best we can do is to cope with error holistically, that is, we interpret so as to make an agent *as intelligible as possible, given his actions, his utterances and his place in the world*. About some things we will find him wrong, as the necessary cost of finding him elsewhere right. As a rough approximation, finding him right means identifying the causes with the objects of his beliefs, giving special weight to the simplest cases [occasion sentences], and countenancing error where it can best be *explained* »: nos italiques. Notons qu'à chacune de ces occurrences, la référence à la vérité et à la cohérence logique est remplacée par des notions plus flexibles proches du sens commun telles que celles d'intelligibilité, de simplicité, etc.

comprise en termes de cohérence logique. Cette formulation, en traitant ainsi du problème de l'attribution de croyances erronées, demeure encore trop forte dans la mesure où elle soulève un problème analogue; celui d'attribuer des croyances irrationnelles au sens de croyances logiquement incohérentes. Le principe d'humanité se rapproche d'une formulation plus faible du principe de charité mais, est encore plus faible. Il prescrit que nous devons interpréter autrui de telle sorte à ce que ses croyances soient raisonnables, douées de sens, ou intelligibles eu égard à sa situation épistémique et à son contexte historique, social et culturel. Avant d'examiner plus en détails l'idée d'un contenu légèrement différent de ce principe en partant d'une compréhension plus large de la rationalité, il est important de déterminer si ce principe, à l'image des autres formulations de la charité, possède un statut constitutif *a priori* au sein du projet de l'interprétation. L'hypothèse de rationalité (dans ce sens plus flexible de rationalité) n'est pas une option mais bien une condition de possibilité de l'interprétation et de la compréhension. La présomption de rationalité ne peut être sujette à révision : nous devons pour comprendre autrui supposer qu'il est rationnel. Le principe d'humanité n'est donc pas une hypothèse empirique – il ne peut être révisable sous prétexte que nous risquons de perdre le fil de l'interprétation. Il n'est pas non plus un impératif pragmatique utile pour contribuer plus facilement à rendre les propos d'autrui intelligibles. C'est une condition transcendante de l'interprétation; il est donc nécessaire pour comprendre les propos d'autrui<sup>735</sup>.

Une contrainte interprétative comme le principe d'humanité prescrit de supposer que nous pouvons rendre compte de manière intelligible des croyances d'autrui, au sens où ce sont des croyances que toute personne raisonnable est censée tenir dans des circonstances épistémiques normales incluant la considération de son arrière-plan culturel ou de sa propre histoire personnelle. Cette contrainte *a priori* d'intelligibilité est, de toute évidence, plus faible que la version large du principe de charité dans la mesure où il peut être raisonnable de tenir des croyances empiriques fausses, mais également plus faible que la version étroite du principe de charité puisqu'elle rend possible l'incohérence logique. Ainsi, dans les cas où on pourrait comprendre qu'une personne raisonnable tienne des croyances logiquement incohérentes, notamment lorsque son *corpus* de croyances est important et complexe, le

---

<sup>735</sup> Lukes, 1982, p. 261-305, nous a été utile pour notre analyse de la relation entre le principe de charité et le principe d'humanité.

principe de charité (la contrainte constitutive d'intelligibilité) semble alors reconnaître la légitimité de l'attribution d'erreur ou d'incohérence. Par conséquent, le principe de charité *via* le principe d'humanité permet d'interpréter autrui comme tenant des croyances irrationnelles (c'est-à-dire logiquement incohérentes) juste aussi longtemps qu'il est raisonnable pour celui-ci de les tenir. De cette manière, l'hypothèse d'intelligibilité contenue dans le principe d'humanité est plus faible c'est-à-dire moins rigide que la présomption de rationalité incorporée dans le principe de charité (dans sa version faible) : elle exige que nous supposions que les croyances d'autrui soient rationnelles dans le sens plus flexible selon lequel ses croyances nous sont intelligibles en vertu d'autres croyances, de leur contexte culturel, de leurs intérêts et pratiques (plutôt que dans le sens strict selon lequel ses croyances sont logiquement cohérentes) au sein desquelles elles s'inscrivent. La notion d'intelligibilité inhérente au principe d'humanité ne peut donc pas être appréhendée en termes de cohérence logique. Ce ne sont néanmoins pas des notions totalement indépendantes si ce n'est que cette notion d'intelligibilité implique la considération d'autres facteurs, à savoir : les lois du raisonnement déductif, les canons du raisonnement inductif (tels que le principe de fiabilité des sens), et des intérêts cognitifs divers comme la prédiction et le contrôle de l'environnement, la compréhension mutuelle, le critère de simplicité dans l'explication, etc. C'est ainsi, en incluant la considération de tous ces facteurs, que l'exigence d'intelligibilité autorise la possibilité d'attribuer des croyances logiquement incohérentes : pour rendre compte du *corpus* de nos croyances ou de celles d'autrui de manière intelligible ou raisonnable, il n'est pas nécessaire de privilégier l'approche qui accordera le plus de poids à la contrainte qui garantira la cohérence au sein de nos croyances, il est peut-être préférable de préconiser l'approche qui mettra davantage l'accent, dans certaines circonstances, sur ces autres facteurs<sup>736</sup>. En fait, « [l]e processus est celui de l'élaboration d'une théorie viable de la

<sup>736</sup> Considérons par exemple le scénario suivant que l'on doit à Grandy, 1973. Kurt arrive à un dîner et dit à un autre convive : « l'homme qui boit un Martini est un philosophe ». Il y a bien un homme dans le champ visuel de Kurt, mais cet homme boit de l'eau dans un verre à Martini et il n'est pas philosophe. Dans une autre pièce, il y a un homme qui boit un Martini (c'est le seul convive à boire un Martini), et c'est un philosophe. Si l'on interprète « l'homme qui boit un Martini » comme référant au premier homme (celui qui boit de l'eau dans un verre à Martini), alors il est difficile de considérer comme vrai l'énoncé prononcé par Kurt. Si l'on interprète « l'homme qui boit un Martini » comme référant au second homme (celui qui est dans une autre pièce), alors on peut considérer comme vrai l'énoncé prononcé par Kurt. L'interprète qui veut autant que possible que les assertions de Kurt soient correctes, conformément au principe de charité, penchera pour la seconde interprétation. Mais celle-ci ne semble pas correcte : Kurt n'a aucune connaissance que ce soit de ce second homme qui boit

croissance et de la signification à partir des phrases tenues pour vraies<sup>737</sup>», qui permettrait de mettre en place une approximation des conditions de vérité de ces phrases. L'accomplissement de ce processus se ferait en répondant à des exigences de simplicité, en utilisant une certaine connaissance du contexte social du locuteur et en faisant appel à des connaissances partagées et des fondements scientifiques et rationnels d'usage. Loin de chercher à uniformiser les différentes interprétations possibles d'un discours, une telle démarche viserait essentiellement l'atteinte d'un accord prenant appui sur les phrases reconnues comme vraies, quand les locuteurs ont une langue commune, ou sur la base d'une théorie de la vérité élaborée par un interprète quand ils ont des langues différentes. Ainsi, il n'y aura jamais d'erreur grave<sup>738</sup> de traduction ou d'interprétation aussi longtemps que les significations et les croyances sous-jacentes aux phrases acceptées comme vraies sont systématiquement corrélées à d'autres phrases reconnues elles aussi comme vraies. La charité au sens d'humanité donnerait ainsi à toute théorie de la traduction ou de l'interprétation une souplesse qui nous éviterait tout excès ou dérive. Une interprétation charitable serait alors justifiée et nécessaire pour optimiser toute interprétation, favoriser l'accord et la communication afin que nos erreurs soient explicables. Le principe d'humanité évite donc le problème de l'erreur ou de l'irrationalité<sup>739</sup>. En rendant le contenu de la présomption de rationalité plus flexible, le principe d'humanité envisage d'un côté, la possibilité de l'erreur ou de l'irrationalité (au sens de cohérence logique) et conçoit, en considérant la contrainte de rationalité, (au sens d'intelligibilité), de lui attribuer un rôle constitutif dans le projet

---

réellement un verre de Martini et tout semble indiquer qu'il réfère à l'homme qui boit de l'eau dans un verre à Martini. Grandy, p. 445, souligne qu'« il vaut mieux attribuer [à Kurt] une fausseté explicable plutôt qu'une vérité mystérieuse ». Par conséquent, si pour Davidson, les attributions d'erreurs doivent toujours se faire sur fond d'accord, l'exemple de Grandy montre très simplement que l'attribution d'erreurs est préférable, dès que nos théories sur la formation des croyances (dans ce cas précis, une théorie assez élémentaire) nous disent qu'il est plus plausible que le locuteur se trompe.

<sup>737</sup> Davidson, 1993, 1974b, p.286.

<sup>738</sup> En plus de son argument de l'interprète omniscient qui, contrairement à l'interprète faillible, est capable de connaître toutes les causes des croyances possibles et particulièrement celles du locuteur et de l'interprète faillible, Davidson écarte la possibilité d'erreur grave, car, à la base, « l'interprète rejettera une interprétation sémantique d'une phrase que le locuteur tient pour vraie si l'interprétation fait de cette phrase une contradiction évidente. ». Voir Davidson, 2005, 1993, p. 44.

<sup>739</sup> Davidson, 1993b, 1974b, p. 286, donne l'exemple d'une croyance erronée attribuée sur la base du principe de la fiabilité de nos sens. Davidson soutient ainsi que si le champ de vision d'un observateur est flou ou s'il s'agit d'une journée nuageuse ou encore si l'observateur est fatigué, etc., et qu'il dise « yawl » alors qu'il observe ce qui est en fait un ketch, alors l'interprète devrait attribuer une croyance empirique fautive au locuteur (à savoir, qu'il y a un yawl) plutôt qu'une croyance vraie (et attribuer au terme « yawl » une référence différente de la nôtre).

d'attribution de la croyance et de la signification. Par conséquent, l'intuition néo-Kantienne d'une dimension normative et rationnelle de l'action linguistique est préservée sans pour autant en faire une contrainte trop forte, ce qui rendrait l'irrationalité impossible, à l'image des autres versions du principe de charité. Si c'est donc ce que Davidson veut vraiment dire au moyen du « principe de charité », alors son approche semble éviter le dilemme que nous soulevons dans cette section. La question est alors : peut-on attribuer, à juste titre, cette version du principe de charité à Davidson ?

Si l'on suggère que nous pourrions nous passer du principe de charité ou que nous devrions le comprendre conformément au principe d'humanité, nous nous heurtons d'emblée à deux difficultés. Tout d'abord, comment le principe de causalité s'accordera-t-il avec le principe d'humanité ? N'allons-nous pas engendrer un problème analogue d'incompatibilité entre d'une part, un principe qui détermine le contenu sur la base de relations causales obtenues entre les énonciations linguistiques et le monde extra-linguistique, et d'autre part, un principe déterminant le contenu de nos croyances et significations sur la base de considérations plus normatives ? De plus, si Davidson entend par principe de charité, principe d'humanité, pourquoi alors faire appel, comme c'est de plus en plus le cas dans ses travaux ultérieurs, au principe de charité ? Nous avons ainsi recours au principe de causalité en tant qu'il rend compte du problème de l'erreur, mais le principe d'humanité, comme nous l'avons vu, remplit déjà ce rôle. Par conséquent, il apparaît que la simple présence du principe de causalité dans la théorie de l'interprétation de Davidson contredise la compréhension de ce principe de charité comme principe d'humanité, et de ce point de vue, son approche est toujours vulnérable face aux problèmes que nous avons décrits.

Ces interrogations peuvent néanmoins être désamorçées. Le principe d'humanité, en tant qu'hypothèse d'intelligibilité peut être lu comme l'exigence d'interpréter autrui de telle sorte à ce que ses croyances soient, pour la plupart, vraies. Nous pouvons également concevoir que le principe de causalité de Davidson poursuit le même but : celui-ci mentionne que nous devons interpréter autrui de telle sorte à ce que ses croyances quant aux objets ordinaires de nos activités quotidiennes, soient, pour la plupart, vraies. Il exprime ainsi cette idée précisément en termes de contenus de ces mêmes croyances généralement causées par son environnement extérieur. Le principe de causalité peut donc être lu dans le sens où l'hypothèse plus faible d'intelligibilité est la seule contrainte *a priori* nécessaire pour

permettre l'interprétation. Le principe de causalité ne contredit pas, en tant que tel, la position de Davidson en faveur du principe d'humanité; il s'agit plutôt de la manière dont Davidson le défend. Si le principe de causalité doit être compris, comme une expression du principe d'humanité, et si le principe d'humanité est la manière dont le principe de charité doit être compris, alors nous obtenons finalement ce que nous voulons. La présomption d'intelligibilité propre au principe d'humanité est suffisamment flexible pour admettre la possibilité de l'erreur ou de l'irrationalité et il n'existe pas d'incompatibilité entre le principe de causalité et de charité dans la mesure où ce sont deux aspects complémentaires d'un même principe, le principe d'humanité. Le principe de charité retranscrit l'hypothèse d'intelligibilité et le principe de causalité représente cette présomption appliquée à l'interprétation de nos croyances ordinaires intervenant dans nos activités pratiques d'interaction avec le monde. L'identification de la cause d'une énonciation (et de ce fait le contenu des phrases occasionnelles) devient une construction de l'ensemble de l'activité interprétative et du principe constitutif d'humanité. La cause d'une énonciation traduira ce qu'il est raisonnable pour l'interprété de croire et ce qu'il est raisonnable de croire sera déterminé en évaluant un certain nombre de facteurs : la cohérence logique, les connaissances que celui-ci peut raisonnablement s'attendre à acquérir, etc. L'identification de la cause se fera donc tant sur la base de la considération des relations logiques et inférentielles entre croyances, que sur ses intérêts et engagements pratiques avec le monde. Le contenu est donc déterminé par les liens causaux que le sujet entretient avec les choses du monde et par sa connexion aux autres croyances. Par conséquent, on ne rencontre pas d'incompatibilité entre un principe causal naturaliste et un principe normatif rationnel. De plus, faire de l'identification de la cause une construction du principe d'humanité ne soulève pas le problème de l'erreur ou de l'irrationalité comme nous l'avons vu précédemment lorsque la cause émanait du principe de charité dans la mesure où il ne s'agit pas ici d'une exigence visant à maintenir la véridicité ou la cohérence logique des croyances de l'interprété, mais d'une exigence veillant à garantir l'intelligibilité de ses croyances, celle-ci incluant déjà la possibilité de l'erreur ou de l'irrationalité.

Le principe de charité (incorporant le principe de causalité) entendu au sens de principe d'humanité semble toujours devoir répondre à un problème récurrent : il s'oppose, en tant que tel, à la thèse de Davidson en faveur d'un faillibilisme global, c'est-à-dire à son rejet des

vérités *a priori*. Nous avons mentionné ci-dessus que, même si le principe de charité de Davidson est interprété sous une forme plus faible au moyen du principe d'humanité de Grandy ou d'une « version améliorée du principe de charité » de Lewis, le principe d'humanité doit toujours être entendu comme contrainte constitutive ou comme une condition *a priori* de l'interprétation : il est donc présupposé par la possibilité même de l'activité interprétative. Néanmoins, en raison de son statut *a priori* transcendantal, on peut être tenté de croire que même la formulation la plus faible du principe d'humanité laisse place au même problème abordé dans la section précédente : il est incompatible avec le rejet de Davidson des vérités *a priori*. Ainsi, bien que le principe d'humanité apporte une solution au problème de l'erreur ou de l'irrationalité généré par les formulations trop fortes du principe de charité, il semble qu'il se heurte toujours à des problèmes internes : il impose de renoncer au faillibilisme global (ou au révisionnisme global) et soutient l'idée de vérités *a priori*, ce qui va à l'encontre des thèses centrales de Davidson héritières de Quine, en faveur d'un faillibilisme global et de l'abandon de la classe des vérités *a priori*. Cependant, cette crainte peut tout aussi être dissipée : la thèse selon laquelle l'intelligibilité est constitutive de la signification et de la croyance est comparable à l'approche selon laquelle aucune croyance particulière ne peut raisonnablement être considérée définitivement à l'abri de la révision dans la mesure où il est tout à fait toujours plausible que, dans certaines circonstances, il soit plus raisonnable de réviser une croyance donnée. Par conséquent, au lieu d'être incompatible avec un faillibilisme global, l'hypothèse constitutive d'intelligibilité en est plutôt une expression : en d'autres termes, dire que le principe de charité au sens de principe d'intelligibilité est constitutif de l'interprétation revient à soutenir qu'aucune croyance particulière ou norme de signification n'est totalement à l'abri de la révision.

#### 4.8. Conclusion

Nous avons ainsi cherché à examiner le sens et la portée d'une normativité forte ou constitutive de la signification via la lecture du principe de charité préconisée par Davidson. Pour analyser la nature de cette dimension d'une normativité constitutive de la signification, il s'est ainsi avant tout avéré nécessaire de comprendre les présupposés et les conditions d'adéquation requises par sa théorie de la signification et de l'interprétation radicale.

Davidson tente ainsi d'établir si une théorie tarskienne de la vérité peut ou non aboutir à une théorie de la signification pour les langues naturelles. Il souligne, comme nous le montrons, l'importance du lien de la vérité à la notion de signification, et sa priorité au cœur d'une théorie satisfaisante de la signification pour les langues naturelles. Cet aspect de la vérité (ce qu'il nomme le « contenu empirique » d'une théorie de la vérité) est, selon lui, laissé pour compte dans les définitions de Tarski, celui-ci restreignant ses définitions de la vérité aux langages formels. Aussi, considère-t-il d'abord qu'une théorie satisfaisante de la signification doit fournir à la théorie de la vérité une assise empirique, en explicitant les conditions empiriques dans lesquelles une phrase est compréhensible par un interprète lorsqu'elle est énoncée par un locuteur. Une théorie de la vérité doit donc être testable c'est-à-dire, se prêter à des attributions vérifiables de signification aux locuteurs d'une langue donnée.

En donnant ainsi à la vérité un rôle central dans la construction d'une théorie de l'interprétation, Davidson renoue alors avec la conception vériconditionnelle de la signification, celle-ci devant répondre à deux conditions d'adéquation, à savoir : 1/ au principe de contextualité selon lequel une phrase ne peut avoir de signification indépendamment du contexte et 2/ au principe de compositionnalité entendu dans un sens holiste selon lequel la signification des parties composantes est déterminée par la signification de la phrase entière. Autrement dit; comprendre un langage, c'est d'abord comprendre l'ensemble des phrases du langage, et ensuite l'ensemble des éléments qui les composent; une phrase n'ayant jamais de signification indépendamment du contexte de l'ensemble des autres phrases d'un langage, et les expressions qui les composent n'acquièrent de sens que dans le contexte de toutes les phrases où elles figurent. Davidson soutient que c'est seulement au niveau des phrases qu'une théorie du langage est confrontée aux données linguistiques c'est-à-dire aux données empiriques accessibles à un interprète radical capables de la confirmer. Une théorie du langage doit donc relier la langue à l'usage.

Davidson en vient ainsi à rejeter une théorie de la signification qui adopterait comme concept primitif la notion intensionnelle de signification d'une part, en raison de l'absence de critère d'identité capable de déterminer lorsque deux expressions ont la même signification et d'autre part, en raison de la difficulté à pouvoir satisfaire la condition d'adéquation de compositionnalité. Il souligne avant tout qu'une phrase est significative, non pas en admettant

des entités abstraites comme significations, mais en s'intégrant au sein d'une activité pratique dans laquelle la phrase ne prend sens qu'en étant corrélée à d'autres phrases. Le problème est alors de trouver une notion appropriée susceptible d'asseoir une théorie de la signification, mais qui, contrairement au concept de signification, soit une notion nous donnant accès aux ressources de la logique des prédicats du premier ordre afin de nous permettre d'expliquer la compositionnalité des significations. La solution de Davidson est alors de substituer à la notion logiquement obscure de signification celle de vérité logiquement claire c'est-à-dire de remplacer l'opérateur intensionnel « signifie que » par l'opérateur extensionnel « est vrai si et seulement si ». Les théorèmes de sa théorie de la signification auront ainsi la même forme que la convention T, et Tarski a montré comment rendre compte de manière compositionnelle de théorèmes de ce genre. Davidson ne cherche néanmoins pas à montrer qu'une théorie de la signification doit être extensionnelle. Il utilise seulement une définition tarskienne de la vérité afin de définir la vérité dans un langage donné et démontrer l'efficacité du calcul des prédicats du premier ordre pour la représentation des langues naturelles. Mais selon Tarski, il semble bien, comme nous le soulignons, que la Convention T suppose, pour pouvoir fonctionner comme critère d'adéquation, que nous comprenions le langage-objet antérieurement à l'élaboration de la théorie. La convention T de Tarski considère donc la signification ou la traduction comme des notions primitives servant à expliquer le concept de vérité. En effet, alors que Tarski présuppose la traduction et la signification pour définir la vérité, Davidson prend la vérité pour base pour accéder à la signification. Selon la théorie de Tarski, employer la vérité pour expliquer la signification équivaut donc à présupposer la signification. Mais, dans la mesure où ces définitions de la vérité devaient, pour Tarski, se restreindre aux langues formelles, ces relations de traduction, en étant purement stipulatives, ne sont pas problématiques. Davidson doit donc modifier l'approche de Tarski afin de pouvoir l'appliquer aux langues naturelles. Tarski présuppose la connaissance d'un manuel de traduction du langage-objet dans la métalangue pour définir la vérité mais pour Davidson, nous ne pouvons présupposer aucune connaissance de la signification des phrases du langage-objet. Davidson propose alors, afin d'asseoir la théorie sur une base empirique, que l'interprète soit capable d'identifier, dans les attitudes et comportements de l'interprété, ceux qui peuvent être tenus pour des manifestations de la croyance en la vérité d'une énonciation.

La théorie de l'interprétation radicale semble bien répondre à cette exigence d'une base empirique. Il s'agit ici de se placer dans une situation où nous ne connaissons pas la langue de ceux que nous avons à interpréter et où nous ne disposons pas d'autres données que les énonciations et comportements des locuteurs. Autrement dit, si nous voulons éviter de présupposer une connaissance préalable de la signification des phrases du langage-objet, nous devons avant tout examiner les conditions requises afin de déterminer empiriquement si une théorie de la signification s'applique à une langue qui demeure totalement inconnue à l'interprète. La traduction du langage-objet dans le métalangage doit donc être vérifiée. Cette vérification consiste en la détermination de l'extension du prédicat « est vrai », c'est-à-dire à établir les phrases-T qui sont vraies, à partir des attitudes intentionnelles que l'interprète manifeste à l'égard des énoncés du locuteur. Nous donnons alors les conditions de vérité des phrases, et par conséquent, leur signification.

Il ne s'agit donc ni d'expliquer la notion de signification exclusivement en termes de conditions de vérité ni d'affirmer que les concepts de vérité et de signification n'ont aucun lien entre eux. Il est vain de chercher à « réduire » les significations dans un langage naturel aux conditions de vérité. Mais, cela n'implique pas pour autant qu'on doive les traiter comme des entités intensionnelles, directement applicables aux phrases et aux expressions. Néanmoins, si nous voulons que nos théories soient encore testables, la forme d'une théorie de la vérité devra être maintenue : pas de sémantique sans conditions de vérité et pas de théorie de la signification sans sémantique. L'idée directrice d'une sémantique vériconditionnelle repose donc sur cette interdépendance de la signification et de la vérité. Celle-ci consiste ainsi à reconnaître qu'en comprenant notre langue, nous savons alors sous quelles conditions nous tenons une certaine phrase pour vraie. C'est ainsi que le concept de vérité est relié au concept de signification : le concept de vérité est ce qui affecte les relations langage-monde de notre propre langage. Nous en venons donc à dire que la signification est donc bien définie dans un sens minimalement social : à partir de la manière dont l'interprète comprend sa propre langue en interagissant avec le monde, et de l'hypothèse d'un monde commun partagé tant par l'interprété que par l'interprète, l'interprète tente de construire les relations langage-monde émanant des propos de l'interprété (et *vice versa*).

Une autre influence majeure se laisse donc bien entrevoir dans les considérations de Davidson quant à son approche globale de la signification : il défend, tout comme Quine, une

position anti-réaliste en théorie de la signification; la signification n'ayant d'autre réalité que ce que nous pouvons tirer des principes normatifs et holistiques de l'interprétation elle-même. Davidson propose néanmoins une perspective nouvelle au holisme de Quine, en étendant la portée du réseau holistique d'implications, à l'origine de la détermination de la signification et du contenu de la croyance d'un énoncé, et en incluant les désirs, croyances et action de l'agent linguistique. Les données dont nous disposons pour comprendre et interpréter le langage et le comportement d'autrui ne peuvent donc être, selon Davidson, purement comportementales. Davidson s'écarte ainsi de Quine en ce qu'il rejette la notion de signification-stimulus et l'idée même de stimulations sensorielles comme intermédiaires entre le langage et le monde et conçoit le langage comme directement en contact avec les objets et événements du monde. Le holisme prend ainsi véritablement une forme positive chez Davidson : le fait que pour reconnaître un élément, il soit nécessaire de recourir à d'autres éléments, notamment à l'attitude de tenir une phrase pour vraie, celle-ci résultant de ce que l'agent croit et de ce que la phrase signifie, est la garantie de la découverte d'une structure. Le holisme de Davidson implique finalement d'étendre l'idée de Quine d'interdépendance des croyances et des significations aux croyances, significations, désirs et actions. Cette thèse découle ainsi de l'extension considérable qu'il donne au principe de charité, mais aussi de sa conception de l'interprétation radicale, celle-ci étant une entreprise de plus large envergure que la traduction radicale de Quine : interpréter et communiquer est possible, non seulement parce qu'un fonds de vérités et de croyances est nécessairement commun à l'interprète et à celui ou ceux qu'il interprète, mais aussi parce que l'objectif de l'interprétation est l'accord ou la compréhension. La vérité n'est donc pas, du point de vue du principe de charité, indépendante de l'accord.

En effet, selon Davidson, l'impératif ou le présupposé charitable visant à interpréter autrui, c'est-à-dire comme respectant (en grande partie) les lois logiques (version étroite) et comme croyant à la vérité de nos propositions empiriques (version large), revêt une autorité beaucoup plus forte que s'il s'agissait d'une force simplement pragmatique. Il ne s'agit pas d'une hypothèse empirique mais bien d'un principe constitutif de l'interprétation intervenant dans l'explication et la compréhension des comportements linguistiques et non linguistiques des agents. Par conséquent, si de tels comportements ne respectent pas le principe de charité, alors il ne nous est plus possible de les interpréter : en effet, s'ils ne se conforment pas au

principe de charité, c'est donc la preuve que nous ne pouvons pas interpréter leur langage et non que le principe de charité est erroné. Le principe de charité a donc un statut *a priori* ou transcendantal dans l'interprétation – il ne peut alors pas simplement être révoqué par l'expérience dans la mesure où il s'agit bien d'une condition de possibilité de toute activité susceptible d'interpréter les pensées, croyances, désirs, et actions d'autrui.

Davidson se trouve donc face au dilemme suivant : l'idée même d'un statut *a priori* de la charité est incompatible avec un holisme global. Davidson conçoit la charité, sur le modèle kantien, comme condition de possibilité de l'interprétation : la charité a le statut d'une proposition synthétique *a priori*. L'interprétation radicale est alors bien la situation à partir de laquelle on peut envisager les conditions de possibilité d'une expérience possible et elle est, en ce sens kantien, transcendantale. Davidson soutient donc bien la thèse d'une normative constitutive du principe de charité : concevoir les agents comme obéissant à des contraintes de rationalité est une condition nécessaire du processus interprétatif. Cependant, considérer le principe de charité comme condition transcendantale de possibilité de l'expérience le conduit à admettre des propositions *a priori*, ce qui va à l'encontre de son engagement pour un holisme global. Un holisme global mène alors à un faillibilisme global c'est-à-dire à la thèse selon laquelle tout énoncé peut se révéler faux (rejet des vérités *a priori*), ce qui s'oppose à l'idée d'un statut *a priori* de la charité.

Il semble donc que, pour rendre compte du sens et de la portée d'une théorie normative de la signification, qu'une lecture plus faible de principe de charité de Davidson au sens d'un principe constitutif d'intelligibilité soit davantage appropriée. Il s'agit donc bien de maintenir la thèse d'une normativité constitutive de la signification mais de formuler de manière plus faible ou moins rigide la contrainte de rationalité afin de permettre la possibilité de tenir ou d'attribuer des croyances fausses ou irrationnelles. En effet, en se référant à la contrainte de rationalité, le principe de charité, en tant que règle méthodologique, est supposé guider l'interprète, et non pas prescrire une assignation définitive de croyances vraies. L'hypothèse d'intelligibilité contenue dans le principe d'humanité semble bien répondre à cette exigence : elle prescrit de supposer que nous pouvons rendre compte de manière intelligible des croyances d'autrui, au sens où ce sont des croyances que toute personne raisonnable est censée tenir dans des circonstances épistémiques normales incluant la considération de son arrière-plan culturel ou de sa propre histoire personnelle (plutôt que dans le sens strict selon

lequel ce sont des croyances qui sont logiquement cohérentes). La cause d'une énonciation traduira donc ce qu'il est raisonnable pour l'interprété de croire et ce qu'il est raisonnable de croire sera déterminé en évaluant un certain nombre de facteurs : la cohérence logique, les connaissances que celui-ci peut raisonnablement s'attendre à acquérir, etc. L'identification de la cause se fera alors tant sur la base de la considération des relations logiques et inférentielles entre croyances, que sur ses intérêts et engagements pratiques avec le monde.

Par conséquent, la thèse selon laquelle l'intelligibilité est constitutive de la signification et de la croyance tend à s'identifier à celle selon laquelle aucune croyance particulière ne peut raisonnablement être considérée définitivement à l'abri de la révision dans la mesure où il est tout à fait plausible que, dans certaines circonstances, il soit plus raisonnable de réviser une croyance donnée. L'hypothèse constitutive d'intelligibilité n'est donc pas incompatible avec un faillibilisme global. Celle-ci contribue à affaiblir la thèse d'une normativité constitutive en posant qu'aucune norme ou croyance ne peut jamais être déterminée une fois pour toutes. Nous en venons donc, à présent, à nous demander comment cette version plus faible d'une normativité forte ou constitutive diffère d'une normativité faible. La charité conçue comme hypothèse constitutive d'intelligibilité impose ainsi une contrainte forte de vérité sur certaines normes de signification, telles que les lois logiques; mais, n'exclut néanmoins pas que ces normes de signification, bien que bénéficiant d'une forte présomption de vérité, puissent être sujettes à révision dans certaines circonstances. Il apparaît alors qu'au moyen de cette version plus faible d'une normativité constitutive, nous en venons à abandonner la thèse d'une normativité forte et à préconiser l'idée selon laquelle toutes les normes de signification sont faibles. Nous soutiendrons cependant dans le prochain chapitre que cette lecture faible du principe de charité implique certes qu'aucune norme de signification ne puisse complètement être à l'abri de la révision mais requiert également que certaines d'entre elles puissent l'être sur la base d'observations purement empiriques et de considérations strictement pragmatiques. Cette lecture fournit, comme telle, un critère de correction qui ne s'assimile pas aux régularités observables dans l'usage et qui s'avère plus fort qu'une simple exigence pragmatique.

## CHAPITRE V

### VERS UNE CONCEPTION SOCIALE ET OBJECTIVE DES NORMES DE SIGNIFICATION

#### 5.1. Introduction

Nous avons vu que Davidson tente de justifier la nécessité de maintenir la thèse d'une normativité constitutive de la signification tout en préconisant une formulation plus faible du principe de charité. Nous reviendrons d'abord sur le sens, la portée et les conditions qu'il entend donner à cette version plus faible d'une normativité constitutive de la signification (§5.2). Une conception sociale adéquate des normes de signification, doit ainsi, à notre avis, établir le sens d'une normativité forte ou constitutive, c'est-à-dire générer des normes objectivement correctes (c.-à-d. faillibles), normes dont la correction est sujette à révision. Nous serons donc amenés à préciser les raisons qui nous poussent à adopter une approche sociale et faillibiliste des normes de signification. Nous introduirons alors le défi d'une telle approche à savoir, celui de formuler une position qui respecte le contextualisme sans sombrer dans le relativisme en resituant cette question de la normativité au cœur du débat entre Rorty et certains contextualistes tels que Habermas et McCarthy (§5.3). Il nous semblera ainsi intéressant de voir, pour donner une cohérence d'ensemble à notre propos, comment McCarthy qui souscrit aux principales options habermassiennes, s'est efforcé de mettre en lumière les difficultés et les ambiguïtés du pragmatisme de Rorty.

Nous nous pencherons ensuite, avant de nous intéresser plus précisément à la thèse d'une normativité sociale, sur deux points préliminaires : le premier point (§5.4) vise à préciser la nature de la relation entre une application correcte des normes de signification et l'objectivité des énoncés prétendant à la vérité ; le second point (§5.5) questionne la portée d'une application correcte et objective des normes de signification et les raisons qui motivent son extension au-delà du vocabulaire logique. Nous montrerons ainsi, en suivant Quine et Davidson, qu'une dérégulation de la signification et qu'un rejet de la distinction de l'analytique et du synthétique justifient d'une part, ce lien entre normes de signification et prétentions à la vérité (§5.4) c'est-à-dire, ce refus de dissocier la signification de la vérité

empirique des propositions et l'application correcte des normes de signification des prétentions à la vérité, et d'autre part, un sens plus faible attribué à cette dimension constitutive d'une normativité forte sans néanmoins en restreindre sa portée (§5.5).

Ces points préliminaires établis, nous pouvons à présent, entrer au cœur de la question du fondement d'une conception sociale des normes de signification (§5.6). Il sera donc question de déterminer, en poursuivant l'analyse de la théorie interprétative de Davidson, si la signification est davantage un phénomène public (§5.6.1), à savoir que c'est avant tout parce que les ressources de l'individu sont publiques et donc accessibles à l'observation et à la révision des autres locuteurs que nous pouvons parler d'une distinction objective entre un usage correct et un usage incorrect d'un terme ou, s'il s'agit plutôt d'un phénomène social (§5.6.2)<sup>740</sup>, à savoir qu'il ne suffit pas que l'usage d'un terme soit simplement observable publiquement par les autres membres de la communauté mais il faut qu'il se construise au moyen d'une interaction partagée entre les locuteurs, c'est-à-dire d'une collaboration entre au moins deux individus. Nous soulignerons également que la triangulation cristallise cette tension en notant que c'est bien le principe de charité qui rend possible les corrélations entre les phrases tenues pour vrai du locuteur et de l'auditeur mais non l'inverse. Nous en viendrons enfin à soutenir (§5.7) que la triangulation, comme l'interprétation radicale, peut certes justifier le bien-fondé d'une conception publique de la signification, mais qu'une approche exclusivement publique ne rendrait pas adéquatement compte de la nature de la signification et de la communication triangulaire, celle-ci présupposant le principe de charité, et donc une conception sociale de la signification.

## 5.2. Sens, portée et fondement de la normativité de la signification

Notre discussion suivra dans ce chapitre trois fils directeurs : le sens, la portée et les conditions nécessaires au fondement de la normativité.

Nous avons vu dans le chapitre précédent que le principe de charité de Davidson porte sur la question du *sens* à donner à la notion de *correction* ou d'une application correcte des normes de signification. Selon Davidson, il est donc nécessaire d'attribuer un arrière-plan de rationalité à autrui pour interpréter adéquatement son comportement intentionnel. Autrement

---

<sup>740</sup> Cette problématique reprend le questionnement abordé au §3.4.1 entre une dimension publique ou sociale de la signification.

dit, le principe de charité, soit la présomption de rationalité, est un standard constitutif à partir duquel nous pouvons attribuer, comprendre et anticiper les croyances et désirs particuliers d'autrui. Cependant, nous constatons qu'ajouter ce standard normatif de rationalité a également pour effet d'affaiblir le sens rattaché à la notion d'une application correcte des normes de signification. Selon Davidson, nous disposons effectivement de normes de signification qui sont *objectivement correctes* ou *constitutives*, mais nous n'entendons pas pour autant par là *normes absolument et universellement vraies* ou *normes immunisées contre toute éventuelle tentative de révision*. La contrainte constitutive de rationalité doit rendre possible l'erreur. La notion d'une application correcte et objective des normes de signification implique alors une démarche faillibiliste: ces normes de signification, bien que pouvant être vraies, sont toujours sujettes à être révisées ou corrigées. Par conséquent, affirmer que ces normes de signification sont fortes ou objectivement correctes ne signifie pas qu'elles soient totalement infaillibles et exemptes de révision. Cependant, en atténuant le sens d'une normativité forte, la question est alors de savoir si nous avons ôté définitivement toute dimension normative ou, si ce n'est pas le cas, si nous glissons à notre insu d'une normativité forte vers une normativité faible. La réponse à cette interrogation dépendra des conditions requises pour fonder un faillibilisme authentique soit, pour arriver à une véritable révision des normes de signification. Nous avancerons que des normes de signification fortes ou constitutives, bien que sujettes à révision, ne sont pas simplement révisables sur la base d'une quelconque régularité observée dans nos usages linguistiques quotidiens à un moment et à un lieu précis, ni même sur celle d'un quelconque objectif nécessaire à la réalisation d'une fin pragmatique. Aussi, dire que les normes fortes de signification ne sont pas totalement infaillibles, ne signifie pas pour autant qu'elles perdent leur dimension constitutive, c.-à-d. que leur fonction demeure bien celle de contribuer à formuler et à confirmer une théorie interprétative, et, selon toute vraisemblance une théorie interprétative qui serve une fin pragmatique, autrement dit, nous voulons dire que ces normes fortes de signification ne sont pas seulement révisables sur la base de considérations strictement ou exclusivement empiriques ou pragmatiques. De cette manière, la notion d'une notion forte (faillibiliste) est ainsi maintenue et n'a pas été éliminée ou réduite à une normativité pragmatique faible.

Le problème de la *portée* du sens d'une normativité forte se pose maintenant quant à la question d'identifier les normes de signification qui sont constitutives du sens et de la

compréhension de nos énoncés, ou celles qui rendent compte d'un arrière-plan constitutif de raisons, de croyances et d'assomptions de rationalité. Nous avons soutenu<sup>741</sup> que les normes du vocabulaire logique revêtent ce sens fort soit que les normes logiques ne peuvent, ni se réduire à de simples régularités comportementales ni être comprises sur le mode d'un *doit hypothétique* dépendant de nos buts ou intérêts pragmatiques particuliers. Nous verrons au §5.5 que, tout aussi bien, la lecture de Davidson dans un sens large<sup>742</sup> ou métaphorique du principe de charité que sa lecture du statut constitutif de la charité dans un sens littéral<sup>743</sup> comprise sous la formulation plus faible d'une présomption de rationalité et de véridicité des croyances appuient l'idée qu'un sens objectif de l'application des termes, en ce qui a trait à leur correction, englobe aussi l'emploi des concepts non logiques. Ainsi, le principe de charité, en nous permettant de comprendre et d'interpréter adéquatement les croyances et propos d'autrui, mobilise donc des normes de signification issues tant du vocabulaire logique que du vocabulaire non logique<sup>744</sup>. Par conséquent, en ce qui concerne la portée d'une normativité forte, celle-ci s'applique certes à la logique mais, selon Davidson (tout comme d'ailleurs, Wittgenstein, McDowell et Brandom), il est justifié de l'étendre au vocabulaire non-logique (§5.5) et aux termes singuliers.

La question qui demeure en suspens est maintenant celle d'identifier les *conditions* nécessaires pouvant fonder ce sens objectif d'une application des concepts, logiques et non logiques, du point de vue de leur correction. Jusqu'à présent, deux réponses ont été envisagées, à savoir : une première approche *platoniste* ou transcendante des normes constitutives (ou objectivement correctes), et une seconde en termes de pratiques sociales

---

<sup>741</sup> Vol. I, Chap.3, §. 3.3.1.

<sup>742</sup> Davidson, 1993b, 1974b, p. 287.

<sup>743</sup> Au sens littéral, charité, qui veut dire : « don de soi dans sa propre chair », trouve une apparente illustration dans les propos suivants de Davidson, 1993b, 1973a, p. 203-204 : « Si nous ne pouvons pas trouver un moyen d'interpréter les propos et autre comportement d'une créature qui soit manifeste d'un ensemble de croyances en grande partie consistantes et vraies par nos propres normes, nous n'avons aucune raison de considérer cette créature comme rationnelle, comme ayant des croyances ou comme tenant des propos sensés. ».

<sup>744</sup> Brandom étend également la portée d'une normativité forte au vocabulaire non logique en privilégiant les propriétés des relations inférentielles matérielles sur les relations inférentielles strictement formelles et en développant sa conception expressiviste d'un vocabulaire de la logique formelle (§6.5). Par ailleurs, l'idée de Wittgenstein selon laquelle les différences entre les cultures reposent sur un accord massif dans nos jugements fondateurs (*bedrock*) de notre langage et de nos formes de vies (ce qui implique des concepts empiriques ordinaires) vient également appuyer cette approche.

concrètes ou de contextes de discussion et d'accord dans nos jugements. Nous chercherons par la suite<sup>745</sup> à approfondir cette seconde alternative et nous proposerons une version d'une approche sociale des normes objectives qui tentera de répondre aux difficultés soulevées dans les chapitres précédents. Néanmoins, nous tenons à mentionner au §5.6 que Davidson ne nous éclaire pas vraiment et ne nous donne pas d'indices significatifs pour traiter cette question du fondement des normes de la signification dans la mesure où son approche se situe de manière ambiguë entre une approche exclusivement *publique* (qui tend vers un point de vue naturaliste éliminativiste et s'avère donc totalement inadéquate pour fonder le sens d'une normativité forte) et une approche *sociale* de la normativité. En effet, une approche sociale adéquate des normes de la signification doit établir le sens d'une normativité forte tel que nous l'avons souligné ci-dessus, c'est-à-dire, qu'elle doit générer des normes objectivement correctes (c.-à-d. faillibles), normes dont la correction est sujette à révision non pas juste sur la base de régularités observables dans l'usage ou d'intérêts pragmatiques adoptés de manière arbitraire et conventionnelle. Autrement dit, une approche sociale doit laisser place à l'amélioration et non seulement à la différence, au progrès et non simplement à la prolifération d'une pluralité de points de vue. Par conséquent, ce qui nous pousse à fonder une normativité ancrée dans nos contextes sociaux d'accord est le caractère tout à fait inapproprié de l'approche *platoniste* ou transcendante des normes de signification. Le défi majeur qu'elle relève est précisément celui de fonder une notion *faillibiliste* de la correction, soit d'éviter toute dérive du contextualisme au relativisme et par cela même, la pente glissante qui consiste à dire que l'application des normes de signification, en ce qui a trait à leur correction, en dépendant certes d'accords socialement constitués pourrait même aller jusqu'à s'y *identifier*<sup>746</sup> totalement ; dans la mesure où le relativisme implique précisément un anti-faillibilisme. En bref, une conception sociale adéquate doit, entre autres, préserver cette dimension faillibiliste de tout accord social, soit de toute affirmation prétendant à une application correcte des termes. Dans la suite de cette section, nous expliquerons comment le danger d'une dérive relativiste émane de ces approches sociales.

---

<sup>745</sup> Vol. II, Chap. VI, VII, VIII.

<sup>746</sup> McDowell rejette explicitement ce sens d'une normativité forte ou cette forme d'une version sociale forte (ou socialité forte) au §3.2.2.

### 5.3. Contextualisme versus Relativisme.

Nous serons donc amenés à expliquer les raisons justifiant une approche sociale des normes de la signification, à savoir que celle-ci nous semble être une réponse plus pertinente à l'échec des alternatives traditionnelles. La conception platoniste d'une application correcte et objective des normes de signification revêt ainsi une forme moderne chez Descartes. La conception cartésienne de la connaissance requiert l'adoption d'un point de départ conçu comme un *explanans* inexpliqué, une position qui échappe à toutes conditions historiques, sociales ou culturelles. Selon Descartes, pour atteindre la connaissance, nous devons d'abord suspendre notre jugement à l'égard de toutes croyances ou idées reçues afin de ne pas nous laisser influencer par des suppositions injustifiées, des préjugés inconscients, des conventions sociales, etc. Le point de vue épistémique cartésien s'avère donc strictement subjectif et sous-entend une conception correspondantiste de la vérité. L'accès à la connaissance est donc assuré par un point de vue solipsiste libre de toute contrainte. Descartes pensait ainsi qu'adopter ce point de vue de Dieu à l'égard de son *corpus* de croyances, c'est-à-dire transcender notre cadre spatio-temporel, nos conditions épistémiques culturelles et historiques, n'était pas seulement possible mais nécessaire pour acquérir une connaissance et une vérité objective. Les historicistes ou contextualistes, tels que Hegel, Wittgenstein, Habermas, Rorty, McDowell, Davidson, Brandom, Michael Williams et McCarthy contestent cette hypothèse de l'épistémologie cartésienne. Ils soulignent au contraire que toutes propositions prétendant à la vérité doivent être situées dans un espace social et dans un temps historique précis. Nous évoluons toujours au sein d'un contexte conditionné par certaines croyances, standards, normes, présuppositions, qui dépendent elles-mêmes de la particularité de ce contexte-ci et de cette tradition (les normes changent au fil des contextes et traditions). Nous voulons donc souligner que la réalité ne nous est cognitivement accessible qu'à partir de nos pratiques sociales et des standards ou schèmes conceptuels qui sont présentement à notre portée.

L'insistance contextualiste de l'impossibilité d'un point de vue de Dieu implique l'échec de la conception réaliste de l'objectivité et de la théorie de la vérité correspondance. Cette conception correspondantiste de la vérité postule que la vérité d'une représentation mentale se dévoile à travers sa conformité avec le monde objectif : un énoncé est vrai s'il est le *miroir* de la nature. En outre, le réalisme prétend qu'il existe Une seule Manière Une et Objective,

indépendante de nos pratiques et accords socialement établis ainsi que de nos procédures de ratification de décrire le monde et, que nos croyances ou représentations correspondent au monde tel qu'il est, que ce soit le monde réel, physique ou le monde platoniste des Formes Idéales (comme le suggère le réalisme métaphysique). Ainsi, avec le rejet d'une théorie correspondantiste de la vérité, nos affirmations prétendant à la vérité et à une application correcte de l'emploi des concepts prennent un sens intersubjectif ou social. Par conséquent, si nous ne pouvons ni regarder autour de nous vers le monde physique réel, ni devant nous vers un royaume platonique de Formes, ni en nous vers les contenus de notre propre conscience alors, la seule option qui s'offre à nous est de regarder à côté de nous et de nous tourner vers autrui pour fonder la notion d'une application correcte et objective des normes de signification. C'est donc bien l'idée d'une conception correspondantiste de la vérité qui semble s'écrouler sous le poids du tournant linguistique et social. Habermas et Rorty parlent ainsi d'un changement philosophique paradigmatique. Dans *Vérité et Justification*, Habermas résume bien ce changement :

[La mutation de l'autorité épistémique vise à remplacer] la relation à deux termes entre sujet de la représentation et objet représenté par une relation à trois termes, celle de l'expression symbolique qui fait valoir un état de choses devant une communauté d'interprétation. Le monde objectif n'est plus ici une réalité qu'il s'agit de reproduire, mais seulement la référence commune d'une communauté de communication, qui s'entendent les uns avec les autres à propos de quelque chose. Les faits communiqués ne peuvent pas être dissociés du processus de communication [...] La connaissance ne se réduit plus à une correspondance entre propositions et faits. Seul, par conséquent, le tournant linguistique, conduit jusqu'à son terme logique, peut surmonter à la fois le mentalisme et le modèle cognitif du miroir de la nature<sup>747</sup>.

Cependant, en adoptant ce tournant social, un nouveau problème apparaît : le relativisme. Aussi, si nous ne pouvons plus soutenir que nos énoncés prétendant à une vérité objective et à une application correcte des concepts correspondent au monde tel qu'il est, il semble alors que la notion d'objectivité n'exprime plus que l'idée d'une solidarité locale au sein d'une communauté : dire qu'une proposition est vraie, ou que l'application d'un concept est correcte, revient juste à dire que nos pairs culturels l'acceptent ou qu'il s'agit bien de la manière dont nous appliquons le concept localement. Rorty pense que c'est tout ce qui peut résulter du concept d'objectivité une fois que nous avons abandonné les présupposés

---

<sup>747</sup> Habermas, 2001, p.173-174.

représentationnalistes forts soit l'idée qu'il existe une réalité en soi que nous pouvons (réalisme) ou serions incapable de (idéalisme) connaître adéquatement par des représentations. Ainsi, selon Rorty, en rejetant une conception objectiviste<sup>748</sup> de l'objectivité, il vaut mieux se débarrasser de l'idée d'objectivité et voir la correction et la vérité comme affaire de *solidarité*<sup>749</sup>. Il choisit alors d'adopter une position ethnocentrique vis-à-vis de la vérité, de la connaissance et de la signification. Cependant, nous ne voulons pas dire, dans cette perspective, que le penseur ethnocentrique, qu'est Rorty, prendrait aveuglément fait et cause pour les seules valeurs de sa communauté, au mépris de toutes les autres. L'ethnocentrisme dont parle Rorty appartient à ceux qui, au contraire, ont intégré l'évidence d'une *pluralité* leur interdisant de tenir *leurs* valeurs pour les *seules* – et *a fortiori* les seules *vraies* –, sans toutefois les autoriser à abandonner tout point de vue qui, de quelque nature qu'il soit, doit bien être ou devenir *le leur*. Nous retrouvons ici une expression de l'illusoire objectivité que la solidarité se verrait opposer s'il fallait souscrire à un quelconque point de vue de Dieu. A cela, le philosophe pragmatiste oppose donc un sens de la solidarité qui risque évidemment de passer, dans le meilleur des cas, pour l'expression d'un provincialisme que les adversaires du pragmatisme ont toujours stigmatisé. Mais ce provincialisme ne s'apparenterait à un authentique esprit de clocher que s'il se pensait dépositaire d'une nature humaine anhistorique, ce que précisément le pragmatisme répudie. En fait, l'ethnocentrisme pragmatiste, conformément à une inspiration faillibiliste, repose sur l'idée que de meilleures convictions sont toujours possibles:

Il y a toujours place pour une croyance améliorée, car de nouveaux faits, de nouvelles hypothèses ou même un vocabulaire entièrement neuf peuvent se présenter. Le fondement est éthique et non épistémologique<sup>750</sup>.»

<sup>748</sup> Nous entendons ici la thèse de l'objectivité comme correspondance à la réalité externe de nos représentations.

<sup>749</sup> Pour Brandom, 2011, p.1053-1055, l'objectivité peut être conservée et reconstruite à partir des propriétés formelles et pragmatiques du discours intersubjectif.

<sup>750</sup> Rorty, 1994, p.226.

Bien sûr, il serait illusoire de penser que l'objectivité peut cesser d'exercer ses charmes, mais :

Le meilleur argument que nous avons, nous partisans de la solidarité, contre les partisans réalistes de l'objectivité, c'est l'argument nietzschéen qui dit que la consolidation métaphysico-épistémologique des habitudes tentée par la tradition occidentale ne marche tout simplement plus, qu'elle ne fait pas son travail. Elle est devenue un artifice tout aussi transparent que le fait de postuler des divinités qui se trouvent, par le plus heureux des hasards, *nous* avoir choisis pour former leur peuple. L'idée pragmatiste que nous donnons à notre sens de la communauté un fondement « simplement » éthique – ou plus exactement, que nous pensons que notre sens de la communauté n'a d'autre fondement que l'espoir partagé et la confiance engendrée par un tel partage – est ainsi avancée pour des raisons pratiques. Elle n'est pas formulée comme le corollaire d'une thèse métaphysique<sup>751</sup>.

Rorty soutient donc que ce que nous représentons ne peut jamais transcender ce que nous décrivons à partir de nos pratiques de description<sup>752</sup> dont l'usage est soumis à des normes de correction, de vérité et de justification. Le tournant linguistique nous suggère ainsi d'adopter une conception holiste et cohérentiste de la justification, puisque rien ne semble pouvoir servir de pierre de touche à nos processus justificatifs sinon des descriptions linguistiques déjà acceptées. C'est également du moins la perspective que défendent certains philosophes du courant post-positivisme logique (notamment le second Wittgenstein et Sellars<sup>753</sup>). Ainsi, l'adoption d'une nouvelle croyance n'est possible que si celle-ci peut s'accommoder à un réseau de croyances déjà acceptées ; nos croyances sur le monde objectif doivent d'abord pouvoir être harmonisées avec nos croyances antérieures.

Nous montrerons, néanmoins, que certains contextualistes, tels que Brandom<sup>754</sup>, Habermas et McCarthy, soutiennent que l'approche de Rorty est intenable ; ce faisant, ils ne préconisent pas pour autant un retour à la théorie correspondantiste de la vérité. En tant que contextualistes, ils pensent que le constat de l'impossibilité d'un point de vue de Dieu marque

<sup>751</sup> Ibid., p.224.

<sup>752</sup> Hors de son usage dans des descriptions particulières, il n'y a rien d'intéressant à dire sur la réalité. Dans un sens causal, il est certes faux d'affirmer que la réalité est toujours le produit d'une description. Mais sur cette réalité causale isolée, il n'y a rien à en dire, elle ne peut entraîner la thèse minimalement objectiviste (rejetée par le pragmatisme) que « la réalité est toujours sous une description », car l'objet d'une description est toujours une description. Hors de sa réalité dans des descriptions, la réalité n'est *épistémiquement* rien.

<sup>753</sup> Comme le remarque James O'Shea, 2007, p.2, dans son livre d'introduction à la philosophie de Sellars: « in place of foundationalism Sellars offered an account of our knowledge as characterized by holism and fallibilism. ».

<sup>754</sup> Vol. II, Chap. VI.

une prise de conscience historique et culturelle. Leur désaccord avec Rorty ne repose pas sur l'idée d'un accord intersubjectif *versus* l'idée d'une correspondance à une manière Objective de décrire le monde, mais porte sur différentes versions de l'approche intersubjective de l'objectivité, à savoir : une intersubjectivité qui remplace l'objectivité ou une objectivité qui s'exprime en termes intersubjectifs. Ce que des auteurs comme Brandom, Habermas ou McCarthy critiquent est l'inférence de Rorty qui part du rejet du point de vue de Dieu pour ensuite remplacer les notions objectives de vérité des jugements et d'application correcte des concepts par les notions sociologiques d'une solidarité locale dans nos jugements et dans les applications de nos concepts<sup>755</sup>. McCarthy, est ainsi d'accord avec Rorty sur l'incohérence du point de vue de Dieu et l'échec d'une théorie correspondantiste de la vérité. Tous deux s'accordent d'une part, pour dire que la recherche de la vérité ne doit plus être conçue comme la recherche d'une adéquation entre nos représentations et la réalité extérieure et d'autre part, pour souligner le caractère faillible de nos processus de justification. Nos limites cognitives nous interdisent toute connaissance certaine et absolue du monde. Or, bien que tous deux partagent cette prémisse, leur approche diffère quant à savoir si cette prémisse appuie la conclusion que tire Rorty. Selon McCarthy, Rorty lui paraît victime d'une ambiguïté, pour ne pas dire d'une incohérence, en déduisant une conclusion normative qui a pour but de prescrire la nature de la validité de la conclusion à partir d'un jugement de fait, soit d'un état de fait ou du caractère inévitablement contextuel de toute enquête<sup>756</sup>. Autrement dit, il récuse cette inférence ou ce glissement qui part du contextualisme – soit de la thèse selon laquelle les normes de vérité objective et les conditions d'application d'un énoncé (ou d'un terme) sont constituées ou dépendent de moments et de lieux particuliers ainsi que de conditions épistémiques spécifiques – au relativisme – soit la thèse selon laquelle la validité des standards ainsi constitués s'identifie aux accords sociaux obtenus sous ces conditions ou dans ces contextes-ci<sup>757</sup>. Par exemple, il ne fait aucun doute que toutes sortes de conditions

<sup>755</sup> Rorty, 1993, p.449, admet que sa conception de la vérité est bien « sociologique », « I view warrant [and truth] as a sociological matter, to be ascertained by observing the reception of S's statement by her peers. ».

<sup>756</sup> McCarthy, 1991, p.18; Hoy et McCarty, 1994, p.41.

<sup>757</sup> Cf. Vol. I, Chap. III. Il nous paraît utile de revenir sur cette distinction entre un contextualisme épistémologique et un contextualisme ontologique. Le contextualisme épistémologique soutient que notre connaissance de certaines vérités et de normes conceptuelles ne peut être effective qu'au sein de contextes sociaux ou de pratiques de discussion, alors qu'un contextualisme ontologique soutient que

sociales, économiques, culturelles et religieuses ont eu une influence sur Galilée, mais cela ne signifie pas pour autant que ces conclusions sur le mouvement des planètes sont seulement valides pour les Italiens du seizième siècle. C'est bien sûr un état de fait que toute recherche et que toutes normes issues de cette recherche dépendent de la conjonction d'un certain nombre de facteurs; c'est en quelque sorte la condition ontologique propre à toute recherche (comme l'affirme le contextualisme ontologique), mais de telles conditions ne devraient pas, par ce biais, restreindre la validité de ces mêmes normes (ce à quoi conclut le relativisme).

McCarthy soutient que l'inférence erronée de Rorty mène à un changement illégitime entre les perspectives de l'observateur (du scientifique qui étudie le monde qui l'entoure) et celles du participant (de l'acteur immergé dans le monde vécu) et aboutit au niveau pragmatique à une position qui se réfute d'elle-même<sup>758</sup>. En partant de la nécessité d'interagir au sein d'une pratique sociale, Rorty infère une conclusion sur la nature de la validité des standards opérant dans cette pratique. Autrement dit, Rorty passe d'une remarque exprimant notre position comme acteur toujours impliqué et plongé dans la *praxis* quotidienne (contextualisme) à une remarque au sujet de (la validité des standards opérant dans) notre position en tant que participant. Cette dernière résulte de la position de l'observateur d'une pratique et non de la perspective d'un participant engagé à utiliser ces standards-ci dans la

---

les vérités et les normes sont elles-mêmes (et non seulement notre connaissance de ces vérités et normes) définies dans de tels contextes sociaux.

<sup>758</sup> McCarthy, 1991, p.17-20; Hoy et McCarthy, 1994, p.40-42; et McCarthy, 1996, p.96. Le dernier article est la contribution de McCarthy à un symposium sur sa *Critical Theory* et celle de Hoy. Les autres participants étaient Benhabib, Hoy et Rorty. L'objection de McCarthy d'une contradiction performative ne vise pas la critique communément attribuée à la thèse relativiste d'un paradoxe auto-réflexif- soit la critique qu'une approche relativiste selon laquelle la vérité (ou la justification) n'est que le fruit d'un consensus communautaire ou d'une solidarité locale se réfute d'elle-même dans la mesure où cette même affirmation est précisément entrevue sous l'angle d'un consensus non communautaire (une solidarité non locale) de la vérité et de la justification par tout un chacun. Cette objection classique a l'apparence d'un *non sequitur*, puisque le relativiste n'agit pas nécessairement ainsi. Le relativiste qui soutient que la vérité, la justification, etc., est une question de solidarité locale, d'un accord provincialement établi, n'a pas besoin de formuler une proposition sur la nature de la vérité, de la justification, etc., c'est-à-dire sur ce à quoi la vérité correspond pour nous tous, à savoir que ce à quoi la vérité correspond pour chacun d'entre nous est l'accord établi au sein de notre communauté. Il pourrait seulement se prononcer sur la nature de la vérité pour sa communauté et soutenir que la vérité est juste ce sur quoi les membres de sa communauté réussissent à se mettre d'accord. Si c'est effectivement ce qu'il soutient, alors il n'est pas question d'énoncer furtivement une proposition universelle sur la nature relative de la vérité mais bien de formuler un énoncé local sur la validité locale des propositions prétendant à la vérité, ce qui ne se réfute pas d'emblée de lui-même. Comme nous le verrons présentement, McCarthy trouve néanmoins cette version plus raffinée du relativisme, qui est précisément celle qui intéresse Rorty, auto-réfutante.

pratique. Cela dit, le fait que nos pratiques de discussion respectent toujours certains standards et que nous ne puissions saisir la réalité sans que celle-ci soit médiatisée par le biais d'une description linguistique ne justifie pas qu'il faille en conclure que ces standards-ci ont un sens relativement à nos pratiques de discussion. En effet, en rester à cette inférence reviendrait à glisser du côté de la position d'un observateur extérieur à notre pratique de discussion, celui-ci ignorant ou dénigrant la perspective du participant ou de l'acteur engagé dans la pratique, et, ce faisant, à commettre l'erreur de tirer une conclusion normative sur la nature de la validité de ces standards opérant dans la pratique (la conclusion relativiste) à partir de l'observation des conditions ontologiques de l'enquête (prémisse contextualiste). McCarthy remarque qu'en partant de la perspective des participants dans nos pratiques culturelles, la vérité n'est pas comprise comme la notion sociologique d'un accord réel de la majorité tel que Rorty, en partant de la perspective d'un observateur extérieur, l'affirme. McCarthy, tout comme Habermas, tient au contraire à rendre compte, en se plaçant du point de vue des acteurs, de l'attitude que ceux-ci doivent adopter devant le monde vécu. Ces acteurs doivent, selon Habermas, présupposer la stabilité du monde objectif ; ils doivent agir sur la base de certitudes pratiques<sup>759</sup>. Selon lui, seul un concept de vérité transculturelle peut nous permettre de rendre compte des certitudes pratiques nécessaires à l'action. C'est ce concept de vérité pragmatiste qui règne dans la vie de tous les jours (dans ce que nous appellerons, à la suite de Habermas, « le monde vécu »). Afin de rendre compte de l'intuition réaliste sans retomber dans la correspondance, Habermas tient à expliquer le *lien fonctionnel* que nous entretenons avec les choses. Il souligne que les acteurs du monde vécu ont toujours tendance à associer leurs interprétations à la réalité en soi, que l'agir quotidien présuppose nécessairement un réalisme fort. D'où vient la nécessité d'un tel présupposé ? Habermas montre que, en tant qu'acteurs du monde vécu, nous avons toujours une créance en la stabilité

---

<sup>759</sup> Cf. Vol. II, Chap.VII : nous rapprochons la notion de « certitudes pratiques » de celle d'idées régulatrices afin d'éviter que le faillibilisme de nos processus de justification nous entraîne vers la voie du relativisme. Dans la sphère de l'action, Habermas affirme que nous évoluons sur la base de certitudes pratiques que nous tenons inconditionnellement pour vraies (concept pragmatiste de la vérité). Jusqu'à preuve du contraire, nous pouvons (et nous devons) avoir une confiance inébranlable dans le monde objectif. Or, lorsque que nous perdons confiance en nos certitudes, celles-ci prennent la forme d'énoncés hypothétiques pouvant être évalués par la discussion. Selon Habermas, ce passage de l'action à la discussion a pour but le rétablissement des certitudes pratiques indispensables à l'action. Dans ce modèle, le monde objectif joue le rôle de référence commune et implicite des acteurs dans le monde vécu; il est, pour Habermas, un correctif permanent transcendant le contexte chaque fois particulier de la discussion.

du monde sans laquelle nos représentations n'auraient aucune raison d'être. En d'autres termes, une tension conceptuelle existe bien entre ce qui est tenu-pour-vrai par notre communauté linguistique et un concept de vérité référant à une réalité extérieure semblable pour tous soit une tension entre ce qui résulte de nos processus de justification toujours locaux et une référence unique et interchangeable à un monde objectif. Rappelons cette importante thèse de Habermas :

Les pratiques du monde vécu reposent sur une conscience imbue de certitude qui, *in actu*, ne laisse aucune place à une attitude réservée à l'égard de la vérité [...] Dans le rapport pratique à un monde objectif supposé identique et indépendant, les acteurs dépendent des *certitudes pratiques*. Or, celles-ci impliquent à leur tour que l'on tienne les opinions qui commandent l'action pour absolument vraies [...] Au réalisme de la pratique quotidienne correspond un concept de vérité absolue<sup>760</sup>.

Ce réalisme de la pratique quotidienne est conçu par Habermas comme une « nécessité pratique ». Il utilise l'expression « venir à bout du monde » pour exprimer cette nécessité. Il veut dire par là que les acteurs du monde vécu doivent avoir une confiance inébranlable en la stabilité du monde pour fonctionner dans la sphère de l'action. Ainsi, le réalisme faible que défend Habermas se découvre à travers le lien fonctionnel entre action et vérité (et non dans une relation de correspondance avec le monde extérieur). Dans ce modèle, à l'opposé des thèses réalistes classiques, Habermas ne conçoit pas le monde objectif comme une *réalité en soi*, mais bien comme la référence implicite des acteurs du monde vécu. Il veut ainsi souligner à travers son concept de certitude pratique l'attitude objectiviste que doivent adopter les acteurs du monde vécu, au niveau de l'action, par rapport à leurs croyances justifiées localement. En tant qu'acteurs ancrés dans la *praxis* quotidienne, nous devons toujours supposer l'existence d'une vérité inconditionnée :

Dans le monde vécu, les acteurs dépendent d'un certain nombre de certitudes pratiques. Il leur faut venir à bout d'un monde supposé objectif et donc opérer avec la distinction entre croyance et savoir. Faire intuitivement confiance à ce que l'on tient pour absolument vrai est une nécessité pratique<sup>761</sup>.

Par conséquent, les notions que nous retrouvons dans notre culture, c'est-à-dire dans les pratiques des sciences naturelles, de la philosophie, du droit, de l'histoire, de la littérature, etc., n'ont pas pour nous une portée strictement relativiste ou sociologique, comme Rorty

<sup>760</sup> Habermas, 2001, p.302-303.

<sup>761</sup> Ibid., p.196.

tente de le montrer. Aussi, la position de Habermas serait, pour Rorty, inacceptable si celle-ci nécessitait que l'on présuppose l'accès à une vérité immuable (c'est-à-dire l'exigence spécifiquement épistémologique au terme de laquelle nous sommes assurés d'aboutir à la certitude). Par conséquent, l'idée de Rorty selon laquelle nos normes de signification et nos prétentions à la vérité ne sont que le fruit d'un accord de la majorité, ne peut pas résulter de la perspective d'un participant plongé dans la pratique, mais doit provenir de la perspective d'un observateur extérieur, qui, s'il n'énonce pas de point de vue de nulle part (impossible en soi), doit lui-même se placer au niveau du point de vue d'un participant engagé dans notre pratique et devra alors adopter des standards transculturels dans la mesure où ces certitudes pratiques, desquelles nous ne pouvons échapper, rendent compte de manière critique de nos processus de discussion.

Néanmoins, le tournant linguistique semble nous suggérer d'adopter une attitude faillibiliste face au monde. Puisque toute tentative correspondantiste est vouée à l'échec, qu'est-ce qui nous autorise à avoir une si grande confiance envers le monde objectif ? Selon Habermas, dans la vie de tous les jours (au niveau de l'action), nous ne pouvons pas constamment tenir une attitude réflexive par rapport au monde extérieur. Avec lui, nous remarquons que, dans l'action, nous ne sommes pas constamment en train de remettre notre rapport au monde en question. Une telle attitude serait tout bonnement impraticable: « le besoin performatif de certitude pratique exclut [...] toute réserve de principe à l'égard de la vérité<sup>762</sup> ». Habermas soutient qu'au niveau de l'action, il n'y a aucune *thématisation* explicite de notre rapport au monde; les acteurs agissent *naïvement* puisqu'ils ne remettent pas continuellement en cause leur rapport au monde. Il affirme que nous avons le droit de tenir pour vrai nos croyances justifiées, car, dans l'action, le monde vécu fait implicitement ses preuves :

Comme les sujets qui agissent doivent venir à bout « du » monde, ils ne peuvent éviter d'être réalistes, du moins dans le contexte de leur monde vécu. Ils ont d'ailleurs le droit de l'être, car les jeux de langage et leurs pratiques, aussi longtemps qu'ils fonctionnent sans provoquer la déception, font leurs « preuves » par le fait même qu'ils fonctionnent<sup>763</sup>.

Habermas met ainsi en lumière l'idée pragmatiste voulant que la vie quotidienne soit traversée par un grand nombre de régularités. Tant que ces régularités ne nous trompent pas,

---

<sup>762</sup> Ibid., p. 188.

<sup>763</sup> Ibid., p. 194.

nous pouvons nous appuyer, sans crainte, sur le tissu des pratiques courantes. La clé de l'argumentation habermassienne et donc cette continuité de l'action s'explique par le fait que les acteurs présupposent nécessairement un réalisme fort (qui nous montre la nécessité de présupposer un concept de vérité absolue). Le problème survient lorsque le monde nous déçoit. Que devons-nous faire lorsque nos certitudes pratiques nous laissent tomber, lorsqu'on ne peut plus adopter une attitude confiante face au monde objectif ? Quelle attitude devons-nous adopter lorsqu'il y a « rupture de continuité » ? Habermas affirme que, lorsque certaines de nos certitudes pratiques s'effondrent, celles-ci se transforment en prétentions à la validité critiquables et doivent être soumises au contrôle d'un processus discursif de justification. Ainsi, lorsque les régularités de la vie quotidienne ne suffisent plus, nous passons du niveau pragmatiste de l'action au niveau réflexif de la discussion. La validité des certitudes pratiques est hypothétiquement suspendue pour être évaluée de façon épistémique (c'est-à-dire que nous devons donner des *raisons* pour ou contre la prétention à la validité remise en cause). Nous devons ainsi expliciter notre rapport avec le monde objectif. C'est à ce niveau, et à ce niveau seulement, que nous effectuons une thématization de notre rapport à la réalité :

Avec le passage de l'action à la discussion, les participants adoptent une attitude réflexive et, à la lumière des raisons avancées pour ou contre, disputent la vérité des énoncés controversés, vérité qui est alors explicitement érigée en thème<sup>764</sup>.

Le but de ce processus discursif est alors de fournir les raisons de l'échec du monde objectif, et ce, afin de revenir vers la sphère de l'action. Voilà pourquoi Habermas affirme :

Du point de vue des pratiques qui ont échoué et des certitudes pratiques qui ont été ébranlées, les argumentations ont une sorte de fonction réparatrice<sup>765</sup>.

La visée « réparatrice » de ce processus de justification est de rétablir la confiance en nos certitudes pratiques. « Les discussions sont ainsi des sortes de machines à laver, filtrant ce qui est rationnellement acceptable pour tous<sup>766</sup> ». Le discours nous permet ainsi d'atteindre ce que Habermas qualifie d'une *vérité épistémique*; c'est-à-dire la vérité qui est justifiée par les meilleures raisons que peut nous fournir le processus discursif (ce qu'il appelait aussi *l'acceptabilité rationnelle*). Bref, si nous quittons, pour un instant seulement, la sphère de la

---

<sup>764</sup> Ibid., p. 303.

<sup>765</sup> Ibid., p. 304.

<sup>766</sup> Habermas, 2003, p.75.

pratique pour expliquer des segments du monde objectif qui nous ont déçus, c'est pour mieux renouer avec celle-ci. Ce mouvement de l'action à la discussion vise donc à recomposer la texture habituelle de la vie quotidienne (de recoudre le tissu brisé de la praxis quotidienne).

L'observation de Habermas sur la nécessité des certitudes pratiques nous permet ainsi de mettre en évidence le caractère faillible de la connaissance (la vérité pragmatiste peut être démentie par le monde vécu, la vérité épistémique par la force d'un meilleur argument encore imprévisible). En effet, le concept de vérité épistémique (qui sert la cause de la praxis en permettant aux acteurs du monde vécu un retour au cours habituel des choses) ne nous garantit pas que le monde vécu ne démente pas, un jour ou l'autre, le résultat de l'argumentation :

Le monde vécu, avec ses conceptions fortes de vérité et de savoir qui sont fonction de l'action, se manifeste à l'intérieur de la discussion et y apporte le point de repère qui, tout à la fois, transcende la justification et rappelle aux interlocuteurs que nos interprétations sont faillibles<sup>767</sup>.

Malgré la reconnaissance de ce faillibilisme, Rorty refuserait sans doute de reconnaître comme valide cette référence nécessaire à un concept transculturel de vérité. Habermas, qui développe sa conception bidimensionnelle de la vérité en partie pour répondre au néo-pragmatisme rortien, tente de deviner ce que Rorty répliquerait à sa nouvelle thèse épistémologique:

Certes Rorty ne contesterait pas ce lien entre discussion et action. Il accepterait aussi d'établir un lien entre les deux perspectives - celle des interlocuteurs qui cherchent à se convaincre réciproquement de leurs interprétations et celle des sujets qui agissent et sont engagés dans leurs jeux de langage et dans leurs pratiques. Mais Rorty ne distinguerait pas ces perspectives de façon à les relativiser l'une par rapport à l'autre. Au point de vue des participants à l'argumentation, sa description emprunte cet aspect des dialogues qui nous tiennent prisonniers et nous interdisent de nous évader des contextes de justifications; à la perspective des acteurs, elle emprunte le mode de la nécessité de venir à bout du monde. La projection réciproque de ces deux perspectives contraires constitue ensuite la certitude ethnocentrique qui mène Rorty à se demander ce qui nous obligerait, d'une façon générale, à nous efforcer de faire converger la prise de conscience contextualiste, que nous devons à nos expériences de l'argumentation, avec le réalisme de tous les jours dont il crédite le monde vécu.<sup>768</sup>

---

<sup>767</sup> Habermas, 2001, p. 195.

<sup>768</sup> Ibid., p.196-197.

Habermas, en supposant toujours ce que Rorty répliquerait à sa conception, continue en affirmant que :

Si les acteurs du monde vécu ne peuvent - jusqu'à nouvel ordre - s'empêcher d'être « réalistes », tant pis pour eux. Il incombe alors aux philosophes de réformer le concept de vérité qui est celui du bon sens, puisque c'est lui qui nous induit en erreur.<sup>769</sup>

Habermas voit certainement juste en supposant que Rorty concevrait cette tendance qu'ont les acteurs à tenir pour inconditionnelles leurs croyances non pas comme un besoin performatif de l'action, mais bien comme le reflet d'une tradition objectiviste qui module notre compréhension du monde. Rorty, en bon pragmatiste, n'essaierait pas d'expliquer fonctionnellement cette tendance réaliste des acteurs : à la suite de Dewey, il exhorterait plutôt les philosophes contemporains à développer de nouveaux vocabulaires afin de changer cette tendance objectiviste du sens commun de notre communauté<sup>770</sup>. En outre, Rorty croit que si nous sommes en mesure de nous libérer de l'emprise de la philosophie occidentale classique<sup>771</sup>, la nécessité de s'appuyer sur des certitudes pratiques (comme la décrit Habermas) se révélera n'être qu'un autre effet pernicieux de l'omniprésence de la métaphysique grecque. Ce point de tension entre les deux auteurs est ainsi clairement résumé par Michael Barber :

« Richard Rorty challenges Jürgen Habermas's belief that every validity-claim raised within context-bound discussions contains a moment of universal validity. Rorty argues that immersion within contingent languages prohibits any neutral, context-independent ground, that one cannot predict the defense of one's assertion before any audience, and that philosophy can no more escape its contextual limitations than strategic counterparts<sup>772</sup>. ».

McCarthy soutient ainsi que Rorty ne peut pas à la fois épouser une position ethnocentriste et rejeter l'idée de valeurs transculturelles; l'ethnocentrisme de Rorty qui s'inscrit dans nos pratiques culturelles et sociales semble contredire sa thèse d'une réforme linguistique :

<sup>769</sup> Ibid., p. 196-197.

<sup>770</sup> Rorty, 1997b, p.17.

<sup>771</sup> Rorty, 1992, p.184: Rorty ne croit pas que la culture occidentale soit « partout structurer autour de notions transculturelles de vérité » : Par bonheur, elle n'est structurée de la sorte qu'en certaines de ses parties. Aussi puis-je avoir recours à des choses qui sont dites et faites dans d'autres secteurs. Je peux opposer certains éléments de notre culture à d'autres [...] Nous appartenons à une culture que n'ont pas seulement nourrie 'la Bible, Socrate, Platon et les Lumières', mais aussi, par exemple, Rabelais, Montaigne, Sterne, Hogarth et Mark Twain ».

<sup>772</sup> Barber, 2004, p.51.

« Dislodging [context-transcending presuppositions] is less a matter of frank self-acceptance than radical self-transformation<sup>773</sup>. ».

Des auteurs comme Habermas et McCarty soutiennent en effet que nous avons besoin de présupposer un moment d'inconditionnalité critique à l'intérieur de nos processus de justification. Présupposer ce moment d'inconditionnalité est nécessaire, selon Habermas, à l'explication de la force critique de nos énoncés ; c'est ce moment qui nous permet d'expliquer la possibilité même de nos processus d'apprentissage. Une question se posera alors. Comment Rorty peut-il expliquer cette force critique s'il renie tout concept transculturel de vérité ? D'où provient la force de révision inhérente aux processus critiques si tout énoncé n'est davantage vrai ? Afin de combler le vide méthodologique laissé par l'abandon du moment d'inconditionnalité, Rorty reconsidère la force critique du langage au moyen de sa thèse des « substitutions langagières ».

McCarthy avance que prôner une réforme du langage n'élimine pas pour autant le problème d'une contradiction performative<sup>774</sup> : nous devons plutôt chercher à expliquer comment concilier une position ethnocentriste et l'idée d'une réforme linguistique. Aussi, comme nous l'avons déjà laissé entendre, Habermas affirme que nos prétentions à la vérité doivent contenir un moment d'inconditionnalité qui transcende le processus de justification à chaque fois particulier. Comme le remarque Rorty :

« Habermas says that every validity claim has « a transcendent moment of universal validity [which] bursts every provinciality asunder » in addition to its strategic role in some context-bound discussion<sup>775</sup>. ».

<sup>773</sup> Hoy et McCarthy, 1994, p.32.

<sup>774</sup> Comme McCarthy le souligne dans *Ideals and illusions*, p.20, et *Critical Theory*, p.42, cette contradiction performative est mise en exergue dans la tentative de Rorty de justifier ses propres prises de position sur la justification comme étant une affaire de solidarité locale ou d'acceptation par nos pairs culturels. Il semble que Rorty reconnaisse (au moins du point de vue du participant) que la justification ne se réduise pas à la pratique d'une solidarité locale de nos pairs culturels. Il tente plutôt de fournir des arguments et des raisons valables pour défendre le bien-fondé de sa position. Aussi, si un point de vue externe à la pratique est impossible, et si les pratiques au sein desquelles nous sommes engagés à notre insu véhiculent par le biais d'un accord – ou de la solidarité – des notions transcendantes de vérité, de justification, et des normes conceptuelles, alors Rorty doit bien faire appel à de telles notions. McCarthy fait également remarquer que Rorty rejette « officiellement » ces mêmes notions et se retrouve pris dans une contradiction performative.

<sup>775</sup> Rorty, 2000a, p.6.

Qu'est-ce que cela veut dire exactement ? Habermas veut ici souligner une autre attitude humaine objectiviste selon lui inévitable. Cette fois, cependant, il n'est pas intéressé par l'attitude des acteurs de la *praxis* quotidienne, mais bien par l'attitude *des interlocuteurs participant à un processus de discussion visant l'entente*. Comme il l'a fait avec les acteurs plongés dans la *praxis* quotidienne, Habermas soutient que les interlocuteurs ne peuvent s'empêcher d'être réalistes dans la sphère discursive. Habermas affirme que ceux-ci doivent considérer les énoncés qu'ils défendent dans les processus de justification comme *inconditionnellement vrais*. Cette attitude objectiviste leur permet, du moins par sa prétention, une assise réaliste qui transcende le caractère contextuel du processus de discussion; *elle permet, en d'autres termes, un moment d'inconditionnalité qui permet d'aller au-delà des limites contextuelles de l'argumentation*<sup>776</sup>. Cette attitude ne présuppose ni de monde en soi, ni de critère infaillible pour atteindre la vérité. Pour Habermas, le rapport que nous entretenons avec la réalité est opératoire et non pas représentationnel :

Il s'agit ici d'une supposition formelle qui ne préjuge d'aucun contenu déterminé et ne nous incite pas non plus à parvenir à cette « image correcte de la nature des chose » que Rorty associe toujours à toute intuition réaliste<sup>777</sup>.

Selon lui, cette attitude que nous adoptons face à nos prétentions à la vérité est une présupposition inévitable des processus argumentatifs. Wellmer résume magnifiquement cette pensée de Habermas dans son livre *Vérité, contingence et modernité*:

À chaque fois que nous élevons des prétentions à la vérité sur la base de bons arguments ou de preuves irréfutables, nous *supposons* [...] qu'aucun argument ou aucune preuve ne surgira ultérieurement, qui remettrait nos prétentions à la vérité en question<sup>778</sup>.

Selon Habermas, lorsque nous soutenons, devant une communauté d'interprétation, une prétention à la vérité – justifiée par de bonnes raisons (Wellmer) ou par la force du meilleur argument (Habermas) – nous présupposons toujours que celle-ci est en mesure de résister à tous les publics dans tous les contextes possibles. Cette attitude que nous devons prendre au

---

<sup>776</sup> Notons ce lien important entre le moment d'inconditionnalité inhérent à nos prétentions à la vérité et le concept de vérité transculturelle. Selon Habermas, présupposer que – nos énoncés sont vraies inconditionnellement (qu'il reflète une vérité transculturelle), nous donne accès au moment où, à l'intérieur même de nos processus de justification, nous allons au-delà des limites contextuelles de l'argumentation.

<sup>777</sup> Habermas, 2001, p.194.

<sup>778</sup> Ibid., p. 192.

sujet de nos énoncés est la condition de possibilité même des processus discursifs visant l'entente. Cependant, pourquoi devons-nous absolument présupposer ce moment d'inconditionnalité ?

Habermas et McCarthy s'opposent tous deux aux conceptions déflationnistes semblables à celle de Rorty<sup>779</sup>. L'une des raisons derrière leur appréhension est qu'ils ne croient pas que de telles conceptions puissent correctement expliquer le fonctionnement de nos *processus critiques de révision*. Selon eux, la conception déflationniste de Rorty ne peut mobiliser les ressources conceptuelles suffisantes à l'explication du déroulement de ces processus d'apprentissage. Pour Habermas et McCarthy, ces processus ont besoin d'un concept fort de vérité servant d'étalon critique, et ce, afin que nos idées reçues puissent être comparées, justifiées ou modifiées à la lumière d'une « idée régulatrice ». Selon ces deux auteurs, cette contre-épreuve critique est nécessaire pour pouvoir mettre en doute nos prétentions à la vérité et les standards mêmes de la vérité dont nous avons hérités culturellement. C'est ici que se dévoile à nous l'utilité pratique du moment d'inconditionnalité dont parlent Habermas et McCarthy. Selon eux, seul ce moment peut nous permettre d'expliquer comment le point de vue critique est possible.

Si Habermas croit nécessaire que nos prétentions à la vérité renferment un moment d'inconditionnalité transcendant le contexte de justification, c'est d'abord dans le but d'avoir une assise nous permettant de mettre en doute les différents points de vue que notre communauté a (ou pourrait avoir) sur le monde. Le moment d'inconditionnalité est une présupposition nécessaire de nos processus argumentatifs, car si les interlocuteurs supposaient toujours que leurs prétentions à la vérité étaient provinciales et faillibles, *il serait impossible d'expliquer ce qui leur permet de critiquer et de réviser leurs propres croyances*. C'est bien ce caractère inconditionné du monde vécu qui permet aux participants à l'argumentation de développer cette conscience faillibiliste. Ce qui génère ce développement, c'est ce rôle d'assise critique que joue le monde vécu dans nos processus discursifs, assise sur laquelle s'éprouvent nos prétentions à la vérité. Le monde vécu est donc objectif en deux sens distincts : d'abord parce qu'il sert de point de référence commun dans le processus d'argumentation, et ensuite parce qu'il nous rappelle que nos croyances bien fondées peuvent se révéler fausses :

---

<sup>779</sup> Ibid., p. 297-299; McCarthy, 1992a, p.77-100.

[Lorsque les jeux de langage et les pratiques échouent], le monde n'a pas joué le jeu comme on l'attendait de lui. C'est ce démenti vécu dans la pratique, celui d'un échec par lequel le monde révoque performativement sa disposition à jouer le jeu, *qui* constitue le concept d'objectivité<sup>780</sup>.

Cette idée régulatrice d'un monde objectif nous permet donc de rendre compte de nos processus d'apprentissage et de nos processus critiques. McCarthy fait la remarque suivante:

Nous formulons généralement des prétentions à une vérité inconditionnelle comme étant contextualisées et faillibles (tel que je viens juste de le faire); et c'est ce moment d'inconditionnalité qui nous ouvre à la critique d'autres points de vue. Sans ce moment d'idéalisation, nous serions incapables de remettre en cause nos idées reçues et nos pratiques et de les confronter à la critique, processus essentiel à l'instauration d'un consensus nous obligeant à élargir notre champ de vision et nous éduquant à regarder les choses sous un angle différent. C'est précisément ce surplus de signification transcendant le contexte, « régulateur » dans notre notion de vérité, qui nous préserve d'être enfermés dans ce qu'il nous arrive d'admettre en un lieu et en temps particuliers, qui nous ouvre à des possibilités alternatives qui recèlent l'altérité et la différence, et qui ont été décisivement invoquées par les penseurs post-structuralistes.<sup>781</sup>

C'est ce « surplus de signification régulateur » contenu dans l'idée d'un moment d'inconditionnalité qui nous ouvre à la critique. Sans cette assise critique, nous ne mettrions jamais en doute les opinions auxquelles nous sommes confrontés<sup>782</sup> puisque nous aurions la conviction qu'il ne sert à rien de s'engager dans le processus critique (chaque point de vue étant local et faillible, toute vérité définitive nous étant inaccessible). Cette attitude objectiviste de l'interlocuteur est donc conçue comme un présupposé fonctionnel permettant la confrontation de différentes interprétations de la réalité ; elle permet l'intégration d'un aiguillon critique expliquant les processus de discussion<sup>783</sup>.

<sup>780</sup> Habermas, 2001, p.221.

<sup>781</sup> McCarthy, 1992a, p 99.

<sup>782</sup> Toute interaction, même conflictuelle, aurait comme horizon ce que Habermas, 1999, p. 179, suivant Mead, 1963, appelle l'« échange idéal de parole » : « Le point de vue moral recherché, qui est préalable et sous-jacent à toutes les controverses, résulte d'une réciprocité fondamentale, incluse dans l'activité orientée vers l'intercompréhension ».

<sup>783</sup> Notons aussi une autre justification de cette attitude objectiviste des interlocuteurs que défend Habermas dans *Vérité et justification*. Dans ce livre, Habermas soutient que les participants aux processus de discussion visant l'entente donnent à leurs prétentions à la vérité une saveur universelle. Dans les processus d'argumentation (dans la sphère discursive), nous devons adopter une attitude imbue de platonisme devant nos énoncés (une position qui rappelle celle des acteurs du monde vécu dans la *praxis*). Qu'est-ce qui justifie cette attitude objectiviste ? Au niveau de la discussion, ne sommes-nous pas justement en train de traiter avec des énoncés problématiques (et n'avons-nous pas ainsi une attitude faillibiliste devant le monde) ? Dans *Vérité et justification*, Habermas affirme que

Étant donné qu'elle rejette un tel moment régulateur, Habermas et McCarthy demandent à la théorie rortienne d'expliquer d'où celle-ci tire sa *force critique*. D'où provient cette force, si aucune croyance n'est davantage *vraie* ? Puisque Rorty affirme que le langage clôt notre champ épistémique en nous imposant un ensemble de catégories inévitables, Habermas et McCarthy doutent que celui-ci soit en mesure d'expliquer d'où vient le pouvoir critique de nos arguments. Notons d'abord que Rorty ne nie pas l'existence d'une réalité extérieure; il ne nie pas que le monde extérieur s'impose à nous indépendamment de nos préférences. Il ne met pas en doute l'idée voulant que nous soyons « en contact avec le monde<sup>784</sup> ». Rorty se rend bien compte que le monde extérieur s'impose à nous et influence nos conceptions de celui-ci<sup>785</sup>. Le point que Rorty met de l'avant en utilisant le terme *antiréalisme*, est que tout accès direct à la réalité « nue » est interdit au sujet de la connaissance. *La vérité ne découle pas d'un rapport correspondantiste entre nos énoncés linguistiques et la réalité*. Étant donné ce rejet du correspondantisme, Rorty suggère que le concept de *réalité unique semblable pour tous* peut et doit être abandonné lors de l'explication du déroulement de nos processus cognitifs; ce concept est, en fait, théoriquement inutile. Les thèses de Rorty mettent ainsi de l'avant l'idée radicale que le tournant linguistique a épuisé la validité conceptuelle du terme de réalité (et, du même coup de celui de vérité) en tant que référence transculturelle<sup>786</sup>. Pour comprendre sa position, il faut souligner le lien qui existe entre sa conception de la vérité et le tournant linguistique. Tout d'abord, Rorty croit que la « vérité » ne peut venir que des descriptions langagières découlant de notre expérience au monde. Le monde extérieur ne peut être qualifié de vrai ou de faux; seuls les énoncés portant sur celui-ci peuvent être ainsi

---

« les raisons pour lesquelles les interlocuteurs, en tant que sujets capables de parler et d'agir, doivent se comporter ainsi [c'est-à-dire présupposer que leurs prétentions à la vérité sont vraies universellement] ne sont pas si difficiles à comprendre, si on décrit leurs discussions, d'un point de vue pragmatique, telles qu'elles sont enchâssées dans le monde vécu. Nous l'avons vu, dans leur pratique, les individus socialisés ont besoin de certitudes qui conservent ce statut aussi longtemps qu'elles se nourrissent d'un savoir accepté sans réserve » (p. 192-193). Cette attitude objectiviste devant nos prétentions à la vérité a donc pour but de *faciliter le retour à la praxis quotidienne*. Si nous ne pouvons nous empêcher d'adopter ce regard empreint d'inconditionnalité devant nos énoncés visant la justification, c'est d'abord pour renouer avec la sphère de l'action où règnent, selon Habermas, les certitudes pratiques.

<sup>784</sup> Rorty, 1995a, p. 34.

<sup>785</sup> Afin de rendre compte de ce contact avec le monde, Rorty, 1995a, p. 34-35, souligne que le monde extérieur exerce sur nous des « pressions causales » qui influencent inévitablement nos croyances. Or, ces pressions causales n'ont pas le potentiel de dévoiler la nature intrinsèque de la réalité.

<sup>786</sup> Comme nous le verrons, Habermas, qui partage des prémisses semblables, utilisera son arsenal théorique pour sauver le concept de réalité d'un sort déflationniste (sans tomber dans la représentation).

qualifiés. Le prédicat de vérité ne peut être associé qu'à des descriptions qui prennent la forme d'énoncés linguistiques:

La vérité ne saurait être là, dehors - elle ne saurait exister indépendamment de l'esprit humain - parce que les phrases ne sauraient exister ainsi, elles ne sauraient être là, devant nous. Le monde est là, dehors, mais pas les descriptions [langagières] du monde. Seules elles peuvent être vraies ou fausses<sup>787</sup>.

Selon Rorty, le monde n'est jamais vrai ou faux; il est ce qu'il est. Le prédicat de vérité ne s'applique qu'à nos seules représentations linguistiques du monde. Cette première observation permet de clarifier la position rortienne :

Dire que le monde est là, qu'il n'est pas notre création, c'est dire, non sans bon sens, que la plupart des choses de l'espace et du temps sont les effets de causes qui n'impliquent pas d'états mentaux humains. Dire que la vérité n'est pas là, c'est simplement dire que sans phrases il n'y a point de vérité, que les phrases sont des éléments des langages de l'homme, et que les dits langages sont des créations humaines<sup>788</sup>.

Rorty ajoute à cette observation l'importante hypothèse du tournant linguistique voulant que le langage nous impose des catégories et des modes de pensée qui structurent nécessairement notre champ épistémique (le champ de raisons et d'arguments possiblement mobilisables dans nos processus de justification). Rorty cautionne d'ailleurs l'idée voulant que toute description langagière présuppose nécessairement l'arrière-plan d'une communauté linguistique. L'influence de cette communauté sur nos processus cognitifs est, pour Rorty, incontestable. Nos descriptions du monde resteront toujours locales et contextuelles puisqu'il est selon lui impossible d'établir un langage qui ne serait pas attaché à un horizon de significations précis. Il appelle cette thèse la *contingence du langage*<sup>789</sup>. Tout langage porte, selon Rorty, l'empreinte d'une communauté particulière et il est donc impossible, selon lui, d'établir un langage universel qui nous permettrait de décrire le monde de façon satisfaisante pour tous et ainsi, de découvrir une vérité transculturelle. Il l'exprime, comme suit :

il n'y a pas moyen de sortir des divers vocabulaires que nous avons employés pour trouver un méta-vocabulaire qui, d'une façon ou d'une autre, tiennè compte de tous les vocabulaires possibles, de toutes les façons possibles de juger et de sentir<sup>790</sup>.

---

<sup>787</sup> Rorty, 1997b, p. 23.

<sup>788</sup> Ibid., p. 23.

<sup>789</sup> Ibid, p. 17.

<sup>790</sup> Ibid., p. 17.

Dans la perspective rortienne, il n'existe pas de vérité au sens platonicien (transculturelle) puisqu'il n'existe pas de langage non humain (transculturel) permettant de décrire le monde de façon unique: *on ne peut établir un métalangage*. Puisque nous sommes prisonniers du caractère contextuel du langage, tous les énoncés que nous émettrons sur le monde le seront aussi<sup>791</sup>. Voilà pourquoi Rorty affirme que:

Il n'existe pas de contact qui, antérieur au langage, permettrait de mettre le doigt sur ce qu'est un objet en lui-même, par opposition à ce qu'il est au regard des descriptions variées que nous en donnons<sup>792</sup>.

Rorty répond à Habermas et à McCarthy en affirmant que c'est par des moyens qui n'appartiennent pas à la philosophie classique que l'on doit rechercher le point de vue critique. L'auteur tourne donc le dos, une fois de plus, à la notion de vérité transculturelle et utilise le pouvoir du langage pour retrouver cette force critique. D'après lui, ce n'est pas un concept fort de vérité référant à un monde objectif unique qui doit nous servir d'idée régulatrice, mais bien ce qu'il appelle les « suggestions alternatives concrètes » permettant une *redescription* de ce qui est communément accepté :

À l'opposé [de Habermas et de McCarthy], je pense que ce qui nous permet de nous livrer à de telles critiques [celles de nos prétentions à la validité particulière et des standards mêmes de la vérité] réside dans des suggestions alternatives concrètes, des suggestions sur la façon de redécrire ce dont nous parlons<sup>793</sup>.

Ainsi, pour Rorty, ce qu'il importe de faire, c'est de confronter des vocabulaires, d'en proposer un meilleur, et non d'effectuer une comparaison avec un monde objectif unique<sup>794</sup>.

---

<sup>791</sup> Certains pourraient mettre de l'avant que cette théorie est autoréfutante : elle semble défendre l'idée qu'il n'existe pas de vérité. Mais Rorty se garde bien d'affirmer que ses propos touchent ou représentent « l'état réel des choses ». Sa mission n'est que de proposer un schéma qui serait plus utile (utilité relative) ou mieux adapté (méliorisme) aux conditions de vie actuelles. Toute sa thèse n'est que « recommandation ». Ce qu'il désire faire, c'est de substituer un vocabulaire à un autre. Il veut remplacer le vocabulaire correspondantiste par un vocabulaire pragmatiste. Il veut, en d'autres termes, « rendre attrayant le vocabulaire auquel vont mes préférences ».

<sup>792</sup> Rorty, 1990b, p.9.

<sup>793</sup> McCarthy, 1992b, p. 179.

<sup>794</sup> C'est pourquoi l'auteur accorde beaucoup d'importance aux métaphores. Rorty abolit la distinction entre *phrase littérale* et *métaphore*. Selon lui, il n'y a que des substitutions de nouvelles métaphores à d'anciennes. Aucune n'est plus proche de la nature intrinsèque des choses (littéralement), il y en a seulement des plus adaptées que d'autres pour une communauté précise (Bouveresse, 1992, p. 38). Rorty s'accorde donc avec Nietzsche pour dire que la vérité est une *armée mobile de métaphores*. Cela montre aussi que l'approche de Rorty ne peut pas être qualifiée de « déconstructiviste » car il n'a aucun

De telles redescriptions sont, selon Rorty, assez puissantes pour nous fournir les « idéaux normatifs » à la base de nos procédés critiques. Il croit que *l'attrait* que nous éprouvons pour ces suggestions alternatives est suffisant à un point de vue régulateur :

Contrairement à ce que dit McCarthy, ce n'est pas un « moment d'inconditionnalité qui nous expose à la critique à partir d'un autre point de vue », c'est *l'attrait* que nous éprouvons pour cet autre point de vue<sup>795</sup>.

Selon Rorty, nous n'avons pas besoin d'une référence qui tienne *au-delà de notre contexte* particulier pour expliquer nos processus critiques (c.-à-d. effectuer des évaluations et des comparaisons<sup>796</sup>). Le penchant que nous éprouvons pour certaines redescriptions langagières est suffisant, selon lui, pour expliquer comment se déroulent nos changements de croyances et, du même coup, combler le vide laissé par l'abandon du concept classique de vérité<sup>797</sup>. Rorty ne fait donc aucune distinction entre le concept de *tenu-pour-vrai* (la façon dont nous et notre communauté comprenons le monde maintenant) et celui de *vérité* (ce qu'est le monde indépendamment de nos préférences). Rorty pense donc que ce qui permet d'expliquer l'adoption d'une nouvelle croyance comme vraie, c'est son acceptation par une communauté particulière. En bref, Rorty ne croit pas qu'une référence à un monde objectif unique soit nécessaire au bon fonctionnement de nos processus de justification. Il croit plutôt qu'il est préférable de débarrasser nos processus de justification du concept de vérité transculturelle servant d'étalon critique. Selon lui, être en contact avec le monde signifie être en contact avec une communauté linguistique particulière. Il prend ainsi parti pour William James et John

---

désir de montrer les problèmes internes aux conceptions philosophiques classiques – mais bien d'en proposer de plus intéressantes.

<sup>795</sup> McCarthy, 1992b, p. 180.

<sup>796</sup> Il tente même d'expliquer certaines grandes révolutions scientifiques à la lumière de cette interprétation (comme « les suggestions de Galilée sur la façon de redécrire l'univers aristotélicien »). Rorty, 1995a p.45.

<sup>797</sup> Il est intéressant de remarquer que cette approche par les redescriptions langagières est intimement liée au désir qu'a Rorty, 1992, p.184, d'encourager certaines traditions de pensée marginales (« il existe des traditions qui s'opposent à ce que nous tenons pour central, ou qui n'occupent qu'une position marginale; ce sont celles que je désire encourager ») – tout particulièrement celle véhiculée par la littérature. Il y a évidemment un rapport direct entre l'apologie que fait Rorty de la littérature, de l'imagination et du besoin de créer et cette nouvelle approche philosophique fondée sur les redescriptions langagières. Comme nous l'avons laissé entendre précédemment, Rorty veut mobiliser les puissantes ressources conceptuelles de l'imagination que déploie la littérature et les utiliser en philosophie. Cette tactique lui permet d'affirmer que le langage n'est pas simplement limitatif et qu'au contraire, l'approche par redescriptions langagières permet de mettre en lumière plus de « choix alternatifs » que ce que permet le concept fort de vérité jouant le rôle de contre-épreuve réaliste.

Dewey et réduit le concept de vérité au concept « d'utilité ». Comme le disait Dewey, « [les catégories de la raison] ne représentent rien de plus que les intérêts d'une certaine race, d'une certaine espèce – leur 'vérité' se résume à leur 'utilité'<sup>798</sup> ». En d'autres mots, Rorty croit que la vérité est simplement ce qui se montre profitable pour nous de croire, selon nos raisons : « [les pragmatistes] considèrent la vérité, conformément à la formule de James, comme ce qu'il est avantageux pour nous de croire<sup>799</sup> ». En ce sens, la vérité n'est qu'une affaire de pratiques sociales dans le sens où ce qu'il est bon pour nous de croire maintenant est toujours sous réserve d'une meilleure compréhension. La redéfinition de Rorty du concept de vérité vise ainsi l'abandon de la compréhension classique de la vérité comme un « but fixe<sup>800</sup> » que nous devrions atteindre. Rorty soutient alors que le processus de rectification publique de nos croyances par une communauté linguistique locale est suffisant à l'explication du déroulement de nos processus de justification.

Il semble donc que c'est l'explication de la force critique et son lien avec le concept d'inconditionnalité qui causent le plus de problèmes à Habermas et à Rorty. Chez Habermas, l'accès à la force critique nécessite que nous tenions nos prétentions à la vérité comme inconditionnellement vraies. Cette inconditionnalité nous assure un point de vue qui transcende le provincialisme du processus de justification. Chez Rorty, l'attrait que nous éprouvons pour une alternative concrète est, selon lui, suffisant pour expliquer le point de vue critique. Pour lui, force et inconditionnalité ne vont pas de pair. Il suggère, au contraire, que nous sommes en mesure de trouver de façon immanente les ressources conceptuelles suffisantes du point de vue critique à la justification.

À première vue, il est tentant de dire que la solution habermassienne est plus intéressante. Il semble, en effet, que le point de vue critique doive, d'une manière ou d'une autre, « transcender » le contexte de justification qui est chaque fois le nôtre. Ce que ce point de vue doit permettre, en fait, c'est d'avoir accès à un ensemble de possibilités qui ne sont pas encore le cas afin de pouvoir critiquer et réviser nos prétentions à la validité<sup>801</sup>. Habermas

<sup>798</sup> Cité par Rorty 1990b, p. 8.

<sup>799</sup> Rorty, 1994a, p. 37.

<sup>800</sup> Rorty, 1995b, p. 298: « We pragmatists think that those philosophers who view the defence of « our realistic intuitions » as an important or moral imperative are held captive by the picture of *getting* closer to a fixed goal. »

<sup>801</sup> Pourrait-on dire que Rorty remplit ce critère contrefactuel grâce à son insistance sur la comparaison et l'imagination ? Peut-être bien. Néanmoins, la perspective rortienne semble incapable de mobiliser

cherche donc à trouver une assise critique universelle puisque « l'aspect inconditionné que nous associons intuitivement aux prétentions à la vérité est ici interprété comme une manière de transcender les contextes locaux<sup>802</sup>». L'inévitabilité d'un concept fort de vérité se dévoile, selon Habermas, à travers deux prétentions transcendantales des acteurs du monde vécu: 1) à travers les certitudes pratiques nécessaires à l'action et, 2) à travers le caractère inconditionné de nos prétentions à la vérité. L'objectivité du monde vécu, quant à elle, se dévoile lorsque nos croyances justifiées ne suffisent plus à expliquer naïvement la praxis quotidienne. Cependant, nous verrons que ce lien interne, supposément nécessaire entre le concept de transcendance (selon lequel la force critique nécessite l'accès à des ressources qui doivent transcender notre contexte de justification) et le concept d'inconditionnalité (selon lequel la force critique nécessite des ressources qui doivent transcender tous les contextes de justification possibles) est loin d'être évident.

Enfin, nous verrons que les véritables oppositions épistémologiques entre Habermas et Rorty concernent l'interprétation que nous devons avoir de l'attitude des acteurs du monde vécu (au niveau de l'action d'un côté, au niveau de la discussion de l'autre). Les présuppositions sont conçues par Habermas comme de véritables conditions de possibilité de l'action et de la discussion et soulignent, du même coup, la nécessité d'un concept transculturel de vérité sans retomber dans un représentationnalisme ontologique ou transcendantal. Rorty préfère, quant à lui, opposer à cette lecture objectiviste une lecture pragmatiste déflationniste. Nous montrerons que cette tension est au cœur d'une approche sociale des normes de signification.

Le débat entre Rorty et certains contextualistes comme Habermas ou McCarthy expose le défi d'une approche sociale, celui de formuler une position qui respecte le contextualisme sans sombrer dans le relativisme. McCarthy et Habermas soutiennent, entre autres, que Rorty, en glissant d'un point de vue contextualiste à un point de vue relativiste, se contredit en affirmant que le participant ou l'acteur ancré dans la pratique est prisonnier d'une position ethnocentrique alors qu'il préconise en même temps un relativisme réformiste et éliminativiste, celui-ci requérant l'adoption de la position d'un observateur extérieur. Nous

---

les ressources conceptuelles suffisantes pour pouvoir expliquer *l'attrait* que nous éprouvons pour un choix plutôt qu'un autre. Chez Rorty, un tel choix semble totalement réduit à l'utilité sociale. Ce problème mériterait, dans un autre contexte, d'être traité plus à fond.

<sup>802</sup> Habermas, 2001, p. 217.

cherchons au contraire à construire une approche sociale qui rende compte de l'équilibre délicat entre un contextualisme et une position relativiste c'est-à-dire une approche sociale qui réussisse à concilier l'idée selon laquelle notre application correcte des normes de signification dépende de nos contextes d'accords sociaux (renvoyant à l'immanence du contexte de l'objectivité des prétentions à la vérité et des normes correctes de signification) avec l'idée selon laquelle la validité des standards ainsi obtenus ne doit néanmoins pas s'identifier avec un accord social (renvoyant à la transcendance des prétentions à la vérité et des normes de signification). Nous visons une approche qui soit contextualiste quant à la vérité et à la signification tout en étant également faillibiliste. Putnam l'exprime clairement ainsi:

« In my view, truth and rational acceptability [...] are relative to the sort of language we are using and the sort of context we are in [...] This does not mean that a claim is right whenever those who employ the language in question would accept it as right [...] There are two points that need to be balanced [...] (1) talk of what is « right » and « wrong » in any area only makes sense against the background of an inherited tradition; but (2) traditions themselves can be criticized [...] Reason is both immanent (not to be found outside of concrete language games and institutions) and transcendent (a regulative idea that we use to criticize the conduct of all activities and institutions) [...] Philosophers who lose sight of the immanence of reason, of the fact that reason is always relative to context and institution, become lost in characteristic philosophical fantasies[...] Philosophers who lose sight of the transcendence of reason become cultural (or historical) relativist<sup>803</sup>. ».

Cet équilibre entre immanence et transcendance dont fait allusion Putnam à propos de la vérité et de la connaissance rappelle le délicat problème soulevé par le défi sceptique de Kripke et de Wittgenstein (et de la discussion de McDowell<sup>804</sup>), si ce n'est que Kripke s'intéressait davantage à la question de l'application objective de nos concepts du point de vue de leur correction. Kripke conçoit également une conception de la normativité de la signification dans laquelle les normes de signification dépendent de nos pratiques *de facto* d'utilisation des termes (c'est-à-dire qu'elles sont immanentes à ces pratiques) sans pour

<sup>803</sup> Putnam, 1985, p.7-8; 1983, p.234-235. Comparons avec Hoy et McCarthy, 1994, p.39: « While we have no ideas of standards of truth wholly independent of particular languages and practices, it remains that 'truth' serves as an idea of reason with respect to which we can criticize the standards we inherit and learn to see things in a different way. Neither the particularity and context-immanence nor the universality and the context-transcendence of truth claims can be ignored without doing violence to our actual practices of truth. We can, and typically do, make contextually conditioned and fallible claims to objective truth. ».

<sup>804</sup> Cf. Vol. I, §3.3.

autant s'y réduire totalement (et dans ce sens, on pourrait dire qu'elles sont transcendantales à de telles pratiques). Nous analyserons et évaluerons la pertinence de différentes formulations d'une approche sociale et contextualiste d'une application correcte et objective des termes : celles de Davidson, de Brandom et de Habermas-McCarthy. Mais, avant d'y venir, nous précisons deux points préliminaires : la relation entre l'objectivité des prétentions à la vérité et l'application correcte et objective des normes de signification ainsi que la portée d'une normativité forte de la signification.

#### 5.4. Les prétentions à la vérité et les normes de signification

Nous traiterons ici de la question de l'application correcte des normes de signification et de la vérité des propositions. Nous nous appuyons sur deux thèses centrales de Quine et de Davidson soit : la *déréification*<sup>805</sup> de la signification et l'abandon de la distinction de l'analytique et du synthétique. Nous verrons par la suite que Brandom, en acceptant les principes posés par Quine et Davidson, considère ces deux points simultanément.

Il semble tout d'abord qu'il faille distinguer entre d'une part, comprendre la signification d'un terme ou appliquer un terme correctement et énoncer, d'autre part, une proposition ou un énoncé prétendant à la vérité. En effet, Davidson<sup>806</sup> souligne que nous pouvons parfaitement comprendre la signification d'un terme et pourtant l'employer de manière erronée. Par exemple, les Anciens pouvaient bien vouloir dire le terme Terre par « Terre » (ou leur mot pour signifier Terre) et pourtant véhiculer des croyances fausses à son sujet. Ainsi, comme nous l'avons mentionné<sup>807</sup>, si ce qui détermine l'objet de notre croyance ici ce sont les autres croyances que nous pouvons avoir, la possibilité d'identifier l'objet d'une croyance comme étant cette Terre est indéterminée ou fausse si nous ne pouvons attribuer de nombreuses autres croyances à propos de corps célestes volumineux, froids et solides tournant autour d'une étoile très grosse et très chaude. Il apparaît alors que l'application correcte de la signification d'un terme soit une chose et que la vérité d'une proposition ou

---

<sup>805</sup> Nous entendons ici le terme de *déréification* par opposition à la notion de *réification* des significations c.-à-d. la position d'entités linguistiques intrinsèquement non relationnelles qui resteraient identiques dans le passage d'une langue à l'autre, et dont l'identité garantirait la réalité. Voir Bjorn Rambert, 1989, p.2 et chap. 8, *passim*. Rorty, 1995a, p.94 souligne que « Bjorn Rambert a tout à fait raison de dire que le principal souci de Davidson est d'éviter toute réification du langage. ».

<sup>806</sup> Davidson, 1993b, 1975, p.247-248.

<sup>807</sup> Cf. Vol. II, Chap. 4, §4.7.

d'un énoncé en soit une autre. La distinction traditionnelle<sup>808</sup> des énoncés en deux classes, à savoir, entre les énoncés analytiques et les énoncés synthétiques, vient aussi appuyer cette remarque. Les énoncés *analytiques*, dont font partie les énoncés logiques et mathématiques, sont des énoncés dont la signification garantit la vérité, alors que les énoncés *synthétiques* sont des énoncés dont la signification ne suffit pas à établir la vérité : il faut en plus recourir à l'expérience pour déterminer s'ils sont vrais ou faux. Carnap avait dans le *Dépassement de la métaphysique par l'analyse logique du langage*, défendu la thèse de la dualité de l'analytique et du synthétique sous la forme suivante :

On peut ranger les énoncés (doués de sens) de la manière suivante : en premier lieu, ceux qui sont vrais en vertu de leur seule forme (ou « tautologie » d'après Wittgenstein. Ils correspondent à peu près aux jugements analytiques kantien). Ils ne disent rien sur le réel. À cette espèce appartiennent les formules de la logique et de la mathématique; elles ne sont pas elles-mêmes des énoncés sur le réel, mais servent à leurs transformations. En second, viennent les négations des premiers (ou « contradictions ») qui sont contradictoires, c'est-à-dire fausses en vertu de leur forme. Pour décider de la vérité ou fausseté de tous les autres énoncés, il faut s'en remettre aux énoncés protocolaires, lesquels (vrais ou faux) sont par là-même des objets d'expérience, et relèvent de la science empirique. Si l'on veut construire un énoncé qui n'appartient pas à l'une de ces espèces, cet énoncé sera automatiquement dénué de sens<sup>809</sup>.

Les énoncés analytiques sont, selon Carnap, des énoncés *a priori relativement* à un système linguistique pourvu d'un certain ensemble de règles sémantiques : puisque ces règles font en sorte que les énoncés analytiques n'excluent rien du point de vue empirique, ceux-ci ne risquent pas d'être réfutés ou révisables par l'expérience sensible<sup>810</sup>, et ne nécessitent pas de

<sup>808</sup> Tout au long du XIXe siècle, la notion kantienne d'analyticité a constitué une pièce maîtresse du discours philosophique établi. Les remaniements successifs dont elle a fait l'objet de la part de Bolzano, Frege, Wittgenstein, Russell et Carnap ont contribué à consolider sa place dans la pensée philosophique du début du XXe siècle. À vrai dire, la démarcation entre l'analytique et le synthétique est plus ancienne que la philosophie kantienne puisqu'elle apparaît déjà chez Leibniz avec l'opposition entre les vérités de raison et les vérités de fait (« relations of ideas », « matters of fact »). Or, dès les années trente du siècle passé, des voix se lèvent contre le concept d'analyticité. Notamment tel qu'il apparaît chez Carnap et le Cercle de Vienne qui s'y réfèrent pour caractériser le statut des vérités logiques. Aux yeux de Carnap, les lois logiques sont *a priori* et analytiques au sens où elles reposent exclusivement sur des définitions linguistiques. Conformément à une idée reçue, il appartient à Quine d'avoir mis au jour les difficultés inhérentes à cette théorie.

<sup>809</sup> Carnap, 1985, p. 173.

<sup>810</sup> Carnap, 1963, p.13 postule ainsi, à l'opposé de Tarski, une différence de nature entre les énoncés logiques et les énoncés empiriques : « there was one problem, however, on which I disagreed with tarski. In contrast to our view that there is a fundamental difference between logical and factual statements, since logical statements do not say anything about the world, Tarski maintained that the

confirmation empirique, tant qu'on s'en tient au même système linguistique. Carnap affirme avoir ainsi obtenu une explication satisfaisante de la nature des mathématiques et de la logique. Quine baptise ce point de vue la *doctrine linguistique de la vérité logique*<sup>811</sup>. Il s'agit d'une doctrine *épistémologique*: les vérités logiques sont vraies uniquement en vertu de la signification qu'on entend donner aux expressions logiques<sup>812</sup>. Par contre, pour ce qui est des propositions synthétiques, nous pouvons comprendre correctement la signification de chacun des termes la composant et énoncer pourtant une proposition erronée. Le fait d'employer les termes correctement ne garantit pas que nous formulions des propositions prétendant à la vérité de telle sorte que nous devons distinguer entre les propositions qui sont fausses parce que nous ne comprenons pas la signification des termes la composant et celles qui le sont parce que nous cultivons une croyance empirique fautive<sup>813</sup>. Autrement dit, il devrait être clair dans la définition d'une proposition synthétique et pour ceux qui soutiennent cette distinction de l'analytique et du synthétique que l'objectivité des propositions prétendant à la vérité présuppose dans l'usage une application correcte des termes. Par exemple, se prononcer sur la vérité de la proposition « Ma table mesure trois mètres de long » requiert que nous sachions

---

distinction was only a matter of degree. This divergence exists even to-day. ». À l'instar de Tarski, Quine se montre hostile à une dichotomie trop rigide de l'analytique et du synthétique. Carnap, 1963, p.35-36, le note ainsi: « I gave several talks on the nature of logic and on the possibility of defining logical truth as a semantical concept. I discovered that in these questions, even though my thinking on semantics had originally started from Tarski's ideas, a clear discrepancy existed between my position and that of Tarski and Quine, who rejected the sharp distinction I wished to make between logical and factual truth. ».

<sup>811</sup> Quine, 1976a, p. 108.

<sup>812</sup> Cf. Quine, 1976a, p. 110; 1986. Pour une défense de ce point de vue, voir Carnap, 1939, p.12-13; 1963, p.47.

<sup>813</sup> Par conséquent, la proposition synthétique « Ma table mesure trois mètres de long » peut s'avérer fautive dans deux cas : soit ma table ne mesure pas trois mètres de long bien que je le croie, soit elle mesure bien trois mètres de long mais ce n'est pas ce que cette phrase exprime. Dans le premier cas, j'ai parfaitement bien compris la signification de « table » et de « trois mètres de long » mais je formule un énoncé erroné – j'ai peut-être fais une erreur en mesurant parce que j'étais fatigué. Il s'agit ici d'une erreur épistémique et non d'une erreur sémantique : je comprends correctement la signification des termes mais j'énonce un jugement erroné en les employant. Mais, un autre cas de figure peut aussi se présenter : il se peut que je comprenne non comme il se doit la signification du terme « table », soit que je veuille dire « chaise » en utilisant le terme « table ». Dans ce cas, il est évident que je n'ai pas compris la signification de « table » et l'ai appliqué par erreur pour désigner « la baignoire ». Le premier type d'erreur est le résultat d'un défaut épistémique, le second type d'erreur est le résultat d'une compréhension déficiente de la signification du terme « table », et il s'agit d'un défaut sémantique. Ceux qui distinguent l'application des normes de signification du point de vue de leur correction de la vérité des propositions pensent que nous pouvons diviser toutes les propositions erronées en celles qui sont fausses en vertu de défauts épistémiques et en celles qui le sont en vertu de défauts sémantiques.

adéquatement ce que les termes de « table » et de « trois mètres de long » signifient et de ce fait, la proposition prétendant à la vérité présuppose la distinction normative entre une signification correcte et une signification incorrecte de ces expressions-ci. Il semble donc qu'il faille établir une distinction entre l'application d'un terme du point de vue de sa correction et la vérité d'une proposition c'est-à-dire que le sens attribué à l'application des normes de signification du point de leur correction devrait être séparé de celui attaché à la vérité des propositions.

Cependant, aboutir à cette distinction entre une application correcte des termes et une assertion revient à séparer la signification des termes de leur application dans la formulation d'énoncés empiriques. Ce point requiert de modeler une dimension de la signification d'un terme qui ne se réduise pas à son usage au sein de phrases empiriques, ce qui nous ramène à la distinction de l'analytique et du synthétique c'est-à-dire à une notion de signification qui s'énonce en termes de règles<sup>814</sup> ou de normes sémantiques et s'exprime par le biais de propositions analytiques vraies<sup>815</sup> en toutes circonstances et antérieures à l'emploi d'un terme

---

<sup>814</sup> Il faut insister sur le fait que la notion de signification proposée par les positivistes logiques possède un caractère *normatif*. La signification des énoncés repose sur les *règles* du langage - aussi appelées « règles sémantiques » ou « postulats sémantiques » - qui *gouvernent* leur usage et servent à *justifier* leur vérité. Les tenants de la distinction de l'analytique et du synthétique accordent donc à cette distinction un caractère *épistémologique*. L'idée est qu'il y a en général deux facteurs qui contribuent à établir la vérité d'un énoncé : les faits (ou l'expérience sensible) et la signification. Les tenants de la distinction de l'analytique et du synthétique affirment que les contributions respectives de ces deux facteurs peuvent être séparées pour chaque énoncé pris individuellement. Rappelons que Carnap définit les énoncés analytiques comme énoncés vrais en vertu de la seule signification des termes qui y figurent et les énoncés synthétiques comme vrais en raison de faits empiriques ou extralinguistiques. Ainsi le règne de l'analytique comprend deux domaines distincts : l'analytique au sens logique, constitué des énoncés vrais en vertu de la seule signification des constantes logiques et, d'autre part, l'analytique au sens large, comprenant les énoncés vrais en vertu de la signification des termes descriptifs. Les deux domaines sont constitués exclusivement de vérités linguistiques. Alors que les énoncés du premier type sont vrais en vertu des règles sémantiques - ils sont dits L-vrais, c'est-à-dire vrais dans toutes les descriptions d'état -, les seconds le sont en vertu de postulats de signification spécifiques dépendant des règles établies pour un langage donné.

<sup>815</sup> Si l'on définit, à la manière de Carnap, la classe des énoncés analytiques dans un langage donné par la classe des énoncés qui sont « vrais par les règles sémantiques », il reste à définir la notion de règle sémantique. Or, conclut Quine, 2003, p. 69 : « du point de vue du problème de l'analyticité, la notion d'un langage artificiel pourvu de règles sémantiques est *un feu follet par excellence*. Des règles sémantiques déterminant les énoncés analytiques d'un langage artificiel n'ont d'intérêt que dans la mesure où nous comprenons au préalable la notion d'analyticité. Elles ne nous aident nullement à acquérir une telle compréhension ». Quine, 1969, p. 322, revient sur la pertinence même du recours à une langue artificielle dans sa réponse à Strawson : « Equally, Carnap's distinction between analytic and synthetic is clear within some artificial diminutive language Lo. My complaint was that his clearly defined class of sentences called analytic-in-Lo [...] it threw no light on analyticity as applied to our

dans les jugements synthétiques soit à la formulation de propositions prétendant à une vérité empirique. Autrement dit, la possibilité de dissocier la signification de la vérité empirique et l'application correcte des normes de signification des prétentions à la vérité, repose sur la distinction de l'analytique et du synthétique. Par conséquent, si nous rejetons, conformément à l'idée reçue de Quine, le point de vue de Carnap selon lequel il existe un clivage entre vérités analytiques, fondées sur la signification indépendamment des faits, et vérités synthétiques, fondées sur les faits, nous éliminons la première distinction.

Notons néanmoins que le point de vue de Quine n'est pas qu'une distinction de l'analytique et du synthétique est impossible à tracer; il soutient plutôt que tout point de vue qui accorde un rôle épistémologique ou fondationnel à cette distinction est dogmatique. Quine critique moins Kant que « l'empirisme moderne » – ce sont les premiers mots de l'article « Des deux dogmes de l'empirisme ». Quine note d'ailleurs que l'analyticité comme « réminiscence kantienne, est souvent présupposée de façon non critique dans la littérature épistémologique moderne<sup>816</sup>». Il s'agit donc bien de dénoncer un simple « présupposé » : à savoir, qu'il existe une classe d'énoncés dont la vérité ne dépend d'aucune donnée empirique, et se décide par les significations de leurs termes – « indépendamment des choses de fait ». Or, le premier problème qui se pose à l'empirisme est celui d'un critère de l'analyticité qui permette de démarquer cette classe d'énoncés : sont analytiques les énoncés logiquement vrais (comme « tous les célibataires sont des célibataires ») et les énoncés réductibles à des vérités logiques par substitution synonymique (comme « Tous les célibataires sont non mariés »). La critique de l'analyticité repose ainsi chez Quine sur celle de la synonymie : une vérité analytique se caractérise en général par le fait qu'on peut remplacer les termes par des synonymes sans préjudice pour la vérité. Or, la notion de synonymie n'est pas si évidente, elle est surtout dépendante de la langue que l'on envisage; et les formules du lexicographe ne doivent pas être avoir « force de loi <sup>817</sup>»; la substituabilité *salva veritate* que la synonymie permet « n'a de sens que relative à une langue dont l'étendue est spécifiée<sup>818</sup>». De plus,

---

own language, nor yet to any full-size substitute language adequate to science. ». On comprendra que le problème de Quine est moins celui de l'analyticité que celui du statut des vérités logiques. Et c'est sur ce point que porte, en réalité, le débat avec Carnap; la place de la logique dans la science, et la possibilité de la « démarquer » à l'intérieur de la science.

<sup>816</sup> Quine, 1981, §14, n.7, p.67.

<sup>817</sup> Quine, 2003, p. 54.

<sup>818</sup> Ibid., p. 61.

comme le fait remarquer Quine, la notion de synonymie est impossible à définir sans circularité, c'est-à-dire sans référence à la notion de signification, ou d'analyticité elle-même. Ce que Quine met donc en doute, c'est la possibilité de définir la différence entre deux catégories d'énoncés. Il affirme qu'il n'y a pas d'énoncés analytiques, au sens où l'entendent les positivistes logiques, ou en un sens philosophiquement intéressant. Et, il n'y a pas de sens à vouloir définir, pour autant, une classe d'énoncés dits analytiques<sup>819</sup>. L'argument n'est pas technique mais philosophique. Par philosophique, Quine veut dire qu'il met en cause la nature de la connaissance, et la manière dont elle s'acquiert. Les questions philosophiques surgissent de la science elle-même. Il en est de même de la notion d'analyticité. Elle ne saurait être au fondement de la connaissance, puisqu'elle est le produit d'interrogations nées à l'intérieur de la science, ou de son évolution :

« The lore of our fathers is a fabric of sentences [...] It is a pale grey lore, black with fact and white with convention. But I have found no substantial reasons for concluding that there are any quite black threads in it, or any white ones.<sup>820</sup>».

De même donc que l'analytique était, pour Quine, toujours synthétique, le synthétique est forcément analytique. L'énoncé « Aucun célibataire n'est marié » dépend tout autant, pour sa vérité, de données du réel que de la signification des mots, et il en est de même pour la vérité logique. « Tous les hommes sont mortels,  $x$  est un homme, donc  $x$  est mortel » qui dépend

---

<sup>819</sup> On pourrait considérer que la logique et les mathématiques sont différentes des sciences naturelles en ceci qu'elles ne portent pas sur un sujet particulier. Une façon de préciser ce point serait de considérer les lois logiques et mathématiques comme *a priori*. (Carnap, 1932, 1939 ; Ayer, 1946). Mais, pour que cette proposition ne soit pas vide de sens, il faut fournir une méthode générale permettant de distinguer les énoncés *a priori* des énoncés *a posteriori* ou empiriques. Les énoncés empiriques peuvent être caractérisés par le fait qu'ils dépendent essentiellement de l'expérience, de sorte qu'il n'est pas possible de les justifier sans faire référence à des expériences spécifiques. Les lois et les hypothèses des sciences naturelles se distingueraient des lois logiques et mathématiques par le fait que seules les premières ont un rapport essentiel, qu'il soit direct ou non, avec les données empiriques. Il faut cependant noter qu'en ce qui concerne les énoncés hautement théoriques des sciences, le rapport avec l'expérience peut être très ténu. En outre, comme des portions importantes de la logique et des mathématiques sont utilisées pour établir des liens entre les énoncés d'une théorie et l'expérience, rien, semble-t-il, ne pourrait nous empêcher de dire que les lois logiques et mathématiques sont elles aussi, en vertu de ce rôle qu'elles jouent, confrontées à l'expérience. Par conséquent, la propriété d'être utilisée pour établir un lien avec l'expérience ne permet pas d'isoler une catégorie particulière d'énoncés : on ne peut au mieux qu'établir une distinction de degré entre les différents énoncés selon leur proximité à l'expérience. Selon Quine, les vérités analytiques n'ont pas de qualité épistémologique particulière ; il n'y a pas de vérités analytiques parce qu'il n'y a pas de vérités *a priori* et il n'y a pas de vérités *a priori* parce que leur absence explique mieux la manière dont le savoir scientifique évolue et se développe.

<sup>820</sup> Quine, 1976a, p.125.

aussi bien de la grammaire que de la nature humaine. À l'inverse, un énoncé comme « Le courant passe<sup>821</sup> » ou « Tache verte maintenant » est théorique. Pour Quine, parler d'énoncé vrai équivaut à dire que cet énoncé parle du monde; il n'y a pas d'énoncé qui ne doive sa vérité qu'au langage, ou à la grammaire. Il n'y a donc pas de sens à vouloir définir les vérités logiques par les propriétés du langage, ni à dire que les vérités logiques sont vraies en vertu du langage. D'un point de vue logique, cette affirmation est plutôt vide et triviale; car une vérité logique est impliquée, par définition, par tous les énoncés. La logique, dit Quine, est vraie « Trivialement, en vertu de toutes les circonstances que vous voudrez – le langage, le monde, n'importe quoi<sup>822</sup> ». De même pour les énoncés factuels, il souligne :

Il est évident que la vérité de l'énoncé le plus platement factuel dépendra en partie du langage<sup>823</sup>.

Quine cherche donc, à partir d'une « épistémologie naturalisée<sup>824</sup> » à combattre « la doctrine linguistique de la vérité logique » selon laquelle les vérités logiques ne dépendent que de transformations d'expressions effectuées au sein du langage. Même des vérités comme «  $p$  ou non  $p$  », considérées par Wittgenstein comme des tautologies qui ne décrivent pas le monde sont dépendantes à l'égard de celui-ci. La vérité de « Brutus a tué César ou Brutus n'a pas tué César » dépend-elle oui ou non de l'action de tuer et des personnes ici envisagées ? Le raisonnement consiste à montrer que de telles vérités logiques dépendent d'une attribution des valeurs de vérité à  $p$  même si elles ne dépendent pas des valeurs elles-mêmes. La confusion dénoncée par Quine est celle qui consiste à assimiler l'attribution des valeurs aux valeurs elles-mêmes : il faut que  $p$  ait une valeur de vérité pour que «  $p$  ou non  $p$  » en ait une; l'indifférence de  $p$  à l'égard du vrai et du faux dans la vérité logique «  $p$  ou non  $p$  » ne signifie pas que «  $p$  ou non  $p$  » ne dépend pas de la vérité de  $P$ . Gochet a bien montré ce point :

De ce que «  $p$  ou non  $p$  » est vrai quelle que soit la valeur de «  $p$  », on ne peut pas conclure que «  $p$  ou non  $p$  » soit vrai indépendamment de la valeur de  $p$ , pas plus qu'on ne peut dire que si un moteur consomme indifféremment n'importe quel combustible, il fonctionne indépendamment de tout combustible<sup>825</sup>.

---

<sup>821</sup> Duhem, 1914, p. 227.

<sup>822</sup> Quine, 1975, p.142.

<sup>823</sup> Ibid., p.149.

<sup>824</sup> C'est le titre de l'un de ses ouvrages.

<sup>825</sup> Gochet, 1978, p. 152.

Si donc le concept de vérité logique présuppose celui de vérité et que la vérité suppose la prise en compte d'une référence à un domaine d'individus, alors les vérités logiques ont une portée ontologique. En général, nous considérons les vérités logiques comme des phrases éternelles, c'est-à-dire déliées des circonstances de l'élocution, mais cela ne vient que du besoin de construire une théorie et de ne plus s'en tenir aux phrases observationnelles. En vérité les phrases éternelles, « dépositaires de la vérité elle-même », ne sont que des projections simplifiées de phrases observationnelles. C'est en ce sens qu'il faut comprendre la simplicité des axiomes logiques ; cette simplicité est une simplification ; elle ne correspond à rien de primitif en soi ; elle ne l'est que pour les besoins de la construction théorique. Plus une discipline scientifique gagne en rigueur, moins elle se rend dépendante du monde : « what was one regarded as a theory about the world becomes reconstrued as a convention of language<sup>826</sup> » ; c'est ainsi que nos définitions prennent l'allure de principes analytiques.

Par conséquent, même si Quine refuse de concevoir la possibilité d'exclure certains énoncés du doute et de la révision et ainsi de déclarer un énoncé vrai « par convention » ou par sa seule signification car ce serait le soustraire au « tribunal de l'expérience », il admet qu'il faut bien des points fixes pour construire une théorie<sup>827</sup>. Aussi, comme semble le dire Quine à la fin de « Truth by convention », le débat contemporain sur l'analyticité ou la convention recouvre peut-être une controverse plus classique, ou plus fondamentalement philosophique, sur l'*a priori*. On a d'ailleurs souvent rapproché Quine et Wittgenstein sur cette question de l'*a priori*. Pour Quine, comme pour Wittgenstein, la notion d'« énoncé de fait » est dépourvue de contenu. Un énoncé comme « le soleil se lèvera demain » possède, pour la plupart des individus, une certitude aussi grande que «  $2 + 2 = 4$  », et occupe une place aussi centrale dans le système de nos connaissances et de nos croyances. La question de l'analyticité devient, dans cette perspective, la question de la place des énoncés normatifs, et de la possibilité même de les mettre en question. Wittgenstein semble dire que des énoncés de type « empirique » peuvent jouer le rôle de « pivot » dans notre connaissance, et que nous sommes en réalité incapables de mettre en doute certains d'entre eux, même si nous ne sommes pas capables pour autant de les démontrer – comme « la Terre a existé avant ma naissance », ou « c'est là ma main » – car cela ôterait tout sens à la notion même de doute ou

---

<sup>826</sup> Quine, 1976b, p. 70.

<sup>827</sup> Quine, 1976a, p.100-125.

d'incertitude<sup>828</sup>. Il est inhérent à la notion de doute que certains énoncés, dans certaines circonstances, soient soustraits au doute, même si, par ailleurs, il n'y a pas de propositions qui soient absolument indubitables. Donc, chez Wittgenstein, ce qui rend le doute possible, c'est aussi le fait que des propositions en soient exclues :

La certitude avec laquelle nous croyons n'importe quelle proposition mathématique est la même que celle que nous avons quand nous savons comment prononcer les lettres A et B, comment s'appelle la couleur de notre sang, et quand nous savons que les autres ont du sang qu'ils appellent 'sang'<sup>829</sup>.

Toutes choses qui pourraient en théorie être mises en doute mais ne le sont pas effectivement, car elles reposent sur un accord social et sur la communauté de l'apprentissage, à savoir :

« Nous en sommes tout à fait sûrs » ne signifie pas seulement que chacun, isolément, en est certain, mais aussi que nous appartenons à une communauté dont la science et l'éducation assurent le lien<sup>830</sup>.

L'apprentissage du langage, comme le processus de recherche scientifique, s'appuient sur des énoncés collectivement reconnus, énoncés qui jouent le rôle de « gonds », ou de « points d'appui », dirait Quine :

C'est-à-dire : les questions que nous posons et nos doutes reposent sur le fait que certaines propositions sont soustraites au doute, comme des gonds, sur lesquels tournent ces derniers [...] Mais ce n'est pas que nous ne puissions entreprendre des recherches sur tout, et que nous soyons contraints de nous contenter de présuppositions. Si je veux que la porte tourne, il faut que les gonds tiennent<sup>831</sup>.

Nous avons d'ailleurs vu<sup>832</sup> que, selon Kripke, Quine et Wittgenstein se rejoignent pour l'importance à donner à la notion d'accord, et sur l'impossibilité de fonder la règle sur l'intériorité c.-à-d., sur un quelconque « langage privé ». Mais, Kripke préfère expliquer la position de Quine par son béhaviorisme radical : tout n'étant, pour Quine, qu'une affaire de dispositions comportementales, on ne peut plus parler de « règle » au sens où l'entend Wittgenstein, et le « paradoxe » disparaît :

<sup>828</sup> Wittgenstein, 1976, §261.

<sup>829</sup> Ibid., §340.

<sup>830</sup> Ibid., §298.

<sup>831</sup> Ibid., §341-343.

<sup>832</sup> Cf. Vol. I, Chap. I.

Puisque Quine formule les questions en termes dispositionnels, le problème ne peut être énoncé à l'intérieur de son cadre conceptuel<sup>833</sup>.

La question de savoir si on suit une règle n'aurait chez Quine, de réponse que comportementale, et il n'y a aucune trace de « normativité » chez lui.

Or, on peut dire ici que cette interprétation de Kripke manque, à notre avis, l'essentiel de l'argument de Quine. L'argument de Quine contre la notion de règle n'est pas béhavioriste au sens strict, il est épistémologique, ou philosophique. Pour Quine, il n'y a pas d'énoncés privilégiés ; les vérités logiques elles-mêmes n'ont pas de qualité épistémologique particulière. La logique n'a jamais pour Quine valeur de règle : il est remarquable que le mot de « règle » ne soit jamais employé dans *Philosophie de la logique*. Quine affirme d'ailleurs à plusieurs reprises que la maxime « Sauver la vérité logique », qui joue le rôle de règle (ou de convention) de traduction, ne donne pas à la logique de statut privilégié, ou de « fondement épistémologique distinctif <sup>834</sup>».

Sur ce point, nous aurions tendance à rapprocher encore Quine de Wittgenstein, puisque ce dernier affirme que des énoncés prétendument « factuels » peuvent servir de fondements à nos interrogations. Chez Quine, ce sont les énoncés « obvies » qui jouent le rôle assigné par Wittgenstein aux propositions « nécessaires ». Le « canon » « Sauver la vérité logique » n'est qu'une forme représentative de la règle générale « sauver l'obvie ». C'est sur ce point qu'intervient l'objection : n'est-il pas risqué et presque contradictoire de la part de Quine d'effacer ainsi toute distinction entre énoncés normatifs et énoncés empiriques ? Ne doit-on pas préserver, dans toute recherche, un « corps » d'énoncés autonomes et « nécessaires » ? Tel est l'argument présenté par Bouveresse<sup>835</sup>, qui note à juste titre que Quine, rejette, de toute façon, la notion de nécessité logique ; celle-ci étant tout aussi infondée que celle d'analyticité, sauf sous une forme socialisée ou « anthropologique » :

« In principle, therefore, I see no higher or more austere necessity than natural necessity; and in natural necessity, or our attributions of it, I see only Hume's regularities, culminating here and there in what passes for an explanatory trait or the promise of it.<sup>836</sup>».

---

<sup>833</sup> Kripke, 1996, §57, p.71.

<sup>834</sup> Quine, 1969, p.318

<sup>835</sup> Bouveresse, 1987, Chap.7.

<sup>836</sup> Quine, 1976c, p. 56.

Par ailleurs, Bouveresse affirme que pour Quine :

il n'y a pas, d'un côté, des vérités proprement dites et, de l'autre, des règles ou des conventions qui, comme dirait Wittgenstein, ne sont en toute rigueur ni vraies ni fausses, mais une seule et unique catégorie de propositions<sup>837</sup>.

Pour Quine, il n'y a pas d'énoncés privilégiés ; les vérités logiques elles-mêmes n'ont pas de qualité épistémologique particulière. Cela vaut pour les énoncés de la logique et des mathématiques, qui occupent simplement une place plus élevée que d'autres dans la hiérarchie des énoncés : ils sont, dit Quine, « plus analytiques » que d'autres<sup>838</sup>, mais ne sont pas pour autant préservés de toute réfutation, et affrontent eux aussi le tribunal de l'expérience. La conséquence épistémologique est alors qu'il n'y a pas de clivage entre les énoncés analytiques et les autres. Quine l'énonce ainsi: « we have here no such radical cleavage between analytic and synthetic sentences as was called by Carnap and other epistemologists<sup>839</sup> ». N'importe quel énoncé peut être, en théorie, tenu pour analytique. C'est la traduction naturaliste d'une déclaration de Quine dans « Deux Dogmes de l'empirisme » : « Tout énoncé peut être tenu pour vrai quoi qu'il arrive<sup>840</sup> ». Mais, s'il n'y a aucun sens à démarquer une classe d'énoncés « maintenus quoi qu'il arrive », les énoncés ne se différencieront « que par leur degré de résistance plus ou moins grand qu'ils sont en mesure d'opposer à une expérience récalcitrante<sup>841</sup> ». Or, comme le remarque Wright, « un verdict de récalcitrance n'est intelligible que sur un arrière-plan de jugements et de principes qui ne sont pas remis en question<sup>842</sup> ».

La position de Quine ne nous paraît cependant pas aussi incohérente que semblent le suggérer Wright et Bouveresse. Quine sait aussi bien qu'un autre que les énoncés qui « codifient » la réfutation d'une hypothèse doivent être tenus pour certains ou fixes au moment de la réévaluation de cette hypothèse ; sinon, cette réévaluation serait impossible. Pour Quine, ces énoncés ne sauraient pour autant être tenus pour révisables. Ils le sont *provisoirement*. Lors d'une révision du schème conceptuel de la science, nous décidons de conserver à certains

---

<sup>837</sup> Bouveresse, 1987, p.99

<sup>838</sup> Quine, 1973, p.80

<sup>839</sup> Ibid., p. 80.

<sup>840</sup> Quine, 2003, p.50.

<sup>841</sup> Wittgenstein, 1983, p 330; cité par Bouveresse, 1987, p.99.

<sup>842</sup> Cité par Bouveresse, 1987, p. 99.

énoncés le statut de point fixe. Mais, aucun énoncé n'est privilégié au point d'avoir ce statut de manière définitive. Quine note alors:

« [...] conventionality is a passing trait, significant at the moving front of science but useless in classifying the sentence behind the lines. <sup>843</sup> ».

Il reconnaît donc bien que nous pouvons attribuer provisoirement à certains énoncés le statut de vérités indubitables - ou conventionnelles : le choix du terme importe peu, puisque ce statut est, de toute manière révisable. Le fait que nous ne mettions pas en doute certains énoncés, dans le cadre de notre recherche scientifique, n'empêche pas qu'ils soient, en principe révisables.

Dans le même ordre d'idées, la remise en question du principe de contradiction repose sur l'idée qu'il n'y a pas en logique d'« immunité intrinsèque et irrévocable<sup>844</sup> » mais seulement des « immunités provisoires » certes « à tout moment il faut qu'il existe une armature logique comme condition préalable à la manifestation d'incompatibilité », mais « il ne découle pas de là qu'il existe une seule et même armature logique à tous les moments<sup>845</sup> ».

Rien n'interdit donc d'adopter des logiques où le principe de contradiction est mis en défaut. Si quelqu'un répond « oui et non » à une question, plutôt que de penser qu'il est stupide, mieux vaut croire qu'il ne donne pas le même sens que nous à l'affirmation et à la négation. Le principe de charité vient prendre la place du principe de contradiction. Que dit ce principe ? Il affirme une présomption de rationalité humaine ; le nier en effet reviendrait à saper les bases mêmes d'une attribution de croyances à quelqu'un et par suite à ne pas pouvoir interpréter un comportement quelconque. Selon ce principe, la hâte à qualifier de prélogique une attitude n'est que le symptôme d'une difficulté à traduire. Ainsi, la logique n'est jamais mise en doute chez Quine, même si elle est révisable en principe, parce qu'elle est, pour l'instant, au centre de toutes nos connaissances, et que la mettre en question consisterait à mettre en doute la connaissance et la possibilité de l'apprentissage dans son ensemble.

En outre, en rapportant le statut des énoncés à leur apprentissage, Quine efface toute démarcation entre les énoncés. La définition par l'assentiment n'était pas tout à fait

---

<sup>843</sup> Quine, 1976a, p.112.

<sup>844</sup> Quine, 1976c, p. 159.

<sup>845</sup> Ibid., p. 159.

satisfaisante; la définition par l'apprentissage de l'obvie sera sans doute la dernière proposée par Quine :

« I have protested more than one that no empirical meaning has been given to the notion of meaning, nor, consequently, to this linguistic theory of logic. But now in terms of the learning process can we perhaps find some sense for the doctrine?<sup>846</sup> ».

Cette définition est proposée à nouveau dans *La poursuite de la vérité* :

Un énoncé est analytique si le locuteur indigène apprend à y assentir dans l'apprentissage d'un des mots qui le constituent<sup>847</sup>.

Même si l'apprentissage produit des usages du langage homogènes, nous avons chacun appris ce langage de façon différente, nous fondant chacun sur des énoncés « obvies<sup>848</sup> » différents :

« If the samples first acquired qualify as analytic, still they gain thereby no distinctive status with respect to the language or the community; for each of us will have derived his universal categorial powers from different first samples<sup>849</sup>. ».

L'analyticité dépendrait, en un sens, de l'histoire personnelle de chacun, et varie avec la diversité des apprentissages. Tout homme s'appuie, dans son apprentissage de la langue comme dans son apprentissage scientifique, sur des énoncés qu'il tient pour « obvies » ou des énoncés « qui vont sans dire », mais qui ne sont pas pour autant immuablement fixés. Pour Quine, l'analytique est affaire de degré, c'est-à-dire :

« In learning our language each of us learns to count certain sentences, outright, as true; there are sentences whose truth is learned in that way by many of us, and there are sentences whose truth is learned in that way by few or none of us. The former sentences are more *nearly* analytic than the latter. The *analytic* sentences are the ones whose truth is learned in that way by all of us; and these extreme cases do not differ notably from their neighbors, nor can we always say which ones they are<sup>850</sup>. ».

Comme dirait Hume, « voici la somme de toutes nos conclusions expérimentales » : il y a des énoncés auxquels nous tenons plus qu'à d'autres. En apprenant le langage, Quine souligne que « chacun de nous apprend à compter certains énoncés, tels quels, comme vrais ». Cette

<sup>846</sup> Quine, 1973, p.78.

<sup>847</sup> Quine, 1993, p.87.

<sup>848</sup> Quine, 1975, p. 122, parle du caractère obvie, l'« obviousness » des vérités logiques, à savoir le fait qu'elles rencontrent, de manière immédiate ou potentielle, l'assentiment général : « Quand je dis '1 + 1 = 2' obvie pour une communauté, je veux seulement dire que chacun y assentirait sans hésiter, et quand je dis 'il pleut' obvie dans certaines circonstances, je veux dire que chacun y assentirait dans ces circonstances. ».

<sup>849</sup> Quine, 1973, p. 79.

<sup>850</sup> Ibid., p.80.

uniformité induite par l'apprentissage est le seul fondement de ce que Quine définit, dans *Roots of Reference*, comme l'analyticité :

In making analyticity hinge thus on a community-wide uniformity in the learning of certain words, we reopen the question of analyticity of logical truths.<sup>851</sup>

L'analyticité est reconstruite en termes de « communauté ». Le seul sens à donner à la notion d'énoncé analytique est social. Quine mentionne qu'il tient à s'écarter du point de vue de Carnap :

« Carnap maintained, and Frege before him that the laws of logic held by virtue purely of language: by virtue of the meanings of the logical words.<sup>852</sup> ».

Quine cherche alors à redéfinir l'analyticité en redéfinissant le statut des lois logiques, tout en sachant pertinemment qu'il ne s'agit plus alors de l'analyticité au sens de Carnap :

« Les énoncés analytiques sont ceux que nous sommes prêts à affirmer quoi qu'il arrive<sup>853</sup> », et ils englobent les énoncés analytiques tels qu'ils sont définis dans *Roots of Reference*. L'analyticité, étant fondée sur l'assentiment, s'avère essentiellement sociale : elle est aussi « naturalisée » :

« Language is social, and analyticity, being truth that is grounded in language, should be social as well. Here then we may at last have a line on a concept of analyticity: a sentence is analytic if *everybody* learns that it is true by learning its words. Analyticity, like observability, hinges on social uniformity<sup>854</sup>. ».

Quine et Davidson<sup>855</sup>, en défendant un point de vue holiste et en attaquant l'idée d'une réification de la signification, rejettent cette distinction de l'analytique et du synthétique et mettent l'accent sur la relation entre l'emploi d'un terme au sein de propositions empiriques et sa signification. Déréifier la signification n'implique pas de séparer la « métaphysique » de la signification de « l'épistémologie » de la signification : nous saisissons la signification des termes que nous employons par le biais de nos pratiques de discussion et d'interprétation et ce sont ces mêmes pratiques (soit celles de la traduction ou de l'interprétation radicale) qui relient la signification à la vérité de nos jugements. Nous avons vu que pour Davidson, les

<sup>851</sup> Ibid., p. 80.

<sup>852</sup> Ibid., p.78.

<sup>853</sup> Quine, 1977a, p.66-67.

<sup>854</sup> Quine, 1973, p.79.

<sup>855</sup> Davidson, 2001b, 1983, p.145-146; 2001, 1990b, p. 194-195, accepte explicitement les critiques de Quine contre la distinction de l'analytique et du synthétique. Nous soulignons ce point au chap. IV, §4.4.

données empiriques de base d'une théorie de l'interprétation sont constituées des phrases tenues pour vraies par les locuteurs en différentes circonstances. L'interprète doit donc trouver un moyen de déterminer les conditions dans lesquelles les phrases sont vraies, à partir de données concernant les circonstances dans lesquelles les phrases sont tenues pour vraies. Il ne peut évidemment pas supposer que les locuteurs soient infaillibles. La proposition de Davidson est de supposer que les phrases tenues pour vraies par les locuteurs sont, autant que possible, vraies. Par conséquent, pour comprendre autrui et déchiffrer la signification de ces émissions linguistiques, nous devons faire appel au principe de charité, soit supposer que ces locuteurs étrangers ont raison dans leurs propos relativement aux normes, croyances et valeurs de leurs propres cultures. Nous pouvons ainsi maintenir que le principe de charité est profondément ancré dans nos pratiques, mais qu'il ne doit pas être conçu comme un principe *a priori* au sens strict. Davidson semble davantage considérer ce principe comme une condition d'adéquation de notre interprétation d'autrui et de nos attributions de croyances. Plus exactement, ce principe indique comment les choses iraient si notre situation épistémique n'était pas aussi limitée. Autrement dit, si nous avions accès à toutes les données relatives aux attitudes d'un agent concernant les phrases de sa langue (c.-à-d. les phrases qu'il tient pour vraies, celles qu'il tient pour fausses, et celles à propos desquelles il suspend son jugement), alors notre interprétation de celles-ci et nos attributions de croyances seraient conformes au principe de charité. Le fait que nous ayons implicitement recours à ce principe en pratique peut être démontré par le fait que nous sommes prêts à ajuster notre interprétation d'un agent conformément à ce principe, lorsque l'on nous fournit des données additionnelles concernant ses attitudes<sup>856</sup>. Nous serons amenés à lui attribuer un ensemble de croyances en grande partie cohérentes et vraies<sup>857</sup> au regard de nos propres normes et nous comprenons, sur cette base, la signification de leurs énoncés<sup>858</sup>.

<sup>856</sup> Supposons que le sujet *X* a une croyance qu'il exprime par « Mon artérite est douloureuse ». Nous serions portés spontanément à interpréter *X* de manière homophonique, et à dire qu'il croit que son artérite est douloureuse. Mais, si nous apprenions que *X* tient aussi pour vraie « L'artérite est une maladie qui affecte les articulations », « l'artérite fait gonfler mes poignets », etc., alors, nous réviserions notre attribution : nous dirons plutôt que *X* croit que son *arthrite* est douloureuse, et que le mot « artérite », dans l'idiolecte de *X*, signifie *arthrite*. Cette nouvelle interprétation fait ainsi en sorte que *X* a davantage de croyances vraies que la première. Nous utilisons ici l'exemple de Burge, 1979 car il illustre bien le phénomène que nous voulons décrire.

<sup>857</sup> Cf. Davidson, 1993a, p. 221-222; 1993b, p. 137, p. 197, p.200-201; 2001b, 1983; 1990b, p. 24-25.

<sup>858</sup> En d'autres termes, bien que la dimension rationnelle du principe de charité affaiblisse la dimension constitutive de la présomption de vérité en déterminant les croyances et les significations, elle implique

Davidson insiste ainsi à plusieurs reprises sur le fait que l'attitude de tenir une phrase pour vraie résulte de deux facteurs : ce que l'agent croit et ce que la phrase signifie. Ce qui est observable, le fait que l'agent tient une certaine phrase pour vraie, est le produit de deux facteurs qui, eux, ne le sont pas, la croyance et la signification. Davidson soutient, comme nous l'avons précédemment souligné, qu'il n'est pas possible de déterminer ce qu'un agent croit sans connaître la langue qu'il parle. Il ne semble en effet pas possible d'individuer précisément les croyances d'un agent sans être en mesure d'interpréter les phrases qu'il utiliserait pour exprimer celles-ci. Une théorie de l'interprétation radicale ne peut donc pas s'appuyer sur des données préalables concernant les croyances des agents. Nos pratiques d'interprétation et de traduction suggèrent ainsi qu'il existe bien une interdépendance entre la vérité de nos croyances ou jugements et la signification de nos jugements (et termes). En effet, soit une certaine phrase tenue pour vraie : si on connaît la signification de cette phrase, on peut savoir quelle est la croyance exprimée par cette phrase, on peut inférer quelle est sa signification. Il n'existe donc pas de séparation entre ce qu'un terme veut dire (qui pourrait d'ailleurs être exprimé par le biais d'un jugement analytique) et la question de savoir si une proposition empirique (synthétique) est vraie, la signification étant élaborée à partir des propositions empiriques qui sont tenues pour vraies.

Par conséquent, pour Quine et Davidson, la question de la signification des termes composant nos jugements n'est pas indépendante et antérieure à toute possibilité de justifier cette vérité par des faits d'expérience ou des propositions empiriques ; il s'agit au contraire d'une activité conjointe. Davidson souligne d'ailleurs qu'une croyance est identifiée par sa position dans un système de croyances, qui détermine ce sur quoi elle porte. Il propose l'exemple suivant : on ne peut supposer que quelqu'un croit que le revolver est chargé, à moins de supposer qu'il croie aussi que le revolver est un objet physique, qu'il est une arme à feu, etc. Il n'est pas nécessaire, poursuit-il que l'agent adhère à une liste précise de croyances pour avoir la croyance en question : néanmoins, il est nécessaire qu'il ait un système de

---

avant tout de relier la vérité à la signification dans la mesure où il semble rationnel qu'autrui, qui fait l'objet d'une interprétation, manifeste en grande partie des croyances vraies. Par conséquent, même si la présomption de rationalité rend possible la formulation de croyances erronées, ces erreurs sont une exception. Une forte présomption de vérité est donc requise afin de lier étroitement signification et vérité. Nous pensons donc que défendre une version faible du principe de charité autorise Davidson à parler simultanément de la vérité des propositions et de l'application correcte des normes de signification.

croyances dans lequel nous trouvons principalement des vérités, pour que les croyances particulières de ce système puissent être identifiées. L'idée est donc que si nous attribuons au regard de nos propres normes un trop grand nombre de croyances erronées à un agent, cela tend à rendre impossible l'identification de ce sujet. Les croyances doivent alors être attribuées en groupes, puisque que c'est la seule manière dont nous pouvons assurer l'identification du sujet sur lequel elles portent. Nous ne pouvons donc dire que telle croyance d'un agent porte sur tel sujet, à moins qu'il n'ait un ensemble d'autres croyances vraies portant sur ce sujet. Cette remarque vise ainsi à combler l'écart entre le fait de dire quelque chose de vrai et celui d'utiliser un terme correctement (conformément à sa signification)<sup>859</sup>. Séparer la vérité et la signification implique donc le maintien de la distinction de l'analytique et du synthétique et nous engage à soutenir une approche réaliste sur la signification c'est-à-dire l'idée qu'il existerait une réalité extérieure<sup>860</sup> à nos significations et indépendante de ce que nous pourrions découvrir par le biais de nos processus d'interprétation et des énoncés que nous pourrions tenir pour vrai.

Il serait donc intéressant de mettre en relation ces observations avec le rejet de la distinction de l'analytique et du synthétique. Si on disposait d'un critère permettant de séparer la signification de la croyance (factuelle), on pourrait alors admettre la possibilité qu'un agent ait peu ou pas de croyances, ou des croyances largement erronées portant sur un certain sujet, puisque l'identité de ce sujet serait déterminée par la signification des mots utilisés par l'agent pour exprimer ses croyances. On dirait que l'agent possède un certain concept s'il connaît la signification du mot exprimant ce concept, et cela, peu importe ses croyances (factuelles) faisant intervenir ce concept. Mais, si on accepte les objections de

---

<sup>859</sup> Davidson, 2001b, 1983, déduit ce résultat de sa théorie de l'interprétation afin d'en tirer une conséquence épistémologique forte. *Via* le lien entre vérité et signification forgé dans son compte-rendu de l'interprétation, Davidson soutient que les croyances ne peuvent pas pour la plupart être fausses et en même temps significatives. Si elles s'avèrent significatives, alors elles doivent en grande partie être vraies. Le principe de charité recommande au linguiste d'interpréter l'agent de façon à ce que tous deux soient largement d'accord. Par conséquent des scénarii cartésiens sceptiques tels que celui du malin génie, qui visent à séparer la signification de la vérité en supposant que les croyances conservent la même signification alors que des changements considérables viennent affecter la valeur de vérité, sont incohérents.

<sup>860</sup> Davidson souligne d'ailleurs : « Nous avons montré qu'il était absurde de chercher un fondement justifiant la totalité des croyances, quelque chose d'extérieur à cette totalité qui puisse être utilisé pour tester ou comparer nos croyances. La réponse à notre problème doit donc être de trouver une *raison* pour supposer que la plupart de nos croyances sont vraies, c'est à dire une sorte de *preuve* » (Ibid., p.139).

Quine contre la notion d'analycité et la possibilité de distinguer les contributions respectives de la signification et de la croyance, on doit reconnaître qu'il n'existe pas de séparation nette entre les croyances « analytiques » et les croyances « factuelles », et de critère strict permettant de déterminer si un agent possède ou non un concept. La solution de Davidson est de dire que le fait d'avoir des croyances principalement vraies sur un sujet constitue le critère pour la possession d'un concept portant sur ce sujet. Ainsi, selon cette proposition, il est contradictoire de dire qu'un agent puisse avoir peu de croyances ou des croyances largement erronées sur un certain sujet. En d'autres termes, la difficulté voire l'impossibilité à attribuer des croyances principalement vraies faisant intervenir un certain concept est davantage une indication que l'agent ne possède pas le concept, qu'une preuve qu'il a une conception presque nulle ou largement erronée du type d'objet désignée par ce concept.

Ainsi, en suivant Quine et Davidson, nous pensons qu'une dérèfification de la signification et qu'un rejet de la distinction de l'analytique et du synthétique justifient ce lien entre normes de signification et prétentions à la vérité. Mais, plutôt que de confondre ces deux questions, Quine et Davidson soutiendraient qu'il est plus pertinent de souligner ce lien entre la vérité des jugements et l'application correcte et objective des normes de signification<sup>861</sup>.

##### 5.5. La portée d'une normativité forte : vocabulaire logique et vocabulaire non logique

Ainsi, en tant que locuteurs, nous acceptons de nous soumettre à des normes qui soient objectivement correctes. Pourtant, la question qui demeure toujours en suspens et que nous soulevions dans notre troisième chapitre est celle de savoir si ce statut normatif fort rattaché à nos normes de signification ne peut être attribué qu'à un vocabulaire strictement logique ou s'il peut tout aussi bien être attribué à un vocabulaire non logique recourant à l'emploi de termes comme « vert », « chaise », « arthrite », etc. Aussi, le concept de correction devrait-il avoir la même signification qu'il s'applique aux concepts logiques comme aux concepts empiriques (ou matériels); autrement dit, comment devrions-nous concilier d'un côté, les

---

<sup>861</sup> Sur la question de la signification des termes et de leur usage dans les jugements empiriques, Brandom accepte la dérèfification de la signification et l'abandon de la distinction de l'analytique et du synthétique de Quine et de Davidson, et parle ainsi sans les dissocier de l'objectivité de la vérité des jugements et de l'application objective des normes de signification du point de vue de leur correction.

distinctions entre une normativité forte et une normativité faible, et de l'autre, l'emploi d'un vocabulaire logique et d'un vocabulaire non logique<sup>862</sup> ? Nous distinguons ici deux sens du terme normatif, à savoir : un sens pragmatique ou l'idée d'une normativité faible et un sens objectif et constitutif ou la thèse d'une normativité forte. Concevoir la signification sous l'angle d'une normativité faible revient à considérer que les normes de signification sont établies en fonction des aspects pragmatiques et pour les fins de la conversation ; ce faisant, le statut normatif de la signification d'un terme a alors un sens proprement pragmatique. Par conséquent, concevoir le statut normatif de la signification comme faible consiste à considérer les normes telles que « 'x' signifie x » comme générant la force normative de « nous devons traduire la langue d'un autre comme signifiant x par 'x' si nous voulons pour les fins de la discussion suivre notre interlocuteur et le comprendre<sup>863</sup> ». Par contre, concevoir la signification sous l'angle d'une normativité forte vise à comprendre la force normative des propositions exprimant la signification des termes non seulement comme temporairement orientée vers la réalisation de fins ou de désirs particuliers mais surtout comme constitutive de la signification. Une lecture de la normativité forte énoncerait ainsi que la norme de signification « 'x' signifie x » véhicule le sens suivant : « nous devons juste traduire la langue

<sup>862</sup> Nous partons de la définition de Quine, 1976b, p. 74, pour préciser la notion de vérité logique : « But if S'' involves only logical expressions, and hence remains true when everything except the skeleton of logical expressions is changed in all grammatically possible ways, the S'' depends for its truth upon those logical constituents alone, and its thus a truth of logic. ». Précisons également comment Quine, 1965, p. 2, définit la notion d'*occurrence essentielle* : « Un mot fait une occurrence essentielle si son remplacement par un autre est capable de transformer un énoncé vrai en un énoncé faux. ». En effet, dans l'énoncé « Socrate est un homme ou Socrate n'est pas un homme » la conjonction « ou » fait une occurrence essentielle et si on la remplace par « et » on rend faux l'énoncé initial. Nous pouvons alors ranger du côté du vocabulaire logique, la classe des vérités logiques en fournissant la liste des connecteurs ou particules logiques qui est donnée par *énumération* dans Quine, 1976a, p. 103 : « First we suppose indicated, by enumeration if not otherwise, what words are to be called logical words; typical ones are 'or', 'not', 'if', 'then', 'and', 'all', 'every', 'only', 'some'. The logical truths, then, are those true sentences which involve only logical words *essentially*. ». En stipulant qu'une vérité logique est un énoncé dont la vérité est assurée par sa structure logique (« un énoncé est logiquement vrai si tous les énoncés qui partagent sa structure logique sont vrais », Quine, 1975, p. 141) et dont les propriétés d'usage sont exprimées au sein de lois logiques, de règles d'inférence et de tables de vérité. Du côté du vocabulaire non logique, nous incluons par exemple les termes de couleurs, les termes médicaux, les objets matériels comme « table », les termes d'espèces naturelles tels que « acide » et « lapin », les expressions mathématiques comme « + », etc.

<sup>863</sup> Ces normes prescriptives régissent notre interprétation de la signification du langage d'autrui, c'est-à-dire s'énoncent à la troisième personne et régissent ainsi l'attribution de la signification de nos propres termes, et peuvent donc s'énoncer à la première personne comme suit : « Si je veux être aisément compris, je dois employer 'x' de cette manière-ci (soit de la manière dont les autres l'utilisent). ».

d'un autre comme signifiant x par 'x' si celui-ci vise à formuler un énoncé porteur de sens (et non seulement si nous sommes enclins à suivre notre interlocuteur afin d'arriver à nous entendre facilement sur les termes que nous employons) ». Cela dit, les normes fortes qui résultent de cette lecture ne sont pas pour autant infaillibles mais bien sujettes à révision dans certaines circonstances (et non juste pour une convenance pragmatique).

Par conséquent, au regard de notre distinction entre une normativité faible et une normativité forte, entre un sens pragmatique et un sens plus objectif de l'application correcte des normes de signification, deux questions se posent : d'une part, celle de savoir si une norme de l'usage linguistique est facultative ou plus ou moins nécessaire pour les locuteurs d'un langage naturel qui cherchent à rendre leur propos intelligible pour les fins de la discussion et d'autre part, celle de savoir dans quel registre (d'une normativité forte ou faible) classer les normes de signification régissant l'usage d'un vocabulaire logique et celles régissant l'usage d'un vocabulaire non logique. Plusieurs positions sont envisageables mais diffèrent tout d'abord sur la question de savoir si l'action linguistique est proprement normative, ensuite sur celle de décider quel sens (fort ou faible) lui attribuer et enfin sur celle de savoir si le même sens s'applique ou non aux deux domaines lexicaux.

Dummett et Davidson soutiennent ainsi que le statut normatif du vocabulaire logique (au moins) doit être compris au sens fort. Pour examiner ce point, il nous faut analyser la critique, explicitement formulée par Dummett<sup>864</sup> et implicitement contenue dans l'œuvre de Davidson, du point de vue naturaliste de Quine sur le langage. La normativité du langage n'est pas une illusion qui peut être éliminée de manière naturaliste, comme Quine le suggère.

Dummett revient d'abord sur la notion d'*expérience récalcitrante* entrevue chez Quine<sup>865</sup> en soulignant qu'une telle notion présuppose bien plus que ce que les ressources naturalistes de Quine prétendent soutenir:

---

<sup>864</sup> Dummett, 1973, p.587-627; la critique implicite de Davidson est discutée dans le chapitre précédent.

<sup>865</sup> Quine, 2003, p.77

« Quine's thesis involves, however, that the principles governing deductive connections themselves form part of the total theory which, as a whole, confronts experience. Presumably [...] we must recognize the total theory as comprising rules of inference as well as logical laws in the form of valid schemata or their instances: but there is no reason to suppose that Quine draws a distinction between the status of such rules as against laws like Excluded Middle; they too must be equally liable to rejection [...] But, in that case, there is nothing for the inferential links between sentences to consist in<sup>866</sup>. ».

Cependant, Dummett souligne, comme nous l'avons précédemment mentionné, que Quine affirme bien qu'il n'existe pas d'énoncés non révisables : « les lois mathématiques et logiques elles-mêmes ne sont pas soustraites à révision<sup>867</sup> ». Ce sont ainsi les interconnexions assurées par l'implication logique entre les énoncés<sup>868</sup> qui rendent n'importe quel énoncé vulnérable à une expérience négative. Or, la notion de récalcitrance requiert que certains énoncés aient un statut normatif c'est-à-dire qu'ils soient à l'abri de la révision, puisque c'est seulement parce que nous pouvons tenir certains énoncés comme exempts de révision que des révisions peuvent être imposées par des expériences inattendues. Les lois logiques et les règles d'inférences sont, selon Dummett, les plus à même de se voir accorder une telle immunité.

Dummett veut ainsi rendre compte de l'incohérence dans le point de vue de Quine entre d'une part, son projet révisionniste global tendant à bannir la normativité et d'autre part, cette notion de récalcitrance faisant appel à la normativité. Dummett insiste sur cette tension interne inhérente à l'approche de la signification et du langage chez Quine : son rejet de la normativité<sup>869</sup> défendue par sa position naturaliste est en contradiction avec son approche de l'apprentissage du langage, de la formation et du changement de croyances, à laquelle s'ajoute subrepticement la normativité *via* la notion de récalcitrance.

Le second point que Dummett critique, et que Davidson souligne implicitement dans son approche de la croyance, du langage et de l'interprétation, est énoncé comme suit :

<sup>866</sup> Dummett, 1973, p. 596.

<sup>867</sup> Quine, 1965, Introduction.

<sup>868</sup> Quine, 2003, p.77: « La réévaluation de certains énoncés entraîne la réévaluation d'autres, à cause de leurs interconnexions logiques – les lois logiques étant, à leur tour, simplement certains énoncés de plus dans le système, certains éléments de plus dans le champ. ».

<sup>869</sup> Comme nous l'avons vu au chap. I, la critique de Quine contre le concept normatif de signification visait à montrer qu'il est vain de rechercher en un tel concept une source de justification des énoncés de nos théories scientifiques.

« As a result of the second thesis [viz. that no sentence is immune from revision] the internal structure of the theory, consisting in the interconnections of sentences with one another, is totally dissolved, and the theory becomes a mere featureless collection of sentences standing in no special relations to each other<sup>870</sup>. ».

Dummett soutient que Quine aplanit toutes les distinctions entre les énoncés. Dans sa métaphore du « web of belief », toutes les croyances sans oublier les croyances logiques ont le même statut épistémologique. Aucune n'est à l'abri de la réfutation par l'expérience, chacune d'elles est un nœud au sein de ce réseau. C'est précisément cette image qui pose problème à Dummett. Les principes logiques ne peuvent pas juste être des nœuds dans ce réseau holistique dans la mesure où ils sont constitutifs de ce « web of belief », en tant que structure ordonnée d'énoncés : ce sont les interconnexions logiques entre les croyances (ou énoncés) qui déterminent qu'il s'agit bien d'un système complexe de croyances liées entre elles et non seulement d'un comportement vide de tout contenu. Les connexions logiques devraient être préservées et occuper une place privilégiée dans le « web of belief », et ce faisant, les lois logiques ne peuvent pas juste être des fragments de ce réseau de croyances au même titre que d'autres. Si les lois logiques n'ont pas ce statut privilégié de normes au sein de ce système de croyance alors il n'est pas nécessaire que le système de croyance soit rationnellement articulé mais il peut seulement l'être causalement, et si la rationalité de la croyance disparaît, il en est alors de même de l'idée que si je crois une proposition alors je dois croire une autre croyance, ou si je crois une proposition alors je ne peux pas en croire une autre. Si la croyance perd cette caractéristique alors cela ne va plus de soi qu'il s'agit bien d'un système de croyances. Il semble bien qu'un système de croyances réponde à la description d'un ensemble de croyances inférentiellement articulées c'est-à-dire une structure organisée et régie par des lois logiques et des règles d'inférences. En naturalisant la logique et en la réduisant à n'être qu'un nœud anodin dans ce réseau de croyances tout aussi sujet à révision que tout autre fragment de l'ensemble, nous perdons la dimension normative de la croyance s'incarnant dans des notions telles que celle de contradiction et d'implication. Une histoire causale ne peut pas rendre compte de ces notions-ci dans la mesure où rien ne m'interdit d'un point de vue causal de tenir des

---

<sup>870</sup> Dummett, 1973, p. 597.

croyances contradictoires<sup>871</sup>. Or c'est irrationnel d'un point de vue normatif d'avoir des croyances contradictoires. Ce n'est que si la logique a un statut privilégié que nous pouvons établir des relations normatives entre les croyances dans ce réseau et que nous pouvons rendre compte non seulement des relations causales mais aussi des relations rationnelles entre les croyances. Dummett soutient ainsi que Quine, en mettant les lois logiques sur le même pied d'égalité que les autres croyances dans ce réseau holistique de croyances, néglige la dimension normative de la croyance.

Nous avons jusqu'ici exprimé le dilemme dans les termes généraux opposant normativité et naturalisme; nous voulons à présent traduire le problème dans ceux plus spécifiques d'une normativité forte *versus* une normativité faible.

Quine semble certes reconnaître une dimension normative au langage à travers le rôle pragmatique qu'il accorde au principe de charité. Le principe de charité recommande ainsi d'attribuer autant que possible au locuteur des concepts auxquels son usage est conforme et il interdit, pour faciliter la compréhension et donc à des fins pragmatiques, tout manuel de traduction selon lequel le locuteur admettrait des lois logiques incompatibles avec les nôtres<sup>872</sup>. Que doit-on donc faire face à un locuteur qui, par exemple, est disposé à accepter un énoncé de la forme «  $p$  ou  $q$  » si, et seulement si, il est disposé à accepter au moins l'un des composants, et à refuser ce même énoncé si, et seulement si, il est disposé à rejeter chacun des composants, mais qui n'accepterait pas certains énoncés de la forme «  $p$  ou  $\neg p$  » ? Une application du principe de charité recommanderait une traduction selon laquelle le locuteur utilise correctement ce « ou » (et ce « non »), autrement dit une traduction

---

<sup>871</sup> Si nous considérons seulement la question de l'émergence des croyances, rien ne nous empêche ici d'avoir des croyances contradictoires puisque nous nous intéressons seulement à connaître la provenance de nos croyances et non à savoir si celles-ci sont cohérentes ou justifiées. Une telle approche peut me montrer que mon professeur, dont je partage les croyances, tient des croyances contradictoires. De cette manière, une approche causale peut « expliquer » que je tiens des croyances contradictoires, mais ne les aura pas, de ce fait, justifiées. On voit ainsi la différence entre des notions normatives (comme celles de justification et de récalcitrance) et des notions causales (comme celle de la provenance causale).

<sup>872</sup> Quine, 1975, semble ainsi osciller entre ces deux interprétations du principe de charité, en admettant tantôt une formulation selon laquelle on doit imputer notre logique à autrui : « nous attribuons à l'indigène notre logique orthodoxe ou nous la lui imposons, lorsque nous traduisons sa langue. Nous incorporons la logique dans notre manuel de traduction » (p.121); tantôt une formulation selon laquelle il ne peut y avoir d'incompatibilité entre la logique des locuteurs étrangers et la nôtre : « Telle est évidemment la difficulté de la position du logicien déviationniste : c'est quand il essaie de nier la doctrine, il ne fait que changer le sujet. » (Ibid., p. 121).

qui révélerait que les inférences qu'il produit à l'aide de ce connecteur sont également valides (et celles qu'il rejette ne le sont pas), et que les lois faisant intervenir ce connecteur qu'il accepte (refuse) sont généralement vraies (fausses). Autrement dit, le principe de charité recommanderait de ne pas traduire ce « ou » (en supposant bien sûr que notre logique est la logique classique et que nous acceptons tous les énoncés de la forme « p ou non p »). Selon cette interprétation, il y aurait différence de langue entre celui-ci et nous-mêmes car son usage de « ou » est conforme à d'autres règles. Ainsi, il n'est pas vrai que le principe de charité nous oblige à imposer nos connecteurs logiques à autrui. Cette distinction est importante : le principe de charité recommande de traduire les connecteurs logiques de la langue étrangère de manière conforme à leur usage, c'est-à-dire en accord avec les lois (contradictions) et équivalences qui sont acceptées (rejetées), pourvu bien sûr que notre langue possède les ressources nécessaires pour exprimer ces lois et équivalences; il ne recommande pas de traduire ces connecteurs par des connecteurs que de fait nous utilisons, sans tenir compte des différences potentielles d'usage entre les connecteurs des deux langues.

Est-ce à dire qu'en pratique, nous aurions tendance à affirmer qu'il n'y a pas de conflit possible en ce qui concerne la logique, et serions toujours prêts à réinterpréter les usages déviants dans ce domaine ? Il semble qu'il n'existe pas de critère strict permettant de déterminer le poids relatif du principe de charité, qui est susceptible de varier selon le contexte. Par exemple, on dit parfois que la logique intuitionniste propose un ensemble de concepts différents de ceux de la logique classique; mais on dit aussi qu'elle est incompatible avec celle-ci. Dans la première affirmation, c'est le principe de charité qui domine, autrement dit, au lieu de voir un désaccord entre la logique intuitionniste et la logique classique, on adopte une traduction non homophonique des connecteurs logiques, qui élimine le désaccord entre le logicien déviant et nous, alors que dans la seconde affirmation, on fera intervenir la contrainte de la communauté<sup>873</sup>, autrement dit, on impose une traduction homophonique, de sorte que les différences de verdicts sont interprétées comme exprimant des désaccords.

---

<sup>873</sup> On peut formuler la contrainte de la communauté comme suit : Tous les individus ayant des dispositions verbales similaires et pouvant échanger verbalement de façon courante doivent être considérés comme des locuteurs compétents parlant une langue commune, de sorte que le manuel de traduction doit, autant que possible, être construit de façon à ce qu'il soit applicable à n'importe lequel de ces individus.

Par conséquent, selon cette approche, il est possible d'attribuer des croyances contradictoires dans la mesure où la norme qui prescrit de ne pas le faire n'a qu'un statut pragmatique et n'est requise que pour faciliter la compréhension. Dummett et Davidson soutiennent au contraire que cette approche pragmatique et normative de la logique ne rend pas compte de la force de la contrainte de la logique sur la signification et l'attribution de la croyance. Ils affirment que la cohérence logique est nécessaire pour l'attribution de croyances et de significations.

Pour reprendre ainsi les termes de Dummett autour de la notion de récalcitrance de Quine, celui-ci affirme que la récalcitrance impose la révision du schème de la traduction. Si certains fragments de notre comportement linguistique s'avèrent récalcitrants au point de rendre notre traduction incohérente, la traduction doit alors être révisée. Il est donc bien nécessaire et non seulement pragmatiquement requis de rétablir la cohérence logique. Par conséquent, la notion de récalcitrance indique que la cohérence logique est une contrainte fortement normative portant sur la traduction. Davidson insiste également sur le statut constitutif de la logique dans son approche de l'interprétation radicale. Nous ne pouvons en effet pas interpréter une croyance sans présupposer d'une part (par le principe de charité comme cohérence) que cette croyance est reliée à un certain nombre d'autres croyances, et sans présupposer d'autre part (par le principe de charité comme correspondance) que la plupart des croyances interprétées sont vraies. Ce point marque une différence cruciale entre la conception du langage, de la compréhension et de l'interprétation de Davidson et celle de Quine, quant aux différents rôles que chacun d'eux assigne respectivement au principe de charité dans la traduction des constantes logiques<sup>874</sup>. Eu égard au rôle constitutif de la rationalité dans l'attribution des croyances et des significations, Davidson donne au principe de charité un rôle constitutif dans l'interprétation et non simplement pragmatique. Le principe de charité dans l'une de ses formulations est un principe de cohérence logique ou de rationalité des croyances, qui nous

---

<sup>874</sup> Cette différence est discutée au Chap. IV, §4.4.2. Affirmer que Quine accorde un rôle strictement pragmatique au principe de charité dans la traduction du vocabulaire logique est discutable : il apparaît parfois qu'il veuille lui donner un usage constitutif, comme ses remarques critiques sur la thèse de Lévy-Bruhl portant sur « la mentalité pré-logique des sauvages » le confirment. Chez Quine, le principe de charité revient à dire que toute traduction doit préserver les lois logiques de la logique usuelle. C'est en ce sens qu'il dit que « la stupidité de notre interlocuteur est, au-delà d'un certain point moins probable qu'une mauvaise traduction » et que « la mentalité prélogique est un trait infusé par de mauvais traducteurs ». (Quine, 1977a, p.59 et Engel, 1989 pour cette critique quinienne de Lévy-Bruhl)

prescrit de tenir celles-ci comme étant, dans l'ensemble, rationnelles, selon les canons de la logique. On peut le résumer ainsi : interprétez de manière à rendre les croyances de ceux que vous interprétez conformes aux lois de la logique. Nous ne pouvons donc pas construire une théorie de la signification sans présupposer et respecter cette maxime.

Nous devrions alors considérer les lois logiques comme ayant une validité indépendante de l'observation des régularités du comportement linguistique d'une communauté particulière ou des buts et intérêts que celle-ci (ou un interprète) peut poursuivre. Les lois logiques ne rendent pas seulement compte de la manière dont nous pensons ou parlons, elles ne prescrivent pas non plus comment nous devrions penser si nous voulons être aisément compris, mais elles prescrivent comment nous devrions penser si nous voulons être considérés comme des individus capables d'avoir des croyances pour avoir des pensées et des locuteurs, capables d'interpréter les croyances d'autrui. Les lois de la logique peuvent certes se justifier par leur utilité et leur efficacité pragmatique mais nous avons aussi de bonnes raisons de croire qu'elles ont bien un statut normatif fort, constitutif et prescriptif soit que la logique exerce une contrainte *a priori* ou « transcendantal » (pouvant néanmoins être mise en défaut) sur la possibilité même de toute interprétation des pensées, croyances et significations d'autrui. La question est alors de déterminer si la signification du vocabulaire non logique est envisageable du point de vue d'une normativité forte.

Nous pensons effectivement que Wittgenstein, McDowell et Davidson préconisent aussi bien pour le vocabulaire logique que pour le vocabulaire non logique le point de vue d'une normativité forte<sup>875</sup>. Cependant, cette approche est-elle justifiée ou ont-ils illégitimement

---

<sup>875</sup> La formulation large du principe de charité interprétative et du réquisit de rationalité semble appuyer la prise de position de Davidson pour une normativité forte du vocabulaire non logique. Une formulation littérale du principe de charité énonce que nous devons considérer les individus que nous interprétons comme suivant des « lois de la logique » et que nous ne pouvons, par conséquent, pas considérer l'interprété comme ayant des croyances, désirs et significations sans le considérer comme rationnel soit en supposant qu'il suit dans sa pensée et dans son action, certains principes très généraux de rationalité. Néanmoins, une version plus large du principe de charité ajoute qu'attribuer des croyances particulières synthétiques à l'interprété soit des croyances impliquant les concepts empiriques tels que pluie, herbe, neige, etc., est également requis pour le concevoir comme ayant des croyances, désirs et significations. Aussi, la contrainte constitutive de rationalité, au cœur de la théorie de l'interprétation, se compose de normes de signification incluant ces deux domaines lexicaux dans la mesure où celle-ci doit satisfaire deux types de condition, à savoir : « (1) le principe de cohérence [qui] pousse l'interprète à découvrir un degré de cohérence logique dans les pensées du locuteur et (2) Le principe de correspondance [qui] pousse l'interprète à considérer que le locuteur répond aux mêmes traits du monde auxquels lui, l'interprète répondrait en de pareilles circonstances. ». (Davidson, 2001b,

supposé qu'il serait possible d'étendre la normativité forte de la logique à la signification des termes non logiques ?

Les principes logiques sont ainsi, dans les termes de Davidson les contraintes *a priori* de notre interprétation des autres dans la mesure où elles semblent rendre compte de ce que nous entendons par normes de rationalité. Cependant, il s'avère bien que nous avons le choix quand nous venons à attribuer les concepts empiriques comme vert, arthrite, etc., aux autres locuteurs. Il semble que nous pourrions tout aussi bien leur attribuer les concepts *tharthrite*<sup>876</sup>, etc., sans que pour cela, ces concepts soient dénués de sens ou dépourvus d'intelligibilité ou même de rationalité. Ces locuteurs emploieraient bien ces concepts-ci mais avec une signification différente de la nôtre ; autrement dit, les normes de signification gouvernant l'usage des termes empiriques tels que « vert », « tigre », « arthrite » diffèrent des nôtres. Ce qui revient finalement à dire que le statut normatif de la signification de ces termes-ci n'est pas fort c'est-à-dire que nos normes de signification ne sont pas des conditions *a priori* pouvant fixer les critères de correction que ces individus utilisent quand ils parlent leur langue, et les concepts que ces termes sont supposés exprimer. Il semble ainsi que nous puissions comprendre l'usage que ces locuteurs font des termes non logiques sans leur attribuer nécessairement nos normes de signification. Par conséquent, si la signification des termes empiriques a un sens normatif alors celui-ci doit être faible et pragmatique : en effet, si nous attribuons à ces termes empiriques des normes de signification différentes des nôtres, alors les comprendre et les interpréter ne devient certes pas impossible mais bien plus difficile que si l'usage de ces termes se faisait dans le respect de nos normes sociales de signification.

Le processus d'interprétation doit ainsi reposer, en vertu du principe constitutif de charité, sur l'idée que la plupart des croyances de ceux que nous interprétons doivent être vraies, et que l'interprète doit partager un ensemble de croyances vraies avec l'interprété. Il s'ensuit que nous ne pouvons pas, en vertu même de l'interprétation, juger que d'autres locuteurs de notre langage, ou des locuteurs d'un autre langage ont des croyances

---

1991, p.211) L'argument de Wittgenstein sur la question de suivre une règle et celui du langage privé tentent également d'établir un sens objectif (faillibiliste) de la notion de correction qui puisse s'appliquer à toutes sortes de concepts non logiques comme les concepts mathématiques ou psychologiques, les objets matériels, les couleurs, etc. McDowell suit Wittgenstein sur ce point. Ce point est discuté dans Vol. I, chap.III, § 3.3.1.

<sup>876</sup> L'exemple le plus discuté de ce phénomène est sans doute celui qu'a proposé Tyler Burge, 1979.

radicalement différentes. Aussi, le but de l'interprétation n'étant pas l'accord mais la compréhension, nous sommes contraints que nous le voulions ou non à reconnaître que l'interprété a raison et que l'interprète et l'interprété partagent un ensemble de croyances en grande partie cohérentes et vraies pour considérer l'interprété comme ayant des croyances et comme tenant des propos sensés :

Minimiser le désaccord, ou maximiser l'accord, est un idéal confus. Le but de l'interprétation n'est pas l'accord, mais la compréhension. J'ai toujours insisté sur le fait que l'on ne peut parvenir à la compréhension qu'en interprétant de façon à obtenir la bonne espèce d'accord. Mais la « bonne espèce » d'accord n'est pas plus facile à spécifier que de dire ce en quoi consiste une bonne raison d'accepter une certaine croyance<sup>877</sup>.

Toute la méthodologie « minimaliste » de Davidson pour la construction d'une théorie de la signification repose donc sur l'idée que, bien que nous ne connaissons pas la signification des phrases de l'interprété, nous devons supposer la vérité comme déjà comprise, ou comme partagée par l'interprète et l'interprété. Le principe de charité, au sens littéral, revient à dire que toute interprétation doit préserver la structure formelle et logique des énoncés; autrement dit, l'interprète doit découvrir un degré de cohérence logique dans les pensées du locuteur. Mais cette lecture du principe de charité ne détermine aucun contenu propositionnel empirique et ne requiert donc pas le partage de certaines croyances empiriques particulières ou de normes de signification. Aussi, le rôle constitutif du principe de charité a également selon Davidson une interprétation plus large dans laquelle l'attribution de certaines croyances synthétiques telles que l'herbe est verte ou il y a une table, etc., (c'est-à-dire des concepts matériels ou empiriques) est une condition *a priori* de l'interprétation des autres. Nous devrions donc attribuer aux autres locuteurs certains de nos concepts comme vert, chaise, etc., (c'est-à-dire des croyances synthétiques en grande partie cohérentes et vraies selon nos propres normes) afin de les identifier comme des individus tenant des propos sensés. Les normes de signification gouvernant ces expressions non logiques-ci doivent donc être comprises comme fortement normatives ou comme s'appliquant à tous les locuteurs au risque de ne pas considérer ces autres locuteurs comme individus capables de formuler des propos sensés.

---

<sup>877</sup> Davidson, 1993b, p. 15.

Cependant, lire le principe de charité dans un sens trop fort, en présupposant des normes *a priori* de rationalité, rend la notion même d'erreur et d'incohérence logique impossible<sup>878</sup>. Concevoir la possibilité de l'erreur affaiblit certes le sens d'une dimension constitutive de la vérité et de la cohérence logique attribué aux croyances et aux significations mais il n'en demeure moins qu'il s'applique bien aux concepts empiriques dans la mesure où ce qui est considéré comme rationnel dépend non seulement des relations logiques que les croyances entretiennent mais également des circonstances et du contexte dans lesquelles nous tenons pour vrai certaines croyances, lesquelles supposent que nous comprenions leur contenu propositionnel. Cette condition *a priori* de rationalité affaiblit certes le sens d'une dimension constitutive de la thèse d'une normativité forte mais elle n'en restreint pas moins sa portée – puisque Davidson soutient que quel que soit le sens logique selon lequel les normes sont fortes ou constitutives (qui est maintenant un sens plus faible), celui-ci s'applique également aux normes de significations des expressions non logiques. Par conséquent, selon la lecture du principe de charité qui part de la rationalité de l'agent, Davidson évite la position intermédiaire d'une normativité forte de la logique et celle d'une normativité faible des concepts non logiques.

Aussi, la thèse d'une normativité forte soutenue par Davidson selon laquelle l'interprétation commence d'abord avec des phrases occasionnelles cristallise bien sa position pour une normativité forte attachée tant au vocabulaire non logique qu'au vocabulaire logique. Les phrases occasionnelles sont, comme nous l'avons vu<sup>879</sup>, des phrases qui, de façon paradigmatique, implique des concepts empiriques, soit des phrases qui expriment des croyances synthétiques, dont l'assentiment requiert la considération d'un contexte particulier. Seules les phrases occasionnelles suscitent ainsi l'assentiment de l'informateur uniquement lorsque l'ethnologue les prononce immédiatement après la stimulation intersubjective appropriée. Les interprètes prennent ainsi comme point de départ ces phrases-ci et tentent de produire une phrase équivalente dans leur propre langue aux concepts empiriques de l'interprété. Selon le principe de charité, il est alors nécessaire de supposer que l'interprété

<sup>878</sup> Si la charité est poussée à l'extrême, elle efface la communication. Dummett, 1978, semble d'ailleurs objecter qu'être trop charitable fera perdre son sens à la notion d'erreur. Mais, il ne s'agit pas, pour Davidson, d'être trop ou pas assez charitable : la charité nous est imposée, ensuite l'interprète « triangule » entre ses jugements de vérité, ceux de l'interprété, et leur interaction avec le monde extérieur.

<sup>879</sup> Chap. IV, §4.4.2.

possède, en gros, les mêmes principes normatifs que l'interprète lui-même, par exemple, le principe de non contradiction selon lequel nous ne croyons pas en même temps une proposition  $p$  et une proposition contradictoire  $\neg p$ . Ce n'est pas que les locuteurs sont, en fait, toujours raisonnables, ni qu'ils souscrivent nécessairement aux mêmes principes normatifs que l'interprète. Davidson soutient seulement que, sans aucune connaissance de l'interprété en question, nous ne pouvons même pas concevoir l'interprétation sans supposer une rationalité semblable à la nôtre. Aussi, nous ne pouvons comprendre les croyances du locuteur sans savoir ce que signifie son énoncé et nous ne pouvons déterminer le sens de son énoncé sans savoir quelle conviction il exprime en l'énonçant :

« The interpreter's problem is that what he is assumed to know - the causes of assents to sentences of a speaker - is, as we have seen, the product of two things he is assumed not to know, meaning and belief. If he knew the meanings he would know the beliefs, and if he knew the beliefs expressed by sentences assented to, he would know the meanings. But how can he learn both at once, since each depends on the other?<sup>880</sup>».

Aussi, les relations logiques et l'attribution de ces mêmes lois logiques aux autres locuteurs sont articulées au sein de relations inférentielles entre concepts non logiques. Autrement dit, les lois logiques semblent satisfaire une contrainte *a priori* portant sur l'interprétation, celle-ci consistant à interpréter autrui comme manipulant des concepts empiriques, c'est-à-dire comme instaurant de manière holistique la signification des phrases occasionnelles au terme d'un processus dynamique de formulations d'hypothèses, de tests et de révisions successives. Aussi, le rejet d'une normativité forte rattachée au vocabulaire non logique semble reposer sur une distinction trop stricte entre normes logiques et normes conceptuelles empiriques. Un tel rejet suppose ainsi que les relations logiques sont indépendantes de toute compréhension du sens des concepts et que les relations logiques peuvent être identifiées et attribuées indépendamment de toute instantiation ou occurrence de concepts. Nous laissons en quelque sorte croire qu'il s'agirait de faire comme si un interprète pouvait discerner qu'un interprété respecte bien le principe de non contradiction, soit que nous ne pouvons pas avoir  $p$  et  $\text{non } p$  à la fois sans pour autant comprendre la signification des concepts composant le contenu des propositions que  $p$ . Mais, nous ne pouvons pas discerner les relations logiques entre les phrases – soit quelles sont les phrases qu'une phrase donnée implique, quelles phrases sont impliquées par celle-ci ou sont cohérentes avec celle-ci

<sup>880</sup> Davidson, 2001b, 1983, p. 148.

– et par conséquent si les principes logiques ont été suivis correctement sans connaître ce que les phrases (soit les concepts les composant) signifient. Ainsi, pour saisir les relations logiques entre les phrases, il faut déterminer conjointement la signification des concepts matériels les composant; aussi, attribuer une croyance et comprendre une proposition soit connaître les conditions qui devraient être satisfaites pour que  $p$  soit vrai, ce qui est, pour Davidson, une contrainte *a priori* portant sur l'interprétation, dépend de l'attribution de lois logiques à certains concepts empiriques.

Par conséquent, selon Davidson, un être ne peut avoir une attitude sans automatiquement avoir un certain nombre de croyances qui sont directement présupposées ou impliquées par celle-ci. Mais la dépendance des attitudes par rapport aux croyances va plus loin, étant donné le holisme soutenu par Davidson. Pour qu'une attitude propositionnelle ait le contenu qu'elle a, il faut qu'elle soit liée à un réseau de croyances qui sert justement à déterminer ce contenu. Autrement dit, les relations logiques entre phrases sont articulées au sein de relations inférentielles entre les croyances synthétiques ou les concepts matériels, de telle sorte que les relations logiques ne peuvent pas faire abstraction de l'attribution de concepts matériels. Mais, il semblerait que ce point ne réponde pas à nos attentes. Une normativité forte des termes non logiques requiert plus qu'un simple engagement au holisme du contenu conceptuel matériel dans la mesure où le holisme<sup>881</sup> exige seulement que, quels que soient les concepts empiriques employés par un autre locuteur, ces derniers soient reliés logiquement les uns aux autres, et ne pose pas que le locuteur que nous interprétons doit disposer (plus ou moins) du même ensemble spécifique de concepts matériels que nous afin d'être reconnu comme un locuteur employant des concepts et tenant des propos sensés, ce qu'une normativité forte rattachée au vocabulaire non logique exige. Autrement dit, même si les relations logiques requièrent l'identification d'un contenu non logique, l'identification n'est

---

<sup>881</sup> McDowell, 1987, p.103, soutient ainsi: « As I understand the holism Davidson accepts, it is the thesis: attributions of content to sentences in a community's language, to their linguistic acts, and to their psychological states are systematically interlocked, in such a way that [...] there is no explaining, 'as from outside' the entire system, what it means to ascribe some specific content to an appropriate item. Clearly this entails repudiating any aspiration to be anything but modest in theories of meaning. ». Du fait que les attributions de contenus soient « enchevêtrées » aux actes de discours des membres d'une communauté et à leurs attitudes, il ne s'ensuit pas que l'on ne puisse pas distinguer le point de vue de l'interprète de celui du locuteur. Davidson doit certes exiger qu'il y ait un certain isomorphisme entre la théorie de l'interprète et la compétence de l'interprété mais il n'a pas besoin d'exiger l'identité des deux.

pas synonyme d'identité; et il semble que l'argument ci-dessus oscille entre identification et identité, entre la thèse selon laquelle nous devons attribuer un contenu propositionnel aux concepts empiriques et celle selon laquelle nous devons leur attribuer le même contenu<sup>882</sup>. Nous parlons ici de l'assomption nécessaire à l'amorce de toute interprétation. À ce niveau, formuler une assomption sur le contenu des croyances revient à formuler une assomption sur la similitude du contenu. Il n'y a pas de contenu sans un minimum de similitude<sup>883</sup> sur le contenu auquel cas nous ne pourrions pas nous comprendre. Par conséquent, si nous voulons déterminer si les concepts utilisés par nos locuteurs possèdent un quelconque contenu, nous sommes contraints de supposer que leurs concepts matériels sont similaires ou qu'ils peuvent être d'une certaine façon ajustés aux nôtres. La distinction entre un contenu identique ou un contenu différent requiert au préalable que nous supposions une certaine similitude du contenu comme condition de possibilité de toute interprétation. Par conséquent, avec le principe de charité de Davidson, nous disposons bien d'une condition *a priori* garantissant une compréhension mutuelle ou partagée (et non simplement l'identification) du contenu non logique, c'est-à-dire que la charité implique qu'une normativité forte s'étende au vocabulaire non logique : le principe de charité pose en quelque sorte la condition minimale que tout concept matériel doit (plus ou moins) partager et satisfaire. Notons que cette similitude dans le contenu empirique des concepts peut bien être d'une portée limitée et permettre d'engendrer des divergences dans le contenu au fur et à mesure que nous procédons à l'interprétation - comme toutes normes fortes de signification, la normativité forte du contenu matériel conceptuel ne signifie pas que ces mêmes normes de signification appartenant au vocabulaire non logique sont fixées de manière rigide et s'appliquent une fois pour toutes. Bien au contraire, nous ajustons nos hypothèses et nos intuitions initiales de façon à les rendre conformes à nos règles normatives et principes de rationalité, jusqu'à ce que nous ayons atteint un état d'équilibre, obtenu par des ajustements successifs. Ainsi, en ayant maintenant une idée plus claire sur la portée d'une application correcte et objective des normes de signification et sur les raisons qui motivent son extension (certes circonscrite) au-delà du vocabulaire logique, nous devons maintenant aborder la question de ce qui, en

<sup>882</sup> En projetant ma logique sur celle de l'interprété.

<sup>883</sup> Davidson, 1993b, 1977a, p. 291 : « Si j'ai raison de vous attribuer la croyance, c'est que vous devez avoir une trame de croyances qui ressemble beaucoup à la mienne. ».

l'absence de toute référence à un royaume transcendant de significations idéales, pourrait être responsable de l'établissement de cette application objective et correcte de la signification.

#### 5.6. Davidson : une approche publique ou sociale des normes de signification

Wittgenstein a souvent été considéré par plusieurs comme ayant établi l'impossibilité d'un langage *privé*. Un langage privé se définit comme un langage qui ne peut (« logiquement ») être compris que par celui qui l'utilise. Wittgenstein invoque l'exemple d'un nom employé par un locuteur pour se référer à une sensation qu'il qualifie lui-même de « privée » : seul le locuteur peut savoir s'il a la sensation et personne n'est en mesure de savoir quelle est sa nature (sa qualité subjective ressentie). Si « S » est le mot choisi pour s'y référer (un signe que le locuteur pourra utiliser par exemple pour noter sa récurrence dans son journal), alors « S » est un nom privé.

Or, un langage (un mot) privé n'est pas possible parce que nous ne pouvons pas faire dans un tel langage une distinction objective entre un usage correct et un usage incorrect d'un mot. Le locuteur du langage ne possède en effet aucun « critère de correction » pour contrôler son emploi de « S »<sup>884</sup>. Un « critère de correction » est un moyen indépendant (et interpersonnel, dit Kripke) pour dire si un mot est utilisé correctement ou non. Nous n'avons pas dans le cas présent de critère de correction parce que la sensation est dénuée de tout accompagnement publiquement observable (S n'a aucune manifestation naturelle, elle n'a ni cause ni effet observable pour autrui<sup>885</sup>) et c'est pourquoi quelqu'un d'autre n'est pas en mesure de savoir quand le locuteur la ressent ni déterminer quand son usage de « S » est correct ou ne l'est pas. Bref, sans critère de correction, il n'y a pas de distinction (objective) entre usages corrects et usages incorrects de « S » et le signe ne réfère pas à quoi que ce soit : il ne joue pas le rôle d'un authentique nom de *S*.

La question qui se pose alors est celle de savoir si la signification revêt une dimension publique ou sociale<sup>886</sup> - soit si la distinction objective entre un usage correct et un usage incorrect suppose et résulte de l'interaction au sein d'une communauté de locuteurs. Ceux qui soutiennent le caractère public de la signification affirment que tout ce que Wittgenstein

<sup>884</sup> Wittgenstein, 2005, §258.

<sup>885</sup> *Ibid.*, §256.

<sup>886</sup> Nous suivons ici, dans les grandes lignes, Fennell, 2000 et Meredith, Me., 2000.

requiert pour une évaluation normative authentique de l'emploi d'un terme par un individu est que son usage puisse être visible et donc observable par les autres membres de la communauté, afin de se définir comme un usage public. Aussi, si les autres locuteurs arrivent à découvrir comment un individu emploie un terme ou, si le locuteur peut montrer à quelqu'un d'autre comment l'employer, alors ils garantissent que l'emploi de ce terme n'est pas arbitraire ou privé et qu'il ne va pas changer indéfiniment. Par conséquent, les partisans d'une approche publique de la signification soutiennent que Wittgenstein s'attaque seulement à la thèse d'un langage privé : soit à l'idée que les faits privés ou personnels de l'individu ne peuvent pas générer une distinction objective entre des usages corrects et des usages incorrects d'un terme ou, entre ce qui semble être le cas et ce qui est le cas quant à la question d'employer un terme de la même manière. Une approche publique de la signification est donc encore individualiste dans la mesure où elle affirme qu'un seul individu est susceptible, en employant un terme, d'être compris et de rendre compte de son usage aux autres locuteurs. Elle rejette certes la dimension privée mais non l'individualisme. Par conséquent, avancer la thèse selon laquelle c'est l'individu seul qui détermine la signification d'un terme c'est-à-dire, adopter un point de vue individualiste sur la signification, doit bien être distinguée de la thèse selon laquelle ce sont les faits purement privés de l'individu qui déterminent la signification des termes c'est-à-dire, préconiser le point de vue du langage privé sur la signification. Un point de vue individualiste, contrairement au point de vue du langage privé, soutient que la signification est un phénomène public, à savoir que, selon le point de vue individualiste, bien que l'individu dispose de toutes les ressources nécessaires lui permettant de déterminer l'application correcte d'un terme, c'est avant tout parce que ses ressources sont publiques et donc accessibles à l'observation et à la révision des autres locuteurs, que nous pouvons parler d'une distinction objective entre un usage correct et un usage incorrect d'un terme.

D'un autre côté, les partisans d'une approche sociale de la signification maintiennent que, pour qu'un individu emploie un terme de manière significative, il faut que l'usage de ce terme ne soit pas seulement visible et donc observable publiquement par les autres membres de la communauté mais qu'il se construise aussi par le biais d'une interaction partagée entre

les locuteurs soit d'une collaboration entre au moins deux individus<sup>887</sup>. Une approche sociale de la signification attaque l'individualisme que nous retrouvons tant dans les approches du langage privé que dans les approches publiques, ne se limitant pas seulement à la dimension solipsiste propre à la thèse du langage privé de la première approche. La constitution de la signification requiert la participation réelle des autres; celle-ci est immanente ou interne à la constitution même de la signification. Alors que selon une approche publique de la signification, la signification d'un individu est externe à celle d'un autre locuteur, l'autre locuteur étant un observateur extérieur qui doit seulement être capable de découvrir ce qu'un individu veut dire en employant les termes de notre langage sans être tenu de participer au jeu de production et de demande de raisons; selon une approche sociale, la participation de l'autre locuteur au sein de pratiques communes partagées avec l'individu est requise pour que l'individu tienne des propos sensés. Par conséquent, ces différentes conceptions de la signification ont des conséquences sur la théorie de la signification. Concevoir la signification comme un comportement publiquement observable suggère l'idée que nous pourrions rendre compte adéquatement de la signification à partir de la perspective désincarnée d'un observateur extérieur<sup>888</sup>. Mais, concevoir, à l'opposé, la signification comme immanente à la pratique d'une communauté et à l'interaction de ses membres, semble justifier une approche méthodologique encourageant la participation et l'engagement de tous ses membres. Autrement dit, préconiser une dimension publique à la théorie de la signification semble engendrer une approche externaliste de la signification privilégiant le point de vue d'un observateur extérieur, alors qu'une dimension sociale semble faire appel à une approche internaliste de la signification préconisant le point de vue du locuteur-acteur.

La différence entre conception sociale et publique de la signification implique différentes conceptions de la communication. Dans la mesure où une approche publique ne considère pas la participation des individus au sein d'une pratique commune comme

---

<sup>887</sup> Baker et Hacker, 1984b ainsi que Blackburn, 1984b, Chap.3, défendent une approche publique de la signification ; Wittgenstein, 2005, §139-142, exige également que cette dimension publique et interpersonnelle soit prise en compte dans l'élaboration de normes de signification objectivement correctes. Brandom, 2011, Kripke, 1996, et Wright 1981, défendent, de différentes manières, cette contrainte sociale portant sur la normativité de la signification. Kripke et Wright incarnent, une socialité forte (cf. Chap.III, §3.2.2, §3.3.2 et §3.3.3), alors que Brandom, McDowell et Williams, Me., préconisent une socialité faible. Établir cette distinction nous semble nécessaire pour ne pas tomber dans l'écueil du relativisme.

<sup>888</sup> Nous risquons de retomber ici dans le béhaviorisme naturaliste de Quine, 1977a.

essentielle à la signification, les individus peuvent alors disposer de leurs propres idiolectes respectifs (dont le caractère significatif exige seulement que ceux-ci soient ouverts en principe à l'approche publique). L'image que nous nous faisons est alors un peu celle d'un groupe d'individus usant de ses propres idiolectes de manière autonome ; la communication et la compréhension ayant lieu entre au moins deux individus quand il est possible d'établir un recouplement entre l'idiolecte de l'un et celui de l'autre. Selon une conception publique de la communication comme « recouplement d'idiolectes », ce phénomène peut *occasionnellement* se produire, et c'est dans cette mesure que la signification s'installe. Cependant, il n'est absolument pas nécessaire qu'un recouplement ou qu'une convergence des idiolectes se produise puisque, comme nous l'avons vu dans le cadre d'une approche publique, un individu peut formuler des propos sensés sans nécessairement avoir besoin de l'implication réelle des autres. Cependant, selon une approche sociale, un idiolecte dépend en premier lieu de l'engagement réel des autres à tenir des propos sensés ; aussi, plutôt que de rendre compte de la communication comme un « recouplement d'idiolectes », nous privilégierons l'idée d'« interdépendance<sup>889</sup> ». La communication émerge donc à partir d'une pratique commune des locuteurs cherchant à se mettre d'accord, à se corriger et à réviser leurs manières d'employer les termes, etc., et c'est cette pratique commune qui joue ainsi un rôle constitutif dans la constitution de la signification de tout idiolecte individuel. La signification de l'idiolecte d'un individu se construit donc au sein d'une pratique partagée avec les autres membres de la communauté et c'est cette appartenance à une communauté au sens large qui est le point de départ de toute communication réussie. Il ne s'agit donc pas d'attendre une corrélation accidentelle et contingente de nos idiolectes respectifs jusqu'ici isolés. Une approche sociale ne suppose donc pas des idiolectes parfaitement autonomes qui n'auraient de sens qu'au sein d'un usage local et clos sur lui-même mais présente davantage la communication comme une activité commune ou partagée qui progresse au terme d'un processus d'instruction, de discussion, de correction, de déférence<sup>890</sup>, etc.

---

<sup>889</sup> Les qualificatifs de « overlapping idiolects » et de « interdependency » qui caractérisent respectivement les conceptions sociales et publiques de la communication sont reprises de Bilgrami, 1995, p.69-71; et Blackburn, 1984b, p.90.

<sup>890</sup> Le sujet délègue en fait implicitement cette connaissance à d'autres membres de la communauté, parfois à certains experts, qui eux connaissent la nature du référent en question.

Nous avons ainsi tenté jusqu'ici de préciser les différences entre une approche publique et une approche sociale de la signification et de la communication. La question est maintenant de déterminer celle à laquelle Davidson donne son assentiment : privilégie-t-il davantage une approche sociale ou une approche publique de la signification<sup>891</sup> ? La réponse de Davidson ne nous paraît pas toujours très claire : sa théorie de l'interprétation radicale ainsi que son refus de recourir à la notion de conventions de significations partagées dans le processus de compréhension et de communication laisseraient croire qu'il adhère à une approche publique de la signification, alors que le rôle constitutif qu'il accorde à la charité, celle-ci impliquant une forte présomption de croyances et de significations partagées, le rapproche davantage d'une approche sociale. L'insistance de Davidson sur l'aspect essentiellement social du langage ne résout pas la question puisque, précisément dans certains de ces contextes, l'emploi qu'il fait du terme *social* s'apparente simplement au terme *public*, alors que dans d'autres contextes, ce même terme est entendu dans un sens plus spécifiquement social. Il semblerait ainsi osciller entre, soit accorder un sens exclusivement public soit, accorder un sens authentiquement social<sup>892</sup>. Or, soutenir en même temps une approche exclusivement publique tout en rejetant toute dimension sociale, nous paraît incohérent<sup>893</sup>.

<sup>891</sup> Davidson, 2001b, 1990; 1990a; 2005, 1994 et 2001, 1992.

<sup>892</sup> Une sélection de passages souligne le fossé entre cette dimension publique ou sociale d'une théorie de la signification, comme suit : dans Davidson, 1990a, p.314 : « Language is *social*. This does not entail that truth and meaning can be defined in terms of observable behavior, or that it is 'nothing but' observable behavior; but it does imply that meaning is entirely determined by observable behavior, even readily observable behavior. That meaning is decipherable is not a matter of luck; *public* availability is a constitutive aspect of language. » : nos italiques. Davidson, 2005, 1994, p.124-125 : « Agreeing with Dummett and Kripke, and perhaps with Wittgenstein, I hold that the answer to the question what it is to go on as before demands reference to social interaction [...] There must be an interacting group for meaning – even propositional thought, I would say – to emerge...It follows that meaning something requires that by and large one follow a practice of one's own, a practice that *can* be understood by others. But there is no fundamental reason why practices must be shared. ». Nos italiques. Notons que dans les deux citations, ce qui débute en faisant référence à l'aspect social et partagé de la signification se traduit à la fin en termes proprement publiques de révision, d'observation et de vérification de la signification du locuteur par les autres membres de la communauté – soit l'exigence que la signification doit d'abord être une activité qui « doit être comprise par les autres » se substitue à l'idée que celle-ci doit être partagée par les autres locuteurs.

<sup>893</sup> Ainsi, alors qu'il n'y a maintenant rien d'incohérent à soutenir que la signification est à la fois publique et sociale, cela peut signifier que la signification, en plus d'être identifiable publiquement aux autres doit aussi se construire par le biais d'une réelle participation au sein d'une pratique commune et partagée, il semble de toute évidence incohérent d'affirmer que la signification est publique mais dénuée de toute dimension sociale (ce qui est bien ce que nous entendons par « exclusivement publique ») et de soutenir que la signification est également sociale.

Nous avancerons alors, qu'indépendamment de la question de savoir quelle approche nous devons, à juste titre, attribuer à Davidson, qu'une approche exclusivement publique de la normativité ne nous semble pas appropriée et qu'une approche sociale s'avère plus cohérente. Nous appuyons d'ailleurs ce point en examinant l'approche publique de la normativité de la signification de Davidson et en démontrant que le succès de celle-ci implique de prendre en considération la dimension sociale de la signification. La dimension publique qu'il confère pourtant à la signification semble s'estomper d'elle-même en laissant entrevoir l'importance d'un point de vue social. Davidson nous met d'ailleurs en garde contre un point de vue transcendant, à partir duquel il serait possible à un être isolé d'évaluer ses impressions sur le monde en les contrastant avec le monde tel qu'il est :

« We cannot occupy a position outside our own minds; there is no vantage point from which to compare our beliefs with what we take our beliefs to be about<sup>894</sup>. ».

Il adhère également à l'argument contre l'hypothèse d'un langage privé<sup>895</sup> (une langue qui ne peut être employée que par une seule personne) défendue par Wittgenstein selon laquelle une langue utilisée par une seule personne ne pourrait jamais être contrôlée pour déterminer si elle représente correctement les pensées ou non, puisque son contenu ne pourrait jamais être partagé. Nous reviendrons sur cet aspect social du langage par la suite. Nous présenterons d'abord le point de vue de Davidson en faveur d'une approche publique de la signification.

#### 5.6.1. Davidson : une approche publique de la signification

Davidson ne nie pas que l'interprétation et la communication prennent nécessairement place dans un contexte d'activités humaines sociales. Il affirme que nous ne pouvons pas atteindre la compréhension des autres et déchiffrer la signification de leurs énonciations en dehors de nos pratiques d'interprétation. Par conséquent, son problème est fondationnel : il s'agit de savoir si nous pouvons parvenir à une conception de la signification, qu'elle soit exclusivement publique ou sociale, à partir des notions comme celles de « convention », de « pratique » ou d' « activités humaines ». Sa réponse est, comme nous allons le voir, négative. Davidson semble ainsi préconiser qu'une approche déréifiée de la signification

<sup>894</sup> Davidson, 2004, 1995, p.7.

<sup>895</sup> Davidson, 2001b, 1992, traite également du sujet de l'impossibilité d'une langue privée.

devrait respecter la théorie de l'interprétation radicale afin de se prémunir contre toute pétition de principe sur le terrain épistémologique : nous ne pouvons en effet pas commencer par présupposer que nous connaissons la signification du langage d'autrui (et partir des notions de « convention ») en élaborant notre théorie de la signification puisque le but de toute théorie de la signification est, précisément, de déterminer en bout de ligne comment nous procédons dans notre interprétation du langage d'autrui. Nous remarquons ainsi que la théorie de l'interprétation radicale encourage l'approche d'un observateur extérieur dans toute construction d'une théorie de la signification ; tout ce qu'il y a à dire sur la signification doit être révélé à partir de la perspective impartiale d'un observateur extérieur ; toute prise de position de la part de l'interprète dans la constitution de la signification du locuteur qu'il tente d'interpréter risquera de présupposer ce qui est précisément en jeu<sup>896</sup>. Par conséquent, selon le modèle de l'interprétation radicale, nous avons l'image de deux individus utilisant et confrontant chacun sur la scène publique leur propre idiolecte (dans l'interprétation radicale, les idiolectes sont bien les premiers objets d'étude à la base de toute théorie de la signification). Nous disposons finalement d'un locuteur dont l'idiolecte (celui du locuteur étranger) semble pouvoir lui permettre de tenir des propos sensés ; propos, comportements et réactions qu'un autre locuteur observe de l'extérieur soit l'interprète radical (doté de son propre idiolecte porteur de sens) qui tente de comprendre le locuteur étranger en reliant de manière systématique la signification associée par celui-ci aux termes qu'il emploie avec une phrase équivalente dans sa propre langue. L'interprétation radicale arbore ainsi toutes les caractéristiques d'une approche publique de la signification et de la compréhension : des idiolectes indépendants et doués de sens, la perspective d'un observateur extérieur pour établir la signification et la compréhension des termes, et une approche de la communication en termes de corrélation ou de recoupement entre idiolectes soit de production de phrases équivalentes.

De plus, Davidson semble bien appuyer l'approche publique de la signification en soutenant explicitement qu'il peut y avoir communication par le langage sans faire appel aux conventions<sup>897</sup>. Bien qu'il ne nie pas l'existence de conventions, Davidson ne croit pas

<sup>896</sup> Cf. Davidson, 1993b, 1976 : « Tout comme Lear gagne en pouvoir en l'absence de Cordelia, je pense que les traitements du langage prospèrent quand ils évitent l'évocation non critique des concepts de conventions, de règle linguistique, de pratique linguistique, ou de jeux de langage. ».

<sup>897</sup> Davidson, 2005, 1986, 2005, 1994.

qu'elles soient nécessaires à la détermination de la signification des énonciations et par conséquent à l'existence de la communication par le langage. En rejetant les conventions, il écarte l'idée qu'un partage préalable des significations soit nécessaire pour rendre possible la communication. Autrement dit, Davidson écarte l'idée que l'interprétation du discours, dans la communication courante, repose en quoi que ce soit sur la connaissance d'un « langage » ou d'une compétence générale de l'interprète, au sens : 1/ d'un ensemble de significations systématiques (déterminées compositionnellement), 2/ partagées par l'interprète et l'interprété, et 3/ apprises et conventionnelles.

Il tente ainsi de justifier sa position en rendant compte de l'interprétation des *pataquès*<sup>898</sup> c'est-à-dire, de l'idée que nous pouvons comprendre et communiquer avec un autre locuteur qui emploie pourtant les termes d'une manière non conforme aux standards ou aux significations communément ou conventionnellement établies. Davidson cite ainsi une magnifique expression non littérale d'une Madame Malaprop en conversation dans l'une des pièces de Sheridan au dix-huitième siècle<sup>899</sup>. Davidson avance que l'explication que nous devons fournir pour parvenir à comprendre ce qui se passe dans le cas de Mme Malaprop peut être généralisée à tous les cas de communication et cette explication fait l'économie de la conception conventionnelle du langage. Mme Malaprop énonce la phrase suivante : « Voilà un beau dérangement d'épithètes » mais elle utilise « dérangement » pour vouloir dire ce que nous signifions habituellement par « arrangement »; c'est-à-dire qu'elle ne respecte donc pas, en employant le terme « dérangement » la règle conventionnelle régissant la signification de ce terme, et qu'elle suit en employant le terme « épithète » la signification communément admise pour le terme « épithète » (plutôt que la règle conventionnelle qui s'applique au terme « épithètes »). Nous pouvons bien parvenir à comprendre que la signification de son énonciation est donnée par la signification de la phrase « c'est un bel arrangement d'épithètes », qui n'a rien à voir avec la signification de la phrase effectivement

<sup>898</sup> La traduction de *Malapropism* par pataquès est proposée par Engel, 1994.

<sup>899</sup> Davidson, 2005, 1986, p.103-107. Dans le même ouvrage, voir aussi sur cet article les commentaires de Hacking, 1986, p. 447-458, et de Dummett, 1986, p. 459-476. Le titre original de Davidson fait allusion au personnage du dramaturge et politicien anglais du XVIIIe, Richard Sheridan, Mme Malaprop, et à ce que le *SOED* appelle son « usage impropre et ridicule des mots, spécialement lors de la confusion de deux mots qui se ressemblent » dans la pièce de Sheridan de 1775, *Les Rivaux*. Nous pourrions peut-être dire que, dans la pièce, le personnage fictif, Mme Malaprop, a ridiculement mal usé des mots, confondant « dérangement » et « arrangement », « épithètes » et « épithètes ». Elle change ainsi la phrase « un bel arrangement d'épithètes » en « un beau dérangement d'épithètes ».

prononcée par Mme Malaprop. Mais, s'il y a des cas dans lesquels il est clair qu'un locuteur diverge involontairement de la norme, comment le déterminer dans tous les cas (le député qui pendant la cérémonie d'intronisation du roi crie « Vive la République ! » peut confondre les institutions, faire un *pataquès* non intentionnel, faire « intentionnellement » une erreur, avoir l'intention de prononcer les mots en leur sens littéral et faire reconnaître cette intention aux auditeurs, etc.) ? Mais, une chose est de dire que la plupart des croyances sur lesquelles porte l'interprétation doivent être vraies pour que l'erreur soit possible, autre chose est de dire que l'on peut déterminer exactement quand les locuteurs se trompent sur le sens des mots ou les emploient de façon non conventionnelle, et quand leurs croyances sont erronées plutôt que leurs significations<sup>900</sup>.

Le *pataquès* est par conséquent précisément le cas d'un usage déviant du langage dans lequel locuteurs et auditeurs ne disposent pas d'un ensemble préalable commun de normes de signification ou de conventions linguistiques sans que cela n'entrave le succès de la communication et de la compréhension. Davidson en conclut que la compréhension ou la communication ne doit pas nécessairement reposer sur un arrière-plan de significations ou de conventions (ou forme de vie, tradition, culture, etc.) préalablement partagées. Le cas du *pataquès* souligne ainsi une conception différente de la signification et de la compréhension : celle d'une conception publique de deux idiolectes parfaitement autonomes et doués de signification d'un locuteur et d'un interprète (un auditeur, un lecteur) où la communication s'installe quand tous deux réussissent à interpréter mutuellement leurs énonciations c'est-à-dire à partir du moment où ils peuvent établir une corrélation entre leurs idiolectes. Davidson se demande « how people who already have a language (whatever exactly that means) manage to apply their skill or knowledge to actual cases of interpretation<sup>901</sup> ». Il propose que le locuteur et l'interprète arrivent à une rencontre linguistique armés d'une *théorie antérieure*. La théorie antérieure de l'interprète inclut toute la connaissance gagnée des rencontres précédentes, raffinée pour l'occasion par ce qu'elle peut remarquer de l'aspect du locuteur devant elle (« knowledge of the character, dress, role, sex, of the speaker, and whatever else has been gained by observing the speaker's behaviour, linguistic or

<sup>900</sup> Engel, 1994, p.136, soulève ce point.

<sup>901</sup> Davidson, 2005, 1986, p.100.

otherwise<sup>902</sup> »). Pendant que le locuteur parle, l'interprète modifie sa théorie initiale en même temps qu'il dégage du comportement du locuteur de nouvelles informations à son sujet. L'interprète peut élargir ou restreindre les possibilités interprétatives pour certains mots, ou remettre en cause des paradigmes entiers de stratégies établies lors des communications antérieures. La nouvelle théorie, celle qui permet à l'interprète de comprendre le sens de l'énonciation en question, est la « théorie passagère » pour cette occasion. Davidson distingue ainsi la théorie antérieure (*prior theory*) de l'interprète, qui exprime comment celui-ci est préparé à interpréter les phrases d'un locuteur particulier notamment en comparant son usage avec celui des membres de sa communauté de la théorie transitoire ou passagère (*passing theory*), qui exprime comment un énoncé particulier de ce locuteur est de fait interprété. Le locuteur suit un chemin d'accès semblable pour se faire comprendre:

« Let's look at the process [of communication] from the speaker's side. The speaker wants to be understood, so he utters words he believes can and will be interpreted in a certain way. In order to judge how he will be interpreted, he forms, or uses, a picture of the interpreter's readiness to interpret along certain lines. Central to this picture is what the speaker believes is the starting theory of interpretation the interpreter has for him. The speaker does not necessarily speak in such a way as to prompt the interpreter to apply this prior theory; he may deliberately dispose the interpreter to modify his prior theory. But the speaker's view of the interpreter's prior theory is not irrelevant to what he says, nor to what he means by his words; it is an important part of what he has to go on if he wants to be understood<sup>903</sup>. ».

Si cette description est exacte, c'est cette théorie passagère qui doit être partagée pour que la communication réussisse, et par conséquent celle-ci satisfait à la condition 2/. Mais la théorie passagère n'est connue ni de l'interprète ni du locuteur avant leur rencontre: chaque théorie passagère est unique, adaptée à une seule occasion de communication qui ne se reproduira plus jamais dans exactement la même configuration. Si le locuteur et son interprète partagent une connaissance, elle doit être un produit de leur rencontre, car la connaissance qu'ils possèdent au départ de leur rencontre n'est pas partagée:

---

<sup>902</sup> Ibid., p. 100.

<sup>903</sup> Ibid., p.101.

« The passing theory is where, accident aside, agreement is greatest. As speaker and interpreter talk, their prior theories become more alike; so do their passing theories. The asymptote of agreement and understanding is reached when passing theories coincide. But the passing theory cannot in general correspond to an interpreter's linguistic competence. Not only does it have its changing list of proper names and gerrymandered vocabulary, but it includes every successful - i.e. correctly interpreted - use of any other word or phrase, no matter how far out of the ordinary. Every deviation from ordinary usage, as long as it is agreed on for the moment (knowingly deviant, or not, on one, or both, sides), is in the passing theory as a feature of what the words mean on that occasion. Such meanings, transient though they may be, are *literal*; they are what I have called *first meanings*. A passing theory is not a theory of what anyone (except perhaps a philosopher) would call an actual natural language. « Mastery » of such a language would be useless, since knowing a passing theory is only knowing how to interpret a particular utterance on a particular occasion. Nor could such a language, if we want to call it that, be said to have been learned, or to be governed by conventions. Of course things previously learned were essential to arriving at the passing theory, but what was learned could not have been the passing theory.

Why should a passing theory be called a theory at all? For the sort of theory we have in mind is, in its formal structure, suited to be the theory for an entire language, even though its expected field of application is vanishingly small. The answer is that when a word or phrase temporarily or locally takes over the role of some other word or phrase (as treated in a prior theory, perhaps), the entire burden of that role, with all its implications for logical relations to other words, phrases, and sentences, must be carried doing by the passing theory<sup>904</sup>. ».

Davidson pense ainsi que nous pouvons effectivement parvenir à établir des corrélations en faisant seulement usage de « théories passagères » utilisées instrumentalement, sans être obligé de postuler l'existence de règles conventionnelles. Bien que les conventions et les régularités interprétatives existent et puissent aider (ou gêner) un effort communicatif particulier, elles ne sont pas l'élément qui assurera le succès de la communication, ni peuvent-elles par leur absence être responsables de son échec :

---

<sup>904</sup> Ibid., p.102-103: nos italiques.

« Neither the prior theory nor the passing theory describes what we would call the language a person knows, and neither theory characterizes a speaker's or interpreter's linguistic competence. Is there any theory that would do better? [...] We could hold that any theory on which a speaker and interpreter converge is a language; but then there would be a new language for every unexpected turn in the conversation, and languages could not be learned and no one would want to master most of them. [...] We may say that linguistic ability is the ability to converge on a passing theory from time to time - this is what I have suggested, and I have no better proposal. But if we do say this, then we should realize that we have abandoned not only the ordinary notion of a language, but *we* have erased the boundary between knowing a language and knowing our way around in the world generally. For there are no rules for arriving at passing theories, no rules in any strict sense, as opposed to rough maxims and methodological generalities. A passing theory really like a theory at least in this, that it is derived by wit, luck, and wisdom from a private vocabulary and grammar, knowledge of the ways people get their point across, and rules of thumb for *figuring* out what deviations from the dictionary are most likely. There is no more chance of regularizing, or teaching, this process than there is of regularizing or teaching the process of creating new theories to cope with new data in any field-for that is what this process involves<sup>905</sup>. ».

Dans l'exemple mentionné, la théorie antérieure de Mme Malaprop et sa théorie transitoire sont que « un joli dérangement des épitaphes » signifie « un joli dérangement des épithètes » (les deux théories coïncident, parce qu'elle fait une erreur sur les sens des mots tels que devrait le comprendre, selon elle, l'auditeur). La théorie antérieure de l'auditeur est que « un joli dérangement des épitaphes » signifie « un joli dérangement des épitaphes » qui est le sens qu'il donne à ces mots avant d'interpréter ceux de Mme Malaprop, et sa théorie transitoire s'accorde avec celles de Mme Malaprop. Nous voyons donc qu'en aucun sens la communication ou le succès de l'interprétation ne peuvent reposer sur la capacité à comprendre un « langage » au sens de 1/-3/. Il est toujours possible de parler ici d'une aptitude linguistique, au sens de « l'aptitude à converger sur une théorie passagère de temps en temps », mais la possession de cette aptitude ne consiste pas en « des règles qu'on utiliserait pour parvenir à des théories passagères, par opposition à des maximes approximatives et des généralisations méthodiques » mais est affaire de « jugeote, de chance, et de sagesse empruntée à un vocabulaire et une grammaire privée, de connaissance des manières dont les gens s'expriment, et des règles grossières permettant d'imaginer les déviations à partir des dictionnaires les plus probables<sup>906</sup> ». Davidson affirme alors que la signification du locuteur n'est rien de plus que « la théorie » à laquelle nous nous sommes

<sup>905</sup> Ibid., p.104-107: nos italiques.

<sup>906</sup> Ibid, p.107.

provisoirement ajustés : le terme « dérangement » dans l'idiolecte de Mme Malaprop est satisfait sous les mêmes conditions que le mot « arrangement » l'est sous nos conditions; la même analyse s'appliquant pareillement au terme « épithaphe » employé par Mme Malaprop et à notre emploi du terme épithète. En outre, la conception de la signification et de la communication développée par l'étude des cas de pataquès adhère à l'argument de Wittgenstein en écartant la possibilité d'un langage privé<sup>907</sup>. C'est bien parce que la théorie de l'interprétation radicale présuppose l'interaction avec autrui qu'un locuteur ne peut avoir de pensées que s'il a été l'interprète du langage d'une autre personne et que son langage a lui aussi fait l'objet d'une interprétation.

En partant de son analyse des pataquès, Davidson montre bien que, pour lui, la communication n'est pas une question de conventions partagées. Il soutient d'ailleurs, comme nous l'avons déjà noté, que communiquer ou arriver à être compris est bien le but ultime du langage<sup>908</sup> et que comprendre un langage, c'est finalement souscrire à certaines normes en vertu desquelles ce langage est déjà compris. Par conséquent, Davidson ne croit pas que les conventions sont essentielles mais ne préconise pas davantage un retour à la thèse d'un langage privé; il semble donc bien engagé à défendre une approche publique de la signification et de la communication. Mais, si, dans la perspective de Davidson, nous n'avons pas besoin pour communiquer de normes socialement instituées et constituées, devons-nous pour autant en conclure que la dimension sociale n'est pas essentielle à la détermination de la signification ?

---

<sup>907</sup> Davidson, 2002, 1994, p. 113, approuve explicitement la validité de l'argument de Wittgenstein contre la possibilité d'un langage privé : « [W]e [i.e. himself and Dummett] both insist that verbal behaviour is necessarily social...Perhaps we even agree on the reason, namely Wittgenstein's, that without a social environment nothing could count as misapplying words in the way that it must be possible to misapply word in speech ».

<sup>908</sup> Il semble que Brandom, 2000a, p. 363, ne partage pas ce point de vue. Nous y reviendrons par la suite.

« Of course I don't mean that there is no reason why we are taught, and why we learn, to speak more or less as others around us do. Nothing could be more obvious: we want to be understood and others have an interest in understanding us; ease of communication is vastly promoted by such sharing...None of this creates a free-standing obligation, however. Any obligation we owe to conformity is contingent on the desire to be understood. If we make ourselves understood while deviating from the social norm, any further obligation has nothing to do with meaning or successful communication...It is absurd to be obligated to a language; so far as the point of language is concerned, our only «obligation», if that is the word, is to speak in a way as to accomplish our purpose by being understood as we expect and intend. It is an accident, though a likely one, if this requires that we speak as others in our community do...The challenge is to draw the distinction Wittgenstein has made central to the study of meaning, the distinction between using words correctly and merely thinking one is using them correctly, without appeal the test of common usage...the challenge can only be met in a social setting. What is needed is a norm, something that provides the speaker with a way of telling (not necessarily always) that he has gone wrong...Speaking in accord with socially accepted usage is such a norm, but one which, I have argued, is irrelevant to communication unless the audience of the speaker happens to speak as he does, in which case the norm is relevant not *because* it is a shared practice or convention, but because conforming to it results in understanding. My proposal takes off from this observation: what matters, the point of language or speech or whatever you want to call, is communication, getting across to someone else what you have in mind by means of words that they interpret (understand) as you want them to. Speech has endless other purposes, but none underlies this one: it is not an ultimate or universal purpose of speech to say what one thinks is true, nor to speak as one thinks others do<sup>909</sup>. ».

Il n'est donc pas requis pour être compris de parler comme les autres locuteurs, c'est-à-dire de parler en se conformant à l'usage social communément reconnu et accepté, même si partager les mêmes conventions facilite de toute évidence la compréhension et la communication. Le problème est alors celui d'une priorité conceptuelle – selon Davidson, être compris est d'une importance primordiale mais ne nécessite pas le respect de normes socialement constituées et partagées. Autrement dit, il se peut que nous parlions en nous conformant à des significations socialement partagées mais seulement parce qu'agir ainsi nous est plus utile et nous permet d'être compris. Il est donc inexact de dire que nous réussissons seulement à être compris parce que nous partageons et déterminons socialement les significations de nos termes. Les mêmes mots peuvent revêtir des significations différentes selon les locuteurs et les auditeurs sans que cela ne fasse obstacle à la compréhension mutuelle de leurs énonciations (pataquès); la seule contrainte requise pour la signification exige, pour garantir la communication, que nous puissions constater une certaine

<sup>909</sup> Davidson, 2005, 1994, p.118-120.

régularité ou cohérence dans l'usage de l'individu de telle sorte que celle-ci soit susceptible d'être observée publiquement, c'est-à-dire d'être visible aux autres locuteurs. De plus, il n'est d'ailleurs pas requis par la suite qu'un locuteur présente la même régularité ou cohérence dans l'usage et lors de son interaction avec autrui, soit qu'il dispose de significations socialement partagées.

#### 5.6.2. Davidson : une approche sociale de la signification

Par conséquent, comme nous l'avons précédemment montré, de solides arguments semblent, selon Davidson, privilégier une approche publique de la signification. Néanmoins, il n'en reste pas moins qu'une approche sociale (perçue comme opposée à cette dimension publique) de la signification et de la communication vient se juxtaposer à cette première approche sans que la cohérence soit évidente entre les deux approches. Dans « *The Social Aspect of Language* », Davidson refuse que des significations constituées au préalable par des normes sociales soient essentielles à la compréhension et à la communication, et pourtant comme son titre l'indique, il semble néanmoins penser que le langage ait une dimension sociale fondamentale. Nous en venons ainsi à nous poser la question suivante : qu'entend-il donc par cette dimension sociale, ou que reste-t-il de cette dimension sociale une fois que nous écartons l'idée de significations préalablement partagées ? Pouvons-nous seulement dire que, par le terme « social » Davidson veut en fait dire « public », c'est-à-dire qu'il commettrait volontairement un *pataquès* et en rester là ? Cette réponse nous semble bien trop facile, dans la mesure où nous pouvons difficilement ignorer d'autres éléments de sa théorie, tel que le rôle constitutif qu'il accorde à la charité, celui-ci exigeant avant tout une lecture plus authentiquement sociale du terme « social ».

Davidson utilise ainsi, comme nous l'avons vu, la théorie de l'interprétation radicale pour identifier la nature de la signification linguistique et pour mettre en lumière le caractère étroitement interdépendant des concepts de signification, des concepts d'attitudes propositionnelles, et des normes de rationalité, mais en raison de leur structure « holistique », soit de l'interdépendance de la croyance et de la signification, l'interprétation radicale est confrontée à une impasse. Ainsi, en vertu du principe holistique essentiel à la théorie interprétative de Davidson, nous ne pouvons pas avoir (se voir attribuer) une croyance sans en avoir (s'en voir attribuer) plusieurs à la fois, inférentiellement liées à la première selon des

critères de cohérence logique, permettant à l'interprète d'attribuer ou de reconnaître une intelligibilité au comportement du sujet<sup>910</sup>. Par conséquent, nous ne pouvons avoir des pensées sans avoir des croyances et nous ne pouvons avoir de croyances sans comprendre son contenu propositionnel. Or, comprendre un contenu propositionnel, c'est connaître ses conditions de vérité, un locuteur ne peut donc avoir de croyances que s'il comprend que celles-ci peuvent être vraies ou fausses. Or, tous les faits pertinents pour déterminer les contenus de croyance et la signification des phrases (ses conditions de vérité), ce qui pour Davidson consiste dans le comportement du locuteur à « tenir pour vrai » certaines phrases, sous-déterminent radicalement ces deux variables. Le principe de charité est ainsi introduit afin de surmonter cette impasse. Nous avons d'ailleurs distingué une formulation étroite ou littérale d'une version plus large du principe de charité. Une version étroite du principe de charité nous prescrit d'interpréter toujours de manière à rendre les croyances de ceux que l'on interprète cohérentes et non contradictoires. Quant à une formulation large du principe de charité, celui-ci prescrit d'attribuer à l'interprété non pas ce qu'il serait, en général, et pour tout individu possible, rationnel et correct de croire, mais lui prescrit d'attribuer ce qu'il est correct ou raisonnable de croire, pour lui interprète. En ce cas, nous nous figurons que l'interprète essaie d'attribuer à l'interprété le maximum de croyances empiriques similaires aux siennes, en projetant son univers doxastique et sa propre psychologie sur celui de l'interprété – par exemple, que l'interprété entretient des croyances empiriques que l'interprète tient pour vraies telles que la neige est blanche, l'herbe est verte, etc. Davidson pose donc *via* le principe de charité cette présomption de rationalité comme une condition nécessaire *a priori* de l'attribution de la croyance et de la signification, celle-ci n'impliquant pas que l'attribution de toutes croyances logiques ou non logiques puisse se soustraire à toute vérification ou révision mais souligne par contre une forte présomption pour que de telles attributions soient vraies. Autrement dit, le cas d'un usage « déviant » ne doit pas être la règle. Cette position est parfaitement décrite dans ce passage de Davidson:

---

<sup>910</sup> Davidson, 1991a, p.66.

« We know in a way no one else can what we believe, fear, want, value, and intend. We know how things seem to us, how they look to us, feel to us, smell and sound to us to be. We know these things in a way we can never know about the world around us. *Whether or not we are sometimes wrong about the contents of our minds, whether or not we can be in doubt about our own sensations and thoughts, one thing is certainly true of such beliefs: they cannot be generally mistaken. If we think we have a certain thought or sensation, there is a strong presumption that we are right*<sup>911</sup>. ».

Par conséquent, la dimension constitutive du principe de charité repose sur deux fortes présomptions interprétatives : l'interprète ne pourra assigner des significations aux phrases de l'interprété que s'il lui assigne des désirs et des croyances et que s'il présume que celui-ci partage, en grande partie, les mêmes croyances logiques et non logiques. Cette assignation se fera sur la base des propres croyances de l'interprète qu'il supposera largement similaires aux siennes propres, et dont il cherchera à « maximiser » l'accord et « minimiser le désaccord » avec celles de l'interprété. Il repèrera ainsi chez l'interprété les actes de jugement ou de « tenir-pour-vrai » de certaines phrases, et sur la base de son propre assentiment aux phrases qu'il tient lui-même pour vraies, et des événements qui peuvent causer ces croyances et sera conduit de proche en proche à déterminer les croyances et les significations, de l'interprété. Nous entretenons et partageons donc, comme le soutient Davidson, des significations et croyances communes qui nous permettent de nous comprendre et de communiquer. La signification de certains termes appartenant tant au vocabulaire logique qu'au vocabulaire non logique semble alors bien socialement constituée<sup>912</sup>. Nous comprenons dès lors pourquoi,

<sup>911</sup> Davidson, 2001b, 1990, p. 193 : nos italiques.

<sup>912</sup> On pourrait bien être tenté d'affirmer que si nous utilisons leur terme « Gavagai » pour vouloir dire « Voici un lapin » alors nous partageons des croyances mais non des conventions linguistiques dans la mesure où nous ne disposons pas, contrairement à ces locuteurs étrangers, du terme de « Gavagai » dans notre langue. Notre convention est que « 'lapin' veut dire lapin » et leur convention est que « 'Gavagai' veut dire lapin », et nous ne partageons effectivement pas les significations (des termes) parce que nous ne partageons pas les termes. Cependant, nous pensons que lorsque nous sommes contraints de leur attribuer les mêmes croyances que nous, nous partageons les significations c'est-à-dire que nous partageons les mêmes concepts. L'expression « significations partagées » ne veut pas dire que ces locuteurs étrangers utilisent le terme « lapin » et que ce terme signifie ce que nous voulons dire par notre terme « lapin », mais plutôt qu'ils utilisent un terme (soit « Gavagai ») qui a la même signification que notre terme « lapin ». Par conséquent, nous considérons que la phrase-T « 'Gavagai' » est vraie si et seulement si voici un « lapin » constitue bien une signification partagée c'est-à-dire, que cette phrase énonce bien qu'un terme dans la langue des locuteurs étrangers veut dire la même chose que le terme « lapin » dans notre langue, soit que nous partageons le concept de lapin. Si la signification est déréifiée (c.-à-d. que quelle que soit la signification de leurs termes, celle-ci doit être déterminée à partir de ce que nous pouvons connaître de ces locuteurs étrangers en appliquant la méthode de l'interprétation radicale), et si une telle connaissance exige que nous supposions que les

selon Davidson, il n'y a aucune situation de communication qui puisse nous amener à supposer non seulement que les autres pourraient « voir » la réalité comme totalement différente de la manière dont nous la voyons, mais encore nous empêcher de supposer que la réalité à propos de laquelle nous communiquons est une réalité objective qui nous est commune. Le rôle constitutif du principe de charité implique que nous partagions un certain ensemble de croyances et de normes de signification, sujettes certes à révision et à changement, afin d'arriver à se comprendre. L'incohérence quant à la question de savoir si le but de la compréhension requiert une conception publique ou sociale de la signification semble alors provenir de cette tension entre l'interprétation radicale et le principe de charité. D'une part, l'interprétation radicale décrit la signification comme publique et intelligible, deux attributs essentiels selon Davidson, à savoir que deux locuteurs pourraient très bien interpréter mutuellement leurs énonciations sans qu'il y ait, en un sens quelconque, un langage commun qu'ils partagent. Autrement dit, le fait de ne pas devoir partager nécessairement des significations communes n'est pas incompatible avec l'idée d'une pluralité d'idiolectes. D'autre part, le rôle constitutif de la charité dans la pratique interprétative exige que la signification ait une dimension sociale, et que la compréhension, comme activité sociale s'insérant dans un jeu de conventions et de règles, implique que nous partagions au préalable certaines significations (un sociolecte) entre locuteurs (interprétés) et auditeurs (interprètes).

À cet effet, un autre élément que Davidson introduit dans sa théorie de l'interprétation radicale et donc dans son approche de la signification est cette dimension triadique ou « triple interaction » de la situation interprétative. Celle-ci ajoute, comme nous allons le voir, à la dimension publique une dimension sociale incontournable. La triangulation requiert une interaction entre un locuteur et un interprète ou plus précisément entre un locuteur, un interprète et un objet, événement ou situation dans le monde. Aussi, tout comme la connaissance du monde dépend de la connaissance d'autres esprits, la connaissance de notre esprit dépend du fait que nous connaissons d'autres esprits et des objets dans le monde. En d'autres termes, c'est l'interaction tripartite de la triangulation qui sert à identifier les objets

---

significations de leurs termes empiriques soient identiques aux nôtres, alors nous pourrions suggérer que le fait qu'ils partagent certains concepts empiriques avec nous joue bien un rôle constitutif dans leur signification des termes. Ceci revient finalement à dire que certains de leurs concepts non logiques sont fortement normatifs.

de la croyance exprimée dans l'énoncé. Par conséquent, c'est la triangulation qui donne un contenu aux pensées et aux énonciations<sup>913</sup>. Sans interaction interpersonnelle dans un monde partagé contenant des objets communs, il ne pourrait y avoir un contenu mental :

Le problème n'est pas [...] celui de vérifier à quels objets une créature réagit ; le problème est que sans une deuxième créature pour interagir avec la première, il ne peut y avoir de réponse à cette question. Et s'il ne peut y avoir de réponse à la question de savoir ce qu'une créature signifie, désire, croit ou vise, il n'y a pas de sens à soutenir que cette créature a des pensées<sup>914</sup>.

La triangulation ne contribue pas tant à résoudre les incohérences dans l'approche de Davidson d'une conception publique et sociale de la signification; elle ajoute à la question l'idée que les conditions de possibilité de la pensée et du discours sont aussi bien causales que sociales et linguistiques. Aussi, en analysant le concept de la triangulation dans une pratique d'interprétation, nous tenterons de dégager un argument en faveur d'une constitution sociale des normes de signification.

L'approche d'une interprétation ou communication triangulaire cherchant ainsi à saisir la relation adéquate de similarité (de réponse) dans le comportement verbal d'un autre locuteur apparaît chez Davidson avec l'introduction du principe causal. Il existe, en effet, un lien causal fort entre les utilisateurs du langage, les événements qui surviennent et les objets du monde et tout ceci détermine la manière dont nous apprenons et communiquons avec les autres. Mieux, cela rend un esprit en principe accessible à un autre<sup>915</sup>. Davidson refuse de faire appel aux conceptions qu'il qualifie de « standards » de la subjectivité<sup>916</sup>. Ces conceptions considèrent uniquement les états mentaux qui occupent l'esprit sans faire référence au monde extérieur. Ce point recouvre les critiques de Davidson à l'égard de l'introspection. Pour l'auteur, il n'existe pas de tels états. Il n'y a aucun mot ou concept qui ne soit pas compris ou interprété directement ou indirectement en termes de relations causales entre les gens et le monde<sup>917</sup>. De manière plus générale, il n'est d'aucune utilité de vouloir séparer ce qui provient de moi et ce qui provient du monde. Pour autant que les similarités de

---

<sup>913</sup> Davidson, 1990a, p. 325: « The ultimate source of both objectivity and communication is the triangle that, by relating speaker, interpreter and the world, determines the contents of thought and speech »; cité dans Rorty 2000a, p. 15.

<sup>914</sup> Davidson, 1989, p.198; cité par Laurier, 2002, p. 151.

<sup>915</sup> Davidson, 2001b, 1988, p.52.

<sup>916</sup> Ibid., p 50.

<sup>917</sup> Ibid., p 51.

réponse renvoient à des croyances, nous comprenons alors mieux la position de Davidson selon laquelle avant la communication triangulaire, il n'est pas possible de dire que les pensées ou les mots d'un locuteur ont des contenus. C'est en effet le triangle qui, en reliant le locuteur et l'interprète au monde, détermine le contenu de ces pensées et de ces mots. Parce que ce qui cause une pensée détermine aussi le contenu de cette pensée, une personne ne peut avoir d'états mentaux avec un contenu empirique déterminé à moins qu'elle ait un contact avec le monde extérieur. Mais, puisqu'un tel contact requiert des similarités de réponses<sup>918</sup> entre interlocuteurs, il en résulte que sans interaction avec une autre personne, rien ne saurait garantir à l'interprète que ses pensées sont objectives tant qu'il n'aura pas effectué la « triangulation » par laquelle il compare les contenus de ses pensées à celles d'autrui, et aux stimulations qu'il reçoit. Il prend au moins deux personnes, dit Davidson, pour trianguler et donc pour donner un contenu à la pensée. Il semble donc que le processus de la triangulation spécifie que la nature de l'interprétation dépend à la fois de l'interprète ou observateur et du locuteur en ce sens que ce qui cause les pensées et les énonciations de chacun est identifié au sein d'une interaction « sociale » et d'un espace public, soit à travers une activité conjointe :

« It is we who class cow appearances together, more or less naturally, or with minimum learning. And even so another classification is required to complete the point, the class of relevant causes is in turn defined by similarity in responses: we group together the causes of someone's responses, verbal and otherwise, because we find the responses similar. What makes these the relevant similarities? The answer again is obvious; it is we, because of the way we are constructed (evolution had something to do with this), that find these responses natural and easy to class together. If we did not, we would have no reason to claim that others were responding to the same objects and events (i.e. causes) that we are... The identification of the objects of thought rests, then, on a social basis. Without one creature to observe another, the triangulation that locates the relevant objects in a public space could not take place. I do not mean by this that one creature observing another provides either creature with the *concept* of objectivity; the presence of two or more creatures interacting with each other and with a common environment is at best a necessary condition for such a concept. Only communication can provide the concept, for to have the concept of objectivity, the concept of objects and events that occupy a shared world, of objects and events whose properties and existence is independent of our thought, requires that we are aware of the fact we share thoughts and a world with others<sup>919</sup>. »

<sup>918</sup> Glüer, 2006, p.1010, mentionne sur ce point Davidson, 2001a, p.4f: « [s]ince any set of causes will have endless properties in common, we must look to some recurrent feature of the gatherer, some mark that he or she has classified cases as similar » et souligne la conclusion à laquelle il en arrive: « This can only be some feature or aspect of the gatherer's reactions ».

<sup>919</sup> Davidson, 2001b, 1990, p.200-201, voir aussi 2001b, 1992, p.118-120.

Par conséquent, pour qu'un locuteur veuille dire quelque chose, il doit considérer certains stimuli comme similaires et réagir à ces stimuli de la même manière; lorsqu'il remarque ainsi que l'observateur réagit aux stimuli de la même manière que lui, il sait par la même occasion que l'observateur a en esprit les mêmes objets auxquels lui, le locuteur pense. C'est donc bien seulement par le biais de la triangulation que les communicateurs peuvent développer un concept d'un monde objectif, et ceci pour les mêmes raisons qu'un langage privé est impossible : selon Davidson, il n'existe effectivement pas de langage privé car à moins que celui-ci soit partagé, il n'y a aucune manière de distinguer entre le fait d'utiliser le langage correctement et de l'utiliser de manière incorrecte<sup>920</sup>. Le langage est l'élément qui nous met en contact avec autrui, et qui nous permet d'attribuer et de se voir attribuer des attitudes évaluatives. Par conséquent, c'est seulement en communiquant avec d'autres personnes, en comparant sa propre réaction à un objet dans le monde avec la réaction d'un autre, qu'un communicateur est capable de remarquer la similarité dans sa réponse et peut être sûr que ses propos signifient ce qu'il veut dire :

Je soutiens que le concept d'une vérité intersubjective suffit comme base de la croyance et par conséquent pour les pensées en général. Et peut-être est-il suffisamment plausible de dire que le fait d'avoir le concept de vérité intersubjective dépend de la communication au sens linguistique plein. Pour compléter l'« argument », néanmoins, je dois montrer que la seule manière par laquelle on puisse arriver au contraste subjectif-objectif passe par la possession du concept d'une vérité intersubjective. J'avoue ne pas savoir comment le montrer. Mais je n'ai pas la moindre idée de la manière dont on pourrait arriver autrement au concept d'une vérité objective. Au lieu d'un argument, j'offre l'analogie suivante. Si j'étais rivé à la terre, je n'aurais aucun moyen de déterminer la distance de nombreux objets par rapport à moi. Je saurais seulement qu'ils se trouvent sur une ligne quelconque menée d'eux à moi. Je pourrais entrer en contact avec succès avec des objets, mais je ne pourrais pas donner de contenu à la question de savoir où ils se trouvent. Comme je ne suis pas rivé au sol, je suis libre de trianguler. Notre sens de l'objectivité est la conséquence d'une autre sorte de triangulation, une qui requiert deux créatures. Chacune interagit avec un objet, mais ce qui donne à chacune le concept de la réalité objective est la ligne de base formée entre les créatures par le langage. Le fait qu'elles partagent un concept de vérité à lui seul donne un sens au fait qu'elles ont des croyances, et sont capables d'assigner à des objets une place dans un monde public. La conclusion de ces considérations-ci est que la rationalité a un trait social. Seuls les communicateurs la possèdent<sup>921</sup>.

---

<sup>920</sup> Davidson, 2001b, 1991, p. 209.

<sup>921</sup> Davidson, 1991a, p. 75.

Cette évaluation de similarité de réponses procède ainsi de manière réciproque, le locuteur ou l'interprète devenant « l'autre observateur » pour l'autre. Si nous pouvons constater des similarités de réponses (en vertu d'un comportement observable publiquement) alors l'interprète sait, à son tour, que le locuteur réagit aux mêmes stimuli et objets auxquels il réagirait lui-même en pareilles circonstances. La source même de l'objectivité réside précisément dans cette intersubjectivité :

« We may think of it as a form of triangulation: each of two people is reacting differentially to sensory stimuli streaming in from a certain direction. If we project the incoming lines outward, their intersection is the common cause. If the two people now note each others reactions (in the case of language, verbal reactions), each can correlate these observed reactions with his or her stimuli from the world. The common cause can now determine the contents of an utterance and a thought. The triangle which gives content to thought and speech is complete. But it takes two to triangulate. Two, or, of course, more<sup>922</sup>. ».

On comprend dès lors que toutes les tentatives visant à fonder les comparaisons intrapersonnelles et interpersonnelles à l'aide de normes ou d'éléments physiques considérés comme objectifs, sont, pour l'auteur, vouées à l'échec car notre connaissance empirique n'a besoin d'aucun fondement épistémologique<sup>923</sup>. Nous avons un accès privilégié à nos propres pensées, mais cela ne veut pas dire que celles-ci sont déterminées uniquement par des éléments internes. Cela ne veut pas non plus pour autant dire qu'elles sont entièrement déterminées par des éléments externes. Ces conditions de l'interprétation montrent, selon Davidson, qu'il peut à la fois y avoir détermination du contenu des pensées par des facteurs extérieurs (causaux) et une autorité de l'interprète sur ses pensées et les significations des mots qu'il emploie. Cette autorité est due en dernière instance au fait que l'interprète peut être lui-même interprété. C'est pourquoi Davidson soutient que l'autorité de la première personne, ou la connaissance qu'un individu a à la première personne de ses contenus mentaux, est expliquée par la connaissance à la troisième personne, par la connaissance que les autres esprits ont de lui dans des conditions communicationnelles<sup>924</sup>. Il y a bien une asymétrie entre les deux types de connaissance, mais c'est en dernier ressort la connaissance à la troisième personne, interpersonnelle, qui fonde la connaissance « personnelle » et

<sup>922</sup> Davidson, 2001b, 1991, p.213.

<sup>923</sup> Davidson 2001b, p. xiv.

<sup>924</sup> Davidson, 2001b, 1984; 2001b, 1991.

« privée ». C'est en ce sens également que Davidson soutient qu'il est possible de parler du caractère « social » de la signification des contenus mentaux<sup>925</sup>. En réalité, pour pouvoir dire que j'ai une pensée ou que j'utilise un langage, il faut au moins être deux. Ceci implique une forte imbrication à autrui. Par conséquent, l'idée fondamentale qui rend la triangulation possible et qui assure la connexion entre « moi », « autrui » et « le monde » est le fait que les systèmes de pensée (le mien et celui d'autrui) doivent concorder. Davidson met bien en évidence ce triangle constitué par l'interprète, la personne interprétée et l'objet commun sur lequel ils doivent être capables de communiquer, comme suit :

La seule voie pour savoir que le deuxième sommet du triangle - la deuxième créature en personne - réagit au même objet que moi, est de savoir que l'autre personne a dans son esprit les mêmes objets. Dans ce cas, la deuxième personne doit également savoir que la première personne constitue le sommet du même triangle dont un autre sommet est occupé par la deuxième personne. Deux personnes doivent être en situation de communication pour que chacun sache de l'autre qu'elles sont reliées de cette manière. Chacune d'elles doit parler à l'autre et être comprise de l'autre.<sup>926</sup>

La situation décrite par la structure de la triangulation suggère ainsi trois relations de « similarités de réponse », à savoir :

1. La similarité dans les réponses du locuteur (en particulier les réponses linguistiques) face à certains stimuli dans le monde.
2. La relation de similarité obtenue entre les réponses linguistiques de l'interprète et les stimuli dans le monde et,
3. La similarité, observée par l'interprète, entre ces deux paires de réponses-stimulus.

Comprendre, c'est-à-dire produire une théorie de la signification pour l'interprété, implique que l'interprète construise de manière systématique des corrélations entre les deux ensembles de paires de réponse-stimulus suivants : les réponses ordinaires du locuteur à certains stimuli, objets ou événements dans le monde et les réponses de l'interprète lui-même. L'interprété ou le locuteur et l'interprète ou l'observateur font donc des classifications différentielles d'objets ou d'événements : tous deux traitent certaines choses plus ou moins semblables les unes aux autres. Le critère objectif d'une telle classification et d'une corrélation réussie de l'interprétation suggère l'idée d'une « similarité de réponse ». C'est ainsi par réaction linguistique aux événements dans le monde que locuteurs et interprètes classent les choses,

---

<sup>925</sup> Davidson, 1992b.

<sup>926</sup> Davidson, 1992a, p.256; cité par Engel, 1994, p. 255.

mettant ensemble celles qui sont semblables et écartant de telles classes celles qui sont dissemblables. Ces structures de réponses peuvent s'expliquer par l'évolution et l'apprentissage. En ce sens, nous avons bien besoin d'une théorie naturaliste des représentations, qui regroupe des structures comportementales et établit des régularités nomologiques. Cependant, le critère autorisant cette « similarité de réponse » ne peut pas provenir des réponses du locuteur ou de l'interprète classifiant, sans quoi il serait circulaire mais plutôt des réponses d'un autre observateur. Le critère de similarité « ne peut venir que des réponses d'un observateur aux réponses de la créature », qu'il observe<sup>927</sup> :

« All creatures classify objects and aspects of the world in the sense that they treat some stimuli as more alike than others. The objective criterion of such classification is similarity of response. Evolution and subsequent learning no doubt explain these patterns of behaviour. But from what point of view can these be called? The criterion on the basis of which a creature can be said to be treating stimuli as similar, as belonging to a class, is the similarity of the creature's responses to those stimuli; but what is the criterion of the similarity of the responses? *This criterion cannot be derived from the creature's responses; it can come only from the responses of an observer to the responses of the creature.* [...] If we discover kinds of objects or events in the world that we can correlate with the utterances of a speaker, we are on the way to interpreting the simplest linguistic behaviour. [...] [W]ithout this sharing of reactions to common stimuli, thought and speech would have no particular content – that is, no content at all. It takes two points of view to give a location to the cause of a thought, and thus to define its content<sup>928</sup>. ».

Du point de vue de Davidson et de Quine, c'est par réaction verbale favorable ou défavorable aux réactions d'une autre personne qu'une créature rationnelle sait objectivement qu'elle fait des classifications différentielles. Cette évaluation des similarités de réponses, où locuteurs et interprètes constatent que leurs comportements respectifs et leurs réponses-stimuli face à des objets ou événements sont identiques, permet, comme nous l'avons dit plus haut, d'envisager la possibilité de connaissances réciproques entre locuteurs et interprètes. Cela dit, pour que tous deux découvrent qu'ils répondent aux mêmes stimuli, chacun doit parler à l'autre et être compris par lui. La source ultime du sens de la communication est ainsi l'objectivité :

---

<sup>927</sup> Davidson, 2001b, 1991, p.159.

<sup>928</sup> Ibid., p.212: nos italiques.

Si nous n'étions pas en communication avec les autres, il n'y aurait aucun moyen pour nous de fonder l'*idée* qu'on puisse se tromper, et par conséquent qu'on puisse avoir raison, dans ce que l'on dit ou dans ce que l'on pense. La possibilité de la pensée en même temps que celle de la communication dépend, à mon sens, du fait que deux personnes ou plusieurs répondent, plus ou moins simultanément, à des entrées d'information venues d'un monde partagé, et que nous nous répondons l'un à l'autre.<sup>929</sup>

Le locuteur doit donc corrélérer les réponses de l'interprété aux objets et événements de son propre univers. Dans le cas de la communication verbale, une régularité du comportement verbal doit être corrélée aux objets que le locuteur perçoit dans son environnement. En s'engageant dans une telle interaction linguistique, le locuteur trouve l'événement, à savoir le fait qu'il pleut, similaire à des événements précédents. Il en va ainsi pour l'interprète qui découvre alors que ses paroles sont similaires en contenu à celles de son interlocuteur. Ces similarités constituent un triangle, un sommet représenté par l'événement, un autre par le locuteur et un autre par l'interprète. Les similarités de réponses sont, au fur et à mesure que le dialogue se poursuit et gagne en intensité, des assertions ou des hypothèses pertinentes. Or toute créature linguistique qui comprend l'acte de discours sait que quelqu'un qui fait une assertion croit à ce qui s'y rapporte. Par conséquent, avoir le concept d'un objet, pour Davidson, c'est être capable de le situer dans un espace partagé par une autre créature, et donc reconnaître qu'une autre créature est capable de réagir au même objet (ceci délimitant la vérité objective du concept)<sup>930</sup>. Il en est de même du concept de croyance et des concepts de vérité et de fausseté : on ne peut les posséder qu'en étant un interprète du discours d'autrui. Les similarités de réponses se rapportent donc bien à des croyances ou autres états mentaux que les interlocuteurs se découvrent avoir en partage. Néanmoins, nous retrouvons ici l'idée que l'environnement doit agir causalement sur les interprètes. Mais, qu'est ce qui justifie un tel point de vue ?

Le problème est ainsi le suivant : qu'est-ce qui détermine que les réponses similaires de l'interprété ne classifient pas différemment les objets ou les événements de l'interprète<sup>931</sup>? En effet, qu'est-ce qui nous garantit, dans le comportement d'autrui, que ce qu'il entend par lapin correspond à ce que j'entends par « lapin » ? Autrement dit, comment l'interprète peut-il

<sup>929</sup> Davidson, 2003, p.419-420.

<sup>930</sup> Laurier, 2002, p.151-152.

<sup>931</sup> Ce problème d'une similarité normative est également soulevé dans notre premier chapitre à propos de la question de la projectibilité des prédicats analysée par Goodman.

garantir que les stimuli auxquels il réagit sont des causes objectives, c'est-à-dire que proviennent bien d'objets du monde ? Il n'y a précisément aucun moyen de le vérifier, puisque, à supposer que l'autre identifie en réalité des parties non-détachées de lapin, l'interaction n'en serait pas altérée pour autant<sup>932</sup>. Aussi, affirmer que les réponses du locuteur sont similaires aux nôtres en ce sens qu'elles traduisent notre réaction face aux objets ou événements dans le monde ne nous paraît pas suffisant pour déterminer ce qui cause une telle similarité de réponse et si celles-ci répondent bien aux mêmes stimuli. En outre, notre similarité de réponses, en étant causalement influencée par l'environnement, renvoie et répond à une pluralité de causes différentes, dont chacune est susceptible d'orienter notre comportement. La régularité dans l'environnement, l'identification de stimuli communs désignant ceux auxquels nous réagissons implique une similarité normative du jugement, à savoir que parmi toutes les causes possibles présentes dans la situation stimulus, nous pouvons tenir certaines pour vraies et d'autres pour fausses. Davidson reproche ainsi à la théorie de Quine qu'il qualifie de *théorie proximale* d'établir ce lien entre signification et expérience et de construire une théorie dans laquelle la similarité des significations et des références est inférée de la similarité des stimuli proximaux qui déclenchent l'assentiment ou le dissentiment. Davidson ne lie pas les stimuli aux sensations ni aux phrases observationnelles : il les lie directement à des propriétés et à des objets distaux qui causent notre croyance. C'est ce qu'il appelle la « théorie distale ». La cause commune de stimulations distales ne peut être identifiée qu'intersubjectivement, comme le point où se rencontrent interprète et interprété. Autrement dit, afin de rendre compte adéquatement d'une similarité normative, l'interprète ne doit pas seulement se limiter, à l'image d'un observateur extérieur, à la description du comportement d'autrui en termes proprement causaux soit en termes de réponse-stimulus. Par conséquent, la triangulation se heurte au problème de l'identification de stimuli communs, c'est-à-dire au problème de la similarité normative; ce problème semble ainsi en suspens si nous analysons la triangulation en termes purement causaux comme la corrélation externe de deux paires de réponse-stimulus.

---

<sup>932</sup> Quine, 1962. La réponse est naturellement qu'il parle le même langage que vous, qu'il participe au même jeu et qu'il partage par conséquent le même schème conceptuel, la même ontologie immanente à ce langage – ce qui présuppose qu'il l'ait appris dans votre communauté linguistique. Mais ce n'est pas un argument dirimant : il est *en principe* impossible de savoir quelle est la référence des termes que l'autre utilise, surtout quand, comme dans le cas de la traduction, on sait que l'autre ne parle pas le même langage, autrement dit ne joue pas au même jeu, du moins pas de la même façon.

C'est d'ailleurs cette considération qui a conduit Quine à la thèse de l'inscrutabilité de la référence ou indétermination de la référence<sup>933</sup>. Voici, résumée par Davidson, en quoi elle consiste :

La thèse de Quine sur l'inscrutabilité de la référence est qu'il n'y a aucun moyen de dire à quoi [précisément] font référence les termes singuliers d'un langage, ou de [*sic*] quoi sont vrais ses prédicats, du moins aucun moyen de le dire, à partir de l'ensemble des données empiriques comportementales, actuelles ou potentielles, et que ces données sont tout ce qui compte dans les questions de signification et de *communication*<sup>934</sup>.

Étant donné les différents comportements linguistiques, nous ne pourrions pas de « l'extérieur » identifier leur référence. En outre, si la signification d'un terme se réduit à la performance des comportements verbaux dans certaines circonstances, nous ne pouvons pas identifier la référence précise d'un terme étranger, car les données comportementales admettent différentes références possibles, ou autant d'ontologies : un comportement verbal qui intervient dans des circonstances où il y a ce que nous appelons un lapin peut identifier tout aussi bien un objet ressortissant à la même ontologie (au même système référentiel) que la nôtre, ou un objet ressortissant à une ontologie totalement différente. Il nous semble alors que les seules données comportementales ne permettent absolument pas de scruter ce à quoi les mots font référence. Il ne s'agit certes pas de nier qu'il existe des régularités nomologiques entre les causes et les structures comportementales. Mais, nous savons aussi que ces régularités ne peuvent jamais être suffisantes pour identifier les objets de référence. Le schème causal n'est jamais unique, et il y a toujours indétermination. Mais, comme nous l'avons vu, l'indétermination n'est jamais telle qu'il ne puisse pas y avoir accord et compréhension. La « mesure du mental » n'implique pas non plus une détermination complète du mental par l'environnement causal<sup>935</sup>. C'est la communication seule qui est la source d'un sens et d'une vérité objective, et par conséquent d'une convergence des opinions. Par conséquent, si le scénario de la triangulation de Davidson vise à répondre au problème d'une similarité normative, alors la triangulation doit fournir plus que ce qui est publiquement disponible ; elle doit tenir compte des jugements de similarité normative. La triangulation ne

<sup>933</sup> Quine, 1993, Chap. I, p. 8.

<sup>934</sup> Davidson, 1993b, p. 327. Le terme « communication » est ici très important.

<sup>935</sup> Davidson ne souscrit pas à l'externalisme défini par Putnam, 1974 et Burge, 1979, pour des raisons qu'il explique dans 2001b, 1987 et 2001b, 1991. Son externalisme est causal, mais il ne soutient pas que l'environnement, physique ou social permette à lui seul d'individualiser les contenus. Nous ne discuterons pas ce point ici. Cf. Seymour, 1994.

doit pas simplement considérer les similarités issues de l'observation des comportements décrits dans le langage causal d'une réponse-stimulus, mais doit prendre en compte la corrélation de telles similarités. Il est donc nécessaire d'introduire la notion de jugement normatif à partir du moment où interprète et interprété découvrent qu'ils répondent aux mêmes stimuli. Dans les termes de Davidson, la triangulation impliquera le principe de charité soit la forte présomption que nous nous accordons ou que nous partageons des croyances communes, et spécifiquement des croyances relatives à la question de l'identification de stimuli communs. L'idée de jugements de similarité normative se retrouve donc, sous l'approche de Davidson, dans le principe de charité, et dire que le principe de charité est constitutif de l'interprétation et de la compréhension revient bien à affirmer que les jugements de similarité normative doivent être partagés, que l'interprète et l'interprété doivent partager un espace de similarité normative afin de réussir à comprendre l'autre et à se faire comprendre de l'autre. Par conséquent, en suggérant le caractère constitutif du principe de charité, il faut renoncer à toute tentative de décrire ce niveau le plus élémentaire de nos énonciations en termes proprement causaux et béhavioristes. Par ailleurs, nous retrouvons à ce niveau la présomption d'identité au sein des jugements de similarité normative entre interprète et interprété. Aussi, si la structure triangulaire de la compréhension requiert bien un partage préalable des espaces de similarité normative entre interprète et interprété alors la compréhension est authentiquement sociale. En attribuant un rôle constitutif au principe de charité, Davidson reconnaît l'importance des jugements partagés de similarité normative pour déterminer la signification.

Cependant, Davidson n'adhère pas, dans son approche de la triangulation, de manière catégorique à l'idée d'une normativité sociale de la signification et de la compréhension ; il souligne que le processus de triangulation, qui est censé être le point d'entrée de l'interprétation, débute avec la corrélation entre deux ensembles publiquement observables de réponse-stimulus. Ce tableau béhavioriste constitue ainsi l'arrière-plan de l'interprétation, soit ce qui est au fondement de l'attribution de la signification et d'un accord dans nos croyances et significations entre l'interprète et l'interprété, celui-ci se traduisant par l'observation de corrélations mutuelles entre les réponses-stimuli du locuteur et de l'auditeur. Pourtant, nous avons d'autre part mentionné que, pour établir des corrélations, nous avons besoin de recourir au principe de charité, soit aux jugements partagés de similarité normative. C'est

donc bien le principe de charité qui rend possible les corrélations entre les réponses-stimuli du locuteur et de l'auditeur et non l'inverse. Par conséquent, au sein de ce modèle de la triangulation, se cristallisent cette tension entre l'interprétation radicale et le statut constitutif de la charité c'est-à-dire entre les approches publiques et sociales de la signification. La triangulation, comme l'interprétation radicale, semble d'abord seulement faire appel et justifier le bien-fondé d'une conception publique de la signification, mais, en y regardant de plus près, nous pouvons noter que la nature de la communication triangulaire présuppose le principe de charité, ce qui explique la nécessité d'une conception sociale de la signification.

En suivant Pascal Engel, nous pourrions alors dire que les bases de l'externalisme transcendantal de Davidson sont causales et sociales :

L' « externalisme » de Davidson tient [...] à deux conditions propres à l'interprétation. La première est l'existence d'une interaction causale entre les objets du monde et nos croyances. La seconde est le caractère public et social des pensées et des significations dans les conditions d'une communication intersubjective. Parce que Davidson considère que ces conditions sont des conditions de possibilité de tout contenu mental et de toute signification, il considère qu'elles excluent *a priori* le scepticisme quant à l'existence du monde extérieur et le solipsisme.<sup>936</sup>

Du point de vue des conditions de possibilité causales du jugement, Davidson a effectivement soutenu, au moyen de son expérience de pensée du *Swampman*<sup>937</sup>, qu'avoir des rapports causaux avec le monde et avec autrui est une condition nécessaire pour être dit capable de penser<sup>938</sup>. Des rapports avec le monde sont donc nécessaires pour avoir certains types de pensées, et non pas pour penser simplement. Ainsi, par exemple :

<sup>936</sup> Engel, 1994, p. 255-256.

<sup>937</sup> Une expérience de pensée introduite par Davidson, 2001b, 1987, met en scène Davidson lui-même qui, partant en randonnée dans des marais est soudain frappé par la foudre. Dans le même temps, à proximité, un second éclair réorganise spontanément toutes les molécules qui constituaient Davidson et par le plus grand des hasards, elles reprennent exactement la même position que celle qu'elles avaient au moment de sa mort. Ce *Swampman* possède néanmoins un cerveau, entièrement identique à celui qu'avait Davidson et se comporte donc exactement comme l'aurait fait Davidson. Alors, suivant à nouveau son chemin, retournant à son bureau à l'université de Berkeley, il reprend le cours normal de sa vie qu'il consacre à écrire des essais philosophiques. Les amateurs de ce type de scénario peuvent également consulter l'histoire de *Instant Louis* (Strawson, 1994).

<sup>938</sup> Rorty, 1998a, p. 149, n. 26, adhère à cette idée : pour pouvoir être interprétable, un comportement linguistique doit être pensé à partir des interactions causales entre l'environnement et le sujet du comportement (c'est là la manière de Rorty de comprendre le *dictum* kantien « les concepts sans les intuitions sont vides »). Pour les limites de l'usage rortyen de Davidson, voir Engel, 1994, p. 262, n.1.

Ce que les mots d'une personne signifient dépend, même dans les cas les plus élémentaires, des espèces d'objets et événements qui ont amené cette personne à supposer que les mots étaient applicables ; il en va de même pour ce à propos de quoi sont les pensées de cette personne<sup>939</sup>.

Ceci vaut pour tous les contenus de pensée *possibles* (cette thèse est encore compatible avec un externalisme transcendantal naturaliste). Mais, selon l'argument de la *triangulation*, une interaction avec autrui est tout autant nécessaire pour être dit penser qu'une interaction préalable avec les objets de nos pensées<sup>940</sup>. Les conditions de possibilité de la pensée et du discours sont aussi bien causales que sociales et linguistiques. Parce que Davidson considère que ces conditions sont des conditions de possibilité de tout contenu mental et de toute signification, il considère qu'elles excluent a priori le scepticisme quant à l'existence du monde extérieur et le solipsisme. C'est en ce sens qu'il dit que « si l'externalisme est vrai, la question de la connaissance du monde extérieur ne se pose pas », et que « si c'est une condition constitutive de certaines pensées que leur contenu soit donnée par leur cause normale, alors la connaissance des événements et des situations qui causent ces pensées ne peuvent pas requérir qu'un sujet établisse indépendamment, ou confirme, l'hypothèse qu'il y a un monde extérieur qui cause ces pensées<sup>941</sup> ».

### 5.7. Conclusion

Nous cherchons donc bien ici à tendre vers une approche sociale qui rende bien compte de l'équilibre délicat entre un contextualisme et une position relativiste c'est-à-dire une approche sociale qui réussisse à concilier l'idée selon laquelle notre application correcte des normes de signification dépende de nos contextes d'accords sociaux c'est-à-dire de nos pratiques d'interprétation et de traduction (renvoyant à l'immanence du contexte de l'objectivité des prétentions à la vérité et des normes correctes de signification) avec l'idée selon laquelle la validité des standards ainsi obtenus ne doit néanmoins pas s'identifier avec un accord social. Nous soutenons alors qu'une distinction trop stricte entre normes logiques

<sup>939</sup> Davidson, 2001b, 1987, p. 37.

<sup>940</sup> Une autre version de l'argument, comme nous l'avons vu, défend la thèse que la connaissance de soi, la connaissance d'autrui et la connaissance du monde sont indissociables. (Cf. Davidson, 2001b, 1991).

<sup>941</sup> Davidson, 1991b, p. 196, 1989, p. 193-200.

et normes conceptuelles empiriques n'est pas souhaitable et que la portée d'une normativité constitutive de la signification s'applique tant à un vocabulaire logique qu'à un vocabulaire non logique. Nous visons ainsi une approche qui soit contextualiste quant à la vérité et à la signification tout en étant également faillibiliste.

Nous en venons donc à préciser, dans la continuité de notre propos<sup>942</sup>, les différences entre deux approches respectives de la signification et de la communication, une approche publique et une approche sociale, en montrant que cette tension se retrouve au cœur de l'approche de Davidson entre d'une part, l'interprétation radicale, qui tend à décrire la signification comme publique et intelligible, et d'autre part, le rôle constitutif de la charité dans la pratique interprétative, qui exige que la signification soit sociale et que nous partageons au préalable certaines significations (un sociolecte) entre locuteurs (interprétés) et auditeurs (interprètes). Nous pouvons donc en conclure que nous avons bien, dans l'analyse de Davidson, un argument en faveur d'une approche de la signification privilégiant le point de vue du participant ou du locuteur-acteur et d'une conception sociale de la signification via une *reductio* du rôle de l'observateur au sein d'une approche publique de la signification. Afin de maintenir une certaine cohérence d'ensemble, la perspective de l'observateur de l'interprétation radicale à laquelle se joint une approche publique de la signification doit finalement s'éclipser devant le point de vue du participant, préconisant ainsi une approche sociale, celle-ci visant à attribuer un arrière-plan de croyances (rôle constitutif de la charité), de significations socialement partagées afin de permettre aux idiolectes individuels d'être dotés de signification et en fin de compte, aux locuteurs, de se comprendre et d'être compris. La triangulation, entendue comme la corrélation externe de similarités observées dans nos réponses-stimuli, requiert que nous présupposions des jugements de similarités partagées. Ce qui fixe le contenu d'une pensée n'est donc pas conféré par le paradigme d'acceptabilité de l'interprète seul ni du locuteur seul, mais procède d'une activité conjointe. Dans une telle perspective, il n'existe alors pas de relations inégales, de relations de manipulation ou d'endoctrinement entre des personnes charitables et celles envers qui il faut montrer de la charité. Dans l'activité conjointe de fixation du contenu de pensée, la pensée de Davidson repose sur une symétrie entre usagers du langage. Le locuteur et l'interprète commencent leurs efforts de compréhension mutuelle sur des bases égales, personne ne pensant détenir les

---

<sup>942</sup> Chap.III, §3.4.1.

normes de similarités de réponses. Nous pouvons donc seulement réussir à comprendre ce qu'un locuteur veut dire si nous supposons que nous partageons bien des jugements, croyances et significations communes ; nous ne pouvons pas nous contenter d'interpréter de l'extérieur le comportement d'autrui en nous fiant seulement sur des données comportementales observables ; autrement, les significations des énonciations de l'interprété et le contenu de ses croyances demeureraient indéterminées. Par conséquent, la tentative de fonder la signification et la compréhension sur une quelconque corrélation observable entre stimulus et réponse, se réfute d'elle-même. Une conception naturaliste implique une interaction sociale, et de ce fait le fondement explicatif de la signification est l'interaction sociale. Nous ne disposons donc de rien de plus fondamental, puisque tout ce que nous posons comme plus fondamental implique implicitement un partage social de nos croyances et significations. Finalement, nous en venons ici avec Davidson, à une conclusion qui rejoint celle de nos premiers chapitres, soit d'une part, que l'activité de suivre une règle ou d'employer un terme de manière significative est une activité régie par une contrainte normative et d'autre part, que cette normativité résulte bien d'une activité sociale.

## CHAPITRE VI

### BRANDOM : UNE VERSION SOCIALE FAIBLE DES NORMES DE SIGNIFICATION

#### 6.1. Introduction

Si nous nous tournons maintenant du côté de l'œuvre de Brandom *Making It Explicit*, nous pourrions la lire, dans la continuité d'une conception transcendantale, sociale et objective des normes de signification, comme une tentative d'élaborer en détail l'approche de Wittgenstein d'une normativité de la signification émergeant de nos pratiques sociales afin de trouver un juste milieu entre immanence et transcendance et parvenir à un point de vue contextualiste (en évitant le platonisme) sans néanmoins tendre vers le relativisme ou le naturalisme. La sémantique pragmatiste de Brandom suit trois axes principaux :

- 1 - elle adhère à un *pragmatisme sémantique*, pour lequel les propriétés sémantiques des contenus sont expliquées à partir des propriétés d'usage des contenus;
- 2 - à un *inférentialisme*, pour lequel les relations inférentielles constitutives de la signification d'un concept sont inter-linguistiques, mais incluent aussi les circonstances et les conséquences non inférentielles de l'usage d'un concept et ;
- 3 - à la thèse d'une dimension *sociale* ou *socialement instituée des normes de signification*.

Ces trois axes reposent sur trois aspirations méthodologiques : il est question ici d'expliquer la dimension représentationnelle du langage et de la pensée par quelque chose de non-représentationnel, mais qui ne soit pas naturel ou non-normatif<sup>943</sup>. Il ne s'agit pas de réduire le vocabulaire sémantique (en premier lieu inférentiel, ensuite, entre autres, représentationnel) à un vocabulaire naturaliste ; le but est ainsi d'expliquer sa structure et son contenu à partir d'un vocabulaire normatif<sup>944</sup>.

Nous partirons donc de la thèse centrale de Brandom selon laquelle, l'origine de nos pratiques linguistiques ne peut être décrite et pensée qu'à partir du normatif, pour expliquer sa pratique assertionnelle du jeu de pointage déontique sur le mode du « je-tu ». Tout

---

<sup>943</sup> Brandom, 2011, p. 185-186.

<sup>944</sup> Ibid., p.34-35.

d'abord, le contenu de ce que je dis et de ce que je pense est déterminé par son rôle inférentiel (§6.2), relevant lui-même des propriétés pragmatico-normatives de mon assertion de ce contenu (§6.3, §6.4); ces propriétés expriment ce que je dois et ce que je peux reconnaître ou faire inférentiellement en ayant énoncé ce contenu; ces propriétés n'existent sous forme de statuts déontiques qu'en étant attribués, explicitement ou implicitement, par les autres participants à la pratique discursive. Les normes pragmatiques qui contribuent ainsi à définir le contenu conceptuel de ces mêmes assertions, dépendent des capacités des individus à exercer implicitement des attitudes normatives, par lesquelles ils s'attribuent mutuellement des statuts déontiques (permissions, engagements, obligations) qui déterminent le contenu de ce qu'ils disent (§6.5, §6.6). Nous verrons alors comment nos pratiques sociales instituent les inférences qui définissent l'usage des contenus assertés et pourquoi cette distinction *perspectivale* entre attribuer un engagement (l'attitude d'un marqueur observateur) et reconnaître un engagement (l'attitude d'un participant) est absolument centrale pour rendre compte de l'objectivité des normes de signification (§6.7). Nous montrerons ainsi qu'un trait majeur de la structure sociale sur le mode du « je-tu » est de générer de nouvelles et différentes perspectives de pointage, la compréhension linguistique dépendant avant tout d'une capacité à évoluer entre sa propre perspective et celle de ses interlocuteurs. Nous soulignerons néanmoins les limites et les difficultés rencontrées par cette approche sociale et perspectiviste, difficultés qui poussent d'abord Brandom à adopter une version de l'externalisme sémantique pour ensuite l'abandonner au profit d'une version sociale et historique des normes de signification (§6.8). Nous montrerons ainsi qu'une tension demeure toujours au cœur de l'approche de l'objectivité de Brandom entre d'une part, une position représentationnaliste externaliste et une position inférentialiste contextualiste. Celle-ci nous conduit ainsi à suggérer, dans les prochains chapitres, la thèse d'une conception sociale des normes de signification, qui conserverait la trame de la pratique discursive décrite par Brandom, tout en y ajoutant la notion d'idée régulatrice, à savoir : l'idée de sujets rationnellement responsables (au sein de ce jeu de pointage déontique), l'idée d'un monde objectif<sup>945</sup> et l'idée d'une vérité objective.

## 6.2. Inférentialisme et pragmatisme

<sup>945</sup> C.-à-d. l'idée, que l'on retrouve d'ailleurs dans la version de l'externalisme sémantique de Brandom, d'un monde objectif « un et toujours le même » qui fonctionne de manière régulatrice afin de guider les pratiques sociales de nos échanges langagiers.

Brandom n'ignore pas la dimension représentationnelle de nos capacités cognitives<sup>946</sup> mais il choisit plutôt de substituer à la notion de représentation celle d'inférence comme notion centrale d'une théorie de la signification<sup>947</sup>. Aussi, si l'inférence est envisagée dans sa dimension sociale et normative, elle peut, selon Brandom, servir de base pour construire une théorie de la signification, de la vérité et de la référence. C'est donc seulement à partir de cette théorie inférentialiste de la référence que la dimension représentationnelle du langage et de la pensée peut être comprise. Les capacités représentationnelles du langage et de la pensée sont, pour Brandom, secondaires par rapport au fait que toute pratique conceptuelle est nécessairement inférentiellement et socialement articulée. Au représentationnalisme, Brandom oppose un inférentialisme, et à l'atomisme Brandom oppose un holisme.

Brandom refuse aussi le naturalisme et l'empirisme : le non-naturalisme c'est-à-dire l'idée qu'il existe une autonomie du normatif par rapport aux faits et aux sciences naturelles n'est pas, pour lui, plus incohérent que le naturalisme. Le naturalisme ne peut pas ne pas tenir compte des engagements ontologiques irréductibles à des faits, en ce qu'il se construit notamment à partir de modalités aléthiques<sup>948</sup> (possibilité, nécessité...). En ce sens, le naturalisme possède lui aussi des éléments irréductibles à de simples faits. Le non-naturalisme considère, selon Brandom, les modalités déontiques comme fondamentales dans l'ordre de la réalité en raison de leur familiarité : nous comprenons toujours déjà, implicitement, ce qu'est une norme. Nous sommes déjà en quelque sorte dans le normatif et ce n'est qu'à partir de cette distinction entre ce qui doit être (le normatif) et ce qui est (le descriptif) que nous pouvons concevoir le naturel ou le descriptif.

À l'opposé des théories naturalistes de la signification, qui visent à faire ressortir le représentationnel du naturel ou du descriptif, Brandom tente d'extraire le représentationnel

<sup>946</sup> Brandom, 2011, p. 170-171.

<sup>947</sup> Ibid., p. 882-883.

<sup>948</sup> Le naturaliste se sert en fait des modalités aléthiques en formulant ses affirmations, même s'il est très douteux que ces modalités puissent être exprimées en termes strictement naturalistes. Par conséquent, plutôt que de servir de modalités aléthiques, Brandom se sert des modalités déontiques (obligatoire, permis, défendu). Brandom, 2011, précise : « [...] il est tentant mais trompeur de comprendre de manière anachronique l'usage kantien de la notion de nécessité dans les termes des débats contemporains sur la modalité aléthique. Cela est trompeur car les préoccupations de Kant sont fondamentalement normatives, au sens où les catégories fondamentales sont la *modalité déontique*, *l'engagement et le droit*, plutôt que la *modalité aléthique*, *la nécessité et la possibilité*, [...] » (Ibid., p. 68): nos italiques.

des pratiques normatives. Tout contenu propositionnel possède une dimension représentationnelle, dont il s'agit de construire la spécificité en la mettant en relation avec l'aspect également expressif de toute activité conceptuelle<sup>949</sup>. Avant de tenter d'expliquer la manière dont un contenu exprimé peut représenter quelque chose, il faut au préalable expliquer ce que c'est qu'exprimer un contenu conceptuel. La forme fondamentale de contenu conceptuel est le contenu propositionnel. L'attribution d'une pensée ou d'une croyance à une personne dépend donc de la possibilité de lui attribuer des contenus propositionnels<sup>950</sup>. Le comportement de cette personne ne peut donc être rendu intelligible sans la mise au jour de croyances et de désirs qui servent de raisons à son comportement : les raisons sont ainsi implicites dans les comportements et sont explicitées par l'usage d'inférences pratiques. Cependant, ce qui peut servir de prémisse ou de conclusion dans un raisonnement doit également posséder un contenu propositionnel et attribuer des états intentionnels à quelqu'un revient *eo ipso* à lui attribuer des contenus propositionnels. La sémantique de Brandom donne donc une priorité logique aux propositions parce que celles-ci sont les éléments privilégiés des inférences. En ce sens, le contenu conceptuel est à comprendre à partir de son rôle dans le raisonnement, et non exclusivement en termes de représentation<sup>951</sup>. C'est ce rôle inférentiel du contenu propositionnel qui le définit :

Être explicite, au sens conceptuel, c'est jouer un rôle spécifiquement inférentiel. Dans le cas le plus fondamental, c'est être doté d'un contenu propositionnel au sens où l'on est prêt à servir à la fois de prémisse et de conclusion dans des inférences<sup>952</sup>.

Ce rôle inférentiel, nous le verrons par la suite, est ainsi dérivé des propriétés pragmatique-normatives des actes d'assertion ou de jugement qui posent le contenu. Attribuer un contenu sémantique à un état intentionnel revient ainsi à déterminer la signification pragmatique de son occurrence dans des contextes variés. C'est donc bien l'emploi d'un terme, tout d'abord au sein d'assertions, qui constitue sa signification<sup>953</sup>. La signification

---

<sup>949</sup> Brandom, 2009, p.15.

<sup>950</sup> Brandom, 2011, p. 207.

<sup>951</sup> Ibid., p.248.

<sup>952</sup> Brandom, 2009, p.24.

<sup>953</sup> Nous ne voulons pas dire que les assertions soient les seules manières d'employer les termes du langage mais que, dans le cadre de la signification littérale du langage représentationnel, les usages assertionnels sont les premiers auxquels nous nous intéressons. Cf. MacFarlane, 2011, p. 80, rend compte du rôle de l'assertion selon les quatre catégories suivantes : « 1. To assert is to express an

n'est pas fixée d'avance en stipulant une règle explicite délimitant l'emploi d'un terme mais est régie par les applications réelles des termes au sein d'assertions. Nous n'avons donc pas deux processus séparés – soit un processus qui consiste à instituer des normes de signification par décret et un processus qui vise ensuite à les appliquer en les insérant dans des jugements – où le premier est indépendant et prévaut sur le second, mais nous avons juste un seul processus à savoir : l'application des termes dans la production d'assertions, celle-ci instituant simultanément la signification des termes apparaissant dans nos assertions<sup>954</sup>.

Le pragmatisme de Brandom est donc étroitement relié à son inférentialisme. Celui-ci confère ainsi une priorité à l'inférence (et plus spécifiquement aux règles d'inférences) sur la représentation pour comprendre la signification. Il s'agit alors, avons-nous dit, de définir le rôle inférentiel de manière pragmatique-normative, en termes d'obligations, de permissions et d'engagements. Les engagements ne sont pas uniquement ce que nous sommes disposés à reconnaître comme conséquences de notre jugement initial, mais aussi les contenus que nous devons reconnaître, que nous le sachions ou non (les règles sont ici des prescriptions). Par exemple, en assertant ou en jugeant quelque chose, nous sommes engagés envers ses conséquences telles qu'elles sont attribuées et instituées par le contexte de pratiques linguistiques dans lequel nous sommes situés. Autrement dit, l'espace logique ou le contexte de pratiques linguistiques ne doit pas être compris relativement à des normes sociales auxquelles les locuteurs devraient se soumettre passivement, dans une forme de relation asymétrique je-nous<sup>955</sup>, les seconds imposant leurs normes au premier. Au contraire, cet

---

attitude; 2. To assert is to make a move defined by its constitutive rules; 3. To assert is to propose to add information to the conversational common; 4. To assert is to undertake a commitment. ».

<sup>954</sup>Nous retrouvons ici l'amorce de la thèse (§5.4) selon laquelle la dérèfification de la signification relie la vérité des jugements aux conditions de l'application correcte des concepts (normes de signification). Aussi, Brandom adhère-t-il d'emblée à ce type de relation tout au long de *Rendre explicite*, et établit un lien similaire entre la vérité des propositions et les conditions de l'application correcte des concepts. Brandom, 2011, p. 1046, l'exprime ainsi : « L'un des défis centraux d'une explication des normes conceptuelles comme implicites à la pratique sociale consiste [...] à rendre compte de l'émergence d'une telle notion objective de correction et d'adéquation des *affirmations et des applications de concepts*. », « Mais alors, comment faut-il comprendre le fait que la vérité de tels *contenus propositionnels*, ou que l'*application correcte* de tels *contenus conceptuels*, renvoie à une façon donnée dont les choses peuvent être ? » (Ibid., p.1047), « La correction objective des *affirmations (leur vérité)* et de l'*application de concepts* est identifiée à ce à quoi on adhère à partir de ce point de vue privilégié. », (Ibid., p. 1055): nos italiques.

<sup>955</sup> Ibid., p.1055: « Mais, traditionnellement, l'intersubjectivité a été comprise à la manière *je-nous*, qui se contraste entre les engagements d'un individu et les engagements de la *communauté* (collectivement), ou les engagements qui sont partagés par *tous* les individus (distributivement). Dans

espace est relatif à une relation symétrique je-tu dans laquelle l'auteur d'une assertion cherche à partager ses convictions vis-à-vis d'un allocataire, qui peut lui demander, par sa simple présence, ou de manière explicite, d'assumer les implications de ses affirmations ou d'en donner les raisons. Cet allocataire permet par là même au locuteur de distinguer ce qui lui apparaît comme vrai de ce qui est vrai relativement à la situation publique d'interlocution. L'une des sources de la théorie de la signification de Brandom peut se trouver dans une idée de Dummett<sup>956</sup> : la signification d'un terme est donnée par les circonstances et les conséquences de l'usage de ce terme (et cette signification s'acquiert et se possède en connaissant ces circonstances et conséquences). Deux aspects fondamentaux d'une expression ou d'un concept définissent ainsi sa signification (Brandom reprend ici des éléments de la sémantique fonctionnaliste de Sellars) :

- a) *les circonstances* ou les situations dans lesquelles il est correct d'appliquer, d'employer ou d'énoncer ce concept ou cette expression – ces circonstances peuvent être linguistiques, dialogiques ou non-linguistiques (circonstances perceptives) ;
- b) *les conséquences* appropriées de son application, de son utilisation ou de son énonciation, qui peuvent aussi être linguistiques ou non-linguistiques (action).

Brandom articule initialement la notion de rôle inférentiel d'une expression à partir des circonstances d'application (linguistiques et non-linguistiques) appropriées de cette expression, et des conséquences (linguistiques et non-linguistiques) appropriées de son usage, pour ensuite complexifier ces notions à partir de celles d'engagements et de permissions<sup>957</sup>. Il s'agit de comprendre ces circonstances et conséquences d'usage en en faisant les objets d'attitudes déontiques attribuées, reconnues (*acknowledged*) et assumées (*undertaken*) d'engagement et de permission. On retrouve ici la conception de Lewis du langage comme jeu de pointage au score<sup>958</sup> (*scorekeeping game*), à savoir que, non seulement le langage est

---

l'explication *grosso modo* davidsonienne [...] au contraire, l'intersubjectivité est comprise à la mode perspective *je-tu*, qui se concentre sur les engagements *contractés* par un marqueur interprétant d'autres personnes et les engagements *attribués* par ce marqueur à ces autres personnes [...] les explications *je-nous* postulent à tort l'existence d'une perspective privilégiée – celle du « nous », ou de la communauté. ».

<sup>956</sup> Dummett, 1973.

<sup>957</sup> Brandom, 2011, Chap.3. Le phénomène de la substitution permet aussi d'étendre l'approche inférentialiste des phrases aux expressions sub-phrastiques (Ibid., Chap.6). L'inférentialisme de Brandom s'applique aux expressions, aux prédicats, mais aussi par exemple aux indexicaux, aux noms propres et aux composants infra-phrastiques (Cf. §6.7.2).

<sup>958</sup> Brandom reprend cette expression et cette idée à David Lewis : penser une pratique linguistique sur le modèle d'un jeu sportif, comme celui du base-ball par exemple, où les joueurs sont à la fois acteurs

un moyen pour produire des raisonnements, mais aussi qu'il est structuré par des actes de discours qui engagent leurs auteurs dans un espace d'interlocution où les étapes de l'argumentation sont tenues sous la forme d'un score, c'est-à-dire d'une archive de « coups » joués par les différents interlocuteurs. Une telle position paraît d'autant plus intéressante qu'elle souligne les liens essentiels entretenus par la sémantique et la logique : comme le langage ne relève pas d'un code, mais d'un « jeu de raisons<sup>959</sup> », la signification des termes et des concepts est déterminée par les inférences tenues par les locuteurs. Autrement dit, Brandom souligne à juste titre que la signification n'est pas un donné ou un objet, mais qu'elle est construite à partir d'obligations et de permissions pragmatiques : le sens est structuré par le jeu normatif des interlocuteurs qui se comprennent comme engagés par des raisons et comme possédant certaines permissions vis-à-vis des assertions qui ont été faites par leurs allocutaires. Aussi, la nature du contenu de l'assertion est déterminée par les permissions et engagements dont il découle (relations d'antécédence) et par les permissions, les engagements et interdictions qu'il produit (relations de conséquences). Plus précisément, il faut distinguer les circonstances appropriées qui nous *engageraient* à appliquer l'expression, de celles qui nous *permettraient* d'appliquer l'expression, et les conséquences envers lesquelles nous sommes engagés en appliquant ou en employant l'expression des conséquences envers lesquelles nous gagnons une permission d'engagement en appliquant ou en utilisant l'expression. Le contenu envers lequel un locuteur est engagé en assertant que *p* est déterminé par les assertions qu'il peut faire pour justifier son engagement envers *p* (et qui portent sur les circonstances d'usage ou d'application de *p*), et par les assertions envers lesquelles il s'engage et qu'il a la permission de produire (ou qu'il n'a pas la permission de produire) à la suite de son engagement envers *p* (conséquences de *p*). Asserter que *p*, c'est, d'un côté accepter la responsabilité de montrer que nous avons la permission d'avoir cet engagement envers *p* (notamment en produisant de nouvelles assertions, et donc de nouveaux engagements), et de l'autre se porter garant de la vérité de *p*, en permettant aux autres

---

et arbitres, en s'attribuant mutuellement des obligations et des permissions. L'essentiel est donc ici de comprendre l'analogie avec le base-ball et avec l'acte d'établir ou de déterminer le score auquel on est parvenu au cours d'une partie. Il s'agit en fait de la pratique de compter le total des points acquis (d'établir le score) lors des différents moments d'une partie. Cette analogie avec le base-ball (que Brandom emprunte, en la modifiant, à Lewis, 1979) est précisément explicite dans Brandom, 2011, Chap.3, section IV.

<sup>959</sup> Brandom, 2011, p. 512 : « Nous sommes des créatures qui donnent et demandent des raisons pour ce que nous disons et faisons. ».

locuteurs de s'engager envers  $p$  (et/ou envers ses conséquences) en les déchargeant de la responsabilité de prouver que  $p$  (ces sujets peuvent en effet se déferer au locuteur initial).

Finalement, les capacités de compréhension conceptuelles résident dans la maîtrise du rôle inférentiel des contenus employés : comprendre ce que nous disons ou jugeons, ou encore ce que notre interlocuteur dit ou juge, consiste à être apte à déterminer ce qui suit d'une assertion ou d'un jugement, ce qui le prouverait et ce qui le réfuterait, les conséquences envers lesquelles nous nous engageons en le produisant, et les raisons qui justifieraient notre engagement<sup>960</sup>. Nous ne pouvons donc exprimer, juger ou asserter un énoncé sans le comprendre et sans l'endosser. Le comprendre, c'est savoir à quoi nous nous engageons en le formulant, ce qui s'ensuit et ce qui peut s'ensuivre, ce qui le corrobore ou peut le corroborer. L'endosser, c'est ainsi s'engager envers ce qui est dit<sup>961</sup>: être prêt à le défendre soit se porter garant de son autorité. Si par exemple, j'affirme que je suis souffrant, je permets à autrui de penser que je souffre (« entitlement »), et je suis obligé de reconnaître que j'éprouve de la douleur (« commitment<sup>962</sup> »).

Pour Brandom, là où les vérificationnistes se penchent davantage sur les circonstances d'application des concepts et des expressions pour définir leur signification, de manière dyadique, les pragmatistes classiques se concentrent sur les conséquences de cette application pour définir la signification. Brandom envisage ces deux dimensions comme complémentaires, en termes d'inférence, pour déterminer la signification. Par conséquent, la signification d'une expression est analysée d'une part, par rapport aux circonstances dans lesquelles il est approprié de l'utiliser dans un jugement (ou une inférence) et d'autre part, eu égard aux conséquences d'une telle utilisation. Comme le souligne Brandom:

Ce n'est que dans la mesure où l'on y fait appel pour expliquer les circonstances dans lesquelles les jugements et les inférences sont convenablement réalisés, ainsi que les conséquences convenables de tels jugements et inférences, que quelque chose qui est associé par le théoricien à des états ou expressions faisant l'objet d'une interprétation peut recevoir le statut d'interprétant *sémantique*, ou mérite le nom de concept théorique de *contenu*.<sup>963</sup>

<sup>960</sup> Brandom, 2002, p.95; 2009, p.101.

<sup>961</sup> Brandom, 2002, p. 360.

<sup>962</sup> Brandom, 2011, p. 317-320.

<sup>963</sup> Ibid., p. 291.

Comprendre un concept, ce n'est ainsi pas seulement connaître pratiquement (en manifestant un certain comportement) les circonstances dans lesquelles il serait approprié de l'appliquer, sans savoir ce qui s'ensuit de cette application. Réciproquement, connaître les conséquences de l'usage d'un concept ou d'une expression ne revient pas à connaître quelles sont les circonstances correctes de son application, qui font aussi partie de sa signification.

Comme nous l'avons laissé entrevoir ci-dessus, Brandom a défini les circonstances et conséquences inférentielles de l'usage d'une expression en termes de permissions et d'engagements inférentiels qui ne prennent forme qu'en étant attribués par d'autres participants au jeu linguistique. Les circonstances et les conséquences d'application d'un concept ou d'une expression peuvent néanmoins être non-linguistiques. Ce ne sont pas nécessairement des prémisses ou des raisons logiques ; il peut s'agir d'expériences perceptives. Mais, l'expression de cette expérience ne devient une assertion ou une éventuelle raison (venant appuyer une autre assertion) que lorsqu'elle peut être comprise en rapport avec ses engagements inférentiels. Un *perceptual report*<sup>964</sup> est un jugement signifiant seulement si son producteur est vu comme engagé – et est engagé – envers ses conséquences inférentielles. Les circonstances d'application ne sont pas nécessairement linguistiques, mais l'application à partir de ces circonstances n'est linguistique que si elle s'accompagne notamment d'engagements inférentiels. Si la signification d'une expression relève notamment de ses circonstances et des conséquences non-linguistiques de son application correcte, cette application, même correcte, n'est une application de l'expression que si elle est aussi inférentielle (il s'agit d'éviter ici l'écueil du mythe du donné) :

---

<sup>964</sup> Ibid., p. 203: « Supposez que les dispositions [...] se rencontrent également chez un perroquet entraîné à produire les mêmes sons face à une simulation identique. »; Brandom, 2009, p.56 : « [...] le thermostat et le perroquet ne comprennent pas leurs réactions, ces réactions ne signifient rien pour eux, même si elles peuvent signifier quelque chose pour nous [...] L'auteur de rapports, contrairement au thermostat et au perroquet, doit avoir le *concept* de température ou de froid [...] Pour le perroquet, « C'est rouge » n'est pas incompatible avec « C'est vert », et ne découle pas de « C'est pourpre » pas plus qu'il n'implique « C'est coloré ». Dans la mesure où la réaction susceptible de se répéter ne s'inscrit pas, pour le perroquet, dans les propriétés pratiques de l'inférence et de la justification, et donc de l'élaboration d'autres jugements, il ne s'agit pas du tout d'une question *conceptuelle* ou *cognitive*. ».

Les conséquences de l'application sont toujours elles-mêmes inférentiellement reliées au concept en question (bien que les inférences impliquées puissent inclure des inférences pratiques, dont les conclusions sont des engagements à agir). Les circonstances d'application n'ont pas besoin d'être elles-mêmes linguistiques. Pour le concept *rouge*, par exemple, elles incluent la présence de choses visiblement rouges. Cependant, même l'usage de concepts de cette espèce peut être considéré comme incluant des engagements inférentiels, comme, par exemple, la propriété d'appliquer les conséquences inférentielles de *rouge* – par exemple, *coloré* – à tout ce à quoi rouge est correctement appliqué<sup>965</sup>.

Il est également nécessaire que la correction de cette application soit validée par la communauté linguistique autrement dit qu'elle soit socialement reconnue et instituée.

Venons-en maintenant *au second postulat* de Brandom :

### 6.3. Sémantique et pragmatique

Brandom estime que la sémantique doit répondre à la pragmatique : « Semantics must answer to pragmatics<sup>966</sup> ». Une théorie de la signification n'est pertinente et valable que si elle rend compte et se base sur les caractéristiques de notre usage des expressions linguistiques. Nous retrouvons ici le slogan wittgensteinien selon lequel la signification d'une expression relève de son usage, c'est-à-dire, selon une certaine lecture, que nous ne pouvons pas séparer la sémantique de la pragmatique. Selon ce critère méthodologique, il n'est pas cohérent d'élaborer une théorie de la signification si cette dernière n'est pas simultanément une théorie des conditions d'usage correct des énoncés et ne répond pas au préalable à des questions comme : Comment pouvons-nous employer correctement les termes de notre langage ? Comment la signification des termes détermine-t-elle la manière dont ils doivent être employés ? Comment l'usage d'un concept s'accompagne-t-il de certains engagements<sup>967</sup> ? Dès lors, ce qu'un sémanticien associe à des états intentionnels ou à des expressions ne peut être qualifié de contenu sémantique que dans la mesure où il joue un rôle dans la pratique qui gouverne ces états et expressions. Il est certes possible d'associer toutes sortes d'objets abstraits à des langages formels, par exemple dans la théorie des modèles. Cependant, une

<sup>965</sup> Brandom, 2011, p. 252.

<sup>966</sup> Ibid., p.193, p. 288.

<sup>967</sup> Brandom, 1997a, p.193: « The concept of meaning is a normative concept. The theoretical explanatory job distinctive of concepts of meanings or conceptual content is to settle how it would be *correct* to use words or to apply concepts, how those meanings determine how they *ought* to be used, how those who use concepts with those concepts are *committed* thereby to apply them. ».

telle association n'est proprement sémantique qu'à la condition de déterminer comment ces symboles sont utilisés correctement. Par exemple, une interprétation du calcul des prédicats ne peut être qualifiée de sémantique que dans la mesure où elle permet de donner une représentation de la notion d'inférence valide, c'est-à-dire une représentation de ce qu'il est correct de faire. Bref, un concept ou une théorie sémantique n'est pertinent que si elle permet de rendre compte de ce que nous faisons ou devons faire lorsque nous parlons ou pensons. McDowell met en évidence l'intérêt de cette idée, comme suit :

« Semantic terms are not intelligible apart from how they pull their weight in enabling us to make sense of what speakers *do*, on the way we make sense of rational behaviour in general.<sup>968</sup> ».

Brandom souligne donc le caractère normatif de la pragmatique, celle-ci relevant de ce qu'il est correct ou incorrect de juger et d'inférer dans un cadre interlocutif. Une telle correction dépend ainsi en dernière analyse des obligations et des permissions qui sont liées aux actes de discours tenus dans un jeu de pointage du score; en conséquence, ces obligations et permissions pragmatiques sont reflétées par les inférences auxquelles est associé cet acte. Dès lors, la signification d'un énoncé doit être analysée relativement au rôle inférentiel qu'elle possède dans un cadre interlocutif et argumentatif.

Par conséquent, l'une des conséquences de cette exigence méthodologique est que l'inférence sera d'abord envisagée comme un mode d'action<sup>969</sup>, et non pas comme une relation logique : sa logique repose sur la signification et la portée de ce que c'est que de faire une inférence. Avec Brandom, il paraît donc judicieux d'appréhender les significations comme des usages, plus précisément comme des rôles inférentiels. Aussi, une théorie de la nature et de l'usage des concepts doit être élaborée à partir de l'activité de raisonner, et non pas à partir de celle de représenter<sup>970</sup>. La frontière entre la sémantique et la pragmatique ne peut donc pas être comprise de manière relative à des objets préexistants aux normes discursives : elle doit au contraire être appréhendée à partir de l'usage et de l'articulation

<sup>968</sup> McDowell, 1997, p.157.

<sup>969</sup> MacFarlane, 2010, p. 85: « [...] can be explicated in terms of proprieties for performing a certain kind of action – the action of inferring, of drawing a conclusion from premises. ».

<sup>970</sup> Brandom, 2009, p. 35.

inférentielle de ces mêmes normes<sup>971</sup>. D'un point de vue inférentialiste, il convient de parler de *pragmatisation de la sémantique*. En conséquence, nous acceptons une forme de *pragmatisation méthodologique* au sujet de la relation entre la sémantique et la pragmatique, au sens où le but des attributions sémantiques entre locuteurs consiste bien, *in fine*, à expliquer la dimension normative des états intentionnels de ces locuteurs et de leurs actes de discours.

Aussi, il n'est pas surprenant de voir le contenu conceptuel défini en termes de conditions d'usage, celles-ci étant déterminées par les pratiques linguistiques. Cette position relève d'ailleurs d'un pragmatisme sémantique, pour lequel les propriétés sémantiques des contenus sont expliquées à partir des propriétés d'usage des contenus. Les concepts ne sont pas seulement appliqués à partir de normes de correction ; ils sont définis par ces normes. Brandom adhère donc à un pragmatisme méthodologique (la sémantique doit répondre à la pragmatique), mais aussi à :

- Un pragmatisme sémantique posant que la signification doit être expliquée par l'usage ;
- Un pragmatisme sur les normes, posant que l'autorité cognitive d'un énoncé ou d'une théorie est instituée par des pratiques et interactions sociales ;
- Un pragmatisme normatif, posant que les normes qui ont la forme de règles (explicites) ne sont intelligibles qu'à partir d'un arrière-plan de normes implicites ;
- Un pragmatisme linguistique posant qu'il ne peut y avoir de pensées sans pratiques linguistiques antécédentes (le jugement est une intériorisation de l'assertion<sup>972</sup>).

Le troisième objectif de Brandom est alors le suivant :

#### 6.4. Rôle expressif de la logique

Aboutir à l'élaboration d'une théorie auto-référentielle du discours, notamment à partir d'un usage expressif de la logique : il s'agit de proposer un modèle explicatif de nos pratiques

<sup>971</sup> Brandom, 2011, p.207 : « inférer, c'est faire un certain type de choses [...] Le statut de l'inférence (comme ce qui peut être fait d'après ces normes) tient la promesse de garantir une relation appropriée entre la pragmatique, l'étude des pratiques, et la sémantique, l'étude des contenus correspondants. »

<sup>972</sup> Dummett, 1973, p. 362; cité par Brandom, 2011, p. 308: « Nous nous sommes constamment opposés à la conception selon laquelle l'assertion serait l'expression d'un acte interne de jugement; c'est bien plutôt le jugement qui est l'intériorisation de l'acte externe de l'assertion. ».

discursives pouvant rendre compte de son propre sens et étant issu de ce dont il rend compte. La théorie de Brandom cherche à rendre explicite la structure de nos pratiques discursives, dans lesquelles elle s'inclut. *Bien qu'il n'y ait pas de point de vue de nulle part sur nos pratiques linguistiques qui puisse les décrire ou les fonder, il est pourtant envisageable de les expliquer de l'intérieur, car il est constitutif de ces pratiques de pouvoir se rendre explicites et de prendre conscience d'elles-mêmes. Il s'agit donc d'expliquer le langage de l'intérieur.* Par conséquent, pour construire et pour comprendre une théorie de nos pratiques discursives, il faut déjà être dans ces pratiques discursives.

La logique dans le cadre d'une théorie expressive est utilisée comme un outil nous permettant d'atteindre une conscience de nous-mêmes en tant qu'agents sémantiques. L'expressivisme s'oppose ainsi à une conception exclusivement formaliste de la logique<sup>973</sup>. Le vocabulaire logique, par le biais de conditionnels, de quantificateurs, ou de concepts sémantiques tels que « réfère », « croit que », « désire que », « est vrai », possède un rôle expressif : il contribue à rendre explicite, sous la forme d'une affirmation, les structures et pratiques implicites à partir desquelles les contenus sémantiques émergent en étant attribués en pratique par des êtres linguistiques qui maîtrisent implicitement des règles d'inférence. Par exemple, nous le verrons par la suite, les inférences matérielles sont constitutives des contenus des concepts. Ce ne sont pas des inférences logiquement c.-à-d. formellement valides ; elles sont maîtrisées et reconnues implicitement par les locuteurs. Les locuteurs reconnaissent ainsi que passer de « Ceci est un chat » à « ceci est un mammifère » mène bien à une bonne inférence, celle-ci pouvant se formuler comme suit : « Si c'est un chat, alors c'est un mammifère ». Le vocabulaire logique des conditionnels<sup>974</sup> permet donc d'énoncer ces relations inférentielles matérielles, voire de les généraliser. Par exemple, en reprenant l'énoncé ci-dessus, nous pourrions dire : pour tout  $x$ , si  $x$  est un chat, alors  $x$  est un mammifère. Ces inférences qui définissent un contenu ou un jugement (« Car donner une raison c'est toujours exprimer un jugement : faire une affirmation »<sup>975</sup>) peuvent alors être rendues explicites comme contenu d'un jugement<sup>976</sup> afin d'être enseignées, critiquées, contrôlées ou révisées pour enrichir ou ajuster nos pratiques linguistiques. Grâce à l'usage de

<sup>973</sup> Brandom, 2009, p.37-38.

<sup>974</sup> Ibid., p.87.

<sup>975</sup> Ibid., p.87.

<sup>976</sup> Brandom, 2011, p.187.

la logique, nous rendons explicite dans la pensée ou le discours ce qui était implicite dans la pratique et dans nos usages d'autres contenus explicites<sup>977</sup>.

Cet usage expressif de la logique est présent à au moins deux niveaux :

À un premier niveau des pratiques implicites, toute activité de conceptualisation à partir des contenus propositionnels est expressive. Énoncer quelque chose, dans le sens de dire quelque chose soit asserter un contenu, c'est produire un énoncé qui est inférentiellement signifiant soit qui s'insère dans ce jeu d'offre et de demande de raisons<sup>978</sup>. Toute dimension représentationnelle, dite ou pensée, est expressive et explicite dans la mesure où elle s'inscrit socialement dans des pratiques inférentielles<sup>979</sup>, à partir des attitudes déontiques (attributions de statuts) de la part de la communauté. Aussi, penser, c'est être capable de justifier ses pensées, de les corriger, d'en rendre compte et de les attribuer à autrui à partir de normes sociales. Penser un contenu, c'est aussi avant tout être capable de faire quelque chose avec lui ; ce n'est pas simplement représenter ou disposer d'une représentation de quelque chose – ce qui fait que généralement, cette conception expressive prend souvent ses distances avec une théorie représentationnaliste dénotationnelle, même externaliste, de la pensée. L'usage de contenus représentationnels, dans le langage et la pensée, est d'abord expressif, et non référentiel.

Ce qui est exprimé en étant conceptualisé possède un contenu en vertu de son articulation inférentielle, dont les propriétés sont implicites dans l'expression. Le rôle de la logique est donc, à un niveau supérieur, de pouvoir expliciter l'articulation inférentielle et pragmatique de ce que nous disons ou exprimons.

Un vocabulaire mérite d'être appelé «logique» dans le cas où il sert à rendre *explicites*, comme contenu d'une affirmation, les règles de convenance relatives à l'usage de l'expression qui demeureraient autrement *implicites* à la pratique, en particulier les règles de convenance en vertu desquelles il possède le contenu conceptuel qui est le sien.<sup>980</sup>

Ainsi, exprimer, c'est d'un côté, dire et signifier inférentiellement (asserter quelque chose), et de l'autre, être capable, en tant que créature logique et rationnelle, de rendre

<sup>977</sup> Brandom, 2009, p.16, p.64.

<sup>978</sup> Ibid., p. 33.

<sup>979</sup> Brandom, 2011, p. 882.

<sup>980</sup> Ibid., p. 244; 2009, p. 65.

explicite cette articulation logique au moyen d'outils logiques. Selon Brandom, une communauté peut formuler des énonciations signifiantes, discursives sans être nécessairement capable de préciser et de détailler comment les contenus conceptuels qu'elle attribue implicitement peuvent être conférés.

Aussi, la conception expressiviste de Brandom semble provenir de son adhésion à l'idée romantique selon laquelle l'esprit aurait plutôt la propriété d'éclairer nos pratiques de l'intérieur que celle de fonder nos pratiques en copiant la structure du monde, position à l'origine de son rejet du représentationnalisme cartésien<sup>981</sup>. Quant à l'irrationnalisme romantique, il semble être pondéré par une conception inférentielle de la raison. Bref, la position de Brandom se base entre autres sur une reprise de l'inférentialisme des Lumières (Leibniz) et de l'expressivisme romantique<sup>982</sup>. Son rejet de l'empirisme explique sa position pour une sémantique rationaliste dans laquelle les contenus propositionnels des énoncés sont décrits comme pouvant servir de prémisse ou de conclusion à des raisonnements. L'idée *expressiviste* selon laquelle le vocabulaire logique peut rendre explicite nos pratiques linguistiques de l'intérieur est une idée centrale chez Brandom. La raison est ce qui nous permet de prendre conscience et de réfléchir sur ce que nous faisons implicitement lorsque nous pensons, parlons ou agissons<sup>983</sup>. Le vocabulaire normatif (« doit », « devrait », etc.) vise à rendre explicite la structure du raisonnement pratique comme le fait le vocabulaire logique pour les énonciations sémantiques.

Il n'y a donc pas de point de départ neutre, hors de nos pratiques, et il n'y pas non plus de point d'arrivée qui transcenderait nos pratiques. Aussi, l'origine de nos pratiques linguistiques ne peut être décrite et pensée qu'à partir du normatif. Brandom soutient en effet que :

---

<sup>981</sup> Brandom, 2009, p. 15 : « À l'image de l'esprit comme *miroir* des Lumières, le romantisme a opposé une image de l'esprit comme *lampe*. ».

<sup>982</sup> Brandom, 2011, p.209 :

Les romantiques sont peut-être mieux connus pour leur rejet non seulement du représentationnalisme des Lumières [...], mais aussi pour leur rejet de l'importance que ce représentationnalisme accorde à la raison [...] Le recul des romantiques [...] en vint à reposer sur une considération forte du sentiment, de l'empathie inarticulée, et de l'enthousiasme. Hegel a vu dans la notion inférentielle de contenu un moyen de rejoindre les romantiques dans leur rejet du représentationnalisme, tout en se distanciant de leur hostilité à la raison. Le résultat est une synthèse de l'inférentialisme des Lumières et de l'expressivisme romantique.

<sup>983</sup> On note la dimension hégélienne de la pensée de Brandom.

Ainsi la tentative théorique de localiser la « source » de la dimension normative dans le discours nous reconduit directement à nos propres pratiques implicitement normatives. La structure de ces pratiques peut être élucidée, mais toujours à l'intérieur de l'espace normatif, c'est-à-dire à partir de – et dans – nos pratiques normatives qui consistent à donner et à demander des raisons<sup>984</sup>.

L'argument est à la fois épistémologique et ontologique. Comme le fait remarquer Gibbard, l'approche phénoménaliste<sup>985</sup> de Brandom renvoie à lui-même:

« the content of the theory is to be explained by explaining the configuration of normative attitudes that constitutes *holding* the theory. This expressivism applies to itself, and so iterates<sup>986</sup>. ».

Autrement dit, d'un point de vue épistémologique, les approches discursives et normatives n'existent qu'à partir des attitudes qui les soutiennent et les reconnaissent. La discursivité étant indissociable des conditions publiques d'usage et d'attribution de statuts pragmatiques, il est impossible pour une telle approche de vouloir concevoir ce qui est avant la norme de ces conditions publiques, puisqu'elle est elle-même un produit de ces conditions. Une origine non linguistique des performances linguistiques ne peut être théoriquement pertinente qu'en étant au préalable définie et envisageable au sein de nos pratiques linguistiques. Brandom insiste pour souligner que, de haut en bas (« all the way down»), ce sont toujours des normes que nous trouvons. De plus, l'argument ontologique met en exergue le caractère non-réductible de la norme soit le fait que la norme n'existe qu'en étant reconnue comme telle par une communauté. L'essentiel, pour Brandom, n'est ainsi pas de réduire la norme à quelque chose de non normatif, mais bien de montrer comment des pratiques sociales peuvent générer des pratiques normatives implicites, mais aussi explicites, et comment, dès lors, la norme – la norme de nos pratiques sociales – peut s'explicitier. Les normes ne sont pas des objets dans l'ordre causal de la réalité<sup>987</sup>, tout comme les engagements

<sup>984</sup> Ibid., p.1138.

<sup>985</sup> Gibbard qualifie la stratégie de Brandom de « phénoménaliste » (*phenomenalist*) : il cherche ainsi à souligner l'idée que nos normes sont le produit de nos sentiments, émotions et états d'esprit (sans tomber dans une forme de *subjectivisme*) tout en soutenant la thèse selon laquelle les normes sont déjà dans l'œil du spectateur, si bien qu'il est impossible d'aborder la question de savoir ce que sont des normes implicites, sans se poser la question de leur *reconnaissance* dans la pratique. Nous retrouvons cette idée chez Brandom, 2011, p. 95 : « L'explication [...] devra rendre compte de l'attitude pratique qui consiste à *tenir* quelque chose pour correct-selon-une-pratique, pour expliquer ensuite, en en appelant à ces attitudes, quel est le statut de ce qui *est* correct-selon-une-pratique. ».

<sup>986</sup> Gibbard, 1996, p.701.

<sup>987</sup> Brandom, 2011, p.1101.

et permissions, qui n'existent pas dans le monde naturel : elles sont plutôt les produits des attitudes pratiques par lesquelles nous considérons que quelqu'un est *engagé à* ou *a la permission* de faire quelque chose<sup>988</sup>. Gibbard<sup>989</sup> note ainsi que le *phénoménisme* de Brandom, pour lequel il n'existe pas autre chose dans la norme que les propriétés issues de nos institutions de la norme, pourrait s'apparenter à une certaine forme d'expressivisme en philosophie du langage, qui exprime la signification d'un énoncé à partir de l'état d'esprit du locuteur qu'elle exprime. Cet état d'esprit est ici un statut déontique (et non pas un état mental interne), qui est lui-même le produit d'une interprétation effective<sup>990</sup>.

Le pragmatisme normatif de Brandom vise aussi à éviter deux stratégies explicatives de la norme : un régularisme, qui identifierait les normes à des règles explicites, explicitement suivies, interprétées et appliquées; et un régularisme<sup>991</sup>, qui identifierait les normes à de simples régularités comportementales. L'approche régulariste soulève, comme nous l'avons vu<sup>992</sup>, un problème de régression à l'infini dans l'application de la norme; la conception régulariste<sup>993</sup>, ne réussit pas, quant à elle, à cerner la spécificité de la norme par rapport à la simple régularité comportementale, et se heurte au problème du découpage arbitraire : une régularité à elle seule est compatible avec une infinité de normes (et les régularités sont partout). Le *phénoménisme* de Brandom l'amène à envisager la norme en fonction des attitudes normatives pratiques qui évaluent ou instituent implicitement des formes de comportement comme étant correctes ou incorrectes.<sup>994</sup>

Ces trois aspirations sont au fondement de la théorie de Brandom de l'usage et de l'application correcte des normes de signification, qu'il convient maintenant de préciser. Son approche peut se formuler ainsi : la dimension représentationnelle ou référentielle du contenu conceptuel se définit à partir de la dimension inférentielle du contenu propositionnel (qui est la forme primitive du contenu conceptuel). Cette dimension inférentielle du contenu propositionnel n'est elle-même que le reflet des propriétés pragmatiques de ce contenu, plus précisément des propriétés qui stipulent ce que nous pouvons faire ou ne pas faire quand nous

<sup>988</sup> Brandom, 2011, p.109.

<sup>989</sup> Gibbard, 1996.

<sup>990</sup> Brandom, 2011, p. 538, n. 1 : Brandom n'hésite pas à rapprocher son *phénoménisme* de l'*expressivisme* de Gibbard.

<sup>991</sup> Ibid., p. 97.

<sup>992</sup> Cf. Chap. I, §1.3 ; Chap. III, §3.3.

<sup>993</sup> Cf. Chap. II, §2.3.4; Chap. III, §3.3.

<sup>994</sup> Heath, 2001.

assertons ce contenu. Par conséquent, la dimension pragmatico-normative de l'assertion ou de la pensée détermine la nature sémantique inférentielle de son contenu. Ces propriétés sont des propriétés socialement instituées.

De manière schématique, nous pourrions reproduire la structure explicative suivante : soit un locuteur  $X$  qui asserte un contenu  $p$ .

Des attitudes normatives des locuteurs (participants aux pratiques discursives)

*Fonction* : Attribuent, en pratique soit implicitement ou explicitement à  $X$ , suite à son assertion que  $p$

*Pragmatique normative*

Des obligations à respecter les normes d'application et d'usage de l'assertion : ce que  $X$  doit, peut, ne peut pas faire en ayant asserté  $p$

*Fonction* : définissent (par transposition)

*Sémantique inférentielle*

Le rôle inférentiel du contenu  $p$  de l'assertion

*Fonction* : détermine le contenu  $p$  de l'assertion (et, notamment sa dimension représentationnelle)

*Dimension représentationnelle du discours et de la pensée*

De bas en haut, nous avons donc : le contenu de ce que je dis et de ce que je pense est déterminé par son rôle inférentiel, relevant lui-même des propriétés pragmatico-normatives de mon assertion de ce contenu; ces propriétés expriment ce que je dois et ce que je peux reconnaître ou faire inférentiellement en ayant énoncé ce contenu; ces propriétés n'existent sous forme de statuts déontiques qu'en étant attribués, explicitement ou implicitement, par les autres participants à la pratique discursive. C'est donc bien à partir de ce que les membres d'une communauté font, peuvent et doivent faire lorsqu'ils produisent ou attribuent des assertions que nous pouvons définir le contenu conceptuel de ces assertions. Par conséquent, les normes pragmatiques qui contribuent à définir le contenu conceptuel de ces mêmes assertions, dépendent des capacités des individus à exercer implicitement des attitudes

normatives<sup>995</sup>, par lesquelles ils s'attribuent mutuellement des statuts déontiques (permissions, engagements, obligations) qui déterminent le contenu de ce qu'ils disent. Comme nous l'avons mentionné plus haut, les assertions et les jugements sont les performances principales qui peuvent être les objets de deux statuts déontiques essentiels : la permission (*entitlement*) et l'engagement (*commitment*). Les actions (soit les engagements pratiques<sup>996</sup> dans le vocabulaire de Brandom) sont aussi des performances qui sont l'objet de permissions et qui fournissent des permissions (et non des raisons, comme peuvent le faire les assertions). Quant aux croyances, en étant exprimées par des assertions (qui sont des engagements), elles sont aussi des engagements. Nous avons donc des engagements assertionnels et des engagements doxastiques. Croire que *p*, c'est avoir un engagement (et l'avoir avoué, ou être disposé à l'avouer) envers *p*; dire que *p*, c'est reconnaître (*acknowledge*) ou assumer (*undertake*) un engagement envers *p*. Analysons à présent chacune des étapes du cheminement (de bas en haut) afin de comprendre la portée d'une normativité sociale des normes de signification.

#### 6.5. De la sémantique représentationnelle à la sémantique inférentielle

Le contenu d'une assertion doit donc être compris à partir de ses propriétés inférentielles soit à partir des inférences que nous pouvons, mais également que nous devons faire, en tant que membre d'une pratique discursive. L'inférence est ainsi une pratique dont la correction n'est pas évaluée à partir du rapport entre les contenus, une réalité objective et le respect des règles formelles, mais à partir de contenus et de règles inférentielles matérielles<sup>997</sup>. Un contenu (et son usage<sup>998</sup>) se définit à partir de la dimension inférentielle, dans le raisonnement (comme offre et demande de raisons) de ce qui l'exprime, soit l'assertion :

<sup>995</sup> Nous rencontrons trois types d'attitudes normatives: attribuer/imputer, assumer et reconnaître (« *attributing/ascribing, undertaking and acknowledging* »).

<sup>996</sup> Brandom, 2011, p. 491.

<sup>997</sup> Brandom, 1983, p.640.

<sup>998</sup> Brandom, 2009, p. 56: « avoir, pour une réaction, un contenu *conceptuel*, c'est simplement pour elle jouer un rôle dans le jeu *inférentiel* des affirmations et de l'offre et de la demande de raisons. Saisir ou comprendre un tel concept, c'est avoir une maîtrise pratique sur les inférences dans lesquelles il est impliqué – savoir, au sens pratique d'être capable de distinguer (une sorte de *savoir-comment*) ce qui découle de l'applicabilité d'un concept et ce dont il découle. ».

« The conceptual role played by the original claim in the social practices of the community is determined by what further assertions that community would accept as appropriate justifications of it, and what assertions they would take to license or justify<sup>999</sup>. ».

Nous verrons alors par la suite comment nos pratiques sociales instituent les inférences qui définissent le contenu. Ce qu'il est important de noter est que la signification pragmatique du contenu de l'assertion ne préexiste pas à la dimension inférentielle de l'assertion :

« The meanings do not determine the appropriate inferences but what inferences are socially appropriate determines the meanings of the sentence involved in those sentences<sup>1000</sup>. ».

La théorie de Brandom est donc bien holiste<sup>1001</sup> dans la mesure où les contenus des assertions sont déterminés par les relations inférentielles qu'ils ont avec d'autres contenus et d'autres actions : il ne peut y avoir de contenus ou de statuts sans d'autres contenus ou statuts. Néanmoins, d'un point de vue contextuel, certaines inférences sont plus appropriées que d'autres quand il est question de définir les conditions de possession et l'identité inférentielle du contenu jugé ou asserté<sup>1002</sup>. C'est donc aussi en contexte que sont déterminées les capacités et les incapacités inférentielles des locuteurs, et donc l'effectivité et la réalité de leurs assertions et jugements – cela notamment en rapport avec leurs capacités inférentielles passées (pour une personne souffrant de troubles neurologiques récents par exemple), et parfois également potentielles (par exemple, le nouveau membre d'un comité syndical ou autre, ne maîtrisant pas encore le réseau inférentiel de concepts utilisés). Nous tenterons ainsi de comprendre la source de ces inférences appropriées et la raison pour laquelle Brandom refuse de se limiter aux inférences strictement formelles pour rendre compte de la dimension normative des relations d'engagement et d'autorisation. Pour ce faire, nous devons nous tourner vers la dimension pragmatique et sociale de l'usage des contenus assertés.

<sup>999</sup> Brandom, 1983, p. 642.

<sup>1000</sup> Ibid., p. 644.

<sup>1001</sup> Brandom, 2011, p.205: « Mais la notion inférentielle de contenu sémantique est essentiellement holiste [...] On ne peut rien savoir du rôle d'un contenu dans des inférences si on ne sait rien des rôles d'autres contenus dans des inférences, des contenus qui pourraient être inférés à partir de lui, ou desquels il pourrait être inféré. »; Brandom, 2009, p.23.

<sup>1002</sup> Brandom, 2011, p.1118.

### 6.6. De la sémantique inférentielle à la pragmatique normative

Les propriétés inférentielles qui définissent la nature sémantique d'un contenu sont fondamentalement des propriétés pragmatiques socialement instituées. La dimension inférentielle du contenu sémantique dérive des statuts pragmatiques associés à la possession de ce contenu. Le contenu dérive de son attribution, de sa reconnaissance au sein d'une communauté discursive, fonctionnant à partir d'activités conceptuelles provenant également d'autres attributions et pratiques, normatives dès le début. Aussi, dans la mesure où les contenus sont des statuts normatifs qui viennent à être institués par des attitudes normatives, la normativité de la signification est « expliquée » par quelque chose qui est déjà en lui-même intrinsèquement normatif et par conséquent se pose le problème d'une certaine circularité explicative, certes parfaitement assumée par Brandom. Nous y reviendrons par la suite.

En effet, les pratiques sociales implicites à l'origine de la signification incluent essentiellement des attitudes déontiques (entretenu par les membres de la communauté) qui attribuent deux types principaux de statuts déontiques ou normatifs à partir desquels les formes d'inférences sont dérivées : les engagements (*commitments*) et les permissions (*entitlements*). L'individu peut donc avoir un engagement envers certaines choses, et avoir la permission de faire certaines choses. Ces deux statuts sont normatifs : il est approprié ou correct de faire ce que l'on peut faire ou ce envers quoi nous sommes engagés. Ces deux statuts n'existent ainsi qu'en étant attribués, implicitement (*attribute*) ou explicitement (*ascribe*) par des membres de la communauté qui exercent des attitudes normatives. L'attribution par un interprète et la reconnaissance (*acknowledgment*) ou l'exercice (*undertaking*) par l'interprété de ces statuts (engagements et permissions) sont les deux mouvements de base de la procédure de pointage déontique du score (*deontic scorekeeping game*). La normativité des engagements et autorisations (et de ce fait, la normativité de la signification) est ainsi instituée par un autre joueur (*the scorekeeper*<sup>1003</sup>) qui comptabilise les points et qui suit le cours (*keep track*) de ce qui est dit par un participant au jeu en lui attribuant des engagements et des permissions. Par conséquent, un locuteur qui se dit engagé à faire telle ou telle action, autorise le marqueur à lui imputer cet engagement. C'est finalement cette pratique qui consiste à reconnaître et à assumer des engagements et

<sup>1003</sup> Il s'agit du joueur qui comptabilise les points dans un match. Il joue le rôle d'un arbitre qui marque les points effectués par chacun des protagonistes. Nous traduirons par marqueur.

autorisations en se les faisant attribuer par un autre joueur marquant les points qui, lui-même, reconnaît et assume à son tour des engagements et autorisations identiques ou différents, qui constitue le pointage déontique. Le pointage déontique se présente donc comme essentiellement social. Dans une conversation, la comptabilité des points correspond juste aux engagements et autorisations associés à chaque participant. Chaque membre de la conversation marque ainsi les points de chacun des participants (incluant lui-même) et les membres d'une pratique discursive suivent les coups des engagements et autorisations de chacun des autres joueurs. Chaque fois que l'un des participants assume (adopte implicitement), reconnaît (adopte explicitement) ou attribue (impute à un autre joueur) un engagement ou une autorisation, cela change le pointage. Brandom fonde ainsi sa théorie pragmatique sur l'idée que l'emploi d'une expression linguistique représente la manière de changer le score d'une conversation. Aussi, Brandom, en mettant l'accent sur la dimension normative du contenu, définit la portée pragmatique d'un acte de discours comme la manière dont une telle pratique de pointage devrait affecter le cours de la conversation dans laquelle s'inscrit cette même performance linguistique. C'est la dimension inférentielle de ces contraintes pragmatiques qui définit le contenu de ce qui est dit. Cependant, l'attribution de ces statuts n'est pas une simple régularité ou une simple disposition, elle se traduit par une activité réglée, reposant sur des normes implicites envers lesquelles nous nous engageons, dans laquelle sont institués et reconnus ces statuts, de manière holiste : par exemple, chaque engagement attribué ou reconnu s'accompagne de la reconnaissance implicite d'autres engagements présumés ou impliqués. Une pratique sociale est donc une pratique discursive dont peut émerger la signification que si elle inclut ce jeu de pointage (incluant l'adoption d'au moins deux statuts normatifs : l'engagement et la permission). Ces précisions étant faites, il s'agit d'expliquer comment le contenu d'une assertion ou d'une croyance relève du rôle inférentiel de l'assertion ou de la croyance, rôle inférentiel lui-même déterminé à partir des statuts normatifs (engagements, permissions) qui sont attribués au locuteur par la communauté.

Les pratiques linguistiques par le biais desquelles les contenus sont institués et attribués se composent essentiellement de trois dimensions pragmatiques en interaction constante : celle relatif aux statuts déontiques adoptés<sup>1004</sup> (parce qu'attribués) quand l'on fait une

---

<sup>1004</sup> Ibid., p. 333-334.

assertion (engagements, permissions à l'engagement, incompatibilité); celle relatif au caractère personnel ou interpersonnel de l'attribution de ces statuts<sup>1005</sup>; et celle relative à la responsabilité cognitive<sup>1006</sup> à l'égard du contenu de ces jugements et assertions. Quant à ce dernier point, nous voulons juste souligner que toute assertion ou jugement que  $p$  est à chaque fois – nécessairement et simultanément – la reconnaissance (*undertaking; acknowledging*) de la part du locuteur d'un engagement qu'il a vis-à-vis de ce contenu propositionnel et de sa vérité.

Nous nous pencherons ici sur la première dimension pragmatique, composée de trois types de normes pragmatiques dont découlent trois types de relations inférentielles définissant le contenu propositionnel. En énonçant une phrase ou en formulant un jugement, nous nous retrouvons engagés envers au moins trois statuts déontiques dont l'objet est déterminé non par la signification du contenu, mais pas ces statuts eux-mêmes, tel qu'un interprète pourrait nous les attribuer. Ces trois types de statuts déontiques se présentent comme suit :

(1) *L'engagement* : asserter que  $p$ , c'est s'engager envers  $p$  et s'engager envers ce que nous pouvons déduire, logiquement ou matériellement, du contenu asserté. L'engagement est, d'une certaine manière, un statut d'obligation : asserter et (donc) s'engager envers  $p$ , c'est reconnaître explicitement (être engagé envers) ce qu'implique déductivement  $p$ . Par conséquent, si nous reconnaissons explicitement que  $q$  est une déduction de  $p$  et que nous sommes aussi engagés à  $p$ , alors nous devons aussi reconnaître un engagement à  $q$ .

(2) *La permission* : asserter que  $p$ , c'est aussi être autorisé, et non obligé de faire telle ou telle action comme dans (1), à s'engager envers d'autres contenus, que le contenu jugé ou asserté supporte inductivement, pour lesquels il peut servir de raison, moyennant d'autres contenus et conditions d'arrière-plan. Il est permis d'asserter  $q$  à partir de  $p$  si nous pouvons inférer  $q$  ( $p$  appuie inductivement  $q$ ,  $p$  peut servir de raison pour  $q$ ). Par conséquent, si nous sommes autorisés à l'égard des prémisses, nous devrions l'être aussi à l'égard de sa conclusion.

---

<sup>1005</sup> Ibid., p. 335.

<sup>1006</sup> Ibid., p. 336.

(3) *L'incompatibilité* : affirmer ou juger que  $p$ , c'est aussi exclure ce qui est *incompatible* avec ce contenu. Autrement dit, deux jugements sont incompatibles si un engagement envers l'un d'entre eux interdit d'être autorisé à s'engager envers l'autre.

Par conséquent, toute pratique sociale doit, pour être considérée comme discursive, reconnaître ces trois statuts normatifs qui régissent ce que nous pensons ou disons. Ces relations sont d'abord des normes pragmatiques qui sont ensuite expliquées en termes de relations inférentielles entre les contenus assertés. Ces relations inférentielles s'apparentent certes aux formes d'inférence classique, à ceci près qu'elles ne se fondent pas sur des propriétés des contenus, mais sur les propriétés normatives de l'acte d'affirmer un contenu. Des normes pragmatiques de l'affirmation dérivent les propriétés inférentielles des contenus des affirmations.

(1) L'engagement se traduit, sur un niveau sémantique et inférentiel, dans des relations d'implication sémantique, et notamment dans des inférences déductives. Quand j'affirme ou je juge « cette tomate est rouge », je m'engage aussi à l'égard de la vérité du contenu « cette tomate n'est pas verte ». Je suis engagé vis-à-vis d'un contenu parce que je me suis engagé envers un autre contenu. Je me suis également engagé envers ce qui se situe – toujours par relation déductive – en amont de mon affirmation, par exemple le contenu « il y a une tomate ». Ces inférences sont des inférences qui préservent les engagements : un engagement envers  $p$  est conservé lorsque je m'engage envers  $q$ , et inversement.

(2) La permission se traduit dans des relations inférentielles de justification, de soutien dans des inférences inductives. Quand j'affirme « ceci est une tomate mûre », je peux me baser sur ce contenu pour arriver à d'autres contenus, par le biais de certaines conditions du monde présumées, comme « si elle est exposée au soleil, elle rougira ». Ici, je suis autorisé à m'engager à l'égard d'un nouveau contenu parce que je suis engagé envers un autre contenu et que je me réfère aussi à certaines conditions du monde, que je dois être en mesure de justifier<sup>1007</sup>.

(3) Les relations pragmatiques d'exclusion se traduisent dans des relations d'incompatibilité entre contenus :  $p$  est incompatible avec  $q$  si un engagement envers  $p$

---

<sup>1007</sup> Brandom, 1983, p. 641.

interdit ou exclut un engagement envers  $q$ <sup>1008</sup>. L'engagement envers  $p$  exclut l'engagement envers non- $p$  ou tout contenu conceptuel duquel on pourrait inférer non- $p$ .

Nous envisageons la relation pragmatique d'exclusion comme un ensemble de paires ordonnées de phrases qui sont incompatibles. Les relations d'engagement et de permission peuvent chacune d'elles être considérées comme des ensembles d'antécédents et de conséquents inférentiels. Un énoncé est l'antécédent de certains conséquents (linguistiques ou non linguistiques) qu'il produit de manière déductive ou inductive, et le conséquent (obtenu de manière déductive ou inductive) de certains antécédents (linguistiques ou non linguistiques). Les antécédents d'un énoncé sont les énoncés à partir desquels nous inférons l'énoncé en question (nommons-le  $p$ ) et les conséquents sont les énoncés à partir desquels nous inférons d'autres énoncés. Les antécédents de  $p$  forment un ensemble dont les membres sont des ensembles à partir desquels nous déduisons  $p$ . Les conséquents de  $p$  forment un ensemble de  $n$ -tuples ordonnées. Le premier membre de chaque  $n$ -tuple est un énoncé qui découle de  $p$  et les autres membres de chaque  $n$ -tuple sont les prémisses (sans compter  $p$ ) qui sont nécessaires pour dériver le premier membre. Par conséquent, le rôle inférentiel de l'énoncé  $p$ , sera constitué comme suit : {les engagements antécédents de  $p$ , les engagements conséquents de  $p$ , les autorisations antécédents de  $p$ , les autorisations conséquentes de  $p$ , les incompatibilités liées à l'engagement envers  $p$ }. Les relations inférentielles déterminent alors la nature du contenu asserté, mais elles-mêmes dérivent des normes pragmatiques. Les relations inférentielles constitutives des contenus n'existent qu'en rapport avec les normes pragmatiques elles-mêmes attribuées et instituées par d'autres attitudes normatives. Brandom insiste donc sur la dimension normative du contenu asserté et définit la portée pragmatique (*pragmatic significance*) d'une phrase par la manière dont son assertion affecte l'évolution du pointage de la conversation au sein de laquelle elle est énoncée. Cette portée pragmatique revêt deux aspects que sont les circonstances et les conséquences d'application d'un contenu de signification, comprises ici en termes de permission et d'engagement : les circonstances

<sup>1008</sup> Brandom, 2011, p. 334; 2009, p. 204: notons que cette sémantique des incompatibilités est « classique » : l'incohérence est le corrélat formel de l'incompatibilité. Il s'agit d'une notion logique qui doit être comprise en termes de négation. Mais, ce qui fait qu'un élément du vocabulaire exprime une négation doit à son tour être compris en référence à sa relation à l'incompatibilité matérielle. La négation d'une affirmation se définit comme l'affirmation inférentiellement la plus faible qui est impliquée (au sens qui préserve les engagements) par tout énoncé incompatible avec l'affirmation initiale.

d'application d'un contenu de signification sont ce qui permet ou ce qui engage l'assertion du contenu – ces engagements peuvent être inférentiels et/ou comportementaux. Comprendre un contenu, c'est connaître les circonstances dans lesquelles on deviendrait engagé ou justifié à adhérer à ce contenu (ce qui justifie le contenu comme *conclusion*), mais aussi connaître ce envers quoi on devient engagé ou justifié en adhérant à ce contenu (ce qui justifie le contenu comme *prémisse*). Ces circonstances et conséquences peuvent être d'autres contenus, mais aussi d'autres paramètres non linguistiques (états de choses, action,...) : les premières visent à relever le pointage des conversations dans lesquelles il est approprié et légitime d'énoncer l'acte de discours en question; les secondes retranscrivent et suivent l'évolution du pointage<sup>1009</sup> – en termes de transformations des attitudes déontiques par lesquelles les interlocuteurs tiennent les comptes des engagements et autorisations des autres – que l'énonciation légitime de ce même acte de discours implique<sup>1010</sup>. La pratique linguistique peut alors être comprise comme une fonction déterminant la façon dont le pointage déontique impose, à chaque étape d'une conversation, d'identifier les performances qui sont appropriées ainsi que les conséquences qui suivent des différentes performances – c'est-à-dire la façon dont celles-ci modifient le pointage. Il est, par conséquent, important de connaître les contenus qui peuvent justifier l'assertion, et non seulement ceux qui découlent d'elle. Le contenu se définit donc à partir de son rôle de prémisse, mais aussi de conclusion dans des raisonnements. Il est alors individué et expliqué en termes d'articulation inférentielle, elle-même expliquée en termes de relations normatives pragmatico-inférentielles entre statuts déontiques. La possession du contenu est, de ce fait, fondamentalement relative à une perspective externe : asserter un contenu, c'est se voir attribuer – potentiellement ou effectivement – la capacité à agir selon des statuts normatifs (engagements et permissions).

Nos assertions ne peuvent donc avoir de contenu que parce que nous sommes engagés dans des pratiques dans lesquelles nous nous traitons mutuellement comme engageant ou engagé par des raisons. L'attribution ou la reconnaissance d'un état intentionnel chez un sujet

<sup>1009</sup> Brandom, 2011, p. 357 : En faisant le parallèle avec ce jeu de pointage au baseball, on parlera par exemple du statut que peut avoir une performance en tant que *balle*, que *prise* ou que *retrait*.

<sup>1010</sup> Ibid., p.357: « Ainsi, le fait de comprendre ou saisir la portée d'un acte de langage exige que l'on soit capable de spécifier en termes de marques de ce type quand cet acte serait convenable (circonstances d'application) et comment il transformerait la marque propre à l'étape de sa réalisation en marque propre à l'étape suivante au sein de la conversation dont il relève (conséquences d'application). ».

va de pair avec l'attribution d'un statut normatif<sup>1011</sup>, qui modifie le pointage déontique du locuteur. Une assertion ou un jugement n'ont de force et de contenu que par rapport à leurs rôles inférentiels, définis par les statuts déontiques attribués implicitement ou attribuables explicitement. Les agents choisissent de se soumettre à des normes pragmatiques, ils choisissent de dire quelque chose, mais ce ne sont pas eux qui déterminent le contenu de ce qu'ils disent, en suivant ces normes. En effet, trois caractéristiques propres aux assertions gouvernent ces actions de marquage au score, à savoir :

- 1) quand un agent formule une assertion, il reconnaît un engagement doxastique. Il assume également tous les engagements et autorisations qui découlent de l'engagement qu'il reconnaît explicitement.
- 2) Une assertion réussie (c'est-à-dire une assertion dans laquelle son auteur est autorisé à énoncer l'engagement reconnu) autorise d'autres membres de la conversation à prendre le même engagement, à l'assumer, à en prendre la responsabilité et à s'engager à son endroit. Des assertions réussies présentent des engagements en vue d'une adhésion publique.
- 3) L'auteur de l'assertion prend la responsabilité de donner les raisons et de justifier son assertion lorsque les autres membres de la conversation le lui demandent. Demander des raisons pour une affirmation, c'est en demander les garanties et demander ce qui autorise son auteur à cet engagement. Ainsi de manière générale, le sens de l'assomption qui détermine la force des actes de discours assertionnels implique, au moins, une sorte d'engagement à l'égard duquel l'autorisation du locuteur peut toujours être défiée c'est-à-dire être toujours potentiellement en question<sup>1012</sup>.

Regardons ainsi de plus près comment ces notions interagissent en prenant l'exemple d'un jeu de pointage déontique.

---

<sup>1011</sup> Ibid., p 79 : « Attribuer un état intentionnel, c'est attribuer un statut normatif. »

<sup>1012</sup> Ibid., p.335-339.

*Un exemple de jeu de pointage :*

Brandom décrit ainsi une conversation dans laquelle trois personnes recensent les points : chacun d'eux pointe l'autre et vice versa<sup>1013</sup>. Prenons trois personnes *A*, *B* et *C*. Nous considérons le pointage à un moment donné comme un ensemble dont les membres sont des ensembles d'engagements et d'autorisations : chaque interlocuteur se voit assigner un nombre de points différent et pour chacun d'entre eux, on assigne, à chaque étape, différents engagements et différentes permissions. Un certain nombre de complications apparaît aussitôt. Puisque chaque membre de la conversation entretient des croyances différentes ou reconnaît des engagements différents, il faudra tenir des livres séparés pour établir le pointage de chaque participant. Non seulement, les points seront comptés pour chaque interlocuteur, mais ils le seront également par chaque interlocuteur. En outre, la pragmatique de Brandom distingue entre trois différentes attitudes : attribuer [*attributing*], assumer/contracter [*undertaking*] et reconnaître [*acknowledging*]. Chaque marqueur [*scorekeeper*] doit séparer les engagements et autorisations associés à chaque autre participant en distinguant les engagements qui sont pris et assumés et les engagements qui sont reconnus. Un marqueur établit ainsi cette distinction afin de tenir les comptes des engagements et permissions des autres joueurs, en attribuant ces engagements et permissions. De même que chaque interlocuteur se voit, à chaque étape, attribuer un ensemble différent de statuts déontiques, de même chaque interlocuteur possède, à chaque étape, un ensemble différent d'attitudes. Par conséquent, la somme des points de chacun des dix ensembles représente le pointage obtenu par chacun de ces trois membres au cours de la conversation. Par exemple, pour *B*, le pointage est : {{les engagements que *B* reconnaît}, {les permissions que *B* reconnaît}, {les engagements que *A* reconnaît}, {les engagements que *A* assume}, {les autorisations que *A* reconnaît}, {les autorisations que *A* assume}, {les engagements que *C* reconnaît}, {les engagements que *C* assume}, {les autorisations que *C* reconnaît}, {les autorisations que *C* assume}}.

Supposons que *A* énonce une phrase dont l'instance (*token*) est *p*. Supposons également que *B* détermine que l'énonciation de *A* a la force d'une assertion, et que *B* comprend *p* (c'est-à-dire que *B* comprend le contenu de *p*). *B* doit d'abord attribuer un engagement

---

<sup>1013</sup> Ibid., p. 369-371. Nous reprenons également les commentaires de Scharp, 2005, p. 211-215.

doxastique à *A* (qui sera inscrit sous les engagements que *A* reconnaît). Nous utiliserons « *p* » à la fois comme un nom pour désigner la phrase que *A* affirme et comme un nom pour qualifier l'engagement que *A* reconnaît en énonçant cette affirmation. *B* attribue ensuite à *A* tous les engagements quant aux conséquences inférentielles qui se rapportent à *p* au titre de conclusion d'inférences préservant l'engagement et ayant *p* pour prémisse (ceux-ci seront inscrits sous les engagements que *A* assume/contracte). Il faudrait alors que *B* ajoute un engagement<sup>1014</sup> à toutes les affirmations *q* qui seraient des conséquences inférentielles et engageantes de *p*, dans le contexte des autres phrases attribués à *A*. Celles-ci varieront en fonction des hypothèses auxiliaires dont on dispose, étant donné les autres engagements que *B* attribue déjà à *A*. Autrement dit, les engagements qui découlent de *p* par le biais d'inférences préservant les engagements (selon *B*) dépendront des autres engagements que *A* est considéré comme avoir contracté et reconnu. C'est là ce qui clôt les attributions de *B* à *A* sous des inférences préservant l'engagement. Ensuite, *B* doit examiner les relations d'incompatibilité dans les engagements de *A* au sein desquelles s'inscrit *p* (et donc également ses conséquences inférentielles engageantes) afin de déterminer lesquelles (à supposer qu'il y en ait), parmi les autorisations que *B* attribue à *A*, sont exclues par le nouvel engagement. Les assertions ajoutent certes de nouveaux engagements; mais, en ce qui a trait aux autorisations, elles peuvent non seulement en ajouter, mais aussi en ôter. *B* doit, par la suite, à la lumière des relations d'incompatibilité associées à tous les engagements attribués à *B*, attribuer des autorisations à toutes les affirmations qui sont des conséquences inférentielles engageantes des engagements auxquels *A* est déjà considéré comme autorisé, ce qui clôt la somme des points attribués selon les inférences préservant l'engagement, là où les autorisations qui en résultent ne sont pas mises en échec par des incompatibilités. Tout au long de ce processus, *B* doit donc continuer à vérifier les relations d'incompatibilités afin de s'assurer qu'il n'attribue pas à *A* des autorisations à deux engagements incompatibles.

La prochaine étape consiste, pour *B*, à attribuer des autorisations à toutes les affirmations qui sont des conséquences inférentielles permissives des engagements auxquels *A* est déjà considéré comme autorisé, ce qui clôt la somme des points attribués selon les inférences préservant l'autorisation (ce que l'on inscrit sous les autorisations que *A* contracte) qui ne sont pas mises en échec par ces incompatibilités. Il est néanmoins important de faire

---

<sup>1014</sup> Pour des développements ultérieurs en communication multi-agents, voir Kibble, 2006.

remarquer que *A* peut être considéré (par *B*) comme autorisé à des affirmations auxquelles il n'est pas engagé – il s'agit des conclusions que *B* peut être autorisé à tirer mais auxquelles *A* ne s'est pas encore engagé. Par exemple, *A* pourrait avoir affirmé que « la température est descendue et le ciel devient nuageux ». *B* pourrait ensuite attribuer l'autorisation de *A* quant à l'engagement à l'égard de l'affirmation « il va bientôt pleuvoir » même si *B* n'attribue pas cet engagement à *A*. Si cet engagement-ci n'est pas déduit à partir des inférences engageantes issues d'une affirmation de *A* (*A* ne s'est donc pas engagé envers cette affirmation) mais bien à partir d'inférences permissives (*A* a bien obtenu, selon celui qui tient le pointage, l'autorisation à asserter cet engagement) alors nous avons de bonnes raisons inductives de croire qu'il va bientôt pleuvoir.

*B* doit alors, pour finir, évaluer l'autorisation de *A* quant à l'affirmation que *p*, en examinant les inférences valides dans lesquelles *p* apparaît comme conclusion et comme prémisse et auxquelles *A* est engagé et autorisé. *A* peut s'avérer autorisé à affirmer *p* de différentes manières. Tout d'abord, *A* pourrait évaluer la possibilité d'une autorisation par défaut à *p*. Il s'agit du type d'autorisation par défaut caractéristique des comptes rendus d'observation où quiconque qui attribue une telle autorité admet implicitement la fiabilité du rapporteur (dans ces circonstances, et à l'égard de ce contenu). Traiter quelqu'un comme un rapporteur fiable, c'est considérer que l'engagement (quant à ce contenu et dans ces circonstances) de ce rapporteur suffit pour que ce rapporteur soit autorisé à cet engagement. Ensuite, *A* pourrait être autorisé à affirmer *p* dans la mesure où, au titre de conséquence d'autres engagements, *p* suit en vertu d'inférences engageantes ou d'inférences permissives issus d'autres engagements que *A* est autorisé à asserter. Enfin, *A* pourrait être autorisé à asserter *p* au nom d'un témoignage. Autrement dit, *A* pourrait reconnaître *p* dans la mesure où *A* attribue l'engagement et l'autorisation à asserter *p* à un autre interlocuteur qui aurait asserté *p* lors d'une étape précédente et de qui *A* hérite l'autorisation à *p*. Si *B* considère que *A* est autorisé à *p* par l'un ou l'autre de ces mécanismes d'héritage ou de défaut, alors *B* considèrera que *A* a avec succès autorisé d'autres personnes (*y* compris *B*) à cette affirmation (en l'absence de facteurs d'échec incompatibles). *B* considère alors les autres membres de la conversation (*y* compris lui-même), *C* dans cet exemple, comme libres de reconnaître le même engagement et, si c'est effectivement le cas pour l'un d'entre eux, celui-ci héritera l'autorisation par voie testimoniale. Cependant, si *C* considère que *A* n'est pas autorisé à *p*,

alors il peut soit se retirer du jeu de pointage et attendre la prochaine assertion, soit il peut formuler une assertion incompatible et mettre au défi l'autorité testimoniale de l'assertion de *A*. La portée propre à un tel défi est qu'il met en cause les attributions d'autorisation par défaut là où l'assertion à l'origine du défi est une assertion à laquelle l'auteur du défi est, au moins à première vue, autorisé. Le fait que *B* considère comme réussi le défi lancé par *C* à l'assertion de *A*, consiste en ceci que *B* y répond en suspendant<sup>1015</sup> son attribution d'autorisation à *A* pour cette affirmation en l'attente d'une justification par *A* de celle-ci. Cela a pour effet de rendre cette assertion indisponible (selon la carte de pointage tenue par *B*) pour d'autres interlocuteurs qui, autrement, auraient pu hériter (sur la base d'un témoignage de *A*) d'autorisations à des engagements à l'égard du même contenu.

Il dépend certes que de moi d'asserter quelque chose de particulier. Mais, contracter un engagement quant à une affirmation dotée d'un certain contenu implique de contracter des engagements quant à des affirmations dont les contenus en sont les conséquences inférentielles engageantes. Ce que je fais en assertant ne dépend donc pas de moi : selon les statuts attribués (et ma capacité à y répondre), je ne peux être engagé envers toutes les conséquences matérielles<sup>1016</sup> de nos contenus. Par exemple, contracter un engagement quant à l'affirmation que Pittsburgh est à l'ouest de Philadelphie est l'une des manières dont je peux contracter un engagement quant à l'affirmation que Philadelphie est à l'est de Pittsburgh. Il est également possible que je ne reconnaisse pas ces engagements *conséquentielles* dans la mesure où « nous ne reconnaissons pas toujours notre engagement quant aux conséquences des engagements que nous reconnaissons réellement<sup>1017</sup> ». Selon le

<sup>1015</sup> Brandom, 2011, p. 373-374, introduit cette idée de rétractation afin de corriger le tableau ou la carte de pointage (*deontic scoreboards*) de chacun des marqueurs. Les désaveux, dit-il, nous permettent de répudier ou de démentir un engagement précédemment contracté, ou de signaler clairement que l'on ne reconnaît pas un tel engagement. Si le désaveu est considéré comme étant réussi et est approuvé par *B*, alors *B* restitue à *A* toutes les attributions d'autorisations qu'il lui avait retirées. *A* reconnaît et répare son erreur afin que la carte de pointage soit réajustée. Mais le désaveu peut aussi échouer en raison des deux façons dont *A* peut contracter un engagement, à savoir : *A* peut acquérir soit directement un engagement à l'égard de *p*, en avouant *p*, soit indirectement, au titre de conséquence d'un engagement à l'égard de *q*, dont *p* suit en vertu d'une inférence valide et préservant l'engagement (selon *B*). Dans ce cas, le désaveu de *p* par *A* ne peut réussir (selon *B*) que si *A* est également disposé à désavouer *q*. Mais si *A* persiste à asserter *q*, alors cet engagement est incompatible avec le désaveu, et le désaveu de *p* ne restitue aucune autorisation à *A*. *A* doit donc se rétracter.

<sup>1016</sup> Ibid., p.377.

<sup>1017</sup> Ibid., p. 377.

contexte, nous sommes seulement engagés envers les contenus inférentiels que les autres participants de la pratique linguistique estiment être pertinents et centraux<sup>1018</sup>. Il n'existe donc pas de formes inférentielles analytiques ou essentielles mais en fonction du contexte, certaines transitions inférentielles sont considérées comme plus centrales que d'autres.

L'une des conséquences de la thèse de Brandom est que la possession d'un concept requiert la maîtrise d'un ensemble d'inférences qui le définit. Ces relations inférentielles articulent le contenu de l'engagement ou de la responsabilité que l'on prend en assertant une affirmation. En dehors de telles relations, il n'y a pas de contenu, et par conséquent aucune affirmation. Comme l'écrit Brandom, dans une formulation qui rappelle les propos de Davidson,

Saisir ou comprendre un tel concept, c'est avoir une maîtrise pratique sur les inférences dans lesquelles il est impliqué – savoir, au sens pratique d'être capable de distinguer (une sorte de savoir-faire), ce qui s'ensuit de l'applicabilité d'un concept, et ce dont découle cette même applicabilité<sup>1019</sup>.

[...] Les concepts font défaut au perroquet et au thermostat, en dépit de leur maîtrise des dispositions correspondantes à répondre de manière non inférentielle et différenciée, précisément parce que la maîtrise pratique de la formulation inférentielle qui définit la saisie des contenus conceptuels leur fait défaut.<sup>1020</sup>

Pour Davidson, comprendre la signification d'un concept ne se réduit pas à une simple capacité de discrimination<sup>1021</sup>, dans la mesure où la première capacité ne requiert pas l'usage de concepts signifiants, inférentiellement liés, publiquement partagés et soumis à des normes de correction et d'usage, là où la seconde capacité l'exige. Comprendre la signification d'un concept requiert, en effet, d'un côté, la compréhension de la signification d'autres concepts, et de l'autre, la capacité d'utiliser explicitement le concept en rapport inférentiel avec ces autres concepts, ainsi qu'une capacité à l'attribuer. Brandom ne mentionne pas des conditions aussi strictes : pour lui, pour avoir des états intentionnels, il est nécessaire de maîtriser inférentiellement le contenu attribué ou asserté, mais il n'est pas nécessaire d'avoir des

<sup>1018</sup> Ibid., p. 378: Brandom reprend cette même idée d'une présomption de rationalité que l'on retrouve dans la formulation du principe de charité de Davidson : « En effet, les explications de ce type ne fonctionnent que dans la mesure où l'individu en question est (considéré) comme *rationnel* – comme n'ayant pas de croyances contradictoires ou incompatibles, et comme croyant aux conséquences de ses propres croyances. ».

<sup>1019</sup> Brandom, 2009, p. 56.

<sup>1020</sup> Ibid., p. 172.

<sup>1021</sup> Davidson, 2004, 2001, p. 137 et sq.

attitudes conceptuelles explicites portant sur les concepts d'autrui. Une interprétation ou une attribution implicite d'autrui suffit.

Aussi, comme nous l'avons vu, pour Brandom, un locuteur ne peut participer à une pratique linguistique que s'il est capable lui-même d'en attribuer, au moins implicitement, aux autres membres de la communauté, en pouvant les corriger ou en pouvant les sanctionner. Par conséquent, pour comprendre la notion d'une application correcte et objective des normes de signification, il faut débiter par une théorie de ces pratiques normatives d'attribution dont émerge *in fine* le contenu propositionnel :

Le fait de parler de contenus propositionnels inférentiellement articulés est une façon de parler des normes implicites qui régissent la pratique du pointage déontique; c'est la conséquence de l'idée que les contenus conceptuels sont *conférés* par une telle pratique<sup>1022</sup>.

Brandom cherche ainsi à répondre à la question suivante : quelles propriétés les pratiques sociales doivent-elles nécessairement exhiber pour mériter le nom de pratiques rationnelles ou de pratiques linguistiques<sup>1023</sup> ?

Ainsi, un ensemble de pratiques peut compter pour un ensemble de pratiques discursives, et largement pour un jeu d'offre et de demande de raisons, s'il inclut dans son fonctionnement la reconnaissance, l'usage et l'institution par ses participants des deux statuts déontiques qui les unissent : les engagements et les permissions aux engagements<sup>1024</sup>. Nous avons précisé que jouer un coup dans le jeu assertionnel doit être considéré comme la reconnaissance d'un certain type d'engagement, articulé par des relations inférentielles conséquentielles reliant la phrase assertée aux autres phrases : c'est donc un mouvement discursif dans le jeu d'offre et de demande des raisons, qui est l'objet d'une série d'engagements inférentiels et cognitifs de la part du producteur<sup>1025</sup>. Nous retrouvons plus spécifiquement six relations entre engagements et permissions qui sont suffisantes pour

---

<sup>1022</sup> Brandom, 2011, p.1103-1104.

<sup>1023</sup> Ibid., p. 33-34 : « [...] à commencer par une élucidation des pratiques sociales, à identifier la structure particulière qu'elles doivent exhiber afin d'être qualifiées de pratiques spécifiquement *linguistiques*, et ensuite à examiner quelles différentes sortes de contenus sémantiques ces pratiques peuvent conférer aux états, aux actions et aux expressions, saisis dans ces contenus de manière appropriée [...] ». »

<sup>1024</sup> Brandom, 2009, p.201.

<sup>1025</sup> Ibid., p. 202-203.

qualifier une pratique de discursive et pour conférer un contenu conceptuel aux expressions, aux performances et aux statuts qui se nichent en elles.

Trois de ces relations sont sémantiques (1,2,3) et concernent exclusivement des mouvements linguistiques dans l'espace des raisons; trois de ces relations (4,5,6) sont pragmatiques et sociales et concernent des mouvements relatifs aux entrées perceptives et aux sorties motrices de l'espace des raisons.

- 1) *Des relations inférentielles* au sein desquelles les engagements sont préservés. Il est question ici des inférences déductives où tout locuteur engagé envers  $p$  est engagé envers  $q$ ;
- 2) *Des relations inférentielles permissives* au sein desquelles des permissions ou des autorisations sont préservées. Il est question ici d'inférences inductives où être engagé envers  $p$  peut être une raison pour s'engager envers  $q$ ;
- 3) *Des relations d'incompatibilité* entre contenus : il est question ici de l'inférence modale où un engagement envers  $p$  interdit une permission pour un engagement envers  $q$ ;

Par conséquent, pour spécifier le contenu inférentiel associé à une phrase, il faut, pour commencer, indiquer le rôle que joue ce contenu (par rapport aux contenus exprimés par d'autres phrases) dans trois types de structures globalement inférentielles : les inférences engageantes, les inférences permissives et les incompatibilités. Faire cela, c'est dire de quoi suit ce contenu, ce qui suit de lui, et ce qu'il exclut ou proscrit. Ces trois structures sont caractérisées comme « globalement » inférentielles dans la mesure où elles impliquent toutes des modifications des statuts déontiques ayant pour conséquences d'autres modifications de statuts déontiques. Aussi, dans l'exemple du jeu de pointage ci-dessus,  $A$  examine les inférences engageantes, les inférences permissives et les incompatibilités qui relient  $p$  aux autres phrases *token* du langage (le contenu de  $p$ ) afin de déterminer les engagements et les autorisations qui devraient être attribués, reconnus et contractés par les participants à la conversation (la portée pragmatique de  $p$ ).

- 4) *Des relations empiriques*, par lesquelles des entrées non linguistiques engagent inférentiellement les agents cognitifs;
- 5) *Des relations pratiques*, par lesquelles des modifications de situations non linguistiques engagent inférentiellement les agents cognitifs;
- 6) *Des relations de confiance (testimony<sup>1026</sup>)*, par lesquelles des engagements et des autorisations sont transmis non inférentiellement entre des personnes (actes de déférence par exemple<sup>1027</sup>).

---

<sup>1026</sup> Brandom, 2011, p.373.

Dans ce dernier type de relations inférentielles, asserter que  $p$ , c'est reconnaître que nous sommes engagés envers  $p$  : cette reconnaissance, aux yeux de la communauté, signifie que le locuteur autorise les autres à contracter également (*undertake*) ce même engagement (par exemple en l'assertant) et qu'il s'engage à démontrer le bien-fondé de son engagement soit qu'il démontre pourquoi il est autorisé à cet engagement, en fournissant des raisons<sup>1028</sup>. Asserter que  $p$  (« il pleuvra demain »), c'est accepter la responsabilité de justifier pourquoi et comment nous en sommes arrivés à cet engagement, mais aussi exercer une autorité envers  $p$ , en autorisant les autres à s'engager eux aussi envers  $p$  ou ses conséquences déductives (« les rues seront mouillées »), tout en les déchargeant de la responsabilité de justifier  $p$ .

Les relations inférentielles, comme nous les retrouvons en 1), sont des relations matérielles entre contenus propositionnels, et non pas seulement des relations logiques : leur validité peut reposer ultimement sur des règles matérielles déterminant des transitions (non enthymématiques) entre des contenus, transitions qui, d'un point de vue strictement logique, ne seraient pas concluantes. Il s'agit ici d'une inférence sémantique au sens « matériel » du terme, c'est-à-dire qui résulte de la signification des mots utilisés dans un langage naturel. Brandom adhère ainsi pleinement à l'idée de Sellars selon laquelle les inférences matérielles ne peuvent être réduites à des inférences formelles, leur correction dépendant également du contenu des prémisses et des conclusions en cause<sup>1029</sup>.

Par exemple, de « aujourd'hui est un dimanche », il est possible de déduire que « demain est un lundi », ou de « je vois un éclair maintenant », je peux inférer « j'entendrai bientôt le tonnerre ». La validité des inférences dépend ainsi du sens des mots ou concepts utilisés, comme « aujourd'hui » et « demain » ou « éclair » et « tonnerre ». Comprendre ces mots, c'est posséder *ipso facto* une maîtrise pratique de ces inférences. Nous ne pouvons en effet

<sup>1027</sup> Ibid., p.344 :

La seconde manière de justifier un engagement en prouvant qu'on y est habilité est d'en appeler à l'autorité d'un autre émetteur d'assertion. La fonction *communicationnelle* des assertions est d'autoriser d'autres personnes qui entendent cette affirmation à l'asserter de nouveau. [...] Ils peuvent invoquer l'autorisation ou l'autorité de l'émetteur, s'en remettant ainsi à l'interlocuteur qui leur a communiqué cette affirmation, et déléguant à cet autre individu toute demande de preuve d'habilitation.[...] Le fait qu'une assertion formulée par  $A$  ait, du point de vue social, pour portée d'autoriser une réassertion par  $B$  consiste dans le fait qu'il est convenable que  $B$  défère à  $A$  la responsabilité de prouver l'habilitation à cette affirmation.  $B$  peut s'acquitter de sa responsabilité en invoquant l'autorité de  $A$ , sur laquelle  $B$  exerce son droit à s'appuyer. C'est à  $A$  que l'on fait porter le chapeau.

<sup>1028</sup> Brandom, 2009, p.204.

<sup>1029</sup> Sellars, 1953; reproduit dans Sellars, 1980a.

maîtriser les concepts d'« éclair » et de « tonnerre » sans être capables par ailleurs de faire l'inférence précédente.

Aussi, Brandom explique la maîtrise de la conséquence logique à l'aide de la maîtrise des inférences matérielles. De manière classique, on considère que l'inférence matérielle « il pleut donc les rues sont mouillées » forme un enthymème<sup>1030</sup>, c'est-à-dire une conditionnelle cachée. Pour un formaliste, par exemple, l'inférence précédente est une conditionnelle cachée (« s'il pleut, alors les objets sont mouillés »). Selon lui, il manque une prémisses au raisonnement, qui devrait être formalisé ainsi :

Il pleut sur la chaussée  
S'il pleut, alors les objets sont mouillés  
 La chaussée est mouillée

Le « dogme formaliste » consiste à estimer qu'une inférence matérielle est une déduction dans laquelle une prémisses a été omise et qu'il est nécessaire d'ajouter cette prémisses pour établir la validité formelle de l'argument. Dans ce paradigme « formaliste », la validité logique ne concerne donc jamais le sens des mots utilisés dans les inférences. Dès lors, les mots du vocabulaire courant peuvent être remplacés par des lettres schématiques.

Cependant, comme le souligne Brandom, l'approche « formaliste<sup>1031</sup> » de la déduction n'est pas nécessairement convaincante, car on peut tout aussi bien juger que le but premier de la logique vise à distinguer les bonnes inférences matérielles des mauvaises<sup>1032</sup> et que le rétablissement des prémisses manquantes constitue alors un moyen de séparer le bon grain de l'ivraie. Aussi, avec Sellars et Brandom, il est possible d'appréhender la logique comme l'explicitation des contenus conceptuels qui demeurent implicites dans les pratiques inférentielles des locuteurs. Les inférences matérielles sont donc valides sans que leur validité

<sup>1030</sup> Brandom, 2011, p. 223 : « Les inférences matérielles [...] sont comprises de manière privative : comme des enthymèmes qui résultent de l'occultation ou du masquage de l'une des prémisses requises pour disposer d'une véritable garantie. »

<sup>1031</sup> Ibid., p.219 : « Cette conception – qui comprend « bonne inférence » comme voulant dire 'inférence formellement valide', postulant des prémisses implicites selon le besoin – pourrait être appelée une approche formaliste de l'inférence. »

<sup>1032</sup> Ibid., p. 219: « La saisie de la logique qui est attribuée doit être une saisie implicite, puisqu'elle ne doit se manifester que par la capacité de distinguer les inférences matérielles comme bonnes ou mauvaises. »

dérive de principes logiques (comme le *modus ponens*)<sup>1033</sup>. Par ailleurs, il paraît légitime de penser que le contenu conceptuel est individué par les inférences matérielles auxquelles il est lié. Par exemple, la signification du mot « rouge » n'est peut-être pas indépendante d'inférences matérielles comme « si ce vêtement est rouge, il est coloré » ou « si ce vêtement est rouge, il n'est pas (entièrement) vert ». En ce sens, nous estimons, avec Brandom que la notion de proposition (ou de phrases énoncées) est première par rapport aux mots qui la composent (termes ou prédicats), car elle est indispensable à la maîtrise du mécanisme de l'inférence. Nous estimons aussi que les inférences matérielles possèdent un rôle cognitif et heuristique<sup>1034</sup> dans le *jeu de pointage*, rôle qui est sous-estimé dans le point de vue « formaliste » sur l'inférence. En conséquence, nous concevons la logique comme une explicitation des normes implicites dans l'usage des concepts du langage ordinaire, et non comme la production de tautologies vides de sens.

De plus, considérer les inférences matérielles comme premières dans l'ordre explicatif présente d'autres bénéfices explicatifs. Tout d'abord, cela permet d'éviter le type de régression à la Lewis Carroll qui s'en suivrait si un conditionnel était requis pour que toutes les inférences soient valides<sup>1035</sup>; ensuite, cela permet de contourner les problèmes relatifs à l'apprentissage d'une langue dans la mesure où si, le contenu était une fonction des relations inférentielles de la logique formelle, alors nous devrions connaître l'ensemble du répertoire des principes logiques et des formes d'argument du système de la déduction naturelle afin de comprendre le contenu des concepts empiriques ordinaires comme « l'est », « l'ouest », « rouge », « la pluie », « mouillé », etc. Finalement, la priorité des propriétés de l'inférence matérielle sur celles de la logique formelle et l'approche *expressiviste* de cette dernière, signifie que quel que soit ce que la normativité de la logique formelle a, ceci doit déjà être présent au sein des propriétés inférentielles matérielles et de ce fait, dans le contenu des

---

<sup>1033</sup> Ibid., p.220-223: l'accent mis par Brandom sur les relations inférentielles matérielles contraste avec la lecture « formaliste » du principe de charité de Davidson dans laquelle le contenu normatif est constitué par les relations logiques obtenues entre les croyances et les énoncés. Cependant, l'approche de Brandom rejoint plus la conception moins stricte du principe de charité (au sens d'humanité) qui conçoit le contenu comme déterminé par le caractère raisonnable ou intelligible des contenus et croyances élaboré à la lumière des croyances et énoncés d'autrui.

<sup>1034</sup> Ibid., p. 220 : il existe peut-être de bonnes inférences non logiques, et notre rationalité consiste peut-être dans la manière dont le fait de les discriminer importe pour nos délibérations et nos évaluations. Pourquoi le fait de « suivre les règles de la logique » devrait-il être soit nécessaire soit suffisant pour cette discrimination ? ».

<sup>1035</sup> Ibid., p.221-222.

concepts matériels ordinaires non logiques tels que « ouest », « est », « pluie », « mouillé », etc. Autrement dit, la normativité forte s'étend également au vocabulaire non logique et bien que la logique formelle soit fortement normative, elle se contente de rendre explicite les propriétés inférentielles obtenues parmi les concepts matériels<sup>1036</sup>. Nous retrouvons bien ici une idée centrale de Sellars, pour qui les propriétés des inférences matérielles dans lesquelles peut figurer un contenu sont essentielles à la signification de ce contenu. Il appartient alors à Brandom de généraliser cette idée, en situant l'origine de ces inférences matérielles (ou plutôt des règles matérielles, qui assurent les transitions entre contenus) dans des pratiques sociales.

Brandom insiste bien en soulignant que la normativité de nos inférences matérielles s'explique par les engagements et les permissions discursifs (*commitments and entitlements*) en termes de statuts déontiques, statuts qui sont eux-mêmes analysés en termes d'attitudes déontiques que les locuteurs entretiennent vis-à-vis de ces statuts<sup>1037</sup>. La rationalité des contenus réside alors dans la reconnaissance explicite des statuts normatifs qui sont présents dans les pratiques discursives. Cette capacité d'attribution et de reconnaissance implicite et pratique de statuts déontiques est suffisante pour définir les êtres linguistiques.

Deux types différents de créatures linguistiques font en effet partie des pratiques discursives : les êtres rationnels et les êtres logiques. Les êtres rationnels peuvent avoir ou exprimer des contenus conceptuels sans avoir la capacité d'attribuer explicitement à leurs pairs des contenus conceptuels sous la forme de statuts déontiques discursifs. Ils ne sont pas capables d'attribuer ou de reconnaître de l'intentionnalité originale, c'est-à-dire un ensemble de pratiques discursives comme jeu d'offre et de demande de raisons. Ils sont cependant tout à fait capables de sanctionner un usage déviant des contenus inférentiellement articulés. Les êtres logiques, pleinement et réflexivement rationnels, ont et expriment des pensées conceptuelles, et possèdent des capacités d'attribution explicite de contenu conceptuel. Ces capacités d'explicitation sont rendues possibles par l'usage des ressources expressives de la logique. Ces êtres sont aptes à formuler des énoncés tels que «  $X$  croit que  $p$  » ou « si  $X$  est engagé envers  $p$ , alors il l'est envers  $q$  ». Ils peuvent reconnaître implicitement que les autres membres de la communauté sont eux aussi des marqueurs (*scorekeepers*), sans pouvoir le conceptualiser et rendre compte de la structure du jeu d'offre et de demande de raisons. Les

<sup>1036</sup> Par conséquent, tout comme Davidson et Wittgenstein, Brandom est partisan de l'extension d'une normativité forte au vocabulaire non logique.

<sup>1037</sup> Ibid., Chap.3, Vol. II, §5.

membres peuvent ici reconnaître et exprimer leurs propres engagements discursifs, mais ne peuvent qu'implicitement en attribuer à autrui. Ces membres linguistiques qui ne maîtrisent pas encore le vocabulaire logique traitent implicitement les autres locuteurs comme étant eux aussi des marqueurs discursifs. Ils ne peuvent pas attribuer des attributions à leurs congénères, ni tenir le pointage des attributions et des attitudes de reconnaissance de ces congénères:

Par conséquent, l'usage des ressources expressives de la logique permet à certains locuteurs de passer d'une posture (*stance*<sup>1038</sup>) discursive implicite de pointage à une posture discursive explicite de pointage (*l'explicit discursive scorekeeping stance*). La maîtrise du vocabulaire logique permet donc de pratiquer explicitement le pointage, en émettant des jugements sur les engagements inférentiels, pratiques et doxastiques des locuteurs. Aussi, être un marqueur discursif et explicite, ce n'est pas seulement assurer le pointage et attribuer explicitement aux autres membres de la communauté des statuts déontiques mais c'est également leur attribuer des attitudes déontiques, c'est-à-dire concevoir qu'ils attribuent et reconnaissent à leur tour des statuts déontiques. Pour Brandom, une attribution ne peut être qu'attribuée explicitement, *via* des contenus propositionnels. Aussi, les membres d'une communauté qui exercent ce marquage discursif explicite considèrent que les autres sont capables d'avoir le même type d'attitude explicite qu'eux : il existe donc une symétrie interprétative explicite pour ces êtres logiques<sup>1039</sup>. Nous n'attribuons pas seulement des statuts déontiques (explicitement ou implicitement) aux autres membres, mais aussi des attitudes déontiques. Les locutions d'imputation d'attitudes propositionnelles permettent d'exprimer explicitement cette attitude de pointage déontique pratique en tant que contenu propositionnel, autrement dit comme contenu d'une affirmation. L'usage des conditionnels rend explicite, comme contenu d'une affirmation, et donc comme quelque chose que l'on peut dire, le fait d'assumer une inférence – attitude que l'on ne pourrait autrement manifester que par ce que l'on fait. Le vocabulaire imputationnel, celui des mots comme « croit » ou

<sup>1038</sup> Nous reprenons ici l'expression « stance » de Dennett, 1971, p. 90, que l'on traduit par posture; Cf. Brandom, 2011, p. 148, parle de la « perspective ou posture intentionnelle » (*intentional stance*) de Dennett comme un moyen pour donner à la fois du sens et prédire le comportement d'une créature quelconque.

<sup>1039</sup> Brandom, 2011, p. 1127 : « Chacun interprète les autres comme s'engageant exactement dans le même type d'activité interprétative, comme adoptant exactement le même type de posture interprétative, comme on le fait soi-même. ».

« déclare » rend également explicite l'attribution d'engagements doxastiques sous la forme de contenus se prêtant à une affirmation. Les êtres logiques sont alors conscients d'eux-mêmes comme créatures normatives et discursives. Seuls des êtres logiques peuvent alors interpréter une communauté (et non un individu particulier) comme engagée dans une pratique discursive.

Les créatures linguistiques ont une intentionnalité dérivée de l'intentionnalité<sup>1040</sup> des pratiques discursives publiques, mais cette intentionnalité est cependant dite originelle, car notamment instituante et implicitement interprétante de statuts déontiques. Les créatures non linguistiques n'ont qu'une intentionnalité dérivée de nos pratiques; elles ne peuvent, en aucun cas, instituer ou attribuer des statuts déontiques. Les êtres logiques sont les seuls capables d'attribuer une intentionnalité originelle (instituante, interprétante) : ils peuvent attribuer de l'intentionnalité aux autres membres et aux autres communautés, en leur attribuant explicitement des attitudes déontiques d'attribution et de reconnaissance d'engagements. Nous n'attribuons donc pas seulement ici la reconnaissance et l'engagement envers des contenus propositionnels, mais aussi des attributions d'engagements :

Seule une créature qui peut rendre explicites des croyances – au sens où elle peut produire des assertions et entretenir un pointage discursif sur des assertions – peut adopter la posture intentionnelle simple et traiter autrui comme ayant des croyances implicites dans son comportement intelligent. De la même manière, seule une créature qui peut rendre explicites ses attitudes envers les croyances d'autrui – au sens où elle est capable d'imputer de pointage – peut adopter la posture explicitement discursive et traiter autrui comme rendant ses croyances explicites, et donc comme ayant une intentionnalité originelle<sup>1041</sup>.

Les créatures linguistiques – rationnelles et/ou logiques – ont une intentionnalité originelle (instituante), bien que dérivée des pratiques sociales d'interprétation. Les êtres rationnels sont

<sup>1040</sup> Cette distinction entre intentionnalité originelle et dérivée est héritée de Searle, 1983, p.44 : « Une énonciation peut être Intentionnelle, une croyance peut être Intentionnelle, mais l'Intentionnalité de la croyance est *intrinsèque*, tandis que l'Intentionnalité de l'énonciation est *dérivée*. ». Searle distingue l'intentionnalité *originelle* ou *intrinsèque* que possèdent les états mentaux et l'intentionnalité *dérivée* que nous attribuons aux objets ou à certains états qui ne possèdent qu'un genre d'intentionnalité « comme si ». Le « chien de mon voisin » représente ce chien particulier, seulement parce que des locuteurs du français utilisent cette phrase pour représenter quelque chose. Les locuteurs possèdent l'intentionnalité intrinsèque, alors que l'intentionnalité du mot est dérivée. Mais, notons, que Brandom, 2011, p.158, rejette la tactique de Searle, choisissant plutôt d'argumenter que l'intentionnalité originelle peut être trouvée, dans les groupes où les agents s'attribuent des états intentionnels les uns les autres.

<sup>1041</sup> Brandom, 2011, p.1123.

aussi capables d'attribuer des statuts déontiques qui instituent l'intentionnalité et le contenu, mais ne reconnaissent toutefois qu'implicitement l'intentionnalité de leurs congénères qu'ils ne font qu'instituer pratiquement, là où les êtres logiques sont capables d'attribuer et de reconnaître de l'intentionnalité originelle.

Chez Brandom, l'intentionnalité originelle (instituant) n'a pas une source interne, mais est produite par nos pratiques sociales. Nous attribuons une intentionnalité dérivée à des créatures non linguistiques, en leur attribuant des engagements, et en expliquant leur comportement en termes de normes conceptuelles<sup>1042</sup>. Cependant, attribuer une intentionnalité dérivée à un système, ce n'est pas le voir comme participant à un jeu déontique. L'intentionnalité originelle est attribuée aux membres d'une communauté, non pas seulement comme activité de tenir le pointage de ces membres, mais en supposant que eux-mêmes marquent le score mutuellement. L'intentionnalité des pensées et des assertions est donc dite dérivée de celle des pratiques publiques en ce qu'elle n'est pas intelligible hors de ces pratiques (en étant attribuée, reconnue). Mais elle est réelle : la posséder, c'est être doté de capacités qui font une différence et qui sont issues de pratiques linguistiques, en considérant chaque créature linguistique comme un membre de cette communauté qui « institue des statuts déontiques et qui confère des contenus conceptuels<sup>1043</sup> ». L'intentionnalité du langage et de la pensée sont au même niveau : l'intentionnalité des contenus linguistiques n'a pas de priorité sur l'intentionnalité des contenus mentaux, ou *vice versa*<sup>1044</sup>. Certes, il y a eu des créatures non linguistiques avant les créatures linguistiques. Mais, l'intentionnalité de ces créatures non linguistiques est dérivée de l'intentionnalité originelle d'une communauté d'interprètes<sup>1045</sup>. Il ne s'agit pas ici de partir de l'intentionnalité élémentaire des créatures non linguistiques pour fonder la possibilité de l'intentionnalité des créatures linguistiques. Il est

---

<sup>1042</sup> Ibid., p.1108, n. 1.

<sup>1043</sup> Ibid., p.1107.

<sup>1044</sup> Ibid., p. 304: il rejoint sur ce point Davidson, 1993b, 1975, p. 229: « Ni le langage ni la pensée ne peuvent être entièrement expliqués dans les termes l'un de l'autre, et aucun d'eux n'a de priorité conceptuelle. ».

<sup>1045</sup> Brandom, 2011, p. 158: « La clé de cette approche est qu'une telle interprétation doit interpréter les membres de la communauté comme se considérant ou se traitant mutuellement dans la pratique comme adoptant des engagements et d'autres statuts normatifs à contenu intentionnel [...] Dans la mesure où leur intentionnalité est dérivée – parce que la portée normative de leurs états est instituée par les attitudes que l'on adopte à leur égard – ils dérivent leur intentionnalité les uns des autres, et non en dehors de la communauté. ».

plutôt question d'expliquer comment une communauté linguistique se constitue et comment son intentionnalité peut se développer et se fonder objectivement.

#### 6.7. Représentation, Perspectivisme - contextualisme et objectivité

Les considérations précédentes peuvent ainsi contribuer à clarifier la portée de la thèse centrale de Brandom, à savoir que :

c'est la dimension *sociale* de la pratique de donner et de demander des raisons qui confère à la pensée et au discours la dimension *représentationnelle* du contenu propositionnel<sup>1046</sup>.

Cette thèse portant sur l'origine sociale et inférentielle du représentationnel serait cependant tronquée si l'on ne considérait pas un autre *desiderata* de Brandom : expliquer comment

La dimension représentationnelle du discours s'enracine dans des différences sociales de la perspective inférentielle, qui correspondent à des répertoires différents d'engagements discursifs<sup>1047</sup>.

On peut donc affirmer que dans ce sens, la stratégie de Brandom est contextualiste : il n'existe pas de contenu sémantique définissable indépendamment d'engagements pragmatiques, d'inférences sémantiques et d'attributions au sein d'un contexte social de communication. La pratique discursive, l'offre et la demande de raisons, implique donc à la fois l'échange de contenus et les relations interpersonnelles. Aussi, la dimension représentationnelle des contenus propositionnels qui jouent les rôles inférentiels de prémisses et de conclusions doit s'entendre sur la base de la dimension sociale ou dialogique des raisons communicationnelles, en supposant que les raisons offertes par d'autres possèdent une signification<sup>1048</sup>. Mais, chaque membre du jeu discursif possède des engagements et des connaissances différents de ceux d'autrui – bien qu'ils soient attribués et reconnus par autrui. Par conséquent, la question que nous nous posons est la suivante : si le contenu n'est relatif qu'à des perspectives inférentielles et normatives subjectives, peut-il être partagé et peut-on

<sup>1046</sup> Ibid., p.883.

<sup>1047</sup> Ibid., p.937.

<sup>1048</sup> Ibid., p.883-884 : « Les contenus conceptuels employés dans le raisonnement *monologique*, [...] se greffent sur, et ne sont intelligibles qu'en référence aux contenus conceptuels conférés par le raisonnement *dialogique*, dans lequel la question de savoir ce qui suit de quelque chose implique essentiellement des évaluations à partir des différentes perspectives sociales d'interlocuteurs ayant des engagements d'arrière-plan différents. ».

parler de normes objectives de la signification ? En utilisant les pratiques sociales d'attribution comme base de survenance du contenu, et en faisant également dépendre la signification inférentielle d'un jugement ainsi que les antécédents et les conséquences d'un engagement doxastique d'un arrière-plan d'engagements auxiliaires précédemment attribués, ne court-on pas le risque de relativiser la dimension représentationnelle du contenu propositionnel ?

Il est erroné, à mon avis, de croire que la thèse d'une normativité sociale de la signification conduirait à une thèse relativiste dans la mesure où Brandom considère clairement, en le justifiant, que l'objectivité est un produit (relatif) de ces pratiques sociales, et que la distinction entre l'apparent et la réalité est au cœur des pratiques sociales discursives<sup>1049</sup>. La dimension sociale et ses différentes perspectives inférentielles sont moins des limites que des ressources pour reconstruire l'objectivité. Bien que le langage et la pensée nous amènent, à notre insu, à regarder le monde à partir d'une certaine perspective, ce ne sont pas les contenus sémantiques eux-mêmes, mais leur expression qui sont propres à une perspective<sup>1050</sup>. Cependant, en raison de la pluralité des arrière-plans d'engagements inférentiels et doxastiques, deux individus, même en employant les mêmes concepts, n'expriment pas la même chose ou le même engagement doxastique. Il n'existe donc pas de manière neutre d'exprimer un contenu conceptuel :

*Il est de l'essence des contenus conceptuels d'être expressivement perspectivaux; ils ne peuvent être explicitement spécifiés qu'à partir d'un certain point de vue, sur l'arrière-plan d'un certain répertoire d'engagements discursifs, et la manière correcte de les spécifier varie d'un point de vue discursif à l'autre.*<sup>1051</sup>

<sup>1049</sup> Dans ce jeu discursif de pointage, nos engagements doivent être examinés à la lumière des relations d'incompatibilité et c'est seulement s'ils ne sont pas mis en échec par des incompatibilités que nous pouvons continuer à participer au jeu discursif; s'ils sont mis en échec, nous sommes alors contraints à des réparations afin d'ajuster nos engagements déviants et être de nouveau admis dans le jeu. Brandom, 2000a, p. 362, sous-entend donc que nos contenus devraient, dans la plupart des cas, converger.

<sup>1050</sup> Brandom, 2009, p.186-187; 2011, p.884: « Sur le versant sémantique, apparaît le caractère socio-perspectival des contenus propositionnels, et par là des contenus conceptuels en général. Les contenus sémantiques des engagements discursifs, des attitudes à l'égard de ces engagements et des performances linguistiques qui expriment ces attitudes ne peuvent en principe être spécifiées qu'à partir de la perspective fournie par un répertoire d'engagements et d'attitudes d'arrière-plan; la manière correcte de spécifier un contenu particulier quelconque varie d'un tel répertoire à un autre. ».

<sup>1051</sup> Brandom, 2011, p.1039 : souligné par l'auteur.

Par exemple, comme le mentionne Brandom, il se peut que l'adepte de Zoroastre *veuille dire* quelque chose d'autre que moi par « soleil<sup>1052</sup> » (ce mot a, dans sa bouche, une autre signification que celle qu'il a à mon oreille) en raison des différences dans ses engagements collatéraux, mais il peut néanmoins parler de la même chose : le soleil. Cela ne signifie donc pas que, d'une certaine façon, nous ne puissions pas néanmoins partager les *mêmes* contenus. Partager un contenu, c'est précisément être capable, pour chaque individu, de concilier des différences dans la perspective sociale et discursive entre interlocuteurs. Ce perspectivisme est certes inévitable dans l'expression des contenus, néanmoins il n'entrave pas une communication et une compréhension inter-perspectivales de contenus; celles-ci nécessitent une interprétation permanente, comme capacité à naviguer d'une perspective inférentielle à l'autre sur la base d'une relation sociale instaurée sur le mode du « je-tu » – d'un arrière-plan de valeurs à un autre, par exemple :

Les contenus conceptuels (de manière paradigmatique les contenus propositionnels) peuvent authentiquement être partagés, mais leur nature perspective signifie pour cela maîtriser le système coordonné de perspectives de pointage, et non pas transmettre de mains en mains (ou de bouche en bouche) quelque chose de non perspectival<sup>1053</sup>.

Des interlocuteurs qui possèdent un arrière-plan d'engagements différents peuvent néanmoins faire les mêmes jugements et exprimer les mêmes concepts inférentiellement articulés<sup>1054</sup>. Une explication de ce phénomène pourrait faire appel à la présence d'un arrière-plan élémentaire stable, incluant d'autres faits sémantiques (à partir desquels nous pouvons discuter de la valeur du terme en question), mais aussi éventuellement les inférences et les critères élémentaires qui peuvent définir la signification minimale (et les conditions d'application) du terme en question (pour en faire un objet de discussion). Cet arrière-plan sémantique élémentaire peut cependant être en perpétuelle révision et discussion, selon le contexte plus général dans lequel il prend place, de nos critères de précisions, et de notre construction d'une perspective commune. Brandom refuse cette alternative : il ne s'agit pas de dire que des individus peuvent exprimer les mêmes contenus parce que chacun de leur arrière-plan inférentiel inclut, entre autres et nécessairement, un ensemble d'inférences

<sup>1052</sup> Brandom, 2009, p.177.

<sup>1053</sup> Brandom, 2011, p.1040.

<sup>1054</sup> Ibid., p.1114 : « Alors, quelque chose de spécial doit être dit à propos du sens où les interlocuteurs, ayant des engagements collatéraux différents, peuvent néanmoins être dits capables de faire les mêmes affirmations et exprimer les mêmes concepts inférentiellement articulés. ».

essentielles qui définissent le contenu en question. Ce serait réduire le contenu à une classe d'inférences privilégiées, et adopter un inférentialisme inégalitaire, non holiste<sup>1055</sup>.

#### 6.7.1. Structure sociale sur le mode du « je-tu » et non du « je-nous »

La stratégie est plutôt la suivante<sup>1056</sup> : en fonction des perspectives doxastiques de chacun, les significations inférentielles des concepts varient, mais le contenu conceptuel, qui *détermine une fonction inférentielle allant de la perspective à la signification*, ne varie pas – bien qu'il soit aussi déterminé et évalué par les normes d'usage de la communauté. Aussi, dans la nature même de l'approche de l'évaluation communautaire, il est nécessaire de ne pas tomber dans le piège de l'erreur d'orientation initiale qui consiste à considérer les relations « je-nous » plutôt que les relations « je-tu » comme la structure sociale fondamentale. Évaluer, approuver, etc., sont des activités que nous faisons en tant qu'individus et que nous nous attribuons mutuellement, ce qui constitue la communauté d'un « nous ». Mais cette façon de voir les choses est déformée par le mode du « je-nous » qui a tendance à identifier ce que certains membres de la communauté (comme les experts par exemple) ou ce que tous les membres de cette communauté tiennent pour vrai en vertu d'une affirmation ou considèrent comme correct en vertu de l'emploi d'un concept avec ce qui est réellement vrai ou correct. Par conséquent, afin d'éviter le relativisme, la relation entre la pratique sociale d'évaluation et l'application correcte d'un terme inhérente à ce jeu de pointage déontique doit considérer la possibilité que tout un chacun, membre de la communauté puisse se tromper en ce qui a trait à l'emploi correct d'un terme ou à la vérité d'une affirmation. Cette relation entre pratique sociale d'évaluation et application correcte des termes doit être telle qu'elle prévienne la dérive communautaire mentionnée plus haut, autrement dit, qu'elle reconnaisse le faillibilisme soit la possibilité de l'erreur communautaire (ou de l'erreur à la première personne)<sup>1057</sup>.

<sup>1055</sup> Ibid., p.1114: « Ce qui, à de nombreux points de vue, pourrait sembler une possibilité de sortir de cette difficulté n'est pas le chemin pris ici. L'approche la plus directe consisterait à adopter une attitude inégalitaire envers les différentes inférences où un concept est impliqué. »

<sup>1056</sup> Ibid., p.1114-1118.

<sup>1057</sup> Ibid., p. 1058-1067, fait explicitement référence à cette condition d'adéquation en soutenant qu'une approche des normes de signification et des énoncés prétendant à la vérité ne peut pas concevoir qu'une erreur communautaire soit impossible. Autrement dit, il affirme qu'une approche adéquate de la vérité objective d'un contenu propositionnel et de l'application correcte et objective d'un contenu

Pour ce faire, Brandom introduit une distinction entre deux types de structure sociale : celle sur le mode du « je-nous » et celle sur celui du « je-tu » qui se calque sur notre distinction entre une socialité forte et une socialité faible<sup>1058</sup>.

La structure sociale sous le mode du « je-nous » fait référence à la relation qui existe entre un individu et la communauté considérée comme un tout. Cette structure sociale accorde un statut privilégié au « nous » communautaire considéré comme l'arbitre qui détient l'autorité (au sens d'évaluateur doué d'autorité) d'évaluer la vérité du contenu propositionnel et l'application correcte du contenu conceptuel (norme de signification) et dont les actions et évaluations ne peuvent pas ne pas être correctes, et dont les membres qui sont sujets à cette autorité sont contraints par les normes instituées par l'accord communautaire du « nous ». Par conséquent, la relation « je-nous » ou celle d'une socialité forte sous ses différentes variantes semble bien ne pas respecter la condition d'adéquation de l'erreur communautaire et mène alors au relativisme<sup>1059</sup>. De plus, elle ne laisse pas de place à la distinction cruciale entre la performance, d'une part, et le statut normatif ou la portée normative d'une performance, d'autre part, c'est-à-dire, entre ce qui est fait, d'une part, et ce qui est correct ou devrait être fait, d'autre part; soit elle échoue à distinguer le fait qu'une affirmation soit correcte (statut normatif) du fait qu'elle est tenue pour correcte (attitude normative) par la communauté dans son intégralité.

La structure sociale sous le mode du « je-tu » décrit, quant à elle, la relation entre un individu ou locuteur et un autre individu membre de la communauté qui tient le pointage. Elle met l'accent sur la relation bipolaire entre deux individus comme lieu de production des standards objectifs et de ce fait, n'accorde pas de statut privilégié au nous communautaire. L'objectivité ne s'apparente ni d'un point de vue sémantique ou épistémique avec le point de vue du « nous », ni ne se trouve éliminée en faveur de l'approche communautaire. Cependant,

---

conceptuel doit satisfaire la condition d'erreur communautaire (ou à la première personne). Il ajoute à celle-ci une seconde condition d'adéquation – la condition d'ignorance communautaire (ou à la première personne) – qui est juste son contraire.

<sup>1058</sup> Ibid., Chap.1, section V et Chap.8, section VI.

<sup>1059</sup> Ibid., p.1055: « les explications *je-nous* postulant à tort l'existence d'une perspective *privilégiée* – celle du « nous », ou de la communauté. La correction objective des affirmations (leur vérité) et de l'application de concepts est identifiée à ce à quoi on adhère à partir de ce point de vue privilégié. L'identification de l'objectivité à l'intersubjectivité comprise de cette façon est fautive en ce qu'elle ne laisse aucune place à la possibilité de l'erreur concernant cette perspective privilégiée; ce que la communauté *tient* pour correct *est* correct. ».

si la structure du « je-tu » ne s'oppose pas à la version à la première personne de la condition d'adéquation, visant à réduire les notions de vérité et d'application correcte des concepts (de la performance linguistique d'un individu) aux dires d'un autre individu (le marqueur), elle doit néanmoins maintenir qu'aucune perspective (pas même celle du marqueur) ne doit être privilégiée<sup>1060</sup>. Aussi, un individu ne peut connaître à lui seul toutes les inférences : il n'en maîtrise que certaines. Le contenu qu'il exprime à partir de son arrière-plan inférentiel est toujours compris et interprété à partir d'un autre arrière-plan inférentiel lui aussi partiel (celui du marqueur). Mais ces deux arrière-plans ont pour caractéristique essentielle de constituer des parties de l'arrière-plan inférentiel général qui détermine le contenu du concept, indépendamment de son expression et de sa compréhension subjectives. Il existe toujours un écart entre l'arrière-plan inférentiel d'un locuteur, l'arrière-plan inférentiel du marqueur, et l'arrière-plan général de la communauté. Personne ne connaît et ne possède cet arrière-plan inférentiel général. Il n'apparaît que dans les différenciations accomplies par les arrière-plans particuliers : il apparaît dans la manière dont je comprends et enrichis, à ma façon, ce que dit autrui à sa manière – cela ne rend pas impossible la communication : je ne modifie pas le contenu de ce que dit autrui, je change la perspective et l'arrière-plan de ce qu'il dit. La perspective d'autrui est placée dans ma perspective inférentielle, elle aussi susceptible d'être interprétée et modifiée différemment (elle peut d'ailleurs être fautive). C'est dans ces changements de perspectives que réside l'objectivité de ce que nous disons. Il est donc important de noter que la structure sociale sur le mode du « je-tu » de Brandom insiste bien sur la « symétrie » des perspectives dans le jeu de pointage déontique. L'application correcte et objective dans l'emploi des termes ou dans la pratique sociale et discursive d'évaluation des engagements ne se réduit pas aux dires du joueur en tant que marqueur, mais se présente plutôt comme la fonction d'une caractéristique formelle ou structurelle du processus – la possibilité permanente inscrite au cœur de la structure même du jeu de pointage de produire

---

<sup>1060</sup> Ibid., p.600: « [...] selon l'interprétation *je-tu* de l'intersubjectivité [...] Mais aucune de ces perspectives n'est privilégiée par avance par rapport à toute autre. [...] à réinterpréter l'objectivité comme consistant en un type de *forme* perspective, plutôt qu'en un contenu non perspectival ou trans-perspectival. Ce qui est partagé par toutes les perspectives discursives, c'est le fait *qu'*il y ait une différence entre ce qui est objectivement correct (dans la façon dont les concepts sont appliqués) et ce qui est seulement considéré comme correct, et non la *nature* de cette différence – c'est la structure et non le contenu . ».

de nouvelles perspectives de pointage, de nouvelles paires de combinaisons sur le mode de la relation du « je-tu » (« locuteur-marqueur ») :

Aussi chaque perspective de pointage maintient-elle une distinction dans la pratique entre statut normatif et attitude normative (immédiate) – distinction entre ce qui est objectivement correct et ce qui est seulement *considéré* comme correct, entre ce à quoi un interlocuteur est effectivement engagé et ce à quoi un interlocuteur est seulement *considéré* comme étant engagé. Pourtant, ce qui, du point de vue d'un marqueur, est objectivement correct – ce à quoi, à partir de cette perspective, un autre interlocuteur est réellement engagé par une certaine reconnaissance – peut être entièrement compris par nous, nous qui interprétons l'activité de pointage, en termes d'attitudes *immédiates*, en référence aux reconnaissances et aux attributions de ce marqueur. Ce qui apparaît au marqueur comme la distinction entre ce qui est objectivement correct et ce qui est seulement considéré ou traité comme correct nous apparaît comme la distinction entre ce qui est reconnu par le marqueur attribuant un engagement et ce qui est reconnu par celui à qui il est attribué. La différence entre le statut normatif objectif et l'attitude normative subjective est interprétée comme une distinction socio-perspectivale entre des attitudes normatives<sup>1061</sup>.

L'idée de base est la suivante : les engagements qu'un interprète ou un marqueur attribue, à partir de son arrière-plan, à un interprété ou à un joueur dépassent les engagements que l'interprété reconnaît ou est disposé à asserter, à partir de son arrière-plan. La différence entre l'objectivité et la subjectivité s'identifie à la différence entre ce vers quoi on s'engage objectivement (à la suite de l'interprétation ou de l'évaluation d'autrui) par une assertion et ce vers quoi on reconnaît être engagé. C'est dans cette différence que se cristallise l'objectivité; cette différence est d'ailleurs exprimée dans les rapports doxastiques *de re*. C'est en dégageant cette objectivité du contenu – cet aspect qui excède sa dépendance inférentielle subjective – que l'on dégage aussi sa dimension représentationnelle. Comprendre ce que représente une expression ou un contenu, c'est justement être capable de discerner ce qui reste constant dans cette expression ou ce contenu quand il change de perspective c'est-à-dire quand il est asserté par d'autres locuteurs. Les membres d'une pratique discursive doivent donc être capables de « transcender leur perspective » afin de reconnaître qu'autrui possède une perspective inférentielle différente de la nôtre et être à même de discerner dans le contenu propre à cette perspective un aspect représentationnel qui nous est commun, mais que nous exprimons différemment (l'interprète, par exemple, substitue une forme à une autre pour exprimer ce contenu).

---

<sup>1061</sup> Ibid., p.1050-1051.

### 6.7.2. Attributions de *re* et de *dicto*

Les formes d'attribution de contenu de *re* sont justement ces formes d'attribution *hybrides* :

Les imputations *de re* constituent l'expression explicite d'attitudes déontiques qui possèdent une perspective hybride : l'attribution, par l'auteur de l'imputation d'un engagement discursif (doxastique ou pratique, et donc doté d'un contenu propositionnel) et la contraction, par l'auteur de l'imputation, d'un autre engagement (substitutionnel)<sup>1062</sup>.

Les attributions explicites (ou imputations) d'engagements peuvent ainsi être *de re* ou *de dicto* : quand un plaignant *S* qui, en caractérisant l'affirmation de son adversaire, dit que « L'avocat de la défense *A* croit que *p* c.-à-d. qu'un menteur invétéré est un témoin digne de foi<sup>1063</sup> », il attribue (explicitement) à *A* un engagement doxastique envers *p*, et il reconnaît (en l'exprimant) son propre engagement doxastique envers cet engagement doxastique de *A* (mais pas nécessairement envers *p*, puisque *S* peut considérer que *p* est faux) soit que « l'homme qui vient de témoigner est un menteur invétéré ». En spécifiant cet engagement doxastique qu'il attribue à *A*, *S* peut utiliser des concepts qui ne font pas partie de l'assertion que *A* serait disposé à exprimer pour exprimer son engagement doxastique envers *p*. Les concepts qu'utilise *S* sont alors liés aux concepts utilisés par *A*, mais par des liens inférentiels que *A* peut ne pas connaître. Le désaccord concerne ici la question de savoir si cet homme est un témoin digne de foi, tandis que le plaignant affirme d'un homme qu'il est un menteur invétéré. L'attribution est alors *de re* ; ce que le plaignant doit dire est « L'avocat de la défense déclare, d'un menteur invétéré, que c'est un témoin digne de foi ». Par conséquent, que quelqu'un soit digne de foi, cela fait partie de l'engagement *attribué* par le plaignant, que cet individu soit en fait un menteur invétéré, cela fait partie de l'engagement *pris* ou *contracté* par le plaignant. L'attribution est alors *de dicto* lorsque l'interprète se base uniquement sur des concepts et des liens inférentiels que l'avocat de la défense ou *A* reconnaît soit « l'avocat de la défense croit que l'homme qui vient de témoigner est un témoin digne de foi ». Les attributions *de re* et les attributions *de dicto* sont deux manières différentes de spécifier le contenu d'un même engagement doxastique, et non deux attributions de deux engagements différents.

<sup>1062</sup> Ibid., p.1033.

<sup>1063</sup> Ibid., p. 895-901; 2009, p.186-187.

Prenons un autre exemple de Brandom<sup>1064</sup> et supposons qu'un agent linguistique soit engagé par ses performances linguistiques à « Benjamin Franklin a inventé le paratonnerre », soit que cet engagement est attribué par un joueur qui marque les points et qui suit ses performances linguistiques. Le marqueur attribue alors, sur la base de cette attribution, au locuteur l'engagement à « Le premier Ministre des Postes et des Télécommunications des États-Unis a inventé le paratonnerre ». Cet engagement est obtenu à partir de la substitution dans une phrase d'une autre phrase qui sert de prémisse et de la substitution d'un terme singulier à un autre. Du fait que Benjamin Franklin a été le premier Ministre des Postes et des Télécommunications des États-Unis, l'inférence de la prémisse à sa variante substitutionnelle en préserve la vérité : dans le contexte approprié, l'engagement envers la prémisse implique un engagement envers la conclusion. Cependant, supposons que non seulement l'agent linguistique ne reconnaisse pas ce dernier engagement mais qu'en plus, il le rejette explicitement. Le marqueur déclarera alors que l'agent n'est pas autorisé à énoncer cet engagement qu'il considère comme erroné « Le premier Ministre des Postes et des Télécommunications des États-Unis n'était pas l'inventeur du paratonnerre », et que la manière dont l'agent applique le concept « l'inventeur du paratonnerre » est incorrecte. L'attribution de *re* est alors « le marqueur croit, du premier Ministre des Postes et des Télécommunications des États-Unis, qu'il est l'inventeur du paratonnerre ». Dans l'attribution de *re*, l'interprète mentionne quel objet est représenté selon son propre arrière-plan doxastique. Il attribue un engagement à l'interprété (« que Benjamin Franklin est l'inventeur du paratonnerre »), et assume lui-même un engagement différent (« que Benjamin Franklin est premier Ministre des Postes et des Télécommunications des États-Unis ») en faisant cette attribution. Le rapport de croyance de *re* exporte ainsi quelque chose de l'interprété à l'interprète, à savoir l'aspect représentationnel du contenu :

Les locutions représentationnelles rendent explicite la distinction des engagements entre ceux qui sont attribués et ceux qui sont assumés – ceux sans lesquels la communication serait impossible, étant donné ces différences de perspective<sup>1065</sup>.

C'est cette différence, inscrite au cœur de la structure sociale du « je-tu », entre les engagements et les autorisations reconnus par le locuteur (c'est-à-dire ceux qui sont attribués

<sup>1064</sup> Brandom, 2009, p.143, p.180.

<sup>1065</sup> Ibid., p. 193.

par le marqueur) et ceux assumés ou endossés par le marqueur qui génère cette distinction entre ce qui *est considéré* comme correct (« what is taken to be correct ») dans la manière d'employer un terme (par le locuteur) et ce qui *est* (« what is correct ») *objectivement correct* (selon le marqueur). Ce envers quoi le locuteur est objectivement engagé est ce qui est rendu explicite dans l'attribution *de re*. Cependant, Brandom précise que le problème ne s'arrête pas là puisqu'il ne s'agit pas de considérer le marqueur comme le seul garant de l'application correcte et objective des termes. Aussi, dans la mesure où la structure du jeu de pointage permet l'itération de marqueurs de telle sorte qu'un premier marqueur qui reconnaît des engagements et autorisations peut aussi être évalué par un autre marqueur qui effectue lui-même des engagements et autorisations de telle sorte qu'il y ait une différence entre les engagements que le premier marqueur reconnaît ou est disposé à reconnaître (attitudes déontiques immédiates) comme correct, et ce envers quoi il est réellement et objectivement engagé (statuts déontiques), sans nécessairement le savoir – cette fois-ci selon le point de vue du nouveau marqueur, et ainsi de suite. La procédure de pointage implique la tenue de deux livres simultanément : un premier livre recueillant les engagements et autorisations que reconnaît l'agent linguistique (c'est-à-dire les engagements et autorisations attribués par un marqueur au score) et un second livre indiquant les engagements et autorisations que le marqueur au score assume ou endosse. De là résulte la possibilité d'une différence entre les engagements et autorisations que le locuteur reconnaît et ceux auxquels le marqueur adhère. Mais, un double livre de pointage peut en retour être tenu sur le premier marqueur par un second marqueur, n'excluant pas encore la possibilité d'une différence entre les engagements et autorisations reconnus par le premier marqueur et ceux assumés par le second, et ainsi de suite pour le second marqueur à l'égard d'un troisième marqueur, de telle sorte que ce qui est vrai ou correct n'est jamais identifié avec ce que l'un ou l'autre des marqueurs considère comme vrai ou comme correct. C'est aussi un trait caractéristique de la structure sociale du « je-tu » de générer de nouvelles et différentes perspectives de pointage, de nouvelles paires de combinaisons sur le mode du « je-tu » : la compréhension linguistique dépend ainsi d'une capacité à évoluer entre sa propre perspective et celle des interlocuteurs, c'est-à-dire d'attribuer explicitement ou implicitement des engagements envers des contenus *de re*<sup>1066</sup>.

---

<sup>1066</sup> Brandom, 2011, p. 910.

Aussi, pour comprendre ce que l'autre dit, nous traduisons ces assertions dans notre réseau d'inférences, à partir de nos pôles référentiels (*de re*) :

Extraire une information des remarques des autres nécessite de mettre en œuvre le genre d'interprétation substitutionnelle qui est exprimé explicitement lorsque l'on propose des caractérisations *de re* des contenus de leurs croyances – il faut, en d'autres termes, être capable de dire ce à *propos de quoi* leurs croyances seraient vraies si elles étaient vraies. C'est saisir le contenu représentationnel de leurs affirmations<sup>1067</sup>.

Nous oscillons alors en permanence entre la compréhension d'autrui à partir de nos critères, à partir de la manière dont les choses sont pour nous, dans le contexte de nos croyances (*de re*), et la compréhension d'autrui à partir de ses contenus de croyance (*de dicto*). Le changement de perspective est rendu explicite lorsque l'on contraste les attributions *de re* avec les attributions *de dicto*.

### 6.7.3. Réponse de Brandom à l'objection d'Habermas

L'objection soulevée par Habermas dans son analyse critique de *Rendre Explicite* à l'encontre de la théorie de Brandom ne nous paraît donc pas justifiée. En effet, Habermas prétend que Brandom ne rend pas adéquatement compte du rôle spécifique attribué à une seconde personne :

He does not really do justice to the position of the specific role of the second person<sup>1068</sup>.

Selon Habermas, Brandom n'a jamais vraiment réussi à relier sa théorie de l'assertion à la perspective sociale sur le mode du « je-tu »; il en reste, selon lui, à une perspective sur le mode du « je-il », où le rôle conféré à une troisième personne est celui d'observateur ou de spectateur et non d'interlocuteur :

« It is no accident that Brandom prefers to identify the interpreter with a public that assesses the utterance of a speaker – and not with an addressee who is expected to *give the speaker an answer*. Every round of a new discourse opens with an ascription that the interpreter undertakes from the observer's perspective of a *third person*<sup>1069</sup>. ».

<sup>1067</sup> Ibid., p.916.

<sup>1068</sup> Brandom, 2000a, p. 362.

<sup>1069</sup> Habermas, 2000, p. 345.

Dans sa réponse à Habermas, Brandom concède ainsi ce point<sup>1070</sup>, mais rejette la prémisse à la base de la théorie de l'agir communicationnel de Habermas selon laquelle le but de la communication est l'accord (envisagé comme une contrainte ou une obligation morale), c'est-à-dire aboutir à une compréhension mutuelle<sup>1071</sup> :

« But I deny this – not because I think that linguistic practice has some *other* point, but because I think it is a mistake to think of it as having a point at all. Linguistic practice is not for something. It does not, as a whole, have an aim or a goal. It may and does, of course, fulfill many functions. But none of them is its *raison d'être*<sup>1072</sup>. ».

Selon Brandom, la pratique linguistique n'a donc pas de *raison d'être* particulière. Ce n'est peut-être pas, énoncé ainsi, la réponse la plus satisfaisante. Cependant, Brandom cherche surtout à souligner que cette notion d'une compréhension mutuelle est quelque peu incomplète et partielle dans la mesure où elle ne rend pas adéquatement compte de l'essence même du processus de compréhension, celui-ci devant être perçu comme le produit d'une coordination partagée de différentes perspectives individuelles contribuant, à l'image des pions (des coups joués) dans un échiquier, à faire avancer ce jeu discursif de *production et de demande de raisons*. Brandom exprime ce point au moyen de la métaphore des danseuses de balai<sup>1073</sup>.

<sup>1070</sup> Brandom, 2000a, p. 362: « I think this is on the whole a fair characterization. There are, to be sure, some elements that are not merely observational in the pragmatics – we might think of the fundamental notion of a *challenge*, which supports the characterization of discursive practice as a 'game of giving and asking for reasons'. And on the semantic side, the *symmetry* of the relation between the one acknowledging a commitment (paradigmatically by producing an assertional speech act) and the one attributing it is a deep-seated feature of the account of scorekeeping, communication, and the perspectival character of propositional and other sorts of conceptual content. ».

<sup>1071</sup> Ibid., p. 363: « From his point of view, the achievement of this sort of mutual understanding is *the* central communicative phenomenon [...] Perhaps so. But I am not so convinced. To begin with, mutual understanding in the strong sense Habermas is insisting upon is *not* required for the undertaking of joint projects. ».

<sup>1072</sup> Ibid., p. 363.

<sup>1073</sup> Strydom, 2006, p.168, n.1 : « [...] that Pierre Bourdieu (1990, n.8) analysis of dance as a model of social integration involving the 'synchronization of the homogeneous and orchestration of the heterogeneous' indirectly supports Brandom's example of Fred Astaire and Ginger Rogers dancing to illustrate how different things are being done at the same time as the participants are engaged in a shared practice [...] Where Bourdieu and, I think, Brandom see a shared situational structure or, rather, a structuring practice which allows participants nevertheless to execute different yet coordinated actions, Habermas (2000, p.347) insists that the participants are at best engaged in 'mutual observation' and that this excludes 'the possibility for them to converge in their intersubjective recognition of the same validity claim and [to] share knowledge in the strict sense of the term' ».

« The participants do not need all to be doing the same thing (sharing) in a narrow sense in order to be engaged in a joint enterprise, and in the broader sense to be doing the same thing (sharing) [...] Conversational partners should not be pictured as marching in step, like soldiers on parade, but more as ballrooms dancers, each making different movements (at any moment, one leads and the other follows, one moves forward and the other back, one sways left, the other right, and so on) and *thereby* sharing a dance that is constituted by the coordination of their individually different movements<sup>1074</sup>. ».

On comprend donc que l'objectif poursuivi par Brandom est avant tout de décrire une pratique assertionnelle plus fondamentale c'est-à-dire une pratique sociale dans laquelle les actions, les réponses et les attitudes de pointage sont appropriées ou correctes (toujours selon le point de vue du joueur qui marque les points). C'est à ce niveau que la pratique assertionnelle, contrairement à l'objection soulevée par Habermas<sup>1075</sup>, revêt bien une dimension dialogique et non seulement monologique, dans la mesure où chaque assertion n'a de sens, comme coup dans ce jeu dialogique, que si elle reçoit une réponse (si elle est validée ou rejetée sous réserve de justification et de correction) et si un accord (succès du jeu<sup>1076</sup>) est envisageable. Par ailleurs, Habermas ne semble pas non plus formuler une véritable objection en reprenant l'exemple de Brandom du procureur général et de l'avocat de la défense<sup>1077</sup> quant à la question de savoir si le témoin est un homme digne de foi ou un menteur invétéré :

---

<sup>1074</sup> Brandom, 2000a, p. 363.

<sup>1075</sup> Habermas, 2000, p.346 : « The intention that a speaker connects with an utterance amounts to more than just the interpreter's attribution to him of a corresponding belief without his being interested in the interpreter's position on this belief. Rather, as a participant in communication, the speaker with his assertion makes a demand on an addressee to say 'yes' or 'no' publicly; at any rate he expects some kind of reaction from her that can count as an answer and that can produce obligations relevant for the sequel of interaction for both parties. ».

<sup>1076</sup> Comme Dummett, 1978, p.2, le souligne en discutant de la finalité des « jeux assertionnels » : « It is part of the concept of winning a game that a player plays to win, and this part of the concept is not conveyed by a classification of the end positions into winning ones and losing ones. [...] Likewise, it is part of the concept of truth that we aim at making true statements. »; cité par Marion, 2012, p.155.

<sup>1077</sup> Brandom, 2011, p. 897.

« [...] the communicative exchange is played out on two different levels: on the one level, both the prosecutor and the defence attorney are speaking to one another in that [...] they reciprocally dispute the correctness of each other's utterance. At the same time, of course, they are aware of the presence of the judges, the jurors and the public who, on a second level of communication, are following their exchange of words and silently assessing it. Interestingly, Brandom singles out the indirect communication of the speakers with the public who is listening to them – and not the communication of those directly involved – as the paradigmatic case<sup>1078</sup>. ».

Habermas pense que Brandom considère ici deux niveaux distincts de communication c'est-à-dire deux niveaux que Brandom tiendrait séparément, à savoir : un premier niveau de communication qui concerne les acteurs directement impliqués et un second niveau de communication (une communication indirecte) mettant en scène les locuteurs (les jurés) et les spectateurs (le public) qui se contentent de les écouter. Il est pourtant clair de constater que Habermas ignore complètement la pratique discursive présente ici. En effet, plutôt que d'être tenus séparément, la pratique discursive et le jeu de pointage déontique exigent, au contraire, que ces deux dimensions soient envisagées simultanément. La trame de la pratique sociale sur le mode du « je-tu » (et non sur le mode du « je-nous », celle-ci préconisant le point de vue du « nous<sup>1079</sup> ») décrit avant tout, tel que Brandom le démontre, la relation entre un auditoire qui attribue dès engagements et qui marque en conséquence les points, et un locuteur qui contracte des engagements, dont la somme des points est établie. L'usage de telles imputations de *re* et de *dicto* d'une attitude propositionnelle tourne essentiellement autour des distinctions sociales de perspective doxastique entre celui qui impute un engagement et celui à qui un engagement est imputé – autrement dit, entre le répertoire déontique associé à celui par qui un engagement propositionnel est *contracté*, et à celui à qui un engagement propositionnel est par là explicitement *attribué*. Ainsi, dans le cas discuté, certainement que

<sup>1078</sup> Habermas, 2000, p. 345.

<sup>1079</sup> Une réponse de Brandom à Habermas sur ce point se lit chez Scharp, 2003, p. 56-57. Brandom considère clairement qu'il n'existe pas de point de vue privilégié (« Every perspective is locally privileged, but none is globally privileged ») ; la perspective du « nous » communautaire ne peut donc avoir un statut particulier (Brandom, 2011, p. 1050-1055; cité n. 1058). L'objection de Habermas, 2000, p. 347, provient précisément, comme le souligne Scharp, de la considération, dans l'élaboration de sa théorie morale (c.-à-d. de son discours éthique), d'une perspective privilégiée accordée aux normes de la communauté : il semble ainsi que ce soit la raison qui justifie son interprétation de la théorie de Brandom comme une forme d'« individualisme méthodologique ».

dans la salle d'audience, les juges qui entendent (*hearing*<sup>1080</sup>) l'affaire et les jurés qui l'écoutent (*listening*) sont ceux qui suivent de plus près la progression de la discussion et qui tiennent les cartes de marquage des parties respectives (c.-à-d. qui recensent les points marqués par l'avocat de la défense et par le procureur) afin d'être capable d'évaluer la fiabilité et la justesse des propos du témoin en question. Il est certain que la perspective des deux locuteurs (avocat de la défense et procureur général) est plus centrale que le point de vue de l'audience ou encore que la perspective des juges est plus centrale que celle des jurés mais il ne faut pas pour autant en déduire qu'une perspective prime arbitrairement sur une autre; aucune perspective n'est en soi privilégiée. Brandom veut simplement souligner le fait « qu'une phrase, dans la bouche d'une personne donnée, n'a pas typiquement la même portée que cette même phrase provenant d'une autre personne, même dans le cas de partage intégral du langage et de parfaite compréhension mutuelle<sup>1081</sup>. ». Le succès final de la communication résulte finalement de ce jeu dialogique<sup>1082</sup> d'attaque et de défense et de cette capacité à naviguer à travers les différences de points de vue<sup>1083</sup>.

<sup>1080</sup> Habermas, 2000, p.345: « Listeners have a different role than hearers. »

<sup>1081</sup> Brandom, 2011, p.904.

<sup>1082</sup> Dans ce processus de justification, une assertion est définie comme un coup dans ce jeu de production et de demande de raisons. La logique dialogique de Lorenz, 1981 et Lorenzen, 1978, l'analyse ainsi : « demander des raisons » correspond à « des attaques » dans la logique dialogique alors que « produire des raisons » correspond à des « défenses ». De plus, les « attaques » sont parfois décrites comme « des droits » alors que les « défenses » sont décrites comme des « devoirs ». Habermas, 2000, p.328, appuie d'ailleurs cette description du processus dialogique: « Naturally, the practice of argumentation lends itself particularly well to a description in terms of rights and duties. The proponent of a truth claim is obliged to offer justifications while the opponent has the right to contradict her ». Nous avons donc, comme le souligne Marion, 2012, p. 148: « Right to attack ↔ asking for reasons; Duty to defend ↔ giving reasons ». Pour arriver à gagner, c'est-à-dire réussir à défendre avec succès nos assertions contre un opposant, nous devons être capables d'appuyer nos assertions d'une justification ou d'en fournir les raisons. Cependant, nous devons souligner le fait que, contrairement aux systèmes dialogiques, qui mobilisent deux joueurs, les systèmes de déduction naturelle sont « monologiques », au sens où ils mettent en scène une personne. Or Brandom ne quitte pas le cadre de la déduction naturelle. Nous ne voulons cependant pas traiter cette question ici. Voir Marion, 2012, p. 149-165, pour une analyse plus détaillée et plus fine de cette théorie de l'assertion en termes d'une sémantique des jeux.

<sup>1083</sup> Il nous faudra examiner, par la suite, si Habermas ne se rend pas aussi coupable de la critique qu'il adresse à Brandom ou plutôt s'il rencontre des difficultés similaires à celles de Brandom. Autrement dit, il nous semble peut-être plus pertinent, plutôt que voir si l'un ou l'autre ne s'enlise pas dans une perspective communautaire préconisant le point de vue du « nous », de regarder en quoi les deux approches peuvent se recouper et si certains éléments de la théorie de Brandom ne pourraient pas enrichir la pratique communicationnelle de Habermas.

L'objectivité n'est donc pas propre à un contenu qui transcenderait toute perspective, elle est plutôt exemplifiée dans le fait que toutes nos perspectives sont capables de faire une différence entre ce qui est vu comme correct et ce qui est objectivement-pour-nous correct. Brandom cherche ainsi à fonder une notion authentique (faillibiliste) d'objectivité – une approche de la vérité et de l'application correcte qui satisfasse la condition de l'erreur communautaire (ou de l'erreur à la première personne) en ne permettant pas d'identifier ce qui est vrai avec ce que tout un chacun (ou ce que tout le monde) considère comme vrai ou comme correct. Notons également que son approche de la vérité et des normes de signification en termes de conditions de correction est contextualiste : elle tente d'envisager la possibilité d'erreurs communautaires (ou d'erreurs à la première personne) dans nos jugements respectifs ou dans la manière dont nous appliquons les concepts en adoptant un autre point de vue dépendamment du contexte (ou point de vue du marqueur) à partir duquel ces jugements ou ces applications-ci doivent être évalués. Il n'existe donc pas de pointage dont les standards ou critères d'évaluation peuvent être fixés sans tenir compte du contexte. Cette forme de contextualisme n'exclut apparemment pas la possibilité de rejeter ou de corriger un jugement émis ou l'application d'un concept qui a communément été ratifié par la communauté, mais seulement la possibilité d'agir ainsi en l'absence de toute perspective de pointage.

#### 6.7.4. De l'externalisme sémantique à l'externalisme transcendantal social

Cependant, Brandom semble penser que l'approche sociale et perspectiviste est insuffisante pour rendre compte de l'objectivité des normes de signification et complète ainsi celle-ci en ajoutant un nouvel élément : l'externalisme sémantique. L'externalisme sémantique met en avant l'idée que les énoncés prétendant à une vérité objective et les normes de signification déterminant une application correcte des concepts *répondent à la manière dont les choses et leurs propriétés sont dans le monde tout court*<sup>1084</sup>. Autrement dit,

<sup>1084</sup> Brandom, 2011, p. 1046: « [...] il reste qu'une part essentielle de la dimension représentationnelle de nos concepts – c'est-à-dire la manière dont ils visent à s'appliquer au monde objectif – consiste en ce que la correction ultime de leur application tient non pas à ce que je ou l'ensemble d'entre nous *considérons* comme étant le cas, mais à ce qui *est* réellement le cas. Il est essentiel au fait que nos concepts portent sur un monde objectif que leur application soit régie par une correction en un sens *objectif* – un sens où leur adéquation est fonction des objets auxquels ils sont appliqués et au monde

l'application correcte et objective de l'emploi des concepts (et de ce fait, des énoncés prétendant à une vérité objective) est déterminée par le monde *lui-même*, plutôt que par (le potentiel permanent des différents points de vue inhérents à la structure sociale et perspective d'évaluation des engagements en jeu sur le mode du « je-tu ») les dires des locuteurs sur le monde. Une relation sémantique allant du langage au monde paraît bien être ce qui garantit, en dernière instance, la vérité objective des jugements et l'application correcte et objective des normes de signification. Par conséquent, au lieu d'expliquer les notions sémantiques représentationnalistes en termes inférentialistes et sociaux et d'inverser l'ordre traditionnel d'explication sémantique, Brandom semblerait finalement offrir une approche représentationnaliste<sup>1085</sup> – les objets et les faits sont ce à quoi nos jugements et nos applications conceptuelles (engagements) répondent et semblent être ce qui détermine si ces mêmes jugements et applications de concepts sont objectivement vrais ou corrects. Ce que nous signifions dépendrait donc, selon Brandom, de la manière dont les choses sont en réalité, que nous (la communauté ?) le sachions ou non<sup>1086</sup> – le contenu d'un jugement est articulé dans des relations inférentielles objectives, qui ne correspondraient à aucune perspective, mais à la manière dont le monde est, *quoi que nous en pensions*.

Cette position va bien évidemment à l'encontre de l'approche anti-représentationnaliste, contextualiste et sociale de l'objectivité des énoncés prétendant à la vérité et des normes de signification exprimée par son engagement à l'égard des thèses de Hegel et de Sellars, selon lesquelles l'idée de la chose en soi indépendante à la fois de toute pratique linguistique et de tous contextes épistémiques sur laquelle viendrait se greffer l'idée d'un point de vue de Dieu auquel nous aurions un accès direct ne fait aucun sens. Pour Brandom, comme pour d'autres contextualistes, le monde et notre accès à celui-ci ne peut qu'être le résultat du déploiement

---

des faits comprenant ces objets. »; « La manière dont les choses et les propriétés dont ils parlent sont réellement détermine la correction des engagements de tous les membres de la communauté pareillement. » (Ibid., p. 1112).

<sup>1085</sup> Nous pensons qu'il s'agit là d'une confusion. L'expression « répondre à des états du monde » n'implique pas nécessairement des présupposés représentationnalistes. Selon Brandom, nos comportements discursifs sont signifiants dans la mesure où ils répondent à des états du monde, ce qui ne veut pas dire que la signification dépende de la capacité d'un comportement discursif à représenter correctement un état du monde. Certains comportements répondent ou sont fondés dans un état du monde sans pour autant dire qu'ils le représentent.

<sup>1086</sup> Ibid., p.1135: « L'externalisme sémantique – la manière par laquelle ce que nous voulons dire dépend de la manière dont les choses sont réellement, que l'on sache ou non comment elles sont – est un trait du caractère perspectival du contenu propositionnel. ».

de certains concepts et d'un jeu de langage socialement articulé. Il n'existe pas de concept indépendant du monde, ni de perspective indépendante de tous déploiements de concepts à partir de laquelle nous aurions un accès au monde. Pourtant, l'externalisme sémantique répond à ces deux idées. Par conséquent, en préconisant la thèse de l'externalisme sémantique, une tension se cristallise au cœur de l'approche de l'objectivité de Brandom entre d'une part, une position représentationnaliste externaliste et une position inférentialiste contextualiste. L'élément représentationnaliste externaliste exprimée par la tournure rhétorique « *answering to the world* » introduit une notion anti-contextualiste d'objectivité – ce qui est vrai ou ce qui est correct est déterminé par le monde un et objectif, indépendant de toutes pratiques sociales ou de toutes procédures d'évaluation de pointage. Autrement dit, Brandom introduit, par le biais de la notion de « *answering to how things are in the world* » une notion de vérité et d'application correcte des normes de signification indépendante de toutes pratiques sociales, une notion que le contextualisme ne peut pas soutenir. Rorty reproche alors au discours sémantique externaliste de Brandom d'être anti-contextualiste et de ce fait non cohérent avec une approche approfondie à la manière de Hegel, de Wittgenstein, de Davidson des normes objectives de la signification basées sur la pratique sociale que Brandom est surtout soucieux de défendre<sup>1087</sup>. Selon Rorty, la position de Brandom s'apparente au représentationnalisme – l'image moderne de l'épistémologie comme l'esprit miroir de la nature et la théorie de la vérité correspondance – auquel s'ajoutent toutes les apories des problèmes sceptiques de la philosophie moderne. Rorty conclut que Brandom (« *should just stop talking about 'answering' altogether*<sup>1088</sup>»), devrait abandonner une telle notion représentationnaliste de l'objectivité plutôt que de tenter d'en rendre compte en adoptant une attitude *conciatrice* (« *conciliatory attitude* »), et rejoindre son approche éliminativiste, que Rorty qualifie d'attitude *oppositional* (« *oppositional*<sup>1089</sup>»). Rorty admet que son alternative oppositionnelle nous laisse avec une notion plus faible d'objectivité mais, il affirme d'une part, que c'est tout ce dont nous avons besoin pour rendre compte de la notion d'objectivité et d'autre part, que c'est tout ce que nous devrions désirer ou espérer avoir.

---

<sup>1087</sup> Rorty, 1997a; 1998b.

<sup>1088</sup> Rorty, 1998b, p.135.

<sup>1089</sup> Rorty, 1997a, p.173: « That is an intuition I cannot accommodate, so I suggest that you try to give it up. »

Brandom réalise la tension entre l'externalisme sémantique et le contextualisme social qu'exploite Rorty et essaie de la désamorcer en tentant de fournir une description pertinente de la notion de « cette manière dont sont les choses » (en faisant appel à son externalisme sémantique) en la présentant comme un artefact social issu de la différence entre deux perspectives quant à la manière dont sont les choses, entre ce qu'un locuteur considère comme correct et ce qu'un marqueur dit être vrai ou correct (ou dans les termes de Brandom, entre les engagements et autorisations qu'un marqueur au score attribue à un interlocuteur et les engagements et autorisations qu'un marqueur au score assume lui-même) :

En réalité, cependant, la compréhension de ce qui est exprimé par les évaluations de la correction objective des applications de concepts exige que l'on ne fasse appel ni à des faits non perspectivaux (= aux contenus propositionnels vrais) ni à des engagements communautaires à des contenus propositionnels. Au lieu de cela, la distinction entre les affirmations ou les applications de concepts qui sont objectivement correctes et celles qui sont seulement considérées comme étant correctes est une caractéristique structurelle de chaque perspective de pointage.<sup>1090</sup>

Ici, l'application correcte et objective des normes de signification et des énoncés prétendant à la vérité n'est pas déterminée par la manière dont sont les choses indépendamment des pratiques d'interprétation ou des attributions et évaluations de pointage, mais résulte des différences entre ces deux perspectives sociales que sont le fait de *reconnaître* (et donc d'assumer) soi-même un engagement et le fait d'*attribuer* un engagement à un autre. Parler de la manière dont le monde est, revient, en définitive, à souligner l'idée qu'il est faux d'opposer « la pratique discursive et un monde de faits et de choses *en dehors* d'elle, lequel contraste est modelé sur le contraste entre les mots et les choses auxquelles ils réfèrent<sup>1091</sup> » : la manière dont les choses sont est une fonction des différences établies sur le mode du « je-tu » dans les évaluations de nos croyances (engagements et autorisations) sur le monde. De

<sup>1090</sup> Brandom, 2011, p.1047.

<sup>1091</sup> Ibid., p.607: « Ce qui détermine les pratiques d'une communauté dépend de ce que les faits sont et à quels objets elles ont réellement affaire pratiquement, et pour commencer, dans la perception et dans l'action. La manière dont le monde est, contraint les propriétés d'engagements inférentiels, doxastiques et pratiques d'une manière directe, à partir de l'intérieur de ces pratiques (note 1 : au sens où il n'y a rien qui soit en principe en dehors de la pratique discursive). ». Habermas, 2000, p.338, commente ainsi l'exemple de Brandom du papier de tournesol afin de montrer que tout découle de l'intérieur de nos pratiques (« [...] of how observation a piece of litmus turning blue provides a reason for correction the hitherto established rules of application for the concept 'acid' ») : « Here, perception is generated through an experimental action. Such an experiment, however, merely exploits the internal connection between perception and action that already exists in everyday practice and that makes possible ordinary 'learning from experience' ».

cette manière, Brandom soutient ainsi que « l'externalisme sémantique est un externalisme perspectival<sup>1092</sup> » en reconstruisant l'objectivité comme le produit d'une pratique immanente ancrée dans nos pratiques linguistiques et sociales. Cependant, la question que nous nous posons est celle de savoir si cette approche contextualiste, proposée par Brandom, réussit à fonder une notion faillibiliste authentique de l'objectivité des normes de signification.

Afin de répondre à cette question, il s'agit d'examiner les raisons qui ont amenées Brandom à introduire un élément externe dans son approche de l'objectivité, c'est-à-dire à expliquer pourquoi le contextualisme pourrait être insuffisant à garantir l'objectivité<sup>1093</sup>. En effet, si le contextualisme cherche à garantir l'objectivité, soit le faillibilisme, alors il doit distinguer entre ce qu'une pratique donnée considère comme « correct » ou comme « vrai » et ce qui est correct ou vrai indépendamment de cette pratique en question, c'est-à-dire qu'il ne doit pas identifier ce qui est correct avec ce qu'une pratique donnée s'accorde à considérer comme correct. Mais, puisque, selon les contextualistes, il n'existe pas de perspective indépendante de nos pratiques, ceux-ci justifient le bien-fondé de cette notion indépendante d'une pratique particulière de l'objectivité en s'assurant de la possibilité de toujours pouvoir adopter un autre point de vue sur cette pratique en question. La question qui se pose alors est la suivante: est-ce que le fait de disposer de la possibilité structurelle de produire, sans aucune limitation, une pluralité de points de vue suffit à fonder le faillibilisme? Le faillibilisme soutient l'idée que tout contexte donné rattaché à l'accord ou à l'évaluation d'une pratique peut s'avérer erroné, dans la mesure où il peut être corrigé par un autre accord contextualisé, qui peut lui aussi être à son tour erroné et sujet à révision et à correction par un autre accord contextualisé, et ainsi de suite. Le faillibilisme sous-entend donc l'idée qu'un point de vue puisse être erroné et qu'un autre puisse venir le corriger c'est-à-dire, que l'un est alors meilleur que l'autre. L'objectivité structurelle génère seulement une pluralité de points de vue ; il est donc nécessaire de distinguer le pluralisme du faillibilisme: le pluralisme implique seulement l'idée de différents points de vue, le faillibilisme, quant à lui, soutient que ces différents points de vue doivent être soumis à une évaluation comparative les situant les uns par rapport aux autres. Si l'on en reste uniquement à constater la divergence dans nos points

---

<sup>1092</sup> Brandom, 2011, p.1113.

<sup>1093</sup> Je ne prétends pas ici que ce paragraphe est un argument qui apparaît dans *Rendre explicite*, c'est plutôt la reconstruction d'une justification possible afin de compléter le contextualisme social du « je-tu » avec l'externalisme sémantique.

de vue, alors il est difficile de distinguer entre ceux qui sont corrects ou erronés et ceux qui sont meilleurs ou moins pertinents que d'autres. Pourquoi alors ne pas conclure que tous sont corrects selon leur propre vision du monde soit finalement qu'ils n'ont aucun rapport les uns avec les autres? Un mécanisme structurel qui garantit la possibilité de produire différents points de vues ou différentes perspectives de pointage ne suggère ni l'idée que certains d'entre eux puissent être erronés et que d'autres puissent être corrects, ni même celle que certains maximisent mieux l'accord que d'autres. Autrement dit, le perspectivisme structurel semble seulement fonder un pluralisme qui ne se laisse pas soumettre à la critique plutôt qu'un faillibilisme, c'est-à-dire une prolifération de points de vue au lieu d'une progression ou d'une gradation allant de l'incorrect vers le correct ou même du pire vers le meilleur. Disposer seulement de la capacité de générer différentes paires de combinaisons sur le mode du « je-tu » ne nous paraît pas suffisant pour justifier le bien-fondé du faillibilisme.

Par conséquent, concevoir l'objectivité comme le produit de différences de perspectives entre le joueur et le marqueur ne permet pas nécessairement d'éviter le danger de préconiser préférentiellement le pluralisme sur le faillibilisme, ce qui est, ce que l'externalisme *via* la réponse au monde peut être amené à faire. Faire appel à un argument transcendantal c'est-à-dire à l'idée d'un monde unique, « un » et objectif, d'une seule manière dont le monde est indépendamment de ce que chacun dit ou pense à son sujet, peut, d'une certaine manière, garantir le faillibilisme en s'assurant que les différents points de vue que nous émettons portent sur la même chose et que l'on ne soutienne pas en même temps une chose et son contraire. Cependant, cette manière de garantir le faillibilisme semble ne pas s'accorder avec le contextualisme qui rejette précisément l'idée d'une manière unique de voir le monde indépendante de nos interprétations et pratiques linguistiques. Compte tenu de cette incompatibilité, Brandom devrait-il considérer le conseil de Rorty et adopter sa position « oppositionnelle » ? Nous n'irons pas jusque-là dans la mesure où la notion plus faible de l'objectivité avec laquelle Rorty nous laisse est juste celle de l'objectivité perspectiviste, laquelle nous paraît insuffisante pour fonder une notion adéquate d'objectivité, soit une notion authentiquement faillibiliste et non simplement pluraliste et indiscutable. Par conséquent, nous aurions tendance à admettre avec Rorty que l'externalisme sémantique de Brandom ne s'accorde pas très bien avec son contextualisme social, mais nous n'adhérons pas à la stratégie de réponse éliminativiste ou « oppositionnelle » de Rorty. Aussi, comme

nous venons de le mentionner, ce type de position n'évite pas nécessairement l'écueil du relativisme, c'est-à-dire ne fournit pas une véritable notion faillibiliste d'objectivité. Par conséquent, nous sommes également d'avis avec Brandom qu'il est nécessaire de fonder une notion plus forte d'objectivité, mais notons que sa manière de procéder risque d'être, soit ambiguë, soit de nous fournir une notion d'objectivité plus faible que celle requise. Nous pensons donc que Brandom est pris dans le dilemme suivant : soit il soutient un contextualisme cohérent qui ne réussit néanmoins pas à fournir une notion adéquate d'objectivité aux énoncés prétendant à la vérité et aux normes de signification, c'est-à-dire qui fonde le pluralisme et non le faillibilisme, soit il réussit à fonder une approche adéquate de l'objectivité en formulant un argument transcendantal c.-à-d. à l'idée ou au concept d'un monde un et objectif – d'un monde indépendamment de la pratique auxquels les concepts et les affirmations répondent (externalisme sémantique) mais cette approche est ambiguë ou risque d'être incohérente avec son cadre explicatif de travail explicitement contextualiste et social<sup>1094</sup>.

#### 6.8. Vers une conception sociale et historique des normes de la signification

Brandom développe une approche plus cohérente de la pratique immanente de l'application correcte et objective des normes de signification dans laquelle toute référence à un externalisme sémantique a été supprimée<sup>1095</sup>. Brandom n'avance pas que nous sommes responsables de l'application actuelle d'un terme, en tant que celle-ci devrait correspondre au monde, il mentionne plutôt que nous sommes tenus responsables de notre application actuelle d'un terme à l'égard de ses applications passées et futures. Autrement dit, c'est l'autorité du passé (les applications passées des concepts) sur l'avenir (les applications futures des concepts), qui fournit un sens dans lequel les applications actuelles et futures sont

---

<sup>1094</sup> Notre intention n'est pas ici de dire que ce dilemme mène à une aporie mais de proposer une manière d'en sortir, ce que Brandom ne semble pas envisager, mais que suggère l'approche de Habermas et de McCarthy. Leur stratégie est de concevoir l'idée d'un monde un et objectif comme une idée régulatrice (dans le sens de Kant) n'obéissant pas aux principes représentationnalistes. Par conséquent, leur approche envisage un compte-rendu de la vérité et des conditions de l'application correcte des concepts qui se fonde sur la pratique sociale intersubjective, celle-ci nous apportant l'objectivité que Wittgenstein, Davidson et Brandom pensent que nous ayons besoin pour fonder nos normes de signification et nos énoncés prétendant à la vérité.

<sup>1095</sup> Brandom, 2000b, p. 231-261.

responsables, pour ce qui est de leur correction, envers les applications passées. Ce faisant, s'ajoute à son approche normative socialement perspective de la pragmatique et à son inférentialisme non pas une relation externe allant du langage au monde mais bien une relation historique instituant les applications des termes et les significations des termes utilisés. L'application correcte et objective d'un terme n'est alors ni comprise en termes de sa responsabilité à l'égard d'un monde objectif de choses indépendant de nos pratiques d'attribution et d'évaluation des engagements, ni comme une simple question de différences dans les engagements que nous devons avoir reconnus<sup>1096</sup> et ceux mêmes qui nous sont attribués par un autre joueur en tant que marqueur (*scorekeeper*) et évaluant le caractère matériellement correct ou approprié de ces engagements-ci (ou un joueur marquant les points et évaluant ce joueur-ci qui marque lui aussi les points, et ainsi de suite) à un temps historique spécifique (soit le présent), mais émerge plutôt des différences entre d'une part nos applications et évaluations actuelles et, d'autre part, celles des autres joueurs de concert avec nos applications et évaluations antérieures et futures et celles de nos semblables. L'application et l'institution de normes conceptuelles émanant de cette activité d'évaluation de pointage n'est donc pas simplement un processus temporel, mais bien un processus historique. La dimension sociale de la reconnaissance ne peut ainsi être pleinement comprise en dehors des dimensions inférentielle et historique. Nous entendons par là qu'elle manifeste une structure de reconnaissance<sup>1097</sup>, qui résulte de l'autorité réciproque exercée, d'un côté, par les applications passées de concepts sur les applications futures et, d'autre part, par les applications futures de concepts sur les applications passées. Il n'y a, pour instituer des

---

<sup>1096</sup> Analyser les engagements et les autres statuts normatifs comme des produits institués par des attitudes de reconnaissance et d'attribution, c'est invoquer l'idée de reconnaissance mutuelle. Reconnaître quelqu'un c'est le considérer comme sujet normatif d'engagements et de responsabilités. En effet, pour que nous soyons engagés, nous devons avoir *reconnu* [*acknowledged*] un engagement, et les autres doivent nous l'*attribuer*. Nous décidons quel pion nous jouerons mais ce n'est pas nous qui décidons ensuite à quoi nous nous sommes ainsi engagés. Quels autres gestes sont appropriés ou obligatoires pour qui a joué ce pion. Notre autorité est réelle, mais elle est partielle. Et, on peut dire la même chose des autres qui jouent le jeu avec nous et simultanément l'arbitrent. Car ils n'ont aucune autorité sur notre reconnaissance d'engagements. Leur autorité n'est opérante que dans l'administration de ces engagements, par laquelle ils nous maintiennent dans un engagement pourvu d'un contenu déterminé dont ils ne sont pas moins responsables que nous. Nous jouissons ainsi d'une certaine indépendance eu égard aux engagements que nous prenons mais nous sommes en même temps dépendant de l'attitude des autres qui nous attribuent et nous maintiennent dans cet engagement et par là en administrent le contenu.

<sup>1097</sup> Brandom, 2000b, p.255.

normes conceptuelles, pour déterminer ce à quoi nous sommes engagés en appliquant un concept, rien d'autre que d'autres applications du concept en question, ainsi que des applications de concepts qui lui sont inférentiellement reliés. Les applications du concept (et de ceux qui lui sont apparentés) qui ont effectivement été faites, ont donc déjà une certaine sorte d'autorité sur les applications futures possibles de ce concept (et de ceux qui lui sont apparentés). Les applications antérieures ont donc une autorité sur la signification ou le contenu du concept. Mais l'autorité a besoin d'être administrée. Les applications des normes instituées par les applications antérieures doivent être évaluées quant à leur correction, en fonction des normes auxquelles elles répondent. Pour que les applications courantes d'un concept soient responsables envers les applications antérieures de ce concept (et de ceux qui lui sont apparentés), elles doivent alors être *tenues* responsables, prises ou traitées *comme* responsables. Aussi, l'engagement que l'on assume [*undertakes*] en appliquant un concept dans un jugement ou dans une action ne peut être considéré comme pourvu d'un contenu déterminé, que s'il doit être administré par d'autres que celui dont c'est l'engagement. Ainsi, en reconnaissant un tel engagement, on reconnaît au moins implicitement l'autorité des autres sur le contenu auquel on s'est engagé. Ceux qui attribuent et évaluent l'engagement sont aussi obligés de reconnaître l'autorité des applications précédentes dans leur administration du contenu qu'instituent ces applications. Par conséquent, l'application correcte et objective d'un concept devra donc être définie comme une fonction des différences perspectiveles résultant de la relation synchronique « je-tu » et des différences perspectiveles résultant de la relation diachronique « je-tu » ; le potentiel structurel de la relation « je-tu » évalue ainsi les applications particulières des termes comme correctes ou incorrectes, comme appropriées ou non en s'étendant non seulement horizontalement à d'autres individus à un moment précis mais surtout verticalement à d'autres individus (en incluant soi-même) au fil du temps. Brandom considère l'application correcte et objective d'un terme comme étant le produit d'un processus de « négociation » entre les « autorités » concurrentes des attitudes d'applications passées, actuelles et futures et d'évaluation de ces applications par les agents linguistiques et les joueurs qui marquent le score. Les relations constitutives du contenu des normes conceptuelles ne sont donc pas seulement sociales et inférentielles mais bien inférentielles, sociales, et historiques.

Brandom utilise l'analogie du développement des concepts de la *common law* pour expliquer et mettre en évidence ce processus inférentiel, social et historique d'application et d'institution des normes conceptuelles. La *common law* diffère du droit écrit [*statute law*], en ce qu'il consiste entièrement en un droit jurisprudentiel [*case law*] et en une suite d'applications de concepts à des ensembles de faits réels. Dans ce processus, chaque juge hérite d'une tradition de cas judiciaires qui peuvent être considérés comme un ensemble de données particulières (les circonstances de l'affaire, etc.) auxquelles ont été appliqués (ou refusés) des universaux juridiques, tels que « délit », « strictement responsable », etc. Le juge est alors confronté à un nouveau cas particulier ( un ensemble de faits ) et doit décider ou refuser d'appliquer l'un de ces universaux – classer les actions en question comme constituant un délit spécifique, ou comme impliquant une présomption de stricte responsabilité. L'autorité de la tradition réside dans le fait que les seules raisons que le juge peut invoquer pour justifier sa décision concernent les précédents : le juge a l'autorité de faire le tri entre les divers rapports de ressemblance et de dissemblance entre les faits du procès actuel et les faits du procès précédemment jugés, de traiter certains d'entre eux comme plus importants que d'autres au égard à la question de savoir si le concept juridique en jeu doit être appliqué ou refusé aux faits qui sont devant lui. Cela fait ainsi en sorte que certains des procès antérieurs qui sont déjà classés comme correctement jugés sont plus pertinents que d'autres pour la décision en cours ; ce qui affecte aussi à son tour l'autorité des applications précédentes de concepts inférentiellement reliés. Ce sont donc les applications antérieures qui, selon le juge, délimitent le contenu du concept.

Mais, il pourrait aussi sembler que le passé n'ait pas de réelle autorité sur le présent dans la mesure où le juge actuel paraît avoir le dernier mot. Après tout, le juge qui a maintenant une décision à rendre peut ignorer ou du moins écarter des décisions antérieures gênantes, en les considérant comme des applications erronées des concepts en question – c'est-à-dire comme des affaires mal jugées – ou comme non pertinentes (et ne constituant pas un précédent) parce que différentes du procès en cours sous les rapports qu'il a décidés d'évaluer comme les plus importants. Ainsi, il semble que le juge actuel n'ait envers le passé que les dettes qu'il décide lui-même de reconnaître. Cela dit, s'il en était ainsi, l'autorité des décisions passées ne conférerait qu'un contenu vide et indéterminé aux concepts juridiques. Or, les applications passées de concepts (les décisions juridiques antérieures) exercent bien

une véritable contrainte et autorité sur les applications réelles actuelles précisément parce que le juge actuel n'a pas le dernier mot mais est tenu responsable devant la tradition dont il a hérité par les juges qui sont encore à venir. Aussi, sa décision n'importe pour le contenu du concept en question, que dans la mesure où la valeur de son application des précédents est reconnue [*acknowledged*] à son tour par les juges du futur. S'ils considèrent son procès comme ayant été mal jugé, à la lumière de leur lecture de la tradition dont a hérité le juge actuel, alors la décision de ce dernier n'a pas d'autorité du tout. Les concepts n'acquièrent ainsi un contenu déterminé que comme le résultat d'une tradition d'application ou d'une négociation perpétuelle entre les trois pôles d'autorités réciproques (ce que le contenu est pour les juges du passé, ce qu'il est pour celui d'aujourd'hui et ce qu'il est pour les juges du futur)<sup>1098</sup>. Les applications passées de concepts exercent une autorité sur les applications futures, car elles fournissent les précédents qui constituent les seuls raisons disponibles pour justifier les décisions ultérieures. Mais réciproquement aussi, les applications ultérieures de concepts par les juges qui héritent de cette tradition exercent une autorité sur les applications antérieures : la portée de l'autorité de la tradition, c.-à-d. le choix du contenu conceptuel que celle-ci est considérée avoir institué, est ainsi décidée par les juges qui rendent aujourd'hui des décisions. Brandom insiste ici sur cette idée d'une symétrie d'autorité en soulignant ce « moment [...] d'autorité constitutive de l'avenir sur le passé, et donc de dépendance du passé envers son avenir<sup>1099</sup>. ». Car ce n'est effectivement que dans la mesure où le juge actuel reconnaît l'autorité d'une décision antérieure que celle-ci revêt une quelconque autorité mais ce n'est qu'en étant reconnue par les juges du futur que celle-ci a une réelle autorité. Comme

---

<sup>1098</sup> Brandom remarque que le processus de négociation ne doit pas nécessairement mener à une résolution unique. Ce qu'un juge actuel estime être une application correcte et une compréhension adéquate des applications antérieures, un juge futur peut le remettre en question (pensons juste aux premiers jugements des tribunaux d'instance qui sont renversés par les cours d'appel et par la suite maintenus ou renversés à leur tour devant des tribunaux de grande instance). Ce processus produit certes un louable faillibilisme ou un potentiel pour une critique objective de nos applications des contenus déterminés des concepts quant à leur correction, mais il laisse aussi entendre que « la négociation des autorités réciproques » implique une indétermination du contenu du concept, ou au moins à une notion plus faible du caractère déterminé et objectif du contenu du concept. Mais, selon Brandom, cela ne signifie pas pour autant que le contenu des concepts soit indéterminé, dans la mesure où la résolution négociée du conflit n'est pas complètement *ad hoc* puisque la tradition d'application exerce une contrainte sur tout un ensemble de résolutions acceptables de telle sorte à ce qu'il n'y ait pas dans chaque cas une réponse unique qui puisse être établie à l'avance. Cf. Ibid., p.250, n.33 et p.256, n.38.

<sup>1099</sup> Ibid., p. 258.

Brandom le dit, « l'autorité du passé sur le présent est administrée en son nom par l'avenir<sup>1100</sup> », autrement dit puisque ce processus n'a en principe pas de terme final, pas d'autorité ultime qui ne dépendrait à son tour d'une reconnaissance (si nous considérons le futur présent comme lui-même un présent dans un temps encore à venir) « c'est une relation dans laquelle [...] le présent reconnaît l'autorité du passé, et exerce en retour une autorité sur lui, et leurs conflits sont administrés par le futur<sup>1101</sup>. ». Les normes conceptuelles déterminées ne sont donc intelligibles et objectives qu'au sein de cette tradition d'application et de cette pratique de reconnaissance où les autorités réciproques du passé, du présent et du futur négocient et administrent dans chacune des trois dimensions de la reconnaissance : sociale, inférentielle et historique.

La question soulevée par cette structure de reconnaissance de la tradition (l'accent mis sur la tradition a pour but de souligner l'importance de la dimension historique) est celle de savoir comment ces évaluations-ci résultant d'une tradition d'application des usages antérieurs et actuels d'un concept (adhésions ou critiques) sont objectives. Cette manière de répondre à la question ne crée-t-elle alors pas sa propre version du problème relativiste ? En effet, les juges futurs sont tenus de statuer sur la tradition d'application d'un concept quant à sa correction mais ils n'ont « pour évaluer la fidélité de sa décision (du juge actuel) au contenu » que « [...] la tradition dont il est l'héritier a effectivement conféré aux concepts juridiques<sup>1102</sup> ». Par conséquent, le juge qui détermine la correction d'une application donnée, en tant qu'ancrée dans la tradition, ne semble pas vouloir faire appel à un quelconque contexte transcendant dans la mesure où les concepts (sur lesquels il est censé porter un jugement) n'acquièrent un contenu déterminé qu'à la lumière des jugements antérieurs. Aussi, puisque le concept n'acquiert un contenu déterminé que comme le résultat de son rôle dans une tradition d'application, il semble que la seule critique que l'on puisse formuler est celle de savoir si l'application actuelle est en accord avec l'application passée et non si la tradition d'application des concepts est sujette, dans sa totalité, à être critiquée. Pourtant, cette dernière notion de critique est ce qu'une véritable conception objective de l'application des concepts quant à leur correction, requiert. Autrement dit, à partir de cette perspective de reconnaissance de la tradition, le juge futur peut seulement critiquer l'application actuelle

---

<sup>1100</sup> Ibid., p.259.

<sup>1101</sup> Ibid., p.260.

<sup>1102</sup> Ibid., p.259.

quant à savoir si celle-ci confirme ou infirme l'application traditionnelle mais il ne peut pas critiquer l'application traditionnelle considérée en elle-même comme correcte.

## 6.9. Conclusion

L'externalisme sémantique de Brandom et la dernière formulation de sa conception inférentielle, sociale et historique de l'application objective des concepts quant à leur correction sont deux façons pour tenter de trouver une juste-mesure entre l'*immanence* et la *transcendance* de l'objectivité et de la normativité dont parle Putnam<sup>1103</sup>. Une approche adéquate de l'objectivité des énoncés prétendant à la vérité et des normes de significations doit respecter le sens de cette pratique immanente dont les standards ont été établis par la critique des notions représentationnistes de la chose en soi (*Ding-an-sich*), de la théorie de la vérité-correspondance et de l'épistémologie du « point de vue de l'œil de Dieu ». Mais, il faut aussi respecter simultanément le sens émanant d'une pratique transcendante de ces mêmes standards, standards qui ne se réduisent pas, dans leur signification ou justification, à des évaluations communautaires (ou à la première personne) de la vérité, et qui ne devraient pas non plus être éliminés en faveur de ce genre-ci d'évaluations. L'équilibre requiert de maintenir le contextualisme – les standards ou normes ne peuvent survenir que dans un contexte de pratiques et de traditions d'application des concepts et d'évaluation de telles affirmations ou jugements résultant de l'application de ces concepts (engagements) – sans tomber dans le relativisme – et d'identifier la validité de ces mêmes standards avec les résultats obtenus dans les pratiques et traditions effectives. L'équilibre approprié entre l'immanence et la transcendance de l'objectivité, l'équilibre entre le contextualisme et le relativisme implique, dans une tournure « continentale », d'atteindre, *la transcendance dans l'immanence*<sup>1104</sup> c'est-à-dire, de générer le standard d'un contexte transcendant de l'intérieur

<sup>1103</sup> Putnam, 1985, p.8: « Reason is, in this sense, both immanent (not to be found outside the concrete language games and institutions) and transcendent (a regulative idea that we use to criticize the conduct of all activities and institutions). ».

<sup>1104</sup> La structure d'équilibre entre l'immanence et la transcendance est celle dont Putnam semble toujours ne pas être satisfait. Sa première phase du « réalisme métaphysique » exagère sans aucun doute la transcendance de la vérité, etc., alors que son « réalisme interne » corrige et souligne de manière excessive l'immanence de la vérité, et tend à la réduire (sémantiquement ou épistémologiquement) à une sorte d'acceptabilité rationnelle (idéalisée).

du contexte de nos pratiques sociales établies sur le mode du « je-tu ». C'est ce que la structure normative d'une faible socialité de l'objectivité de Brandom tente d'atteindre.

Cependant, si celle-ci ne vise pas à introduire un élément externe à la pratique, elle ne peut alors que générer un pluralisme de perspectives différentes, duquel ne découlent pas les notions faillibilistes d'erreur, de correction, d'amélioration, etc. La version inférentielle, sociale et historique de Brandom développe, de manière peut-être plus cohérente, la structure d'objectivité inhérente aux normes de signification, mais nous nous demandons si celle-ci peut vraiment produire une notion de correction plus forte que celle de « fidélité à la tradition ». Nous pensons donc que sa dernière approche des normes de significations et des propositions prétendant à la vérité eu égard à leur objectivité s'enlise dans le premier volet du dilemme. Nous aborderons alors, en conclusion, une autre approche contextualiste, sociale et historique de l'objectivité, celle proposée notamment par Habermas et McCarthy. Elle ajoute ainsi à la structure normative de l'objectivité de Brandom la notion kantienne d'*idées régulatrices*, et plus spécifiquement l'idée régulatrice d'un monde objectif. Par conséquent, au lieu d'invoquer l'idée d'un monde objectif indépendant de toutes pratiques linguistiques ou contextes d'usage qui nous donneraient accès au monde tel qu'il est comme l'externalisme sémantique le suggère, il nous paraît plus approprié de le concevoir comme une idée régulatrice au sens de Kant, c.-à-d., une notion qui, bien qu'inaccessible dans toute pratique réelle de discussion et contexte social de négociation, etc., règle néanmoins la manière dont de telles pratiques sont menées, en jouant le rôle d'un foyer imaginaire ou *focus imaginarius*<sup>1105</sup> vers lequel nos pratiques discursives sont orientées. Une telle idée fonctionnant de manière régulatrice est peut-être ce qui manque à l'approche de Brandom et ce qu'il tente de saisir avec la notion de responsabilité (*answerability*) : celle-ci fournit une notion faillibiliste d'objectivité mais en étant régulatrice elle implique certainement des engagements incompatibles avec le contextualisme.

---

<sup>1105</sup> Kant, 1974, §46, p.108-109.

## CHAPITRE VII

### HABERMAS-MCCARTHY : UNE SOCIALITÉ FAIBLE À LAQUELLE S'AJOUTE LA NOTION D'IDÉE RÉGULATRICE<sup>1106</sup>

#### 7.1. Introduction

Les conceptions sociales de l'objectivité des normes de signification et des énoncés prétendant à la vérité rencontrent deux problèmes fondamentaux, à savoir : le relativisme et une forme de circularité explicative. Une approche contextualiste entendue au sens d'une socialité faible nous paraît être un bon départ pour répondre au premier problème mais si nous voulons définitivement éviter totalement le relativisme, il semble que celle-ci doive laisser place à une approche faillibiliste et non seulement pluraliste de l'objectivité. Quant au second problème, les approches anti-réductionnistes ne peuvent y répondre que partiellement puisqu'elles ne conçoivent pas ce point comme un problème dans la mesure où elles considèrent la normativité comme une notion primitive. Par conséquent, le mieux que nous puissions attendre comme partisan d'une approche anti-réductionniste vise à clarifier cette dimension normative et à examiner la nature de ces pratiques *de facto* afin de fonder une notion objective de vérité et d'application correcte des normes de signification.

Nous tenterons donc, dans ces deux derniers chapitres, de modéliser une approche de l'objectivité des normes de signification en nous basant sur l'approche de l'objectivité des énoncés prétendant à la vérité proposée par Habermas et McCarthy. Nous définirons d'abord la notion d'idée régulatrice (§7.2) et examinerons la manière dont ces présuppositions idéalisantes, nécessaires à l'instauration de nos pratiques sociales, d'un monde objectif et d'une vérité transcontextuelle donnent sens aux assertions formulées dans les contextes communicationnels ordinaires, discursifs et critiques (§7.3). Nous montrerons ainsi comment la structure normative de l'objectivité et l'approche inférentielle, sociale et historique de la vérité et de l'application correcte des normes de signification de Brandom peuvent venir enrichir la théorie de l'agir communicationnel de Habermas (§7.4). Nous démontrerons finalement comment et pourquoi cette idée régulatrice d'une vérité et d'un monde objectif est

---

<sup>1106</sup> Friedman, 1997, 2002.

une condition nécessaire à l'élaboration d'une conception sociale objective et faillibiliste des normes de signification (§7.5).

## 7.2. La notion d'idée régulatrice : Kant versus Habermas-McCarthy

Habermas et McCarthy ne préconisent pas un retour au paradigme représentationnel de la connaissance et à la conception de la vérité comme correspondance entre nos pensées et un monde indépendant de nos interprétations et de nos pratiques. Ils sont contextualistes et s'accordent à dire avec Rorty que la vérité et la connaissance doivent être considérées dans un contexte historique et culturel mais précisent également qu'il est nécessaire d'introduire un élément objectif dans nos approches respectives. Notre approche tentera donc de minimiser et de concilier l'accent mis sur le contexte avec la nécessaire transcendance des approches contextualistes afin d'éviter l'écueil du relativisme. En raison de l'échec du « point de vue de Dieu » et de la théorie de la vérité correspondance, parler de la « manière dont le monde est », en tant que tel, ne peut pas être la source de l'objectivité. Cependant, le problème est que, si l'on ne peut pas faire appel au monde, qu'est-ce qui va alors déterminer la vérité et l'objectivité du contenu de nos énoncés et de nos normes de signification ? La solution proposée ici en lien avec les propositions prétendant à la vérité repose sur l'idée d'adopter et d'adapter « Kant's bifocal vision of the ideas of reason<sup>1107</sup> ». Selon Kant, les Idées de la raison, par exemple les idées de Dieu, de l'âme, d'un monde objectif donnant accès aux choses telles qu'elles sont, ont deux facettes. Elles sont d'une part, des conditions nécessaires de la pensée, et ne sont, d'autre part, ni tirées de l'expérience sensible ni directement applicables à cette expérience mais ont, selon Kant, un pouvoir de régulation sur notre activité discursive. Dans le vocabulaire de Kant, on dirait qu'elles n'ont pas un rôle constitutif c'est-à-dire, qu'on ne doit pas les considérer comme des descriptions de la réalité. S'il en était ainsi, elles conduiraient à des illusions métaphysiques que Kant qualifie d'antinomies de la raison pure. Or leur emploi en tant que tel est purement régulateur, puisque ce ne sont pas des objets à connaître mais des postulats qu'il faut admettre pour diriger et donner sens à notre vie pratique : elles régulent ou gouvernent nos pensées, celles-ci représentant des objets dans le monde de l'expérience. Pour ainsi dire, d'une part, la raison

---

<sup>1107</sup> McCarthy, 1991, p.2.

fixe le but de la science et, d'autre part, l'entendement en détermine la forme générale, tandis que l'expérience sensible fournit le contenu empirique de la connaissance. Par exemple, l'idée d'un monde déterminé indépendant de notre expérience et de nos ressources cognitives n'est pas une idée que l'on pourrait tirer du monde de l'expérience ou encore de l'usage de nos ressources conceptuelles mais si elle sert à guider, structurer et réguler nos pensées, elle permettra une appréhension du monde plus complète, plus cohérente et plus intelligible. Par conséquent, lorsque les idées de la raison se limitent à leur rôle légitime régulateur<sup>1108</sup>, elles sont source de connaissance et non d'illusion<sup>1109</sup>.

<sup>1108</sup> Les idées de la raison, contrairement aux principes de l'entendement, sont simplement régulatrices et non constitutives; elles ont donc un usage régulateur en ce sens qu'elles ne déterminent pas d'objet, mais qu'elles servent de règle à l'esprit. Kant, 1944, p.453, dans l'*Appendice à la dialectique transcendantale*, décrit cet usage régulateur des idées de la raison comme suit : « elles ont un usage régulateur excellent et indispensablement nécessaire : celui de diriger l'entendement vers un certain but qui fait converger les lignes de direction que suivent toutes ses règles en un point qui, pour n'être, il est vrai, qu'une idée (*focus imaginarius*), c'est-à-dire un point d'où les concepts de l'entendement ne partent pas réellement, - puisqu'il est entièrement placé hors des bornes de l'expérience, - sert cependant à leur procurer la plus grande unité avec la plus grande extension. ». Ces idées de la raison sont ainsi des concepts auxquels ne correspond aucun objet donné par les sens comme l'Idée de l'Âme, du Monde ou celle de Dieu; elles n'ont qu'une valeur heuristique, comme Kant le souligne : « Cet usage n'est que régulateur, c.-à-d. qu'il sert à mettre, autant qu'il est possible, de l'unité dans les connaissances particulières et à rapprocher ainsi la règle de l'universalité. » (Ibid., p. 455). Elles posent alors une nécessité méthodologique qui nous impose d'essayer de nous rapprocher d'un idéal; nous devons faire « comme si » nous parvenions à une compréhension meilleure, l'idéal étant une fin heuristique, alors que nous ne savons pas si nous pouvons compter sur la réalisation effective, factuelle de l'idée régulatrice. L'idéal est simplement une mesure qui oriente, règle et motive en fournissant l'archétype (*Urbild*) qui est la seule chose dont nous disposons pour diriger nos actions et pensées. Kant, 1944, *Dialectique transcendantale*, Livre II, Chap. 3, p. 414, l'exprime ainsi : « Ces idéaux, quand bien même on ne peut leur attribuer de réalité objective (existence), ne doivent pas pour autant être considérés comme de simples chimères (*Hirngespinnste*), mais donnent à la raison une mesure (*Richtmass*) indispensable ; celle-ci a besoin du concept de ce qui est absolument parfait (*vollständig*) en son espèce pour pouvoir apprécier et mesurer suivant cela le degré et les défauts de ce qui est imparfait. ».

<sup>1109</sup> Kant montre que le problème est de les restreindre à ce domaine, car la raison a une tendance abusive à quitter le terrain de l'expérience et à les considérer comme des représentations du monde de l'expérience plutôt que comme des idées qui régulent les idées qui sont des représentations du monde de l'expérience. La raison voudrait atteindre la sphère de la pensée pure ou « point de vue de Dieu » soit un domaine inconnaissable par nos formes *a priori* de la sensibilité et nos catégories de l'entendement (soit les conditions de possibilité de la connaissance) se croyant à tort indépendante du domaine phénoménal et empirique. Selon Kant, la raison a cette maladie ou prétention illégitime de tenter de vouloir connaître le monde tel qu'il est indépendamment des conditions nécessaires de possibilité de toute connaissance Cf. Kant, 1944, *Théorie transcendantale de la méthode*, Chap. I, Sect. 3, p.507 (« La discipline de la raison pure par rapport à son usage polémique. ») et p.519 (« Le premier pas de la raison pure est dogmatique, le second sceptique, le troisième critique. »).

Habermas et McCarthy conçoivent comme des corrélats modernes des idées de la raison de Kant l'idée de sujets rationnellement responsables, l'idée d'un monde objectif connu en commun et l'idée d'une vérité objective. Ces deux dernières idées se traduisent de manière intersubjective en termes d'un « accord universel rationnellement motivé », c'est-à-dire un accord dont la validité est reconnue par tous. Afin de respecter le contextualisme, McCarthy considère ces idées comme « social-practical analogues of Kant's ideas of reason<sup>1110</sup> ». Par conséquent, ces idées ne sont pas envisagées comme nécessaires à la pensée de sujets transcendants qui seraient isolés et détachés de toute contrainte sociale, culturelle et historique tel que Kant le pensait ; mais, ce sont, en suivant Habermas, des présuppositions idéalisantes nécessaires à l'instauration de nos pratiques sociales réelles – par exemple, l'activité d'une compréhension mutuelle dans nos pratiques communicationnelles quotidiennes ainsi que dans les discours théorique et critique de la recherche scientifique et de la délibération morale – afin qu'elles acquièrent leur sens dans la pratique. Habermas reconstruit les idées de la raison de Kant à la fois comme des présuppositions pragmatiques de nos pratiques sociales concrètes, et comme des idéalizations : elles ne sont actualisées dans aucune pratique donnée pour lesquelles elles sont des présuppositions. Elles sont vraiment effectives mais comme les idéalizations sont contrefactuelles, elles n'ont pas besoin d'être actualisées en permanence dans les pratiques qu'elles contribuent à *réguler* et à *constituer* en partie<sup>1111</sup>.

Il faut donc s'entendre sur ce qu'on entend par « idées constitutives ». Deux sens doivent ainsi être distingués : d'une part, celui de Kant et d'autre part, celui d'Habermas et de McCarthy. Les Idées de la raison sont, dans le sens de Kant, régulatrices, et non constitutives c.-à-d. qu'elles régulent nos pensées et représentations du monde sans être elles-mêmes des représentations du monde issues de l'expérience. Habermas et McCarthy soutiennent cette fonction régulatrice des idées de la raison et rejettent ce caractère constitutif dans le sens de Kant. Cependant, ils avancent que les idées de la raison dans leur fonction régulatrice sont des présuppositions pragmatiquement nécessaires de nos pratiques prétendant à la vérité c.-à-d., des observations empiriques ordinaires et des énoncés plus théoriques issus de la recherche scientifique sur la nature des choses. Par conséquent, ils affirment bien que

<sup>1110</sup> Hoy et McCarthy, 1994, p.38.

<sup>1111</sup> Le caractère constitutif des idées régulatrices de nos pratiques prétendant à la vérité sera illustré par l'exemple du langage ordinaire sur le monde et par celui du discours scientifique.

ces pratiques-ci n'auraient pas le sens qu'elles ont sans avoir été guidées par de telles idées et que ces idées sont constitutives en vertu du sens que ces jeux de langage-ci ont dans la pratique. Ce qualificatif de constitutif a donc pour Habermas et McCarthy un sens pragmatique – ces idées sont constitutives du sens pratique émanant de nos pratiques linguistiques et sociales. Autrement dit, ces idées-ci ou normes du monde et de la vérité objective ne sont fondées sur aucune pratique transcendante comme le Monde ou un royaume platonique de formes, mais se hissent toutes seules à partir de la pratique elle-même au rang de conditions nécessaires donnant sens à la pratique. La manière dont certaines pratiques sont réellement conduites révèle et maintient le fonctionnement régulateur de ces idées-ci en les situant elles-mêmes dans les pratiques comme conditions de possibilité de nos pratiques sociales. Dans la mesure où ces idées sont à la fois, régulatrices et constitutives (dans le sens de Habermas et de McCarthy et non dans celui de Kant) de la signification pragmatique de certaines activités, elles sont immanentes et transcendantes, réelles et idéales (contrefactuelles), limitées au contexte et le transcendant en même temps<sup>1112</sup>. Les idées de la raison ouvrent ainsi nos pratiques réelles contextualisées et nos jugements à l'évaluation et à la révision objective, ce qui leur confère une dimension authentiquement objective (faillibiliste). Habermas illustre ce point en parlant d'une tension, « désormais située dans le monde vécu des acteurs communicationnels eux-mêmes, entre, d'une part, le caractère inconditionné des prétentions à la validité transcendantes, et faisant éclater les limites de chaque contexte, et de l'autre la factualité des prises de position par oui ou par non,

---

<sup>1112</sup> Ibid., p.73: « [T]he idea of an objective world knowable in common functions not only « constitutive » – that is, as an assumption that actually informs our practices – but also, « regulatively » – that is a presupposition that normatively constrains and guides our practices...It is just this constitutive-regulative duality that Habermas wants to capture in his conception of ideas of reason as unavoidable *suppositions* of communicative interaction ». Également à la page 38: « Habermas attempts to identify and reconstruct social-practical analogues of Kant's ideas of reason. They appear there as *pragmatic presuppositions* of communication...this has the effect of relocating the Kantian opposition between the real and the ideal within the domain of social practice. Cooperation interaction is seen to be structured around ideas of reason which are neither fully constitutive in the Platonic sense not merely regulative in the Kantian sense. As *idealizing suppositions* we cannot avoid making while engaged in the processes of mutual understanding, they are *actually effective* in organizing communication and at the same time *counterfactual* in ways that point beyond the limits of actual situations. As a result social-practical ideas of reason are both « immanent » in and « transcendent » to practices constitutive of forms of life ».

dépendantes du contexte et significatives pour l'action, qui engendrent des faits sociaux localisés<sup>1113</sup>».

### 7.3. Conditions de possibilité de nos pratiques communicationnelles

Cette tension se retrouve d'ailleurs dans la notion de rationalité communicationnelle de Habermas soit dans son approche d'une prétention à la validité visant tout aussi bien le langage ordinaire quotidien que le discours scientifique et moral<sup>1114</sup> plus théorique. Pour Habermas, les contextes communicationnels ayant pour but l'accord et l'entente mutuel, qu'il soit question de contextes conversationnels quotidiens ou de discours plus théoriques et critiques de la recherche scientifique ou de la délibération morale, doivent être compris comme impliquant (implicitement ou explicitement) des énoncés prétendant à la validité. Aussi, à travers la communication linguistique c.-à-d., à travers l'acte de discours, le locuteur cherche à s'entendre à propos de quelque chose avec son auditeur. Cette finalité illocutoire comporte deux niveaux :

- d'abord l'auditeur doit comprendre l'acte de discours c.-à-d. comprendre la signification des termes, comprendre de quoi il s'agit.
- puis, autant que possible, il doit l'accepter.

Il y a donc compréhension et acceptation. La rationalité inhérente à la pratique communicationnelle renvoie ainsi à la pratique de l'argumentation qui recouvre un type de discours où les parties prenantes thématisent des prétentions à la validité qui font l'objet de litiges<sup>1115</sup>. D'où la formulation de Habermas : l'agir communicationnel se noue autour de prétentions à la validité critiquables<sup>1116</sup> émises par les locuteurs. Celui qui parle prétend que ce qu'il dit est valable. Dans quelle mesure ces actes de discours, ces prétentions, peuvent être, soit contestés, soit réfutés, soit finalement acceptés par son auditeur et en retour

<sup>1113</sup> Habermas, 1999, p. 86.

<sup>1114</sup> Cf. Habermas, 1979, 1987b, p. 273-337, 1992b.

<sup>1115</sup> Habermas, 1987b, p. 34 : « un *argument* contient des raisons qui sont systématiquement reliées à la prétention à la validité d'expressions problématiques. La « force » d'un argument se mesure, dans un contexte donné, au bien-fondé des raisons; ce bien-fondé se montre entre autres, dans la capacité d'une expression à convaincre les participants d'une discussion, c.-à-d. à motiver l'admission d'une prétention à la validité. Sur cet arrière-plan, nous pouvons aussi juger de la rationalité d'un sujet capable de parler et d'agir d'après la façon dont il se comporte dans un cas donné en tant que partie prenante de l'argumentation. ».

<sup>1116</sup> Ibid., p.34.

défendus, affaiblis ou abandonnés par celui-ci : tel est précisément l'enjeu de la discussion, de l'agir. Pour que les prétentions à la validité soient acceptées par le ou les auditeurs, il faut qu'elles soient honorées par des justifications ou des raisons discursives, c'est-à-dire argumentées. En clair : nous, en tant que locuteur, nous affirmons que ce que nous avançons est valable, et nous argumentons pour le prouver et le soutenir (face à des contre-arguments, des contestations de la part des autres participants, d'autres points de vue). C'est ainsi par l'argumentation, et non par la force ou par la contrainte ou l'influence, que nous obtiendrons une entente (solide) sur les énoncés émis. Cette entente aura pouvoir d'engagement pratique, pour coordonner les actions. Un accord solidement établi est alors source de solidarité dans l'action. Si, au sein de la discussion, il y a exercice d'une force, c'est simplement la force du meilleur argument qui finit par s'imposer au cours de la discussion.

La rationalité est donc comprise comme une disposition propre à des sujets capables de parler et d'agir, se traduisant dans des modes de comportement pour lesquels de bonnes raisons peuvent à chaque fois être exhibées. Les expressions rationnelles sont accessibles à une appréciation objective et toute mise à l'épreuve explicite de prétentions controversées à la validité requiert la forme ambitieuse d'une communication remplissant les présuppositions de l'argumentation. Celle-ci peut in fine être conçue comme une « poursuite de moyens réflexifs de l'activité orientée vers l'intercompréhension<sup>1117</sup> » dans la mesure où ceux qui prennent part à l'argumentation sont obligés de présupposer généralement que la structure de leur communication exclut toute contrainte hormis celle de l'argument le meilleur et écarte également tous les motifs hormis celui de la recherche coopérative de la vérité. L'argumentation, en tant que procédure, renvoie à une forme d'interaction, réglée spécifiquement, entre proposant et opposants. Le procès discursif de l'entente est normé de telle sorte que les participants thématisent une prétention à la validité rendue problématique et, en tant qu'ils sont *ex hypothesi* déchargés des pressions de l'action, contrôlent uniquement avec des raisons si la prétention défendue par le proposant est juste ou non. La notion de rationalité communicationnelle concerne donc bien cet exercice intersubjectif d'une contestation et d'une justification rationnelle. La communication est en soi une activité rationnelle et la rationalité est un phénomène communicationnel issu de nos pratiques

---

<sup>1117</sup> Ibid., p. 41.

sociales : elle s'incarne dans nos pratiques communicationnelles entre locuteurs et auditeurs au sein d'un contexte historique et social précis.

Par conséquent, comme pour Brandom, parler un langage (dans les contextes communicationnels) est bien pour Habermas un jeu social de production et de demande de raisons. Brandom s'accorde ainsi avec Habermas sur l'idée que, pour accomplir un acte de communication, un agent doit fournir des raisons appuyant ses prétentions à la validité. Par conséquent, la notion de « prétention à » permet à Habermas, tout comme à Brandom, de mettre en lumière les conditions nécessairement sous-jacentes à nos pratiques communicationnelles et à cet effet, de formuler un argument transcendantal<sup>1118</sup>. L'argument transcendantal ainsi formulé ne dit rien d'autre que ceci : soit un jeu de langage défini, par exemple, celui de la philosophie ; ce jeu de langage comprend « des prétentions à », sans lesquelles il ne serait pas ce jeu de langage mais un autre. Il ne s'agit donc pas de classer les jeux de langage à partir de leur contenu propositionnel mais bien de penser les discours à partir de la notion de « prétention à ». Habermas développe, dans ce sens, une *pragmatique universelle* qui définit la communication comme une action orientée vers une compréhension mutuelle, c'est-à-dire :

« The speaker must have the intention of communicating a true proposition or propositional content [...] so that the hearer can share the knowledge of the speaker<sup>1119</sup>. ».

Le locuteur énoncé alors une prétention à la validité<sup>1120</sup>. Autrement dit, il s'engage à fournir des raisons justifiant l'acceptabilité de son affirmation, à savoir :

« We understand a speech act when we know the kinds of reason that a speaker could provide in order to convince a hearer that he is entitled in the given circumstances to claim validity for his utterance –in short, when we know what makes it acceptable. A speaker, with a validity claim, appeals to a reservoir of potential reasons that he could produce in support of the claim<sup>1121</sup>. ».

<sup>1118</sup> L'argument transcendantal est un argument qui montre que les opérations quotidiennes des êtres humains ne seraient pas possibles sans un certain nombre de conditions que le chercheur doit déterminer. Ce sont ces conditions qui justifient le recours au terme kantien de « transcendantal », terme conçu de manière très générique comme recherche des conditions nécessaires de possibilité d'une opération cognitive.

<sup>1119</sup> Habermas, 1998, p. 22.

<sup>1120</sup> Pour plus de détails sur cette notion, voir Heath, 1998.

<sup>1121</sup> Habermas, 1998, p. 232-233.

L'approche de Habermas est donc proche de celle de Dummett<sup>1122</sup> ou de Brandom, sans y être complètement équivalente; elle en est quelque sorte une généralisation en ce qu'elle intègre celle-ci au sein d'une théorie plus large de l'action sociale. Habermas articule alors le concept d'activité communicationnelle à la théorie des actes de discours<sup>1123</sup>. Les actions sociales peuvent être distinguées en fonction de l'attitude adoptée par les participants selon que cette attitude est orientée vers le succès ou l'intercompréhension. C'est grâce à une analyse des actions langagières que l'on est capable de délimiter la frontière entre des actions orientées vers le succès et les actions orientées vers l'intercompréhension. Habermas a recours pour ce faire à la distinction posée par Austin entre actes locutoires par lesquels le locuteur énonce des contenus objectifs, actes illocutoires par lesquels le locuteur accomplit une action en disant quelque chose et les actes perlocutoires qui visent un effet chez l'auditeur. La lecture de la théorie austinienne vise ainsi à montrer en quoi l'usage du langage orientée vers les conséquences n'est pas un usage originaire mais la subsomption, sous les conditions de l'agir orienté vers le succès, d'actions langagières qui servent des objectifs illocutoires. Les attitudes orientées vers l'intercompréhension peuvent être élucidées seulement au regard des actes illocutoires. Les perlocutions peuvent, elles, être conçues comme une classe spéciale d'interactions stratégiques, les illocutions étant, dans cette perspective, mobilisées en tant que moyens dans des contextes d'actions téléologiques. Habermas compte dans l'agir communicationnel les interactions médiatisées par le langage où tous les participants poursuivent par leurs actions langagières des objectifs exclusivement illocutoires et il considère comme un agir stratégique médiatisé par le langage les interactions dans lesquelles l'un des participants au moins veut susciter par des actions langagières des effets perlocutoires chez un vis-à-vis<sup>1124</sup>. Par la référence au lien illocutoire, Habermas démarque les actions communicationnelles de toutes les autres actions sociales. Seules sont constitutives pour l'agir communicationnel les actions langagières auxquelles le locuteur relie des prétentions à la validité critiquables.

<sup>1122</sup> Dummett, 1973, p. 298: « There is a general convention whereby the utterance of a sentence, except in special contexts, is understood as being carried out the intention of uttering a true sentence. ».

<sup>1123</sup> Austin, 1991.

<sup>1124</sup> Si Habermas s'inspire d'Austin, il s'en écarte en distinguant l'action langagière et le contexte d'interaction qu'elle constitue en réalisant la coordination des actions ; cette distinction est occultée par Austin dans la mesure où ce dernier est resté trop fixé sur le cas paradigmatique d'actions langagières liées à des institutions.

Tout comme Brandom, les outils d'analyse de Habermas sont les actes de discours et la prétention à la validité de ce langage est conçue comme une activité pratique : l'unité de base de signification des actes de discours est l'énonciation, celle-ci traduisant l'action concrète d'un locuteur, et non la proposition atemporelle abstraite. On tient là la base du déploiement de la démarche associée à la *pragmatique universelle* qui tente de reconstruire les conditions de possibilité universelles de l'intercompréhension : tout agent communicationnel qui accomplit un acte de discours quelconque est forcé d'exprimer des prétentions universelles à la validité. Il ne peut alors éviter d'émettre les prétentions suivantes : s'exprimer de façon intelligible; donner quelque chose à entendre; se faire comprendre; et s'entendre l'un l'autre. Le but de l'intercompréhension est de parvenir à un accord conduisant à une communauté réciproque, de savoir partagé, de confiance réciproque et de convergence de vues. À ces quatre dimensions de la communauté intersubjective correspondent quatre prétentions à la validité, l'intelligibilité, la vérité, la sincérité et la justesse. Habermas distingue surtout trois genres de prétentions associées à l'énonciation linguistique, la prétention à la vérité, la prétention à la validité ou à la justesse ou encore la prétention à la sincérité, dont dépendent ce sur quoi nous cherchons à nous mettre d'accord. C'est cette démarche qu'adopte Habermas au début de la *Théorie de l'agir communicationnel* lorsqu'il se réfère à la théorie des trois mondes de Popper<sup>1125</sup>. Selon Habermas, il existe ni plus ni moins trois domaines d'objets auxquels se rapportent les propositions : le monde objectif, le monde social et le monde subjectif. Le monde objectif correspond à l'ensemble des faits que l'on appréhende au moyen des cinq sens. Le monde social correspond aux normes ou aux attentes intersubjectivement partagées au sein d'une communauté donnée. Enfin, le monde subjectif inclut tous les phénomènes psychiques auxquels on accède au moyen de l'introspection. Il existerait donc autant de types de prédications assertives. Habermas établit alors pour chacun

<sup>1125</sup> Habermas s'inspire de la théorie des trois mondes de Popper en la modifiant légèrement ; cf. à ce sujet Habermas, 1987b, p. 68, p. 92-100, ainsi que 1987a, p. 408-409. La théorie des trois mondes est exposée par Popper, 1998, p. 245-293. Dans la suite de la *Théorie de l'agir communicationnel* et dans d'autres ouvrages, Habermas ne suit pas cette méthode. Habermas cherche en effet à établir les prétentions *universelles* à la validité, c'est-à-dire les prétentions à la validité qui valent pour toute proposition indépendamment de leur contenu sémantique. Habermas ne peut donc se contenter d'une simple analyse sémantique. La méthode de Habermas repose alors sur le principe selon lequel les prétentions universelles à la validité correspondent aux différents types de points de vue à partir desquels on peut critiquer ou nier une même proposition.

de ces mondes la prétention à la validité correspondante. Il distingue ainsi du point de vue sémantique trois types de prétentions à la vérité<sup>1126</sup>, à savoir :

- *Les prétentions à la vérité* : celles-ci sont relatives aux actes de discours constatifs soit aux faits et événements dont nous affirmons une explication ou une interprétation « vraie », en nous référant au monde objectif. S'il s'est produit une panne, j'affirme qu'il s'agit de telle panne : énoncer un diagnostic, c'est honorer une prétention à la validité sous forme de vérité à propos d'un événement. C'est encore plus vrai, bien entendu, d'un énoncé scientifique, qui prétend dire le vrai sur tel ou tel phénomène (physique ou social). Pour Habermas, les discussions et débats scientifiques illustrent, du mieux possible, ce type de prétention. C'est ainsi que progresse le savoir.

- *Les prétentions à la justesse (des normes) et des commandements* (actes de discours prescriptifs) qui méritent reconnaissance dans un monde social intersubjectivement partagé. Pour Habermas, ces normes sont morales ou socio-morales dans la mesure où elles sont élaborées collectivement, dans un contexte social de recherche de solidarité, et ont un contenu moral. Comme tout contenu moral, ces normes définissent le bien et le mal, ce qu'il est bon de faire, ce qu'il est mal de faire, et s'imposent par leur valeur d'obligation. Il faut bien distinguer la procédure (un débat démocratique, argumenté, un véritable agir communicationnel) et son résultat soit des obligations, des impératifs moraux (donc des commandements, au sens moral du terme). Sur ce plan, Habermas est ouvertement kantien.

- *Les prétentions à la sincérité* des actes de discours (expressifs), qui manifestent (expriment) des expériences vécues subjectives auxquelles le locuteur a un accès privilégié. Elles portent donc sur le monde subjectif prétendant soit à l'authenticité – j'éprouve bien ce que je dis – soit à la véracité – je dis bien ce que j'éprouve. La validité consiste ici à faire accepter comme sincère ce que le locuteur affirme de son expérience propre (qui lui appartient en propre subjectivement et qui reste opaque pour autrui).

Par conséquent, dans cette triple relation ou trois cas de figure, il y a recherche de compréhension et d'adhésion par argumentation. Comprendre un acte de discours, c'est déterminer ce qui le rend acceptable, les conditions d'acceptabilité étant identifiées aux

---

<sup>1126</sup> Voir Habermas, 2001, p.53 : « Nous rencontrons alors, *ni plus ni moins*, trois sortes de prétentions à la validité [...]. » ; 1993, p.149 : « Nous rencontrons, *ni plus ni moins*, trois prétentions à la validité [...]. » Voir aussi 1987b, p.54, où il est question d'un « *système* » des prétentions à la validité.

conditions de succès illocutoire. C'est donc bien dans le discours et, de ce fait dans l'argumentation que se manifeste la rationalité propre inhérente, ou du moins présumée, à l'action communicationnelle. Dans le langage ordinaire, la prétention à la validité soulevée par chacun de ces trois types d'actes de discours n'apparaît pas explicitement dans les propositions mais elle serait en quelque sorte implicitement émise par les locuteurs : elle est offerte par le locuteur et acceptée par l'auditeur sans aucune justification supplémentaire. Les raisons ne sont normalement pas exigées par l'auditeur et ne sont pas fournies par le locuteur. Cependant, ce n'est pas toujours le cas, un auditeur peut à n'importe quel moment défier son interlocuteur et lui demander de justifier la prétention à la validité de l'énonciation émise. L'agir communicationnel n'est alors mis en jeu, que parce qu'il y a un problème dans le discours, donc nécessité de formuler un accord explicite et de défendre ou de rejeter la prétention à la validité d'une énonciation. Ainsi, pour qu'il y ait communication et pour que deux personnes échangent et se comprennent il faut un minimum d'accord, ne serait-ce que sur la signification des mots, leurs implications pratiques et sur le fait que l'interlocuteur est, implicitement, habilité à contester ce qui est dit. L'agir communicationnel ne se déploie donc vraiment que lorsqu'il y a un élément de mise en cause (de mise en interrogation) du monde vécu et la recherche de formulation d'un nouvel accord. Le monde vécu est en effet un réservoir d'évidences que nous partageons implicitement, dès lors que nous faisons partie de la même culture, de la même société, de la même structure sociale. Ce monde vécu se particularise partiellement au niveau des groupes sociaux : chaque groupe possède ses référents, ses valeurs, sa culture, etc. C'est lorsque ces évidences du monde vécu ne suffisent plus ou cessent d'être des évidences partagées, que la nécessité d'une activité communicationnelle apparaît pour recréer de la solidarité. En ce sens, il y a toujours un facteur de mise en crise du monde vécu : un événement nouveau, un problème non-résolu, un défaut dans le savoir social, une crise dans l'adhésion commune aux mêmes valeurs... Bref, lorsque la création d'une entente de type nouveau se fait sentir comme nécessité. Les évidences du monde vécu restent mobilisées : celles qui ne posent pas problème font partie de l'héritage commun, que les participants à la discussion mobiliseront, sans même avoir à y penser. Ce seront, par contre, les éléments de crise ou de carence qui seront mis en avant, thématiques, pour trouver ou retrouver rationnellement un accord. S'engager dans une activité communicationnelle linguistique est donc bien pour Habermas une activité rationnelle vécue

intersubjectivement. De plus, ce processus argumentatif, pour être rationnel, doit satisfaire certaines conditions ou règles procédurales et demeure ainsi fondamentalement « démocratique » dans sa forme c'est-à-dire que tous les participants disposent d'une égale opportunité de faire valoir leurs points de vue et de remettre en question ou de critiquer les opinions des autres sans que la force, la contrainte, la tradition ou l'influence n'entrent en ligne de compte dans la recherche d'une entente. Cette perspective pragmatique permet ainsi de dégager des prétentions qui restaient invisibles si l'on se contentait du point de vue sémantique. Ainsi, un accord rationnel ne peut être réalisé que dans des conditions de liberté et d'égalité de parole<sup>1127</sup>. Si tout énoncé considéré du point de vue de la validité présuppose un locuteur idéal qui prétend obtenir un accord rationnel, et si un accord rationnel présuppose de telles conditions de liberté et d'égalité, alors tout énoncé considéré du point de vue de la validité présupposerait un locuteur qui prétend que de telles conditions sont ou doivent être réalisées. Toute prétention à la validité s'accompagnerait donc d'une prétention à une situation idéale, c'est-à-dire à des conditions sociales idéales d'égalité et de liberté<sup>1128</sup>. La notion de situation idéale de parole constitue un présupposé immanent aux actes de discours courants, elle en définit l'idéal régulateur et établit bien l'un des points d'articulation entre démocratie et agir communicationnel. Le concept d'argument est donc, d'une certaine manière, intrinsèquement pragmatique : ce qui compte comme bonne raison se manifeste seulement dans le rôle qu'elle a dans le jeu de l'argumentation, c'est-à-dire, dans la contribution qu'elle offre, et en accord avec les règles du jeu, pour décider de l'acceptation ou de la non-acceptation d'une prétention à la validité. En cela, nous pouvons trouver le caractère procédural de la rationalité, et cela même parce que les raisons sont un système d'échange discursif<sup>1129</sup>.

<sup>1127</sup> Cf. Habermas, 1987a, p.323-324, 1999, p.110-111, où Habermas expose en détail les conditions de possibilité d'une discussion rationnelle, 1995, p.111 sq., 2001, p.191, p.300 et 1987a, p.406-407, qui résume l'ensemble des conditions, à la fois langagières et non langagières, d'un consensus rationnel.

<sup>1128</sup> C'est dans la notion de *communauté de communication idéale* que semble résider l'une des clés de l'articulation entre le concept d'agir communicationnel et le concept de démocratie en ce sens que la démocratie pourrait apparaître comme une préfiguration politique concrète d'une situation idéale de parole.

<sup>1129</sup> Habermas, 1997. Dans ce sens, on trouve des points de comparaison avec Wittgenstein, 1976. En fait, on peut lire chez lui au §105: « Toute vérification de ce qu'on admet comme vrai, toute confirmation ou information prennent déjà place à l'intérieur d'un système. Et assurément, ce système n'est pas un point de départ plus ou moins arbitraire ou douteux pour tous nos arguments; au contraire,

De plus, comme nous ne pouvons pas trouver de critères indépendants possibles, nous ne pouvons que juger selon la perspective du participant au discours. Il n'existe donc pas de critères indépendants à la perspective d'un participant pour apprécier les raisons :

[...] procédures et raisons sont toutefois tellement imbriquées les unes des autres qu'il ne peut pas y avoir d'évidence ni de critère d'appréciation qui précéderaient l'argumentation, donc qui à leur tour ne seraient pas fondés dans des argumentations ni ne devraient être validés, sous des présuppositions d'argumentation, par un assentiment discursivement acquis et rationnellement motivée [...] dans les questions substantielles, il n'y a pas d'évidence 'dernière' ni d'argument 'dirimant', nous devons recourir à l'argumentation comprise comme procédure, afin d'expliquer, à partir de la pragmatique qui le sous-tend, pourquoi nous pouvons, d'une manière générale, nous sentir capable d'élever des prétentions à la validité transcendantale, et de les honorer<sup>1130</sup>.

Ceci manifeste le caractère faillible et la limitation structurelle de la communication. Même les critères qui, éventuellement pourraient être établis, devraient être compris par le discours. De cette façon, même la distinction entre le consensus vrai ou faux doit être décidée, en cas de doute, par le discours.

La thèse de Habermas d'une conception théorique de la prétention à la validité d'une action linguistique requiert l'application de certaines suppositions idéalisantes, c.-à-d. que ces idéalizations-ci doivent être présupposées par le locuteur et l'auditeur afin d'établir tout échange discursif. McCarthy identifie également ces idéalizations comme suit : l'idée d'un monde objectif, l'idée d'une notion transcontextuelle de la vérité et l'idée que locuteurs et auditeurs sont des agents rationnellement responsables de leurs actes<sup>1131</sup>. Dans ce qui suit,

---

il appartient à l'essence de ce que nous appelons argument – le système n'est pas tant le point de départ des arguments que leur milieu vital. ».

<sup>1130</sup> Habermas, 1992a, p. 148-149.

<sup>1131</sup> Selon Habermas, les suppositions d'un monde objectif et d'une notion transcontextuelle de la vérité sont seulement nécessaires à l'élaboration du sens de certains types d'actes de discours comme les assertions scientifiques et les commandements moraux. (Bien que nous nous concentrons exclusivement sur les premiers, Habermas s'intéresse tout autant au raisonnement moral dans la mesure où une des motivations centrales de son approche de la rationalité communicationnelle est de fonder une éthique discursive, c.-à-d. de défendre le point de vue selon lequel les questions éthiques peuvent être déterminées discursivement). Les prétentions à la validité de ces actes de discours requièrent que ces idéalizations formellement présupposées soient considérées comme des *idées régulatrices*. Les actes de discours expressifs ne font par contre, pas appel à ces idéalizations-ci ; formuler avec succès un énoncé prétendant à la sincérité revient à dire que ce que nous éprouvons doit bien exprimer ou refléter ce que nous disons ressentir. De tels énoncés ne décrivent pas le monde mais expriment l'état d'esprit et les expériences subjectives d'un sujet ; ils ne formulent pas des jugements de réalité sur le monde et ne présupposent donc pas des idéalizations sur le monde et sur la vérité. Ces dernières idéalizations sont nécessaires seulement pour le sens des actions linguistiques qui prétendent

nous nous limiterons aux actes de discours assertifs et à la prétention à la validité qui leur est rattachés, soit la proposition prétendant à la vérité, et nous illustrerons la manière dont les présuppositions idéalisantes d'un monde objectif et d'une notion de vérité transcontextuelle sont censées donner un sens aux assertions formulées dans des contextes communicationnels ordinaires, discursifs et critiques.

Aussi, les divergences dans les propos et observations des interlocuteurs soulignent bien l'idée sous-jacente, émanant de ces échanges communicationnels du langage ordinaire, d'un point de vue unitaire, cohérent et objectif sur le monde présupposé par ces échanges communicationnels du langage ordinaire. Supposons ainsi la situation conversationnelle suivante : je vois au loin une tour ronde mais mon interlocuteur la voit carrée. Je vais tenter de justifier son erreur de diverses manières, comme par exemple en disant qu'il n'a pas une bonne vue de loin, qu'il n'utilise pas les termes « rond » et « tour » de la même manière que moi ou encore qu'il essaie de me jouer un tour, etc. Mais une explication logiquement possible de notre divergence est que la tour est à la fois ronde et carrée et que les objets dans le monde peuvent revêtir simultanément différentes formes. Autrement dit, je ne dis pas que nous avons tous les deux raison : en effet, pour résoudre nos divergences d'observation, il suffit de reconnaître que l'une des deux est fautive (pour telle ou telle raison), plutôt que d'avancer qu'il existe une « erreur » ou une incohérence dans le monde lui-même. Par conséquent, notre comportement de résolution de nos divergences dévoile et maintient la présupposition d'un monde objectif et cohérent formellement et intersubjectivement présupposé par les interlocuteurs<sup>1132</sup> (mais au sujet duquel on peut parfois émettre des jugements de vérité erronés). Tout comme le monde est considéré comme une unité cohérente et objective, les observations correctes désignant des faits ou des événements à propos de ce monde-ci sont censés se rejoindre au sein d'une approche univoque soit celle de la vérité sur le monde. Aussi, dans la mesure où l'on abandonne le mythe d'un point de vue de Dieu et la théorie correspondantiste de la vérité, l'objectivité du monde dépend intrinsèquement du rapport intersubjectif que nous entretenons les uns les autres avec le monde tel qu'il est.

---

à la validité sous forme de vérité à propos d'un monde objectif de faits et d'un monde social de normes (conditions d'application correcte des normes)

<sup>1132</sup> Brandom adhère certainement à la formulation de cette condition de possibilité de la pratique linguistique et sociale dans la mesure où son externalisme sémantique soutient qu'il n'existe pas de concept indépendant du monde, ni de perspective indépendante de tous déploiements de concepts à partir de laquelle nous aurions un accès au monde.

Par conséquent, si nos pratiques supposent le point de vue unitaire d'une réalité indépendante, elles présupposeront de la même manière que si nous rendons tous compte de la même réalité alors de tels comptes-rendu concorderont ensemble et viendront renforcer l'idée d'un point de vue cohérent et unitaire sur le monde. Énoncer une proposition prétendant à une vérité par rapport aux faits du monde objectif vise donc à attendre ou à anticiper l'accord d'autres propositions. Par conséquent, la relation fonctionnelle qui nous maintient dans un monde qui répond normalement à nos anticipations et qui semble confirmer nos interprétations suscite une confiance en la stabilité de ce monde. Mais, une fois encore, cet accord ou consensus rationnel peut ne pas se produire puisque même sous des conditions procédurales proches de la perfection, rien ne permet en effet d'exclure qu'une entente puisse ultérieurement être infirmée par la secousse d'un argument meilleur, mais imprévisible dans l'état actuel de la discussion. Cela dit, un consensus rationnel erroné n'élimine pas l'idéalisation d'une prétention à la vérité en termes d'une acceptation rationnelle universelle. Cependant, il n'empêche que nous considérons toujours cette activité comme un processus d'apprentissage intersubjectif incluant certes essais et erreurs mais maintenant néanmoins des conditions idéales de justification et conduisant progressivement à un accord universel rationnel. Ce qu'est un fait, nous ne pouvons le déterminer qu'en recourant à des énoncés auxquels nous assignons une valeur de vérité, et la vérité de ces énoncés, nous ne pouvons en définitive la fonder qu'à partir d'autres énoncés ou interprétations, à la lumière de nos expériences. Nous n'affirmons pas ainsi de deux saisies contradictoires qu'elles sont toutes deux vraies mais que l'une est vraie et que l'autre est fausse ou encore que les deux sont fausses. L'idée d'un point de vue unitaire et objectif et de l'unanimité d'une vérité sur le monde sont donc des idéaux qui régulent normativement nos interactions communicatives quotidiennes.

De manière similaire, le discours théorique de la recherche scientifique présuppose ces deux idéalizations au sein d'une relation intersubjective. Dans le discours scientifique, un scientifique formule une assertion sur le monde et affirme une proposition nouvelle ayant prétention à la vérité. Le monde est ainsi considéré comme étant une totalité déterminée qui a toujours été indépendante des croyances entretenues à son égard. Aussi, alors que nos justifications rationnelles peuvent être infirmées, rejetées ou corrigées, nous comprenons la vérité comme une propriété inaliénable de nos énoncés qui, par définition, ne saurait être

invalidée, parce qu'inséparable de la référence à un monde objectif en soi indépendant de nos descriptions. Ainsi, formuler une proposition prétendant à la vérité revient à formuler une proposition dont la validité s'étend au-delà du contexte actuel de l'ici et du maintenant. Habermas l'exprime ainsi sous une autre forme au chapitre « L'unité de la raison au sein de la pluralité de ses voix » dans *La pensée postmétaphysique* :

[...] la validité revendiquée pour des propositions et pour des normes transcende les espaces et les temps, mais cette prétention est chaque fois élevée ici et maintenant, dans des contextes déterminés [...] <sup>1133</sup>

Le même problème est soulevé dans *Vérité et justification* :

[...] comment concilier, d'un côté, la normativité, incontournable du point de vue des participants [...] et, de l'autre, la contingence du développement historico-naturel que connaissent les formes de vie socioculturelle <sup>1134</sup>. Plus généralement : On ne voit toujours pas *ce qui nous autorise* à tenir pour vrai un énoncé supposé idéalement justifié dans les limites d'un esprit fini <sup>1135</sup>.

Dans le même ouvrage, Habermas décrit et résout en partie ce problème en s'appuyant sur la *Phénoménologie de l'esprit* de Hegel <sup>1136</sup>. D'après Hegel, le sujet perçoit le monde, raisonne et agit non seulement à partir des données de l'expérience sensible, mais aussi à travers les concepts issus de sa raison. Or, ces concepts ne sont pas déjà donnés dans la conscience, mais acquis dans l'espace social et le temps historique. Ce qui signifie que les concepts dont se sert le sujet pour organiser son expérience sensible, raisonner et agir de façon responsable ne sont pas absolument fondés, mais au contraire faillibles <sup>1137</sup>.

<sup>1133</sup> Habermas, 1993, p.179.

<sup>1134</sup> Habermas, 2001, p.264.

<sup>1135</sup> Ibid., p.193. On trouve d'autres formulations de ce problème chez Habermas, cf. par exemple *ibid.*, p.153 : « Même s'il obéit à l'idée de validité absolue, l'esprit fini, prisonnier de son présent et de son passé, reste provincial au regard du savoir futur, qui sera mieux instruit. » ; voir aussi Habermas, 1993, p.58 : « Ainsi, il est impossible d'échapper aux conséquences qui, sans cesse, contestent à une raison située sa prétention à l'universalité. » ; *ibid.*, « Une raison située » p.47 sq.; p.59-60, p.155, p.170 et surtout p.174-175. Voir enfin Habermas, 1988, p. 382.

<sup>1136</sup> Habermas, 2001, p.125-163.

<sup>1137</sup> Habermas s'appuie ici sur Hegel, 1941, t.1, Chap. IV, p.174 : « C'est pourquoi elle en fait elle-même l'aveu, elle se confesse d'être une conscience tout à fait *contingente, singulière*, – une conscience qui est *empirique*, qui se dirige d'après ce qui, pour elle, n'a aucune réalité, obéit à ce qui, pour elle, n'est aucunement essence, fait, et élève à la réalité effective, ce qui pour elle n'a aucune vérité. ». Habermas, 2001, p. 292, évoque à nouveau le caractère pour ainsi dire « terrestre » – et non purement « transcendantal », indépendant de l'expérience sensible et la conditionnant – de l'entendement : « Les structures qui rendent transcendantale possible les expériences de contact avec quelque chose se trouvant dans le monde – pour nous – objectif, et qui permettent également

Aussi, c'est parce que nous n'avons pas un accès direct au monde auquel se réfèrent nos affirmations prétendant à la validité que le scientifique ne peut pas faire appel au monde directement. De plus, ce sont ces prétentions à la vérité qui prêtent à discussion, mais qui en même temps, peuvent permettre de faire avancer la connaissance. Il faudra alors passer l'épreuve des justifications discursives soit des contre-argumentations. La mise à l'épreuve pratique ne suffit pas. Elle permet de falsifier un énoncé, mais non de produire une vérité. Pour qu'une vérité soit (provisoirement) admise au sein d'une communauté et pour arriver à une acceptation rationnelle intersubjective, il faut qu'elle soit reprise, critiquée et défendue dans une argumentation convaincante, vis-à-vis d'autres propositions. Habermas avance la solution d'une discussion rationnelle entre le plus grand nombre de points de vue, la discussion permettant selon Habermas de décloisonner les points de vue particuliers et d'accéder à un point de vue plus impartial sur le modèle des forums scientifiques. Il décrit ainsi le discours à l'image de ce jeu de langage décrit tant comme un « forum de justification<sup>1138</sup> » que comme un « processus de construction<sup>1139</sup> » : c'est l'épreuve, soit des vérités déjà établies, qu'il faut soumettre à la critique et « déloger », soit des autres points de vue, qui prétendront, eux-aussi, dire le vrai à propos des mêmes faits ou événements. Les locuteurs n'ont plus d'autre choix en effet que de chercher à s'assurer, *par la seule force des bonnes raisons*, de la validité de leurs positions respectives, puisqu'ils ne disposent d'aucun accès direct à des conditions de vérité qui seraient soustraites au langage et qui pourraient échapper à la médiation interprétative. Mais pour ce faire, avance Habermas, ces locuteurs doivent en même temps admettre des *conditions idéales de justification et d'argumentation*, qui seules leur permettent de transcender les contextes immédiats de l'action, puisque ce qui peut être rationnellement considéré comme vrai dans l'espace réflexif du discours doit pouvoir valoir non seulement dans un contexte déterminé, mais dans tous les contextes, en tout temps et devant tout interlocuteur futur. Cela dit, il est bien sûr impossible de justifier une affirmation à la totalité des auditeurs rationnels en tout temps et devant tout interlocuteur

---

qu'on en parle, peuvent être dès lors comprises comme les résultats d'un processus de formation qui témoigne d'une portée cognitive. » Habermas se réfère également aux travaux de Piaget. Voir par exemple Piaget, 1970, p.5 : « [...] la connaissance ne saurait être conçue comme prédéterminée ni dans les structures internes du sujet, puisqu'elles résultent d'une construction effective et continue, ni dans les caractères préexistants de l'objet, puisqu'ils ne sont connus que grâce à la médiation nécessaire de ces structures et que celles-ci les enrichissent en les encadrant [...] ».

<sup>1138</sup> Habermas, 2001, p. 102.

<sup>1139</sup> Ibid., p.44, p. 47.

futur. Par conséquent, ici l'arbitre, c'est comme l'affirme Habermas, « ils [les arguments qui] ont la force de convaincre ceux qui prennent part à une discussion de la prétention à la validité, autrement dit de *motiver rationnellement* la reconnaissance des prétentions à la validité<sup>1140</sup>» ou encore c'est « le bon argument », la démonstration la plus convaincante, celle qui l'emportera. Il en résulte une nouvelle vérité, un élargissement ou un déplacement du savoir, une remise en cause des savoirs préexistants (et donc des évidences du monde vécu). Mais cette prétention à la validité d'un savoir est toujours précaire et provisoire dans la mesure où elle pourra ultérieurement être remise en question. En ce sens, l'idée d'une acceptation rationnelle est toujours hypothétique, toujours falsifiable et n'est alors jamais complètement actualisée dans la pratique (c'est une idée régulatrice). Elle est surtout basée sur cette contrainte de décloisonnement (*the decentering*) des perspectives des participants au discours ou sur ce travail de distanciation qu'exige la pratique de ce jeu de langage. Ce décloisonnement de la part des locuteurs indiquent alors que ce processus d'une construction discursive d'un monde commun est, en quelque sorte, un processus d'apprentissage<sup>1141</sup>.

#### 7.4. Habermas et Brandom : idée régulatrice d'un accord rationnellement motivé et structure sociale sur le mode du je-tu.

Or, si l'on se tourne du côté de l'approche dialogique de Brandom, on note qu'au contraire, l'identité de la signification, l'interprétation herméneutique de la signification des actes de discours, et le suivi de la compréhension ne laissent pas place à une telle contrainte de décloisonnement ou d'apprentissage. Autrement dit, il nous semble que chez Brandom, c'est moins les compétences individuelles des agents que le scénario ou le jeu de pointage déontique qu'ils suivent, qui importe. Habermas met l'accent sur le caractère central de ce processus d'apprentissage au moyen des interprétations générées par une communauté de communication, mais ni la dynamique, ni le profil décousu du processus communicatif n'est clairement expliqué. Il aurait peut-être pu tenir compte, comme le fait Brandom dans son approche sociale et historique de la normativité, de la dimension temporelle et séquentielle du déroulement du processus discursif, celui-ci débutant d'abord à partir des contributions

<sup>1140</sup> Habermas, 1978, p. 148-149.

<sup>1141</sup> Habermas, 2001, p. 105, p. 161.

individuelles, impliquant ensuite une pratique conjointe coopérative mais également compétitive et conflictuelle mais se dirigeant, par coups successifs et au moyen d'un pointage déontique, vers une évaluation des différentes perspectives jusqu'à aboutir à la coordination de nos perspectives sociales selon des pratiques communes.

Par conséquent, nous aurions ainsi tendance à donner moins d'importance à cette thèse d'une vérité consensuelle dans la mesure où elle semble forcer à tout prix l'accord intersubjectif sur le mode du « nous ». Pourtant, tout agent s'engageant dans une activité communicationnelle devrait se faire comprendre dans des termes raisonnables sans que cela implique une obligation d'ordre impérative<sup>1142</sup>. Habermas aurait en quelque sorte confondu l'obligation relative aux prétentions à la vérité et la nature impérative des obligations d'ordre morale. Nous adhérons à l'idéal régulateur d'une situation idéale de parole dans laquelle toute communication présuppose la *possibilité* d'un accord entre les participants à la discussion mais seulement si celle-ci est bien fondée sur la seule valeur des arguments échangés. De ce point de vue, nous estimons nécessaire de redonner toute sa place aux processus liés aux négociations et aux compromis dans ce processus dialogique.

Nous suggérons peut-être à Habermas d'adopter la solution de Brandom. Celle-ci peut se découper en trois étapes<sup>1143</sup> :

- Brandom substitue à l'approche traditionnelle de la communication conçue comme transmission de quelque chose (signification) et comme le fait de partager une relation à une seule et même chose (de saisir une signification commune), un modèle de communication qui reflète sa conception inférentialiste de la pratique discursive en termes de pointage déontique, c'est-à-dire un modèle dans lequel les marqueurs coopèrent au sein d'une activité conjointe et coordonnent des perspectives sociales en marquant les points déontiques selon les pratiques communes. Ce holisme inférentiel (qui exige que la portée pragmatique de l'adhésion doxastique à un contenu propositionnel soit rapportée à un répertoire d'engagements concomitants) doit ainsi être compris dans le contexte d'un holisme social. Ce holisme social exige que la saisie du contenu sémantique dépende de la capacité des marqueurs à exploiter

---

<sup>1142</sup> Swindal, 2007, p.118: « So while Brandom sees communication as a self-sufficient process of informing, Habermas views it as guided by the imperative of social integration – rational coordination of action plans. Certain beliefs of others are to be not merely acknowledged, but also believed such that stable moral obligations can be built from them for an action sequence. ».

<sup>1143</sup> Nous suivons ici la suggestion de Scharp, 2003, p. 55-57.

les relations entre les différentes perspectives constituées par les différents engagements contractés et attribués à ceux dont ils enregistrent les points déontiques. Chaque marqueur doit donc pouvoir tenir deux registres de pointage afin de suivre le cours et de recenser les points de chaque engagement et autorisation reconnu par chacun des marqueurs. La communication s'ordonne ainsi entre deux marqueurs, *A* et *B*, quand un énoncé *p*, contracté par *A*, devient disponible à *B* comme prémisse dans le système inférentiel de *B*. *B* comprend alors l'assertion de *A* si *B* attribue le bon engagement à *A*. Autrement dit, le type de compréhension d'une affirmation requise pour la communication réussie consiste en ceci que les auditeurs représentent correctement la performance en question dans leurs décomptes : qu'ils attribuent le bon engagement à l'auteur de cette assertion. Ce qui fait de l'engagement un engagement discursif, c'est son articulation inférentielle. Par conséquent, une condition nécessaire pour que *B* comprenne l'assertion de *A* est ainsi que *B* connaisse les conséquences que cette assertion de *A* aurait s'il y adhérerait lui-même. Il faut donc que la compréhension par l'auditoire d'une affirmation donnée détermine la portée inférentielle qui serait attachée au fait d'adopter ou de croire en cette affirmation – qu'elle détermine ce à quoi on s'engagerait si on y adhérerait, quels autres engagements sont incompatibles avec elle (et excluent toute autorisation à une telle adhésion), etc. En d'autres termes, *B* doit être capable d'évaluer la portée inférentielle de l'énonciation de *A* eu égard à sa paire ordonnée d'engagements. La communication exige que *B* soit capable de sélectionner une phrase qui (si énoncée par *A*) ait la même portée inférentielle (ou une portée inférentielle comparable) eu égard à son répertoire d'engagements que l'énonciation de *A* aurait par rapport à son propre répertoire d'engagements. Le marqueur doit alors substituer l'affirmation de *A* à la sienne afin de déterminer le bon engagement dans son attribution. Il est donc nécessaire qu'à chaque fois qu'un marqueur au score attribue un engagement à un autre joueur ou contracte un engagement en énonçant une assertion, que le contenu de l'engagement soit spécifié dans la mesure où la portée inférentielle et de ce fait la compréhension et la signification d'une affirmation dépendent des engagements d'arrière-plan et des hypothèses auxiliaires dont on dispose comme prémisses collatérales lorsque l'on en tire les conséquences inférentielles.

- Brandom propose ensuite de rendre compte de la manière dont les marqueurs coordonnent leurs répertoires d'engagements. En interagissant avec un étranger, un marqueur attribue un noyau central d'engagements inférentiels (et peut-être doxastiques) qui sont

supposés être approuvés par tous les marqueurs censés saisir le contenu, le concept ou la signification et un ensemble de croyances périphériques, qui pourraient différer sans que le contenu en soit affecté. Saisir le concept impliquerait alors uniquement la maîtrise de ces inférentielles essentielles, et celles-ci se résumeraient à ce que les interlocuteurs doivent partager sous peine de malentendus réciproques. Or, dans la mesure où ce dispositif vise à contribuer à l'analyse des langues naturelles, il est nécessaire de spécifier quelles caractéristiques de la pratique discursive donnent lieu à une telle différence de statut entre engagements conceptuels et engagement pratiques. Il sera donc nécessaire d'expliquer en termes de pointage les différents rôles que jouent les attitudes pratiques consistant à considérer des engagements comme des engagements conceptuels et empiriques. Les engagements seront alors ajoutés ou retranchés de ce pointage initial au moyen de l'usage de chaînes anaphoriques<sup>1144</sup> reliant le répertoire d'engagements d'un marqueur au livre d'engagements qui est tenu sur lui par un autre marqueur.

- Brandom adhère enfin à une version de l'externalisme qui lui permet de modéliser le contenu inférentiel exprimé par une instanciation de phrase comme une fonction assignant une portée inférentielle donnée à chaque répertoire d'engagements concomitants. On pourrait se représenter les portées inférentielles comme des paires ordonnées de circonstances et de conséquences d'application. Le premier élément pourrait consister en ensembles d'affirmations antécédentes inférentiellement suffisantes (celles dont l'affirmation en question peut être inférée), et le second en un ensemble d'affirmations conséquentes inférentiellement nécessaires (celles qui peuvent être inférées de l'affirmation en question). Étant donné que ce qui vaut comme preuve pour une affirmation ou nous y engage, ainsi que ce pour quoi elle-même vaut comme preuve ou ce à quoi elle nous engage, dépend des engagements d'arrière-plan dont on dispose comme hypothèses auxiliaires, alors les contenus

---

<sup>1144</sup> Voir Brandom, 2011, p. 564, qui s'explique sur son utilisation de la notion de chaîne anaphorique au sens de Chastain. Sur la question des chaînes anaphoriques, voir le travail pionnier de Chastain, 1975. Voici par exemple comment celui-ci introduit ses « chaînes anaphoriques » : « Les grammairiens parlent parfois d'*anaphore*, terme par lequel ils désignent le genre de rapport qui vaut entre, par exemple un pronom et son antécédent. Appelons « *chaîne anaphorique* » une suite de termes singuliers apparaissant dans un contexte tel que si l'un d'eux réfère à quelque chose, alors tous les autres réfèrent à cette chose. » (Chastain, 1975, p. 204-205; cité et passage traduit par Corblin, 1985, p. 175). Chastain défend ainsi avec vigueur la thèse selon laquelle les indéfinis ont parfois valeur de pronoms référentiels et permettent d'initier des chaînes anaphoriques (à la différence des noms propres, qui en sont incapables).

inférentiels pourraient être considérés comme des fonctions. Le contenu de chaque affirmation serait représenté par une fonction qui prendrait pour argument des ensembles d'engagements d'arrière-plan concomitants et qui produirait comme valeurs les portées inférentielles. Les interlocuteurs se comprennent les uns les autres à partir du moment où ils saisissent le contenu de leurs énonciations; autrement dit, quand ils déterminent la fonction qui est associée à l'emploi de leurs expressions. Quant à savoir en quoi consiste, pour une expression, le fait d'être employée de façon à se voir associée telle signification plutôt que telle autre, demeure matière à débat. Par conséquent, la version de l'externalisme de Brandom vise à expliquer le statut d'une communauté comme une pratique discursive en termes d'un interprète attribuant ce statut. Le statut d'une communauté comme pratique discursive est défini par un ensemble déterminé de contenus conférés aux expressions linguistiques par ses membres. Les normes qui régissent le choix d'interprétation de l'interprète dictent la manière dont les contenus de la pratique discursive dépassent les conceptions que ses membres peuvent en avoir. Selon l'interprète, le contenu d'une expression employée dans la pratique discursive (ce à quoi un interlocuteur est engagé) excède les dispositions des membres à l'appliquer. Mais, même si les dispositions des membres ne déterminent pas un contenu particulier, elles leur permettent bien de saisir des contenus particuliers. Ce qui est partagé par le locuteur et les auditeurs ne sera donc pas un contenu compris comme fonction, mais une pratique de pointage. Les contenus compris comme fonctions depuis des répertoires vers des portées inférentielles peuvent être considérés comme implicites dans les pratiques en question, mais il reste que la pratique peut conserver son identité bien que des fonctions différentes y soient implicites (à différents moments, et depuis des points de vue doxastiques différents). Depuis chaque point de vue doxastique sur un acte de discours, il peut en effet y avoir un contenu commun à celui qui contracte un engagement et aux marqueurs qui l'attribuent, et cependant il se peut que ce que l'on considère comme partagé soit différent selon les différents points de vue des marqueurs. Il est donc nécessaire d'adhérer à la conception de la pratique sociale sur le mode du « je-tu », telle que Brandom la décrit, afin de pouvoir distinguer entre ce qui est considéré comme correct et ce qui est correct. Aucune perspective n'est alors privilégiée, à la différence de la pratique sociale sur le mode du « je-nous », telle que décrite plus haut par Habermas, qui semble bien préconiser une perspective, celle du « nous » ou l'obligation (au sens fort d'un impératif) d'un accord

communautaire. Nous pensons donc que l'idée (régulatrice) d'un accord intersubjectif rationnellement motivé n'est pas incompatible avec cette idée que les contenus inférentiels sont essentiellement perspectivistes. Comme le souligne Brandom :

Ce qui est partagé, c'est une capacité à naviguer et à traverser les différences de points de vue, à spécifier des contenus depuis différents points de vue [...] correspond à différents aspects d'une activité interprétative unique consistant à comprendre, à saisir une communication – à différents aspects de cette saisie cognitive de la communication qu'est le pointage déontique<sup>1145</sup>.

Aussi, c'est parce que l'idée du monde objectif un et toujours le même est à la fois fondée sur l'accord intersubjectif que la pratique de la science est un processus continu toujours ouvert à confronter et à justifier ses croyances avec celles d'autrui. L'importance constitutive de ces idées-ci de la raison vis-à-vis du sens pratique de l'activité scientifique peut être illustrée en opposant les discours de la recherche scientifique et les discours subjectifs portant sur un jugement de goût personnel. Un jugement de préférence tel que « les fraises ont une saveur agréable » ne correspond à aucun fait objectif sur le monde et l'on ne peut pas non plus déterminer la vérité objective quant au goût des fraises en argumentant au sein d'un débat. La discussion s'épuise dans l'impasse du « Je pense / Tu penses » : les jugements de préférence, en faisant appel à des goûts subjectifs, ne mènent à aucune vérité objective ou discussion rationnelle. Par conséquent, le discours subjectif, à la différence du discours scientifique, est porteur de sens sans qu'il soit nécessaire de recourir aux présuppositions idéalisantes d'un monde et d'une vérité objective : les désaccords en matière de goût ne maintiennent pas ces idées régulatrices de la raison. À l'opposé, le discours scientifique requiert l'idée d'un point de vue unique et objectif sur le monde et la vérité.

#### 7.5. Vers une conception transcendantale et sociale de l'objectivité

La conception de Habermas et de McCarthy d'un monde et d'une vérité objective conçue comme idées régulatrices et comme conditions de possibilité de nos pratiques communicationnelles et théoriques déploie un certain nombre de caractéristiques lui permettant de remplir les conditions d'adéquation propres à une conception transcendantale et sociale de l'objectivité. Nous en dégagerons quatre, à savoir :

---

<sup>1145</sup> Brandom, 2011, p. 865-866.

1. En fondant la vérité objective sur nos jugements réels et concrets, leur conception ne fonde pas l'objectivité sur un quelconque ordre ontologique antérieur et transcendant. La notion d'objectivité ne saurait s'identifier à une réalité transcendante.

2. En reliant les idées d'un monde et d'une vérité objective au processus de l'argumentation rationnelle entre les participants au sein d'une pratique discursive, la notion d'objectivité implique, comme nous venons de le voir, la réelle participation des autres locuteurs : il s'agit donc bien d'une objectivité sociale et non publique. Et, dans la mesure où ces idées sont intimement liées aux pratiques linguistiques et sociales, la notion d'objectivité est contextualiste et non représentationaliste.

3. Les idées d'un monde et d'une vérité objective ne prennent ainsi tout leur sens pratique qu'en s'incarnant dans nos pratiques contextuelles de l'échange langagier. Elles sont certes conceptuellement liées à l'accord rationnel de tous les participants aux pratiques communicationnels, mais Habermas et McCarthy veillent à ne pas réduire la vérité objective, dans ce travail discursif, au critère d'acceptabilité des justifications rationnelles (l'acceptabilité rationnelle universelle ou l'acceptabilité rationnelle sous des conditions idéales) ou à ne pas l'éliminer en faveur d'une acceptabilité rationnelle<sup>1146</sup>. Leur intérêt ne se focalise pas sur la nature de la vérité ; *a fortiori*, ils ne préconisent pas une conception épistémologique de la vérité, mais s'intéressent à l'épistémologie des propositions prétendant à la vérité : il s'agit d'honorer les prétentions à la validité que renferment les énoncés par la critique, la contestation ou la justification<sup>1147</sup>. Ils notent le lien interne entre la justification des affirmations prétendant à la vérité et l'acceptabilité rationnelle dans nos pratiques langagières. Ce lien interne entre vérité et accord rationnel ne se ramène donc pas à une définition épistémologique de la vérité mais est bien l'expression d'un contextualisme épistémologique : critiquer ou justifier des jugements prétendant à la vérité doit être médiatisé par l'évaluation rationnelle et intersubjective des croyances des autres locuteurs. Cependant, dans la mesure où ils s'intéressent à l'épistémologie des affirmations prétendant à la vérité, le

<sup>1146</sup> Dans sa phase initiale, la théorie consensualiste de Habermas de la vérité dans « *Théories relatives à la vérité* » (Habermas 1987a, p. 322 et sq) peut être lue comme réductionniste, mais ses dernières remarques sont plus précises. Voir à ce sujet sur Hoy et McCarthy, 1994, p.73-74, p.239, 1989, p. 321-322.

<sup>1147</sup> Par conséquent, ils ne confondent pas la vérité avec la justification, même la justification idéale, au risque de commettre un sophisme naturaliste.

danger est que leur approche réduise la justification des prétentions à la vérité à l'accord rationnel d'une communauté de locuteurs<sup>1148</sup>. Habermas et McCarthy cherchent certes à éviter ce type de réductionnisme. Aussi, comme nous l'avons dit, puisque l'acceptation rationnelle intersubjective des prétentions à la vérité est toujours un processus ouvert et provisoire, la justification de ces affirmations-ci ne se réduit jamais à un simple accord rationnel. De plus, si l'entente se réalise en suscitant l'accord des autres, ce n'est pas leur accord en tant que tel qui constitue l'entente, mais plutôt ce qui les a amenés à se mettre d'accord : l'appui de preuves, d'arguments crédibles et la formulation d'inférences valides. Par conséquent, une proposition n'est pas rationnellement justifiée parce que tout le monde est (jusqu'à présent) d'accord mais tout le monde est d'accord parce que la proposition est rationnellement justifiée :

« It is not the reported agreement that warrants the claim, but the warrant for the claim that grounds the agreement<sup>1149</sup>. ».

Ils insistent sur l'importance du rôle constitutif des notions de monde et de vérité objective pour nos pratiques communicationnelles. Ainsi, bien que l'argumentation demeure pour nous le seul moyen de nous enquérir de ce qui peut être tenu pour vrai une fois suspendues les évidences de la pratique, la vérité ne perd pas pour autant le sens objectif et inconditionnel qu'elle revêtait dans l'agir quotidien. Par conséquent, leur conception sociale de l'objectivité est envisagée dans un sens faible.

4. De 1 à 3, la conception sociale de l'objectivité de Habermas-McCarthy s'apparente de près à la structure de l'objectivité de Brandom et à son approche inférentielle, sociale et historique de la vérité et de l'application correcte des normes. Comme nous l'avons vu, le point de vue de Habermas-McCarthy est procédural en ce que l'objectivité est garantie à partir du moment où les conditions nécessaires (telles l'ouverture du discours à une publicité généralisée, l'accès égal de tous à la discussion, etc.) d'un dialogue illimité et sans contrainte

<sup>1148</sup> Nous mentionnons ce point précédemment. Scharp, 2003, p. 57, souligne également cet écueil: « Habermas's theory of communicative action is an I-we account because it defines truth and rightness as what is taken to be correct by the members of a community. For Habermas, there is no distinction between what this privileged perspective takes to be correct and what actually is correct. ». Nous pensons que l'approche de Habermas peut éviter cet écueil en envisageant ce processus de justification communicationnel en termes de pratique de pointage déontique.

<sup>1149</sup> Hoy et McCarthy, 1994, p.42, p. 238-239.

instaurent une situation idéale de parole réalisant pleinement le déploiement de différents points de vue. Cette approche procédurale est formelle ou structurelle au sens où la vérité objective ne dépend pas du contenu d'un point de vue particulier, mais à l'image de la socialité du « je-tu », elle émerge une fois qu'il existe une garantie structurelle d'une multitude de points de vue ou repères d'évaluation. De plus, cette approche est *sociale* en ce qu'elle implique, comme Brandom le préconise, une confrontation et une évaluation réelle des croyances des différents locuteurs et auditeurs (affirmations prétendant à la validité ou engagements) ; elle est également *historique* en ce que cette confrontation et évaluation doit pouvoir valoir non seulement dans un contexte déterminé, mais dans tous les contextes, en tout temps et devant tout interlocuteur futur c.-à-d., qu'elle entend convaincre un groupe d'interlocuteurs, l'un après l'autre (plus ou moins de la même manière que les applications d'un terme par un locuteur actuel ou antérieur et ceux d'un joueur, en tant que marqueur, sont ratifiées par les marqueurs du futur), et elle est *inférentielle* en ce que le processus d'argumentation est rationnellement articulé : les locuteurs cherchent à s'assurer par la seule force des bonnes raisons ou arguments de la validité de leurs positions respectives. Cependant, en ajoutant le statut régulateur de l'idée d'une vérité objective, un principe supplémentaire est introduit dans leur approche d'un processus continuellement ouvert d'évaluation rationnelle et d'accord (provisoire) des affirmations prétendant à la vérité. Celle-ci permet ainsi de fonder l'idée de l'inconditionnalité de la vérité et de souligner l'idée d'une argumentation toujours en progrès (mouvement) plutôt que celle plus réductrice d'une prolifération de perspectives dans laquelle aucune perspective n'est correcte ou erronée, meilleure ou moins bonne que l'autre. Ils soutiennent finalement que cette idée régulatrice d'une vérité objective n'est pas seulement un anathème jeté sur le pluralisme faillibiliste ou critique (requis pour un point de vue véritablement objectif) mais en est une condition nécessaire.

5. Par conséquent, dans la mesure où l'idée d'une vérité objective oriente le processus épistémique des propositions prétendant à la vérité, leur contextualisme épistémologique semble pertinent et échappe au risque d'un retour au platonisme<sup>1150</sup> ou au réductionnisme

---

<sup>1150</sup> Nous accordons ici au terme « platonisme » ou « platoniste » la signification qui, ordinairement, est donnée aussi au terme « fondationalisme ». Le terme « platonisme » lui-même est fourvoyant (aussi, le mettrai-je en italique) parce qu'il n'a guère affaire avec Platon.

naturaliste. Le concept *pragmatiste* de la vérité n'est donc pas annulé dans la discussion, tant s'en faut ; au contraire, avertit Habermas, lui seul est en mesure d'assurer qu'un contact pourra éventuellement être rétabli entre les sphères de la pratique et du discours. Cette idée régulatrice d'une vérité objective ne vient pas de nulle part et Habermas insistera tout particulièrement sur l'enchâssement des discussions dans le monde vécu, et sur la résurgence, en celles-ci, du point de repère que représente la référence à une vérité objective. Dans la modalité réflexive du discours, il n'en va donc pas tant de réduire la vérité à l'autorité, forcément relative, de nos raisons, mais plus simplement de produire des raisons convaincantes, en principe universellement acceptables, qui pourront à nouveau venir éclairer l'action quotidienne, en restaurant la continuité signifiante. Par là, le concept pragmatiste de la vérité remplit bel et bien une fonction correctrice et vient se substituer au concept épistémique : l'enjeu déterminant auquel il reste suspendu, dans l'immédiateté de l'action quotidienne comme dans la distance réflexive de l'argumentation, est en fin de compte celui d'une pratique qui ne doit pas s'effondrer<sup>1151</sup>.

Par conséquent, l'élément régulateur de la conception sociale faible de Habermas-McCarthy de l'objectivité est ce qui la distingue d'autres approches sociales faibles (comme celles de Brandom et de Wittgenstein-McDowell) et ce qui lui permet donc de satisfaire les conditions d'adéquation soutenues par un tel point de vue c'est-à-dire, qu'il fonde une notion de vérité non réductionniste, non relativiste (c.-à-d. faillibiliste) à savoir une notion qui a une certaine force explicative. En hissant cette idée régulatrice ainsi comme condition de possibilité de nos pratiques linguistiques, celle-ci s'avère bien contextualiste, et sa dimension régulatrice, contrefactuelle d'une situation idéale de parole implique qu'elle est susceptible de ne pas se réaliser complètement au sein de toute pratique réelle d'une argumentation rationnelle, par conséquent, ce n'est pas d'un point de vue relativiste qu'elle rend compte de la structure et de la dimension faillibiliste de nos pratiques sociales : l'idée régulatrice met en évidence le fait que le monde un n'est pas en soi assimilable à nos

---

<sup>1151</sup> Habermas, 2001, p.183 : La réserve faillibiliste et hypothétique qui affecte tout le savoir conquis par l'argumentation (c'est-à-dire son exposition de principe à la force d'un argument meilleur, mais pour le moment imprévisible) ne saurait par conséquent être maintenue indéfiniment, sans quoi la pratique ne pourrait jamais recouvrer ses assurances. Aussi, sous peine de condamner la vie à une sorte d'impuissance pathologique (où chacun serait condamné à argumenter sans fin !), le processus discursif devra-t-il être interrompu dès que les interlocuteurs, ayant épuisé toutes les explications et les preuves *disponibles*, seront parvenus à un accord rationnellement motivé.

perspectives langagières, même s'il n'acquiert de signification pour nous que dans cet élément langagier soit intersubjectivement. Cette conception satisfait donc notre *desideratum* : elle est anti-réductionniste, faillibiliste (soit relativiste), *anti-platoniste* et explicative.

Notre discussion de l'approche de Habermas et de McCarthy s'est ainsi concentrée sur l'objectivité des prétentions à la vérité puisque cette question est au centre de leurs réflexions mais nous nous intéressons davantage à l'application des normes de signification du point de vue de leur objectivité et de leur correction. En raison du lien entre la vérité et la signification forgée par la dérèfification de la signification et l'abandon de la distinction de l'analytique et du synthétique de Quine-Davidson, les pratiques tels que la science empirique et les contextes communicationnels descriptifs du langage ordinaire peuvent être analysées sur le modèle des propositions portant sur le monde objectif, prétendant ainsi à la vérité ou à l'application correcte des termes empiriques théoriques tels que « eau », « acide », « lumière », etc. ainsi que des termes empiriques non théoriques tels que « vert », « neige », « herbe », etc. Ainsi, nous pensons qu'à l'image de l'objectivité des énoncés prétendant à la vérité, un traitement similaire pourrait être envisagé pour l'objectivité des normes de signification. Autrement dit, nous proposons que l'idée d'un monde objectif « un et toujours le même » qui fonctionne de manière régulatrice afin de guider les pratiques sociales de nos échanges langagiers soit une condition nécessaire pour fonder l'idée de normes objectives (soit faillibles) de la signification.

Ainsi, afin de rétablir le lien entre la discussion de Habermas et de McCarthy sur les énoncés prétendant à la vérité et notre discussion sur les normes de signification, rappelons que nous exigeons d'une conception sociale des normes de la signification qu'elle puisse asseoir l'idée de normes constitutives ou objectivement correctes de la signification, c.-à-d. faillibles, et pour établir un concept faillibiliste de normes de signification, nous devons fonder une notion de correction ou de révision qui ne s'appuie pas seulement sur l'observation d'une régularité dans l'usage ou encore sur l'adoption arbitraire de préférences ou intérêts quelconques. Le faillibilisme<sup>1152</sup> conçoit cette notion de révision sur la base de ce qui existe soit comme une révision guidée par la vérité d'un phénomène. Par conséquent, le

<sup>1152</sup> Ibid., p.188 : « Lorsque nous adoptons une attitude réfléchie, nous savons parfaitement que tout savoir est faillible; néanmoins, dans la vie quotidienne, il nous est impossible de nous contenter d'hypothèse et de mener une vie de part en part faillibiliste. »; cf. aussi *ibid.*, p.196 et p.303.

faillibilisme, dans le cadre d'une approche sociale, ne doit donc pas simplement permettre de générer différents points de vue mais surtout différents points de vue qui viennent corriger ou corriger les points de vue précédents. Nous pensons ainsi qu'une socialité faible ou un contextualisme social guidé par l'idée régulatrice d'une vérité objective est nécessaire pour asseoir cette notion de révision, pour fonder le faillibilisme et non seulement le pluralisme, et pour fonder une notion de normes constitutives ou objectivement correctes de la signification. Dans le prochain chapitre, nous défendrons ce point de vue face à certaines objections.

Il est important ici de rappeler ce que nous n'entendons pas par l'idée de normes de signification objectivement correctes : il ne s'agit pas de normes de signification absolues et universelles, établies une fois pour toutes et détenues par tout locuteur. Nous voulons plutôt sous-entendre qu'il existe une forte présomption, certes non complètement à l'abri des critiques (non critiquable), pour que celles-ci soient vraies. C'est ainsi en fondant cette notion de la signification du point de vue de sa correction et de son objectivité que la notion kantienne d'idées régulatrices est utile dans la mesure où la vérité en tant qu'idée régulatrice ne signifie pas que la vérité sur un phénomène soit toujours établie une fois pour toutes dans tous les contextes réels de l'enquête, ni qu'il faille comprendre qu'appliquer les normes de signification conformément à leur correction et d'un point de vue objectif nous engage à dire que la signification est déterminée une fois pour toutes dans tous les contextes réels d'usage ou de discussion. Aussi, l'histoire de la science empirique devrait s'identifier à l'histoire de la révision des soi-disant « vérités », soit l'histoire qui va de réponse faillible en réponse faillible jusqu'aux normes de signification objectivement correctes attribuées aux termes scientifiques. Le faillibilisme souligne d'ailleurs bien que c'est cette idée envisagée comme idée régulatrice qui donne sens à la recherche scientifique perçue ainsi comme une enquête faillible, un processus perpétuellement en cours d'établir les normes de signification de certains termes empiriques. Autrement dit, la pratique de la science comme investigation faillible portant sur la nature du monde révèle et maintient la fonction régulatrice de cette idée de vérité ; la pratique des scientifiques illustre cette démarche dans la mesure où ces derniers ne se contentent pas d'en rester à des résultats expérimentaux conflictuels mais procèdent plutôt à de nouvelles expériences afin d'invalider les positions concurrentes et de se rapprocher du domaine de la connaissance soit de la manière correcte d'employer les termes théoriques. Par conséquent, parler de normes constitutives ou objectives de signification n'est

pas synonyme de normes nécessairement ou universellement fixées d'avance; elles sont au contraire sujettes à révision (dans des circonstances particulières).

#### 7.6. Conclusion

L'application objective des termes en vertu de leur correction, en étant comprise comme une idée régulatrice, fait en sorte que tout accord sur l'emploi d'un terme est provisoire, c'est-à-dire sujet à une possible évaluation, critique ou correction. Mais, en concevant la signification correcte comme le point de repère orientant une telle discussion, évaluation ou critique, on laisse place à un pluralisme critique des points de vue, soit aux notions faillibilistes d'erreur et de correction, dans la mesure où nous suggérons en quelque sorte que ces points de vue rivalisent pour obtenir « le prix » de l'application correcte (et deux applications d'un terme ne peuvent pas toutes deux être correctes). Par ailleurs, il s'agit d'une conception anti-platoniste des normes de signification dans la mesure où elle situe la normativité là où elle devrait être localisée – dans l'évolution actuelle, sociale et historique des usages linguistiques. Les normes constitutives ou objectivement correctes de la signification émergent donc de nos pratiques d'usage au sein de nos contextes sociaux de discussion, de critique, etc. Elles assurent la fonction d'« idée régulatrice » : bien qu'on ne puisse actuellement garantir sa réalisation effective soit la manière adéquate d'employer un terme, tendre vers cet idéal introduit des pratiques fécondes. Cette approche non relativiste (faillibiliste) de la normativité est ainsi alimentée par une idée régulatrice orientant l'articulation réelle de nos pratiques sociales. Elle est aussi non réductionniste dans la mesure où elle ne réduit pas la normativité à de simples régularités *de facto* et elle a une fonction explicative en ce qu'elle dévoile la manière dont nos pratiques sociales doivent être structurées afin de fonder une dimension normative, à savoir : nos pratiques linguistiques sociales doivent être guidées par cette « idée régulatrice » dévoilant la manière correcte d'appliquer un terme. Cette approche de la normativité de la signification est donc bien non-réductionniste, non-relativiste, non-platonicienne bien que justifiant (dans une certaine mesure) l'idée d'une normativité primitive.

## CHAPITRE VIII

### OBJECTIONS ET RÉPONSES

#### 8.1. Introduction

La position de Habermas et de McCarthy fait face à deux objections majeures : tout d'abord, celle (§8.2) selon laquelle l'idée régulatrice d'une vérité unique et objective énoncée sous sa forme intersubjective comme accord universel rationnel est ni nécessaire ni suffisante pour fonder le pluralisme faillibiliste qui constitue la signification pragmatique des énoncés prétendant à la vérité, et corrélativement, celle (§8.3) selon laquelle notre extension de cette affirmation à la normativité de la signification, soit l'idée que l'application objective de certaines normes de signification fonctionnant de manière régulatrice, n'est pas constitutive du sens de certaines pratiques linguistiques, notamment de celles qui décrivent le monde. Nous aborderons donc dans ce chapitre ces deux points successivement.

#### 8.2. Première objection et réponse: contre l'idée d'une vérité objective sur le monde.

Habermas et McCarthy cherchent d'abord à éviter l'écueil du relativisme, et pour ce faire tentent de fonder le faillibilisme c.-à-d. une approche soulignant l'idée que certains points de vue peuvent être meilleurs, plus adéquats ou plus correctes que d'autres. Ils considèrent ainsi que nos pratiques sociales de discussion, de critique, etc., organisées autour de l'idée régulatrice d'une vérité objective sur le monde sont des conditions nécessaires et suffisantes pour asseoir le faillibilisme. Aussi, en se voulant contextualistes, ils s'efforcent de fonder la notion faillibiliste de vérité à partir des situations concrètes émergeant de nos pratiques sociales. Aussi, Habermas, par son articulation des deux niveaux différents à partir desquels se situent l'action et la discussion, parvient à maintenir la référence à un monde objectif servant d'horizon indispensable aux participants à l'action, ce qui permet, lorsque l'action a été interrompue pour que soient examinées des prétentions à la validité contestables, d'orienter les discussions réparatrices en direction d'une référence à la fois commune, unique et « indisponible » à savoir l'existence d'un monde et d'une vérité objective (formulée de manière intersubjective en termes d'acceptabilité rationnelle

universelle de tous les juges compétents) « existant de façon indépendante et identique pour tous<sup>1153</sup> ». Cette référence à un monde et à une vérité objective – ou, dans le domaine moral, à un « monde social idéalement projeté » - tient au fait que les sujets qui parlent et agissent ne peuvent éviter d'obéir à des normes et de recourir à un savoir intuitif. En faisant de ce monde objectif et de ce monde social idéalement projeté les horizons qui permettent à la fois de stabiliser les actions et de conduire les discussions, le pragmatisme de Habermas permet donc de rendre compte du processus d'apprentissage en général. En effet, sans ces références transcendantes, et sans les idéalizations propres à toute discussion, la connaissance serait tout simplement impossible. Une discussion n'est donc à proprement parler rationnelle que si elle se déroule dans certaines conditions, à savoir : « (a) un déroulement public et l'inclusion complète de tous les intéressés, (b) une égale répartition des droits de communication, (c) le caractère non violent d'une situation n'admettant que la force non coercitive du meilleur argument et (d) la sincérité des déclarations faites par tous les participants.<sup>1154</sup> ». Comme nous l'avons mentionné plus haut, Habermas désigne ces conditions sous l'expression de « situation idéale de parole<sup>1155</sup> ». Il s'ensuit qu'un énoncé ne peut être véritablement fondé et considéré comme valide que s'il obtient un consensus au terme d'une discussion rationnelle ayant lieu dans une situation idéale de parole. Dans le cadre de la fondation des normes, Habermas appelle « principe discussion » ou « principe D » ce critère de validité<sup>1156</sup>. Ce principe peut tout aussi bien s'appliquer aux énoncés empiriques ordinaires tels que « Il y a une tour ronde au lointain » qu'aux propositions scientifiques plus théoriques comme « La lumière est une onde et non un ensemble de particules », puisqu'elles prétendent également à la reconnaissance intersubjective. Ainsi, un énoncé qui prétend à la vérité empirique sera considéré comme valide s'il permet non seulement de rendre compte de ce qu'il prétend décrire mais aussi s'il fait l'objet d'un accord de tout être rationnel en tout lieu au terme d'une discussion rationnelle. Mais, ce processus complexe de l'apprentissage dépend aussi du

<sup>1153</sup> Habermas, 2001, p. 279 : « Empêtrés dans leurs pratiques, les sujets, à partir de l'horizon de leur monde vécu, se réfèrent à quelque chose dans le monde objectif qu'ils présupposent, sur le mode soit de la communication soit de l'intervention, comme un monde existant de façon indépendante et identique pour tous. ».

<sup>1154</sup> Ibid., p.300.

<sup>1155</sup> Habermas, 1987b, p.275-328, 1995, p.101-122, en particulier p.115-118.

<sup>1156</sup> Habermas, 1997, p.123 : « D : Sont valides strictement les normes d'action sur lesquelles toutes les personnes susceptibles d'être concernées d'une façon ou d'une autre pourraient se mettre d'accord en tant que participants à des discussions rationnelles. ».

fait que des « attentes régulatrices de l'action sont devenues problématiques<sup>1157</sup> ». Or, comment « résoudre » le tissu du monde sans s'appuyer sur une référence à la fois ancré dans celui-ci et le dépassant ? Habermas reprend ainsi « le naturalisme atténué » de Putnam qui, à la question de savoir comment des processus d'apprentissage sont possibles au-delà des limitations linguistiques des différentes époques et des différentes cultures, donne la réponse suivante :

Pour qu'une interprétation, que ses propres conditions de connaissance ont rendue rationnellement acceptable, puisse être identifiable comme étant erronée dans une autre situation épistémique, il faut, en effet, que le phénomène à expliquer ne se perde pas au cours du passage d'une interprétation à l'autre. Il faut que, en dépit de descriptions différentes, on puisse maintenir la référence à un même objet<sup>1158</sup>.

Le « primat épistémologique » propre à un pragmatisme fondé sur une *détranscendantalisation* de la théorie de la connaissance et assumant des présupposés contextualistes ne doit donc pas supprimer le « primat ontologique », c'est-à-dire cette référence à un monde unique et indépendant. Par conséquent, sans la référence à un monde objectif ou à un monde social – même idéalement projeté – il n'existerait pas de motivation rationnelle incitant à élargir le champ d'appartenance et à opérer un décentrement sans cesse plus important des perspectives morales et cognitives. Le sujet cherche donc à convaincre et à défendre ses opinions devant un public en étant guidé par cette idée régulatrice l'orientant vers une compréhension toujours accrue de ses raisons et des raisons des autres. Car il y a bien lors du processus de discussion, transformation, modification, amélioration, amendement, rectification, restauration, etc. Cette idée régulatrice envisagée comme un idéal à atteindre est ainsi nécessaire pour donner à ce processus une dimension faillibiliste et non pluraliste.

L'objection de Rorty consiste alors à dire que les désaccords qui peuvent se produire dans le cas d'énoncés empiriques comme celui relatif à la forme d'une tour ou d'énoncés scientifiques comme celui relatif à la nature de la lumière peuvent éviter de dégénérer en un relativisme de la forme « tous deux sont également vrais » sans faire appel à l'idée régulatrice d'une vérité une et objective sur le monde (ou un consensus rationnel universel) – cette idée n'étant donc pas nécessaire pour expliquer le sens non relativiste de ces propositions-ci. Sur

<sup>1157</sup> Habermas, 2001, p.275.

<sup>1158</sup> Ibid., p.297.

l'exemple cité plus haut de l'énoncé « Il y a une tour de forme ronde au loin », Rorty ferait remarquer que de se contenter de faire reposer ce désaccord sur des positions différentes et contradictoires ne sert pas nos intérêts. Il est ainsi difficile de permettre aux locuteurs de coordonner leurs actions et de coopérer avec succès s'ils entretiennent des croyances contraires ou différentes sur le monde. Certes, les locuteurs peuvent par la suite rendre compte du monde différemment mais selon Rorty, ces processus d'apprentissage ne sont pas rendus possibles par la présupposition de l'idée régulatrice d'une vérité unique et objective sur le monde ou d'un accord rationnel universel, mais par la poursuite d'un intérêt pragmatique comme celui de la recherche du maximum de bonheur possible. Par conséquent, la dimension faillibiliste de la discussion et du débat argumentatif ne devrait pas faire appel à la notion d'idées régulatrices comme le préconise l'approche de Habermas et de McCarthy; le critère pragmatique tel que la satisfaction de nos intérêts ou la promotion du bonheur serait suffisant. Ainsi, selon Rorty, la notion de révision, d'amélioration, de rectification inhérente aux processus de discussion pourrait bien être fondée – une réponse serait alors meilleure qu'une autre si elle augmente la quantité de bonheur et moins bonne si elle en réduit la quantité - mais l'idée d'une référence à un monde et à une vérité objective n'est pas nécessaire. De plus, dans l'exemple des débats théoriques sur la nature de la lumière, il semble que nous puissions seulement faire appel aux standards et critères pragmatiques ; suggérer que nous disposons de l'idée d'un monde un et objectif ne paraît pas s'appliquer ici puisqu'il s'agit d'une question dont le critère objectif reste indécidable. Les preuves sous-déterminent la théorie ; aussi, décider quelle est la meilleure position à soutenir sur la nature de la lumière n'est pas déterminée sur la base de ce qui se rapproche de la vérité quant à la nature de la lumière. Celle-ci est plutôt déterminée par le critère pragmatique qui nous permet de poursuivre nos buts scientifiques de prédiction et de contrôle. Par conséquent, Rorty conclut que ce faillibilisme qui critique et évalue les différents points de vue afin d'arriver à la meilleure réponse possible, ne requiert pas les idéaux Kantiens de vérité objective ou d'accord rationnel universel.

Habermas admet d'ailleurs lui-même la pertinence de cette objection portant sur la prétention à l'universalité comme venant à l'encontre de son principe de discussion. Un énoncé universellement valide est donc un énoncé faisant l'objet d'un accord rationnel universel. Il faudrait donc, pour s'assurer de la validité universelle d'un énoncé, le confronter

à tous les êtres rationnels dans l'espace et le temps. Cependant, si la réalisation d'un accord rationnel factuel dans une situation idéale de parole n'est pas fondamentalement irréalisable, ce n'est pas le cas d'une audience universelle rassemblant tous les êtres rationnels. Dans l'état actuel des choses, tout ce que la procédure dialogique peut obtenir, c'est un accord rationnel « concret », c'est-à-dire un accord obtenu à partir d'un nombre limité d'individus<sup>1159</sup>.

Or, un accord rationnel concret, même obtenu dans des conditions idéales, n'équivaut pas à un accord rationnel universel. En effet, comme le dit Habermas lui-même :

dans la mesure où toutes les discussions réelles, se déroulant dans le temps, sont provinciales par rapport à l'avenir, nous sommes incapables de savoir si les énoncés qui sont aujourd'hui rationnellement acceptables, serait-ce dans des conditions approximativement idéales, résisteront également, à l'avenir, à toute tentative de réfutation<sup>1160</sup>.

Dans les termes tranchants de Rorty, dire « Je vais essayer de défendre ceci contre tous les opposants » est louable; mais dire que « Je peux défendre cela contre tout opposant » n'a pas de sens. Il est impossible de dire que l'on pourra défendre sa position contre tous les contradicteurs, car on ne sait pas comment seront les futurs contradicteurs<sup>1161</sup>.

Par conséquent, la prétention discursive permet sans doute de satisfaire à la prétention à la reconnaissance factuelle, mais non à la prétention à l'universalité proprement dite. Rorty mentionne ainsi que l'idéal d'une vérité objective et unique définie comme *prétention à l'acceptation rationnelle universelle* n'est pas suffisante pour fournir un cadre contextuel aux participants adoptant une attitude faillibiliste à l'égard de leurs propres jugements et

<sup>1159</sup> Cf. Rorty, 2000b, p. 56-64, p. 60: « Our finitude consists in the fact that there can never be an ideal audience, only more spatially, temporally, and socially restrained audiences. ».

<sup>1160</sup> Habermas, 2001, p.192 ; cf. aussi *ibid.*, p. 190 : « [...] l'acceptabilité rationnelle, même dans ces circonstances idéales, n'exclut pas l'erreur et ne simule donc aucune propriété "définitivement acquise" [...] » ; *ibid.*, p.304 : Rien ne garantit que dans « d'autres contextes, sur la base d'autres expériences, de meilleurs arguments pourraient survenir, pouvant affaiblir l'ici et le maintenant de ce que toi et moi tenons pour vrai. » ; *ibid.*, p. 301 : « Même les arguments qui, ici et maintenant, nous convainquent irrésistiblement de la vérité de 'p' peuvent, dans une autre situation épistémique, se révéler faux. ».

<sup>1161</sup> Cf. Rorty, 2000b, p. 56 où Rorty formule l'exemple du « champion du village » ; cf. aussi *ibid.*, p.56-57: « It is not the case, as Habermas says, that 'What we hold to be true has to be defensible on the basis of good reasons, in all possible contexts.' If it were, I would, whenever I acquired a belief, be tacitly making an utterly unjustified empirical prediction about what would happen in a potentially infinite number of justificatory contexts before a potentially infinitely diverse set of audiences. I find this as implausible as the suggestion notoriously made by the logical positivists: that every empirical assertion is an empirical prediction about a potentially infinite number of future sense-data. ».

standards de justification. La discussion est sans doute une condition nécessaire pour garantir la validité universelle des énoncés, mais non une condition suffisante. Dans les termes de Habermas, la vérité – définie comme accord rationnel universel – transcende toute justification. D'où cette formule paradoxale de Habermas : « Le but des justifications est de découvrir une vérité qui dépasse toute justification<sup>1162</sup>. » Plus précisément, si le *désaccord* de la communauté scientifique garantit que la proposition en question est mal justifiée et donc insuffisante, en revanche, un *accord*, même obtenu dans une situation idéale de parole, ne garantit pas une validité universelle. Cela concerne *a fortiori* les normes de la situation idéale de parole. Si celles-ci doivent être établies au moyen de la discussion, et si le résultat de la discussion ne garantit pas une validité absolue, alors les conditions idéales de parole ne peuvent être absolument fondées. Rorty affirme ainsi que ce n'est pas l'idée régulatrice d'un accord rationnel universel qui encourage les participants au dialogue à admettre qu'un autre point de vue puisse être meilleur que le leur, mais c'est bien en se confrontant concrètement aux différents points de vue qu'ils reconnaissent que leurs points de vue méritent d'être rectifiés ou réexaminés. Pour préserver le pluralisme faillibiliste – soit l'examen critique de nos propres opinions de telle sorte que ce qui nous tenons pour vrai ne se réduise pas à ce que nous pensons être vrai à un moment donné – nous devons garder à l'esprit que nos points de vue et standards dépendent de nos besoins et intérêts et qu'il est probable que d'autres aient des besoins et intérêts différents pouvant ainsi générer des points de vue différents des nôtres.

Aussi, pour éviter l'écueil du relativisme, les approches sociales choisissent de défendre le faillibilisme, celui-ci impliquant non seulement l'idée de différents points de vue mais surtout celle de leur évaluation : c'est l'idée qu'un point de vue n'est pas simplement différent d'un autre point de vue mais qu'il peut être meilleur (ou moins bon) ou le fruit d'une version améliorée de celui-ci. Par conséquent, le faillibilisme fonde aussi la notion d'amélioration ou de progrès, et la vraie question est alors de savoir quelles sont les conditions nécessaires et suffisantes pour arriver à un accord socialement reconnu afin d'évaluer le point de vue de tout un chacun comme un progrès ou un recul dans la discussion.

---

<sup>1162</sup> Habermas, 2001, p. 303 ; cf. aussi *ibid.*, p. 304 : « [...] même les meilleures raisons sont faillibles ; ainsi, au seul endroit où la vérité et la fausseté des énoncés peuvent faire l'objet d'un discours, l'abîme entre acceptabilité rationnelle et vérité est insurmontable. » ; *ibid.*, p.302 : « Bien que le lien entre vérité et justification soit pour nous irréductible, il ne faut pas que ce lien épistémiquement incontournable soit stylisé – au sens d'un concept épistémique de vérité – en lien conceptuellement indissoluble. ».

Finalement, nous nous demandons si nos pratiques sociales orientées en fonction des standards pragmatiques immanents au contexte sont suffisantes pour rendre compte de la notion de progrès ou si ces mêmes pratiques sociales mais orientées, cette-fois, par les idées de raison transcendantes au contexte, tels que la vérité objective et l'accord universel rationnel, sont nécessaires ? Autrement dit, la question pourrait se résumer ainsi : avons-nous besoin de l'idée de perfection (une référence à un standard idéal) pour asseoir la notion de progrès, ou est-ce que des standards plus locaux et provisoires pourraient suffire ? Rorty et Michael Williams pensent que des pratiques comme les sciences naturelles génèrent la notion d'amélioration sans faire appel à des idéaux – le développement interne de la pratique de la science, qui découle et sert des intérêts et buts pragmatiques, semble suffisant pour instituer des normes d'évaluation, générer une notion d'amélioration ou de progrès et de ce fait, le faillibilisme; nous n'avons pas besoin d'ajouter l'idée d'une vérité objective sur le monde, celle-ci agissant à l'image d'un *focus* imaginaire vers lequel tous nos points de vue convergeraient afin de définir ce qu'est le progrès. L'histoire de l'évolution de la science, la reconfiguration continue des problèmes et des solutions scientifiques détermine les questions qui sont actuellement pertinentes de celles qui ne le sont pas, distingue les problèmes qui ont été traités avec succès de ceux qui n'ont pas abouti, c'est-à-dire qu'elle définit les standards en évaluant le progrès effectué ou la régression réalisée. Être contextualiste ne veut donc pas dire qu'il faille se prévaloir de normes indépendantes de tout contexte d'usage pour évaluer le progrès mais bien au contraire, que de telles évaluations se basent avant tout sur nos intérêts pragmatiques et nos pratiques sociales. Rorty avance alors que de tels fondements, qui sont finalement les seuls fondements que nous puissions avoir, fournissent un sens tout à fait satisfaisant de la notion de progrès et sont donc suffisants pour éviter l'écueil du relativisme :

« [T]he only criterion we have for applying the word « true » is justification, and justification is always relative to an audience. So it is also relative to that audience's lights – the purposes that such an audience wants served and the situation in which it finds itself. This means that the question « Do our practices of justification lead to the truth? » is [...] unanswerable because *there is no way to privilege our current purposes and interests* [...] But surely, it will be objected, we know that we are closer to the truth. Surely we have been making both intellectual and moral progress. Certainly we have been making progress, *by our lights*. That is to say, we are much better able to serve the purposes we wish to serve, and to cope with the situations we believe we face, than our ancestors would have been. But when we hypostatize the adjective « true » into « Truth » and ask about our relation to it, we have absolutely nothing to say [...] [I]t seems enough simply to define scientific progress as an increased ability to make predictions [and] to define moral progress as becoming more like ourselves at our best (people who are not racist, not aggressive, not intolerant, etc.)<sup>1163</sup>».

Prenons un exemple discuté par Rorty<sup>1164</sup> : la position de Bellarmine n'est pas « aussi pertinente que » celle de Galilée. Selon Rorty, la position de Galilée est plus pertinente, mais, seulement dans la mesure où le point de vue de Galilée contribue favorablement à servir nos intérêts (prédiction et contrôle), que celle de Bellarmine, intérêts que le débat entre Galilée et Bellarmine a permis d'instituer. L'histoire de la pratique scientifique a privilégié certains intérêts en avançant que certains seraient meilleurs que d'autres ; cependant, ces derniers ne seraient pas meilleurs ou plus pertinents que d'autres en tant qu'ils représenteraient une manière unique et objective de décrire le monde ou en tant qu'ils détiendraient le vocabulaire correct et adéquat pour saisir la réalité. Rorty distingue donc deux sens du terme « meilleur » – un sens absolu qui n'admet aucun comparatif et un sens relatif rattaché à nos intérêts et buts pragmatiques :

« I want to recommend explaining 'better'[...] as 'will come to seem better to us' [rather] than 'closer to the way the world really is' or 'closer to the way God sees it' or 'closer to the facts of the matter' or 'closer to idealized rational acceptability.'<sup>1165</sup>».

<sup>1163</sup> Rorty, 1998a, p.4-5: nos italiques; Rorty, 2000b, p.58: « The whole idea of context-independence, in my view, is part of an unfortunate effort to hypostatize the adjective 'true'. »

<sup>1164</sup> Rorty, 1979, p.328f.

<sup>1165</sup> Rorty, 1993, p.455.

« What should [my and] Brandom's strategy be when charged with a relativistic disbelief in the reality of intellectual progress? The obvious option is Kuhn's: « to substitute evolution-from-what-we-do-know for evolution-toward-what-we-wish-to-know », thereby switching from progress toward a *focus imaginarius* to improvement on the historical past. This amounts to switching from pride in being closer to reality to pride in being farther from the cavemen<sup>1166</sup>. ».

Le qualificatif « meilleur » signifie d'une part « proche de la Vérité », et d'autre part, « meilleur que certains autres points de vue en raison des buts et intérêts visés ». C'est l'idée même de vérité qui devient suspecte pour Rorty, puisqu'elle ne peut que contribuer à rééditer le modèle de la philosophie traditionnelle, celui qui aperçoit dans la subjectivité et dans le langage (réduit à sa fonction sémantico-référentielle) une simple structure de représentation et de reproduction mise au service d'une objectivité pure<sup>1167</sup>. Pour une philosophie comme la sienne, qui a pris acte du caractère essentiellement interprétatif de l'expérience humaine, et qui n'admet plus qu'une raison incarnée et située dans les pratiques contextuelles de l'échange langagier, la référence à une telle réalité fondamentale se révèle trompeuse ; c'est même la fonction thérapeutique de la philosophie que de nous guérir de cette illusion, en concevant la connaissance non plus comme une affaire de représentation, mais comme une affaire de pratique sociale et de conversation, par où elle se voit dépouillée de son ancien privilège...<sup>1168</sup> Cela dit, refuser l'idée d'un standard absolu n'implique pas pour autant que tous nos jugements ou prises de position soient arbitraires ou également pertinents; nos prises de position sont le reflet des buts pragmatiques que nous poursuivons : si ces buts sont utiles pour nous, nous les recherchons mais s'ils nous sont nuisibles ou désavantageux, nous les évitons<sup>1169</sup>. Rorty soutient par la suite que toute référence transcendant les contextes locaux

<sup>1166</sup> Rorty, 1997a, p.175.

<sup>1167</sup> Rorty, 1979.

<sup>1168</sup> Habermas, 2001, p.175.

<sup>1169</sup> Williams, Mi., 1984, p. 260, esquisse une image très similaire : « [T]he thought behind the idealization theory is that, although philosophers like Rorty can pay lip service to such ideals as openness to correction, they have no real motive for taking them seriously. Lacking the idea of an 'ideal terminus' to our human conversation, they have no way of seeing changes as improvements or steps backward. Remarks come and go, but the conversation doesn't get anywhere [...] However, to say this would be arguing that, in politics, those who reject the idea of utopia lack the resources for advocating or even recognizing reform. But, in theory as well as in practice, improvements are made by addressing things that are unsatisfactory to us now [...] And as for openness to correction, past experience has provided enough instances of a well-entrenched theory's getting into trouble to deter us from identifying truth *simpliciter* with current acceptability ». Après avoir déterminé les motifs de l'approche de Kant (la nécessité de rendre compte du faillibilisme ou d'être ouvert à toute procédure de

doit être éliminée, c.-à-d. : puisque tout acte humain, toute réalisation humaine passent par le langage, ce qui advient est toujours à construire au moyen du langage et des hasards imprévisibles de l'interlocution. Ce sont donc les usages du langage qui priment sur tout point de vue prétendument divin, lequel suggérerait à tort qu'il existe un espace clos où tout aurait déjà été dit et qu'il s'agirait seulement de déconstruire et de reconstruire éternellement. En effet, si l'on suit Rorty, le concept de vérité est obsolète, car il suggère qu'il existerait des « truth-makers » en dehors de la réalité, et surtout en dehors du langage. Or, il n'y a rien, pour l'homme, en dehors du langage. Le vrai n'est rien d'autre que ce qui peut être justifié au regard de croyances propres à une communauté donnée. Le vrai se confond avec l'avantageux, l'utile. La célèbre phrase de Hegel sur la rationalité du réel peut dès lors s'énoncer en des termes nouveaux. La justification de ce qui est tient dans son utilité, utilité pour la vie des hommes. Ce qui nous semble rationnel est ce qui est de fait utile pour la vie des hommes. On perçoit que le risque de réduction naturaliste et/ou fonctionnaliste n'est pas loin. Cependant, Rorty pense que ce « frank ethnocentrism » est suffisant pour évaluer nos jugements selon les critères du meilleur et du pire ou de l'utile et du désavantageux.

Pourtant, si l'on suit Rorty, le problème auquel se heurte, selon nous, cet ethnocentrisme est de sombrer dans le relativisme. En effet, à la question de savoir entre les points de vue de Galilée et de Bellarmine, celui qui est considéré comme le meilleur, Rorty affirme que bien qu'il n'y ait pas de point de vue neutre sur lequel nous puissions émettre un jugement, cela n'implique pas que le dernier point de vue soit aussi bon que le premier dans la mesure où certains buts comme la prédiction et le contrôle préconisés par les approches de Galilée sont indéniablement supérieurs. Aussi, nous pouvons maintenant appliquer et poser de nouveau la question : sur quoi se base-t-on pour affirmer que les intérêts de Galilée de prédiction et de contrôle sont meilleurs que ceux de Bellarmine qui sont de « préserver l'Église et la structure culturelle et religieuse de l'Europe » ? Autrement dit, quels sont les critères qui nous permettent de décider entre des intérêts que nous jugeons meilleurs et utiles et des intérêts que nous jugeons moins bons et désavantageux ? Rorty laisse entendre que nous procédons

---

correction), Williams soutient qu'un *focus* imaginaire (« ideal terminus ») n'est pas nécessaire, et pour la même raison que Rorty – la notion faillibiliste de « meilleur » peut être obtenue à partir de standards locaux et d'intérêts cultivés à travers l'histoire de la résolution de problème scientifique. Williams conclut donc que nous n'avons pas besoin de l'idée de perfection pour comprendre la notion de « meilleur » (« we do not need the idea of utopia to fund a notion of social reform »); nous pouvons la saisir par le biais d'une pratique immanente qui admet la comparaison.

comme suit : nos intérêts sont meilleurs que ceux d'autrui selon nos standards et les intérêts d'autrui sont meilleurs que les nôtres selon leurs standards. Finalement, quand nous essayons de savoir quels intérêts il serait préférable de servir, les nôtres ou ceux d'autrui, nous n'avons pas de critères fixant ce qu'il est meilleur ou moins bon de privilégier. Nous pouvons décider parmi différents discours sur la réalité que selon nos valeurs scientifiques modernes de prédiction et de contrôle, le point de vue de Galilée est celui qu'il faut préconiser. De manière similaire, nous pouvons considérer le type d'organisation sociale dans lequel nous vivons comme meilleur ou moins bon selon notre propre critère « libéral » de liberté, d'égalité, de tolérance, etc. Selon ces mêmes intérêts libéraux, les sociétés démocratiques sont meilleures pour nous que le totalitarisme, mais si certains intérêts incluant la richesse et le pouvoir l'emportent sur d'autres, alors une société fasciste serait meilleure. Par conséquent, d'après Rorty, nous ne pouvons pas établir que certains intérêts sont meilleurs que d'autres puisque cette évaluation dépend des intérêts particuliers de l'auteur du jugement. C'est précisément cette réserve que McCarthy partage avec Rorty (et Hoy) : leur approche, en ne permettant aucune évaluation des intérêts ou buts suivis, conduit à « an uncritical pluralism of 'whatever serves your purposes, whatever they may be'<sup>1170</sup> ». Hoy et Rorty, en offrant dépendamment du contexte des perspectives différentes visant des intérêts spécifiques, proposent en fin de compte une conception du discours scientifique (aussi bien que du discours moral et politique) sur le mode peu critique (c.-à-d. relativiste) du Je pense (en raison de mes intérêts) et du Tu penses (en raison de tes intérêts).

Finalement, pour Hoy et Rorty, lorsque l'on nous demande ce que nos intérêts et valeurs peuvent bien avoir de spécial, nous ne pouvons que répondre en disant « ce sont nos intérêts ». Nous avons effectivement été élevés au sein d'une culture et d'une tradition, ce qui implique, par le fait même, l'acceptation d'un certain nombre d'intérêts et de valeurs bien spécifiques; McCarthy précise que nous ne devons pas croire pour autant que nos traditions sont meilleurs que celles d'autrui, munies d'autres valeurs et intérêts. Aussi, si nous disposons de nos propres valeurs et qu'autrui dispose des siennes, quelles raisons pourrions-nous bien faire valoir pour mettre en avant nos valeurs comme les meilleures ? Si nous n'avons pas de raisons pour justifier que nos intérêts sont meilleurs que d'autres, ne serait-il pas alors plus rationnel de les abandonner ? Cette question n'est pas sans rappeler le

---

<sup>1170</sup> Hoy and McCarty, 1994, p.233.

point de vue de Carnap sur les questions externes, soit les questions portant sur le cadre linguistique à adopter<sup>1171</sup>. Carnap pense ainsi qu'adopter un cadre linguistique relève d'un choix conventionnel guidé par nos intérêts pragmatiques d'utilité et non d'une délibération rationnelle. La rationalité suggère donc un cadre de travail relatif, une fois que nous adoptons un cadre de travail et que nous posons un certain nombre d'intérêts, certains points de vue seront objectivement meilleurs que d'autres (eu égard à ces mêmes intérêts visés). Pour Carnap, le choix d'un cadre de travail relève d'une décision conventionnelle ; il en est de même pour Rorty, pour qui le choix des intérêts que nous poursuivons (ce qui, pour Carnap, revient à choisir un cadre de travail) n'admet pas de justification. Par conséquent, Carnap et Rorty ne peuvent pas justifier les raisons qui nous poussent à préconiser certains intérêts sur d'autres.

Pour eux, la « rationalité » de la science se résume en définitive à rien de plus qu'un choix arbitraire qui nous pousse à poursuivre certains intérêts et à en négliger d'autres. McCarthy et Friedman pensent que si la science est bien une pratique rationnelle, alors l'image de valeurs et d'intérêts non justifiés dénature la pratique scientifique ou au moins certains épisodes cruciaux de la pratique scientifique. La distinction employée par Kuhn entre la science « normale » et la science « révolutionnaire » permet d'éclairer leur objection. La science normale renvoie à ces périodes de l'histoire de la science où les praticiens sont d'accord sur les problèmes qui ont été résolus et sur ceux qu'il reste à résoudre, sur les standards ou normes formelles qui doivent fournir des réponses acceptables, sur la nature des intérêts à rechercher, sur les dispositifs expérimentaux à mettre en place, sur les « règles de recherche » à respecter afin de définir le cadre dans lequel s'effectue la mise à l'épreuve des théories, etc. Dans le vocabulaire de Kuhn, nous dirions que nous nous accordons sur le paradigme<sup>1172</sup>, ou dans les termes de Carnap qu'un cadre linguistique de travail a été établi

<sup>1171</sup> Friedman, 1997, 2002.

<sup>1172</sup> Le concept de « paradigme scientifique » est d'une imprécision constante chez Kuhn, cela de son propre aveu, cf. par exemple 1990, Chap XII, p.391-392 : « Aucun aspect du livre n'est autant responsable de cette excessive plasticité que l'introduction du terme 'paradigme', un mot qui y apparaît nettement plus que les autres, à l'exception des particules grammaticales. [...] Un commentateur, ayant trouvé qu'il valait la peine de voir la chose de près, a préparé un index partiel des matières, et a trouvé au moins vingt-deux usages différents, allant de 'découverte scientifique véritable' (p.11) à 'ensemble caractéristique de croyances et d'idées préconçues' (p.17), ce dernier sens regroupant des adhésions (*commitments*) instrumentales, théoriques et métaphysiques (p.39-42) » ; voir aussi Kuhn, 1983, p. 238-239. L'imprécision de ce concept est telle que les « malentendus » entre paradigmes évoqués par Kuhn peuvent tout autant s'appliquer à sa propre théorie des paradigmes

afin de permettre aux experts de la discipline de poursuivre leur processus de résolution de problèmes. Cependant, la poursuite de la science normale avec son pouvoir accru de précision des prédictions mène à des anomalies. Kuhn ne comprend ainsi les « anomalies » qu'en termes d'observations ou de « phénomènes », comme on peut le constater dans le passage suivant : « La découverte commence avec la conscience d'une anomalie, c'est-à-dire l'impression que la nature, d'une manière ou d'une autre, contredit les résultats attendus dans le cadre du paradigme qui gouverne la science normale.<sup>1173</sup>» Actuellement, la théorie des cordes est certes une « théorie élégante » qui permet d'unifier les lois de l'infiniment grand et de l'infiniment petit, mais elle ne bénéficie d'aucune vérification empirique, ce qui constitue une raison suffisante, selon la communauté scientifique, pour la considérer comme une simple hypothèse<sup>1174</sup>. Par conséquent, à ce stade, il n'y a pas d'accord sur le contenu théorique ou l'ensemble des postulats à adopter, sur les problèmes, sur les normes et intérêts de la science, tout comme sur les procédures visant à s'engager dans la pratique scientifique. Il en résulte des « révolutions scientifiques » ou « des changements de paradigme » où les problèmes que la science pose sont remodelés, tout comme les procédures et les critères utilisés pour évaluer leur validité sont revisités (le débat entre Bellarmine et Galilée est un exemple d'une révolution scientifique).

Le cadre de travail de Rorty ou sa conception de la rationalité et des évaluations du progrès orientée vers nos intérêts pragmatiques rend au mieux justice à la structure justificative de la science normale. Autrement dit, selon son approche, ce n'est que la science normale qui est décrite comme rationnelle puisqu'elle vise des intérêts qui permettent d'évaluer les théories comme meilleures ou non adéquates, rationnelles ou irrationnelles, vraies ou fausses. La révolution scientifique, où se joue la mise en concurrence de positions radicalement différentes, n'est pas une entreprise rationnelle mais vise simplement l'adoption *ad hoc* de certains intérêts au détriment d'autres. Par conséquent, l'objection adressée par McCarthy et Friedman à Rorty et Carnap revient à dire qu'au niveau de la révolution

---

scientifiques. Kuhn a bien tenté de préciser sa position dans ces textes en réduisant la notion de paradigme aux « ensembles communs » et en l'incluant dans le cadre plus large de « matrice disciplinaire ».

<sup>1173</sup> Kuhn, 1983, p.83. À vrai dire, Kuhn évoque à certains moments l'insuffisance de ce critère de l'adéquation aux faits, par exemple *ibid.*, p.203-204 : « Pour l'historien tout au moins, dire que la vérification est établie par l'accord des faits avec la théorie ne signifie rien. Toutes les théories ayant une importance historique ont été d'accord avec les faits. ».

<sup>1174</sup> Smolin, 2007.

scientifique, nous n'avons pas de critère rationnel, non relativiste pour qualifier une position de meilleure; il ne nous reste qu'un critère relatif d'évaluation basé sur la convention pour suggérer du bout des lèvres l'idée de progrès ou du moins de position ouverte à une éventuelle ou future révision (par opposition à la reconnaissance de points de vue complètement différents).

McCarthy et Friedman proposent du moins, à la différence de Rorty et de Carnap, des principes de rationalité qui régiraient l'ensemble des changements de paradigmes afin que ces derniers soit le fruit d'une décision rationnelle éclairée et non le résultat de l'adoption arbitraire de nouveaux points de vue. L'idée de progrès scientifique devrait ainsi refléter l'idée que la science évolue toujours vers des théories meilleures que les théories antérieures. Kuhn le reconnaît lui aussi implicitement lorsqu'il pose la question suivante : « Comment se fait-il qu'un changement de paradigme produise toujours un instrument plus parfait, en quelque sens, que ceux que l'on connaissait auparavant ? <sup>1175</sup> ». Il est vrai certes que lorsque l'on se situe au niveau de la notion de révolution scientifique, nous ne disposons plus de principes rationnels communs ou de formes *a priori* de l'entendement (comme Kant le pense) pour décider si une théorie est meilleure qu'une autre. Aussi, lorsque des paradigmes rivaux entrent malgré tout en discussion, chacun se réfère alors à ses propres critères de validité pour défendre sa position. Ce qui peut conduire à un dialogue de sourds :

Dans la mesure [...] où deux écoles scientifiques sont en désaccord sur ce qui est problème et ce qui est solution, elles s'engagent inévitablement dans un dialogue de sourds en discutant les mérites relatifs de leurs paradigmes respectifs. Dans la discussion proche du cercle vicieux qui en résulte régulièrement, il apparaît que chaque paradigme satisfait plus ou moins les critères qu'il a lui-même dictés et reste incapable de satisfaire certains des critères dictés par son concurrent<sup>1176</sup>.

<sup>1175</sup> Kuhn, 1983, p.235.

<sup>1176</sup> Ibid., p.156: « Quand les paradigmes entrent – ce qui arrive forcément – dans une discussion, leur rôle est nécessairement circulaire. Chaque groupe se sert de son propre paradigme pour y puiser ses arguments de défense. [...] Pour ceux qui refusent d'entrer dans le cercle, il ne saurait être rendu contraignant sur le plan de la logique ou même des probabilités. Dans une discussion concernant les paradigmes, les prémisses et les valeurs communes aux deux parties ne sont pas suffisantes pour permettre une conclusion sur ce plan. » ; voir également, Kuhn, 1990, p.448 : « Ces mêmes limites font qu'il est difficile, voire impossible, à un même individu d'avoir présentes à l'esprit les deux théories en même temps pour les comparer point à point et avec la nature. C'est pourtant bien ce genre de comparaison qui constitue le processus dont dépend la justesse d'un mot comme 'choix' ». ».

S'il est impossible de mettre en concurrence des paradigmes différents et donc d'établir de manière rationnelle une préférence pour l'un ou l'autre, alors un paradigme ne peut s'imposer à l'ensemble de la communauté scientifique que par la force ou la persuasion. Ainsi, le paradigme de la relativité générale n'aurait été préféré à la physique de Newton et admis par la communauté scientifique sur la base d'arguments rationnels, mais en raison des talents rhétoriques de ses défenseurs et de la pression socio-économique des institutions de recherche. Dès lors, le secteur scientifique ne serait pas cette société idéale où règnent les plus sages, mais au contraire une sorte d'« état de nature » où ceux qui survivent sont les plus agressifs et les plus retors<sup>1177</sup>.

Cependant, ce que McCarthy et Friedman proposent pour maintenir la rationalité de telles transitions sont les idées régulatrices Kantiennes, et tout particulièrement l'idée d'un accord rationnel à long terme parmi les chercheurs. Cette idée institue ainsi la notion procédurale de rationalité : elle ne nous incite donc pas à croire en l'idée de vérités absolues ou universelles tout en se basant sur quelque chose de moins arbitraire que la convention. Les positions orientées vers la recherche d'un consensus sont tout de même bien meilleures et plus rationnelles que celles qui en sont dépourvues, le consensus étant un processus toujours en cours de révision et d'amélioration et de ce fait jamais totalement achevé. Ainsi, avec cette idée, la structure perspectiviste selon laquelle, pour toute position discursive guidée par la recherche de nos intérêts, une autre mue par d'autres intérêts peut tout aussi être valable, gagne une dimension rationnelle.

Néanmoins, la transition vers de nouveaux paradigmes et le consensus qui finit par s'établir procède de bonnes raisons et non d'une convention arbitraire. Ce n'est certainement pas par

---

<sup>1177</sup> Cf. Kuhn, 1983, p.133-156 où Kuhn met en parallèle révolution politique et révolution scientifique, cf. par exemple p.135 « [...] les partis face à face dans un conflit révolutionnaire doivent finalement recourir à des techniques de persuasion de masse et souvent même à la force. » ; cf. aussi *ibid.*, p.269 : « Étant donné que le vocabulaire dans lequel ils discutent se compose toutefois, en grande partie, des mêmes termes, ils doivent établir entre ces termes et la nature un rapport différent, ce qui rend leur communication inévitablement partielle. En conséquence, la supériorité d'une théorie sur l'autre ne peut se prouver par la discussion. J'ai insisté sur le fait qu'au lieu de prouver, chaque parti doit essayer de convertir l'autre par la discussion. » ; voir également Feyerabend, 1987, p.97 : « Ceux qui critiquent l'idée selon laquelle les débats scientifiques sont tranchés de manière objective ne nient pas qu'il existe 'des moyens de décider' entre différentes théories. Au contraire, ils soulignent que de tels moyens sont nombreux ; qu'ils entraînent la possibilité de différents choix ; que le conflit qui en résulte est fréquemment résolu par les jeux de pouvoir soutenus par des préférences populaires et non par le raisonnement ; et que le raisonnement en tout cas n'est admis que s'il n'est pas seulement valide, mais aussi plausible, c'est-à-dire en accord avec des hypothèses et des préférences non raisonnées. ».

hasard que la grande majorité de la communauté scientifique finit par être convaincue de la supériorité d'un paradigme sur un autre : c'est le paradigme le plus performant dans la résolution de problèmes<sup>1178</sup>. Par exemple, la théorie d'Einstein est meilleure que celle de Newton parce que la théorie d'Einstein comble une partie au moins des problèmes que rencontre la théorie de Newton, sans que celle-ci soit capable de résoudre les problèmes que rencontre la théorie d'Einstein. De fait, la théorie d'Einstein permet de prédire des événements que ne peut prédire la théorie de Newton, notamment le périhélie de Mercure, alors que la théorie de Newton est incapable de résoudre un seul des problèmes que rencontre la théorie d'Einstein, par exemple le problème de son application à l'échelle quantique. Par conséquent, le consensus n'est jamais atteint parfaitement; il n'est que provisoire et aucune position ne peut prétendre délivrer la manière unique de décrire le monde. Nous avons donc ici une conception faillibiliste, soit non relativiste de la rationalité du discours scientifique sur le monde, notamment à travers les révolutions scientifiques<sup>1179</sup>.

### 8.3. Seconde objection et réponse : contre l'idée d'absence d'une dichotomie entre les vérités portant sur le monde et celles se rapportant aux normes de signification.

La seconde objection soutient que le sens de correction attaché aux normes de signification des termes du vocabulaire non logique ne requiert pas l'idéal d'une manière une et objective de décrire le monde. Aussi, quelle que soit notre réponse à la première question, rien ne semble justifier qu'un sens analogue de l'objectivité s'applique à la correction des normes de signification des termes empiriques. Cette objection laisse sous-entendre que les énoncés prétendant à la vérité portent sur le monde tandis que les normes de signification ont pour objet le langage et dans la mesure où les secondes sont, d'une certaine manière, le produit de l'activité humaine et que les premières ne le sont pas, le même degré d'objectivité ne peut être attribué aux normes de signification et aux énoncés prétendant à la vérité sur le

<sup>1178</sup> Kuhn, 1983, p.212, *ibid.* p.46 : « Les paradigmes gagnent leur rôle privilégié parce qu'ils réussissent mieux que leurs concurrents à résoudre quelques problèmes que le groupe de spécialistes est arrivé à considérer comme aigus. Réussir mieux, ce n'est pourtant pas réussir totalement dans tel problème unique, ni même réussir bien dans un grand nombre de problèmes. »

<sup>1179</sup> Or, Kuhn, 1983, affirme: « Bon nombre d'énigmes de la science normale contemporaine n'existent que depuis les révolutions scientifiques les plus récentes. Il n'en est qu'un tout petit nombre qui remontent aux débuts historiques de la branche scientifique qu'ils concernent. » (*Ibid.*, p. 195). Ce qui signifie que des théories nouvelles peuvent rencontrer davantage de problèmes que les précédentes.

monde. Par conséquent, si nous venions à accepter (pour la commodité du raisonnement) que les énoncés prétendant à la vérité peuvent requérir l'idée régulatrice d'une vérité une et objective sur le monde, ou d'un accord rationnel universel, rien n'implique que cette exigence s'applique aux normes de signification. Notre réponse à 1. n'a pas d'incidence sur notre réponse 2.

Cependant, si cette objection semble reposer sur le fait de vouloir séparer vérités portant sur le monde et vérités se rapportant à nos normes de signification, elle ne nous paraît pas justifiée dans la mesure où nous admettons que la signification est déréifiée dans nos pratiques d'usage. Il est au contraire évident que l'on ne peut séparer la notion d'application des termes du point de vue de leur correction et celle des énoncés prétendant à la vérité : plus nous formulons des énoncés erronés en employant les termes, moins il est probable que nous en usions correctement. Nous croyons donc que notre réponse au point 1 est pertinente pour notre point 2. Ainsi, en raison de ce lien entre propositions prétendant à la vérité et normes de signification, les échanges conversationnels du langage ordinaire et du discours scientifique sur le monde peuvent être tous deux tout aussi bien s'apparenter à un processus de discussion, d'évaluation, de défense, de rejet, etc. des énoncés prétendant à la vérité sur le monde ou à l'évaluation de normes en vue de l'application correcte de certains termes empiriques. Aussi, puisque nous avons soutenu que les énoncés prétendant à la vérité doivent faire appel à l'idée régulatrice d'un monde un et objectif ou celle d'un accord rationnel universel pour acquérir ce sens objectif qu'ils revêtent dans l'agir quotidien, nous avançons également qu'il en est de même pour les normes de signification des termes empiriques (aussi bien que logiques).

#### 8.4. Conclusion

Nous retrouvons finalement deux éléments issus de la conception d'une normativité forte des normes de signification. Tout d'abord, l'idée que les normes de signification sont constitutives de certaines pratiques linguistiques dans le sens où le terme « constitutif » signifie « susceptible d'être révisées ». Ensuite, l'idée que le contenu des normes de signification est régi par l'idéal d'un accord rationnel universel des locuteurs quant à l'application correcte d'un terme. Jointes ensemble, ces deux points soulignent que parler d'une normativité forte de la signification ne signifie ni pour autant que les normes de signification soient une fois pour toutes fixées d'avance, ni non plus qu'elles soient

réductibles à des régularités observables dans l'usage de nos pratiques linguistiques ordinaires. Autrement dit, la thèse d'une normativité forte rejette ni le fait anthropologique de différentes communautés linguistiques mobilisant différents concepts, ni le fait issu de l'histoire des concepts de la science au sein de laquelle une culture peut subir des changements au fil du temps ; mais elle tente de concilier ces faits-ci sans tomber systématiquement dans un relativisme conceptuel ou un réductionnisme naturaliste. Elle ne nie pas le fait que certaines normes de signification sont relatives à notre environnement culturel, à nos intérêts spécifiques et ponctuels, etc. Mais, ces normes ne sont pas universellement et absolument valides, elles demeurent révisables mais cette révision est régie par ces idées rationnelles régulatrices, et non par un simple intérêt arbitraire ou par l'observation d'un comportement.

La première caractéristique que cette notion de normes constitutives de la signification cherche à éviter est de donner un sens trop fort à la notion de normativité, comme celle qui lui associe l'idée de normes universellement admises et fixées d'avance ; la seconde, quant à elle, veut se prémunir du danger de lui donner un sens trop faible et d'éliminer toute dimension normative. La fonction régulatrice de cette idée d'un accord rationnel universel, laisse ainsi entendre que la norme forte de correction dont sont dotées les normes de signification ne se réduit pas à un accord communautaire, ou socialement fondé sur l'emploi des termes. Jointes ensemble, ces deux caractéristiques-ci fondent objectivement (de manière faillible) des normes de signification du point de vue de leur correction, normes qui sont sujettes à révision sur la base d'accords socialement motivés sans pour autant s'y réduire. Par ailleurs, elles contribuent favorablement à faire progresser le « problème » d'une circularité explicative en spécifiant le sens faillibiliste d'une forte normativité et en révélant la structure que des normes socialement constituées doivent avoir afin d'être porteur de ce sens-ci: accords fondés sur l'usage et orientés par l'idéal d'un consensus rationnel universel. En outre, leur dimension sociale constitutive veut dire que des normes fortes de signification ne se calquent pas sur un quelconque royaume d'entités platoniques, mais émergent (toujours de manière provisoire) de nos accords sociaux réels. Nous pensons donc que cette approche répond au critère d'adéquation *a priori* pour fonder une approche objective des normes de signification : non-platoniste, non réductionniste, non relativiste et explicative (jusqu'à un certain point).

## CONCLUSION GÉNÉRALE

Nous avons, tout au long de ce travail, cherché à établir le sens, la portée et les conditions nécessaires au fondement d'une conception sociale adéquate des normes de signification.

Notre premier volume vise d'abord à situer cette question de la normativité de la signification au sein du débat qui oppose le réalisme à l'anti-réalisme sémantique et social. Celui-ci se cristallise ici autour de la discussion sur les différentes formes de socialité défendues par Kripke, Wright et McDowell. Kripke, Wright et McDowell peuvent ainsi tous trois être considérés comme préconisant un « point de vue communautaire » ou une interprétation sociale se rattachant aux considérations de Wittgenstein quant à l'argument de « suivre une règle ». Kripke et Wright semblent se rejoindre en ce qu'ils attribuent à Wittgenstein une conception anti-réaliste de la signification. Cependant, ces points de comparaison ne vont pas sans cacher d'importantes différences. Leurs similarités et différences semblent bien jeter un pont entre deux niveaux d'analyse, à savoir : un premier niveau qui a trait à ce que chacune de ces approches considère être le point de vue de Wittgenstein quant à la signification; et un second niveau qui examine les conditions requises auxquelles doit répondre une approche adéquate de la signification, qu'elle soit fidèle ou non à la lecture de Wittgenstein.

Kripke attribue à Wittgenstein, comme nous le montrons, une conception sociale anti-réaliste de la signification. Il soutient que Wittgenstein endosse un argument en faveur du scepticisme sémantique, récusant l'idée même d'une application correcte et objective de la signification de toute expression linguistique donnée. Kripke pense ainsi que Wittgenstein arrive à la conclusion que la signification n'est pas objectivement déterminée, qu'elle ne peut être établie au moyen d'un fait ou de conditions de vérité, c'est-à-dire par l'obtention ou non de conditions objectives portant sur le monde. Il soutient alors que Wittgenstein, en acceptant le paradoxe sceptique, affaiblit le sens selon lequel le langage est porteur d'une signification déterminée – il n'existe pas de fait objectif (*fact of the matter*) pouvant déterminer la signification – dans la mesure où celui-ci est déterminé non par le monde mais par nous, par

les conditions sous lesquelles nous sommes justifiés à affirmer les énoncés que nous formulons. La solution sceptique attribuée à Wittgenstein vise ainsi à substituer à une conception réaliste de la signification – reposant sur l'idée de conditions de vérité objectives qui peuvent être ou ne pas être réalisées sans que nous soyons dans tous les cas en mesure de déterminer si elles le sont ou non – une conception anti-réaliste, appuyée sur l'idée de conditions d'assertabilité justifiée pour une proposition. La signification d'une phrase du point de vue de sa correction doit donc être comprise en termes d'une sémantique des conditions d'assertabilité relatives à la communauté. Mais si la signification ne consiste pas en des conditions de vérité objectivement déterminées, la « solution sceptique » ne veut pas non plus dire qu'il n'existe pas de signification correcte mais seulement, que la signification n'est pas objectivement déterminée par une condition émanant du monde c'est-à-dire, indépendamment de notre habileté à la connaître. Les énoncés sont au contraire employés et assertés dans certaines circonstances par les membres d'une communauté linguistique. La signification d'un énoncé est donc rattachée à l'observation de régularités dans le comportement d'un groupe de locuteurs, et notamment de régularités dans ses conditions d'assentiment au sein d'une communauté linguistique. Selon une telle conception, pour arriver à déterminer si «  $58 + 67 = 125$  » veut dire  $58 + 67 = 125$  ou  $58 <+> 67 = 5$ , il faut pouvoir déterminer les dispositions comportementales particulières d'une communauté linguistique spécifique c.-à-d., savoir dans quel cas la proposition peut être légitimement assertée c.-à-d. à quelle condition un agent est justifié de juger qu'une énonciation d'un autre agent est correcte. Par conséquent, la signification n'est pas objectivement déterminée mais plutôt relativement déterminée, en vertu des conditions d'assertabilité relatives à une communauté de locuteurs, c'est-à-dire de conditions selon lesquelles nos assertions ont un rôle et une utilité dans nos vies. Le sceptique soutient ainsi clairement que nous ne pouvons pas rendre compte d'un sens fort et objectif de la notion de correction mais que nous devons envisager, eu égard au problème de la détermination de la signification, une conception relativiste plus faible.

La solution sceptique concède alors au sceptique ses conclusions « paradoxales », mais montre que notre pratique quotidienne qui consiste à attribuer la maîtrise de l'usage de certains mots ou énoncés peut être autorisée. Elle admet, d'une certaine manière, que la position sceptique est, dans un sens philosophique profond, correcte; elle tente, tout en

reconnaissant la vérité profonde de la conclusion sceptique de vivre avec la conclusion problématique en offrant une version plus faible du phénomène mis en discrédit par le sceptique. Selon Kripke, Wittgenstein ôterait ainsi toute dimension objective à la notion de normativité de la signification et envisagerait la signification d'un point de vue épistémique, c'est-à-dire comme étant déterminée par ce que nous savons, assertons et faisons. Cependant, la solution sceptique réduit, comme nous le soulignons dans un second chapitre, la normativité de la signification, la dimension de « ce qui doit ou devrait être » contenu dans les expressions explicites de la signification d'un terme ou d'une phrase, à de simples régularités observables dans le comportement de groupes de locuteurs<sup>1180</sup>. Or, une conception de la correction relative à la communauté ne se présente pas comme une conception adéquate des normes de signification (c.-à-d. qu'elle ne permet pas de distinguer entre ce qui est correct et ce qui semble correct pour une communauté de locuteurs) dans la mesure où elle efface toute idée d'une application correcte et objective de la signification. Le danger d'une telle approche dispositionnelle est alors de tomber dans une forme de naturalisme sémantique et de réduire l'application correcte des normes de signification à notre usage *de facto*. McDowell ne manquera pas d'ailleurs de formuler une telle critique.

Quant à Wright, il s'accorde avec Kripke, en attribuant également à Wittgenstein une conception sociale anti-réaliste de la signification. Cependant Wright considère, dans le prolongement de Kripke, une approche de la normativité de la signification basée sur l'usage social ou communautaire comme tout à fait adéquate. Il n'envisage néanmoins pas cette approche anti-réaliste sur le modèle d'une « solution sceptique ». Wright considère plutôt l'argument du langage privé et la discussion autour de l'activité de suivre une règle de Wittgenstein comme une critique du platonisme sémantique. Le glissement vers une conception anti-réaliste de la signification ne constitue alors pas une solution de second choix venant supplanter, faute de mieux la solution initialement privilégiée c'est-à-dire une conception réaliste de la signification, celle-ci menant à un certain scepticisme sémantique. Selon Wright, c'est bien parce qu'une telle approche conduit inéluctablement au scepticisme

---

<sup>1180</sup> Kripke, comme McDowell, soutient que des normes de signification authentiques doivent être objectivement valides, c.-à-d. irréductibles à un accord communautaire, ce qui explique pourquoi, selon lui, l'anti-réalisme social de Wittgenstein, en rejetant l'idée d'une application correcte et objective des normes de signification, s'apparente à une solution sceptique quant à la normativité de la signification.

sémantique qu'une conception réaliste de la signification est une périlleuse illusion. Par conséquent, alors que Kripke croit que l'argument sceptique récuse la notion d'une application correcte et objective des normes de signification et interprète Wittgenstein comme offrant une « solution sceptique », celle-ci laissant la question d'une application correcte et objective des termes en suspens; pour Wright, une application correcte et objective des normes de signification est une illusion platoniste qui mène au scepticisme sémantique. Il n'existe rien qui réponde à l'idée d'une application correcte et objective des normes de signification et adhérer à une telle thèse implique un certain scepticisme sémantique. Le scepticisme sémantique se présente donc comme un faux problème engendré par une notion illusoire d'objectivité. Wright soutient ainsi que Wittgenstein, afin de nous écarter de toute conception platoniste (illusoire) de la signification, propose de substituer à cette notion d'une application correcte et objective des normes de signification celle, plus fidèle à une approche sociale anti-réaliste de la signification, d'un accord ou d'une « ratification dépendant de la communauté ». Wright avance que nous n'avons pas besoin de critère indépendant ou absolu pour évaluer la dimension objective de nos pratiques, celle-ci émane directement d'une distinction *immanente* à nos contextes ou pratiques communautaires. Un critère absolu ou indépendant de nos pratiques communautaires est non seulement *pas nécessaire* mais surtout *pas plausible* : nous devons nous résoudre à utiliser les outils à notre disposition à savoir, nos procédures d'enquête ou d'accord. Autrement dit, en affirmant qu'il n'existe rien au-delà de l'accord communautaire, Wright pense ne pas offrir une définition de l'objectivité sur le mode d'un consensus communautaire, mais plutôt insister sur le fait que l'idée même d'objectivité est dénuée voire vide de signification si elle fait référence à un consensus obtenu indépendamment de l'accord communautaire. Par conséquent, il soutient ne pas défendre une approche sémantique réductionniste : il n'affirme pas que la « vérité » et « l'objectivité » veulent dire ou se réduisent à un accord ou à un consensus communautaire. Plutôt que de préconiser une conception réductionniste de la normativité et de l'objectivité de la signification, Wright entend proposer une approche révisionniste : celle-ci ne prend en compte que l'accord communautaire, la dépendance aux contextes d'accord communément ratifiés et exclut toute notion de vérité et d'objectivité entendue dans un sens radicalement indépendant de nos pratiques communautaires, celles-ci n'étant, selon lui, d'aucune utilité.

McDowell s'oppose ainsi à l'interprétation que donne Wright de Wittgenstein et à l'anti-réalisme social comme théorie adéquate de la signification. Il pense que Wittgenstein doit préserver une conception « non anti-réaliste » de l'objectivité des normes de signification. Par conséquent, McDowell refuse d'admettre que Wright puisse établir une distinction entre les approches réductionnistes et révisionnistes de l'objectivité et de la normativité, et soutient que l'anti-réalisme de Wright s'apparente à une forme sociale de réductionnisme c.-à-d. qu'il réduit la notion d'une application correcte et objective des normes de signification à celles de régularités comportementales communautaires. La critique de McDowell ne consiste donc pas à montrer que Wright cherche à produire un critère objectif d'une application correcte quant à la signification à partir de ressources sociales et contextualistes mais à souligner que la conception de Wright en termes de « ratification dépendant de la communauté » s'identifie, comme nous l'avons montré dans notre chapitre III, à une forme de naturalisme réductionniste. Or, la position de McDowell vise à privilégier une conception non réductionniste des pratiques linguistiques et sociales afin de rendre compte de normes de signification objectivement correctes. Mais, selon McDowell, les pratiques normatives de la signification et de la compréhension deviennent, dans les termes de Wright, rien de plus que des régularités naturelles au sein de groupes d'individus répondant aux stimuli de leur environnement. McDowell tente ainsi d'expliquer la raison qui le pousse à qualifier la conception de Wright de naturalisme réductionniste, en mettant l'accent sur l'usage que fait celui-ci de la condition de *manifestation* : la signification doit être *manifestable* et compréhensible aux autres locuteurs ne disposant pas de connaissance préalable du langage, faisant ainsi du comportement le paradigme de cette dimension publiquement observable de la signification. Kripke formule également cette objection à l'encontre des approches s'appuyant sur les conditions d'assertabilité ou des conceptions sociales anti-réalistes de la signification. Mais si McDowell adhère à une telle objection, il s'oppose néanmoins à l'interprétation de Kripke qui présente Wittgenstein comme un anti-réaliste. En effet, McDowell est en désaccord sur ce point avec Kripke et ne croit pas que Wittgenstein préconise une conception de la signification basée sur les conditions d'assertabilité; il considère davantage que Wittgenstein vise à nous sensibiliser au danger d'une élimination naturaliste inacceptable de la dimension normative de la signification à laquelle conduit un anti-réalisme social. McDowell conçoit plutôt que Wittgenstein envisage

le paradoxe de l'argument des interprétations comme un *a reductio* à la fois des conceptions platonistes réalistes et des conceptions sociales anti-réalistes de la signification : si la signification est perçue comme un fait objectif déterminant d'avance et une fois pour toutes l'emploi des expressions, il en résulte alors un certain scepticisme sémantique; si la signification est conçue à l'image de régularités observées dans les réponses d'un groupe de locuteurs, la signification et la normativité sont alors naturalisées, et la portée significative et normative de toute expression donnée s'efface. McDowell soutient donc avec Wright que le réalisme platoniste mène au scepticisme sémantique et admet avec Kripke que l'anti-réalisme social réductionniste ne peut pas rendre adéquatement compte de l'objectivité et de la normativité de la signification. McDowell met ainsi en avant une position qu'il situe entre le réalisme et l'anti-réalisme et qu'il qualifie de « non-anti-realism ». Autrement dit, il défend une position anti-réaliste en ce qu'il soutient que l'objectivité d'une application correcte des normes de signification dépend de nos pratiques linguistiques et sociales (habituellement associée à l'anti-réalisme social); mais se réclame aussi d'une position non-anti-réaliste pour laquelle l'objectivité d'une application correcte des normes de signification (habituellement associée au réalisme platoniste) ne peut se réduire à nos pratiques et à l'idée d'un accord ou d'un consensus communautaire. McDowell propose ainsi pour sortir de cette impasse une conception sociale plus faible quant à l'objectivité et à la portée de l'application des normes de signification, en s'inspirant des notions Wittgensteiniennes de coutume et de pratique et de l'idée aristotélicienne de seconde nature. Ces notions-ci privilégient la socialisation de l'objectivité et de la normativité dans un sens non-réductionniste : elles visent à redonner aux concepts d'objectivité et de normativité leur irréductibilité essentielle requise pour fonder la signification sans en faire des entités métaphysiques transcendant la pratique sociale de l'usage du langage. Il s'agit donc, selon McDowell, d'introduire cette dimension normative, sociale et historique de nos pratiques linguistiques afin de ne pas réduire l'objectivité à des régularités d'usage observées au sein de groupes d'individus.

Cependant, nous ne pouvons néanmoins affirmer que McDowell échappe à son propre dilemme. En effet, si d'un côté, nous insistons sur la constitution sociale des standards normatifs à l'intérieur de nos pratiques et coutumes, une forme de naturalisme brut ou de réductionnisme social peut représenter un danger et si, de l'autre côté, nous soutenons que les standards normatifs sont objectivement corrects et indépendants de nos procédures de

ratification, nous nous heurtons au platonisme rampant. Par ailleurs, l'approche contextualiste de McDowell n'est pas à l'abri d'une certaine circularité explicative : la dimension normative de la signification, étant expliquée en termes d'autres notions contenant elles aussi une dimension normative, il semble que la normativité ne soit pas expliquée mais présupposée. La question sera alors d'évaluer si une telle position du juste milieu est possible.

Au terme de ces trois premiers chapitres, nous avons montré que trois dangers menacent d'entraver une approche sociale acceptable de la normativité de la signification : celui d'un réalisme platoniste ou d'un platonisme rampant, d'un naturalisme brut ou d'un réductionnisme social et d'une circularité explicative, correspondant aux trois conditions d'adéquation suivants : une condition d'adéquation anti-platoniste, non-réductionniste, non-relativiste et non-circulaire. Le second volume tente de savoir si une conception sociale de la signification peut répondre à ces trois conditions d'adéquation.

Davidson nous permet d'abord d'affiner ce que nous jugeons être une conception sociale adéquate des normes de signification. Le chapitre IV cherche donc à examiner le sens et la portée d'une normativité forte ou constitutive de la signification *via* la lecture du principe de charité préconisée par Davidson. Le principe de charité a bien, selon Davidson, un statut *a priori* ou transcendantal dans l'interprétation : il ne peut alors pas simplement être révoqué par l'expérience dans la mesure où il s'agit d'une condition de possibilité de toute activité susceptible d'interpréter les pensées, croyances, désirs et actions d'autrui. L'impératif ou le présupposé charitable visant à interpréter autrui, c'est-à-dire comme respectant (en grande partie) les lois logiques (version étroite) et comme croyant à la vérité de nos propositions empiriques (version large) revêt alors une autorité forte. Le principe de charité est donc un principe constitutif de l'attribution de croyances et de significations. Ainsi, Davidson, en mettant l'accent sur le rôle constitutif du principe de charité, adhère à l'idée d'une normativité forte de la signification; autrement dit, à l'idée que les phrases exprimant explicitement la signification d'expressions logiques et non logiques ne sont pas simplement que des hypothèses empiriques ou des alternatives pragmatiques. Il soutient également que les attributions de significations doivent obéir à des principes normatifs de rationalité n'ayant pas d'écho dans la théorie physique et conçoit donc l'activité linguistique comme une activité exclusivement rationnelle. Autrement dit, concevoir les agents comme obéissant à des

contraintes de rationalité ou comme entretenant des croyances, pour la plupart, vraies est une condition nécessaire du processus interprétatif. Cependant, considérer le principe de charité comme condition transcendantale de possibilité de l'expérience le conduit à admettre des propositions *a priori*, ce qui va à l'encontre de son engagement pour un holisme global. Or, un holisme global mène à un faillibilisme global c'est-à-dire à la thèse selon laquelle tout énoncé peut se révéler faux (rejet des vérités *a priori*), ce qui s'oppose à l'idée d'un statut *a priori* de la charité. Nous en venons alors à la conclusion que, pour rendre compte du sens et de la portée d'une théorie normative de la signification, une lecture plus faible de principe de charité de Davidson au sens d'un principe constitutif d'intelligibilité soit davantage appropriée. Il s'agit donc bien de maintenir la thèse d'une normativité constitutive de la signification mais de formuler de manière plus faible ou moins rigide la contrainte de rationalité afin de permettre la possibilité de tenir ou d'attribuer des croyances fausses ou irrationnelles. Par conséquent, au lieu d'être incompatible avec un faillibilisme global, l'hypothèse constitutive d'intelligibilité en est plutôt une expression : en d'autres termes, dire que le principe de charité au sens de principe d'intelligibilité est constitutif de l'interprétation revient à soutenir qu'aucune croyance particulière ou norme de signification n'est totalement à l'abri de la révision.

Nous reviendrons plus précisément dans un chapitre V, avant d'arriver dans les prochains chapitres à une formulation plus raffinée de la thèse d'une conception sociale des normes de signification, sur le sens, la portée et les conditions que Davidson entend donner à cette version plus faible d'une normativité constitutive de la signification. Il s'agit bien ici de défendre une approche sociale et faillibiliste des normes de signification : 1/ une approche sociale au sens où elle n'est pas exclusivement publique (il ne suffit pas que l'usage d'un terme soit simplement observable publiquement par les autres membres d'une communauté linguistique) et où la compréhension, comme activité sociale conjointe, s'insérant dans un jeu de conventions et de règles, implique que nous partageons, en tant que locuteurs-acteurs, au préalable certaines significations entre locuteurs (interprétés) et auditeurs (interprètes). 2/ une approche faillibiliste au sens où l'idée de normes constitutives ou objectivement correctes (et non absolument et universellement valides) est maintenue sans pour autant exclure la possibilité de soumettre celles-ci à des processus critiques de révision. Cette approche d'une conception sociale répond à notre premier critère d'adéquation : la réalité des normes. En

outre, dans la mesure où les normes sont instituées au sein de nos activités dialogiques, elle répond également au second critère d'adéquation, à savoir que l'origine de nos pratiques linguistiques ne peut être décrite que de l'intérieur de nos pratiques discursives c'est-à-dire qu'à partir du normatif. La normativité de la signification résulte donc bien d'une activité sociale mais qu'en est-il de l'objectivité ? Il faut encore satisfaire la troisième condition d'adéquation : une conception sociale doit pouvoir établir de manière objective (non relativiste) une application correcte des normes de signification. Les chapitres VI, VII et VIII se concentrent sur ce point.

Nous montrons dans un chapitre VI comment la conception sociale faible des normes de signification s'efforce de trouver un équilibre en maintenant un contextualisme – les standards ou normes ne peuvent survenir que dans un contexte de pratiques et de traditions d'application des concepts et d'évaluation de telles affirmations ou jugements résultant de l'application de ces concepts (engagements) – sans tomber dans le relativisme – et d'identifier la validité de ces mêmes standards avec les résultats obtenus dans les pratiques et traditions effectives. En effet, si le contextualisme cherche à garantir l'objectivité, soit le faillibilisme, alors il doit distinguer entre ce qu'une pratique donnée considère comme « correct » ou comme « vrai » et ce qui est correct indépendamment de cette pratique en question, c'est-à-dire qu'il ne doit pas identifier ce qui est correct avec ce qu'une pratique donnée, en l'occurrence celle de la communauté linguistique, s'accorde à considérer comme correct. Le faillibilisme soutient l'idée que tout contexte donné rattaché à l'accord ou à l'évaluation d'une pratique peut s'avérer erroné, dans la mesure où il peut être corrigé par un autre accord contextualisé, qui peut lui aussi être à son tour erroné et sujet à révision et à correction par un autre accord contextualisé et ainsi de suite. Le faillibilisme implique donc que ces différents points de vue doivent être soumis à une évaluation comparative les situant les uns par rapport aux autres. C'est ce que la structure normative d'une faible socialité de l'objectivité de Brandom tente d'atteindre : l'application correcte et objective des normes de signification n'est pas déterminée par la manière dont sont les choses indépendamment des attributions et évaluations de pointage déontique, mais résulte des différences entre ces deux perspectives sociales que sont le fait de reconnaître (et donc d'assumer) soi-même un engagement et le fait d'attribuer un engagement à autrui. La manière dont les choses sont est donc une fonction des différences établies sur le mode dialogique du « je-tu » dans les

évaluations de nos croyances (engagements et autorisations) sur le monde. Aucune perspective ne prime arbitrairement sur une autre et n'est donc en soi privilégiée. Brandom cherche ainsi à fonder une notion authentique (faillibiliste) d'objectivité – une approche de la vérité et de l'application correcte qui satisfasse la condition de l'erreur communautaire (ou de l'erreur à la première personne) en ne permettant pas d'identifier ce qui est vrai avec ce que tout un chacun (ou ce que tout le monde) considère comme vrai ou comme correct. Son approche de la vérité et des normes de signification en termes de conditions de correction est ainsi contextualiste : elle tente d'envisager la possibilité d'erreurs communautaires (ou d'erreurs à la première personne) dans nos jugements respectifs ou dans la manière dont nous appliquons les concepts en adoptant un autre point de vue dépendamment du contexte (ou point de vue du marqueur) à partir duquel ces jugements ou ces applications-ci doivent être évalués. Il n'existe donc pas de pointage dont les standards ou critères d'évaluation peuvent être fixés sans tenir compte du contexte. Cette forme de contextualisme n'exclut pas la possibilité de rejeter ou de corriger un jugement émis ou l'application d'un concept qui a communément été ratifié par la communauté, mais seulement la possibilité d'agir ainsi en l'absence de toute perspective de pointage. Brandom propose également d'affiner son cadre explicatif de travail explicitement contextualiste et social en y ajoutant une dimension historique : l'application correcte et objective d'un concept devra donc être définie comme une fonction des différences perspectiveles résultant de la relation synchronique « je-tu » et des différences perspectiveles résultant de la relation diachronique « je-tu ». Brandom considère ainsi l'application correcte et objective d'un terme comme étant le produit d'un processus de « négociation » entre les « autorités » concurrentes des attitudes d'applications passées, actuelles et futures et d'évaluation de ces applications par les agents linguistiques et les joueurs qui marquent le score. Les relations constitutives du contenu des normes conceptuelles ne sont donc pas seulement sociales et inférentielles mais bien inférentielles, sociales, et historiques.

Cependant, même si cette version inférentielle, sociale et historique semble présenter, avec cohérence, une structure d'objectivité inhérente aux normes de signification, nous nous demandons si celle-ci peut vraiment produire une notion de correction plus forte que celle de « fidélité à la tradition ».

Nous en venons donc dans les deux derniers chapitres à aborder une autre approche contextualiste, sociale et historique, à savoir celle proposée par Habermas et McCarthy, celle-ci ajoutant la notion kantienne d'idées régulatrices. Nous suggérons ici une conception sociale des normes de signification, qui conserverait la trame de la pratique discursive décrite par Brandom, tout en y ajoutant la notion d'idée régulatrice, à savoir : l'idée de sujets rationnellement responsables (au sein de ce jeu de pointage déontique), l'idée d'un monde objectif et l'idée d'une vérité objective (qui se traduisent de manière intersubjective en termes d'un « accord rationnel universel »). Par conséquent, au lieu d'invoquer l'idée d'un monde objectif indépendant de toutes pratiques linguistiques ou contextes d'usage qui nous donneraient accès au monde tel qu'il est comme l'externalisme sémantique de Brandom le suggère, il semble plus approprié de le concevoir comme une idée régulatrice au sens de Kant, c.-à-d., une notion qui, bien qu'inaccessible dans toute pratique réelle de discussion et contexte social de négociation, etc., règle néanmoins la manière dont de telles pratiques sont menées, en jouant le rôle d'un foyer imaginaire vers lequel nos pratiques discursives sont orientées. Par ailleurs, si comme pour Brandom, parler un langage est bien pour Habermas un jeu social de production et de demande de raisons, il pourrait être plus adéquat de tenir compte dans sa théorie de l'agir communicationnel de la dimension temporelle et séquentielle du déroulement du processus discursif décrit par Brandom, celui-ci débutant d'abord à partir des contributions individuelles, impliquant ensuite une pratique conjointe coopérative mais également compétitive et conflictuelle mais se dirigeant, par coups successifs et au moyen d'un pointage déontique, vers une évaluation des différentes perspectives jusqu'à aboutir à la coordination de nos perspectives sociales selon des pratiques communes. Ce qui est partagé par locuteurs et auditeurs est donc la pratique du pointage. Depuis chaque point de vue doxastique sur un acte de discours, il peut en effet y avoir un contenu commun à celui qui contracte un engagement et aux marqueurs qui l'attribuent, et cependant il se peut que ce que l'on considère comme partagé soit différent selon les différents points de vue des marqueurs. Il est donc nécessaire d'adhérer à la conception de la pratique sociale sur le mode du « je-tu », telle que Brandom la décrit, afin de pouvoir distinguer entre ce qui est considéré comme correct et ce qui est correct. Aucune perspective n'est alors privilégiée, à la différence de la pratique sociale sur le mode du « je-nous », telle que Habermas semble la décrire, qui semble bien préconiser une perspective, celle du « nous » ou l'obligation (au sens fort d'un impératif)

d'un accord communautaire. L'idée (régulatrice) d'un accord intersubjectif rationnellement motivé ne nous semble alors pas incompatible avec cette idée que les contenus inférentiels sont essentiellement perspectivistes.

Nous pensons donc que cette approche sociale répond au critère d'adéquation *a priori* pour fonder une approche objective des normes de signification : non-platoniste, non réductionniste, non relativiste et explicative (jusqu'à un certain point). Il s'agit bien d'une conception *anti-platoniste* des normes de signification dans la mesure où elle situe la normativité là où elle devrait être localisée – dans l'évolution actuelle, sociale et historique des usages linguistiques. Les normes constitutives ou objectivement correctes de la signification émergent donc de nos pratiques d'usage au sein de nos contextes sociaux de discussion, de critique, etc. Elles assurent la fonction d'« idée régulatrice » : bien qu'on ne puisse actuellement garantir sa réalisation effective soit la manière adéquate d'employer un terme, tendre vers cet idéal introduit des pratiques fécondes. Cette approche *non relativiste* (faillibiliste) de la normativité est ainsi alimentée par une idée régulatrice orientant l'articulation réelle de nos pratiques sociales. Elle est aussi *non réductionniste* dans la mesure où elle ne réduit pas la normativité à de simples régularités *de facto* et elle a une fonction explicative en ce qu'elle dévoile la manière dont nos pratiques sociales doivent être structurées afin de fonder une dimension normative, à savoir : nos pratiques linguistiques sociales doivent être guidées par cette « idée régulatrice » dévoilant la manière correcte d'appliquer un terme.

## BIBLIOGRAPHIE

- Allen, Barry. 1989. « Gruesome Arithmetic: Kripke's Sceptic Replies ». *Dialogue*, vol. 28, no 2, p. 257-264.
- Aristote. 1959. *Éthique à Nicomaque*. Paris: Vrin.
- , 1989. *La décision du sens. Le livre Gamma de la Métaphysique*. Paris: Vrin.
- , 2002a. *Politique, livre I et II*. Paris: Les belles lettres.
- , 2002b. *Les parties des animaux*. Paris: Les belles lettres.
- Austin, J. L. 1994. *Écrits philosophiques*. Paris: Seuil.
- Ayer, A. J. 1946. *Language, Truth, and Logic*. Londres: Gollancz.
- Baillargeon, Normand. 2005. *Petit cours d'autodéfense intellectuelle*. Montréal: Lux.
- Baker, G. P., et P. M. S. Hacker. 1984a. « On Misunderstanding Wittgenstein: Kripke's Private Language Argument ». *Synthese*, vol. 58, no 3, p. 407-450.
- , 1984b. *Scepticism, Rules and Language*. Oxford: Blackwell.
- Barber, Michael D. 2004. « A Moment of Unconditional Validity ? Schutz and the Habermas/Rorty Debate ». *Human Studies*, vol. 27, no 1, p. 51-67.
- Bennett, Jonathan. 1985. « Critical Notice ». *Mind*, vol. 94, no 376, p. 601-626.
- Berkeley, George. 1993. *Les principes de la connaissance humaine*. Paris: Garnier Flammarion.
- Bilgrami, Akeel. 1995. *Belief and Meaning: The Unity and Locality of Mental Content*. Oxford: Blackwell.
- Blackburn, Simon. 1984a. « The Individual Strikes Back ». *Synthese*, vol. 58, no 3, p. 281-301.
- , 1984b. *Spreading the Word [Part I]: Groundings in the Philosophy of Language*. Oxford: Clarendon Press.
- , 1993. *Essays on Quasi-realism*. Oxford: Oxford University Press.

- , 2000. *Ruling Passions*. Oxford: Oxford University Press.
- Boghossian, Paul A. 1989. « The Rule-Following Considerations ». *Mind*, vol. 98, no 392, p. 507-549.
- , 1990. « The Status of Content ». *The Philosophical Review*, vol. 99, no 2, p. 157-184.
- Bouveresse, Jacques. 1987. *La force de la règle*. Paris: Minuit.
- , 1992. « Sur quelques conséquences indésirables du pragmatisme ». In *Lire Rorty*. Combas: L'Éclat.
- , 1995. « Règles, dispositions et habitus ». *Critique*, p. 573-594.
- , 1996. *La demande philosophique: que veut la philosophie et que peut-on vouloir d'elle ?* Paris: L'Éclat.
- , 2005. « On the meaning of the word 'Platonism' in the expression 'mathematical platonism' ». *Proceedings of the Aristotelian Society (Hardback)*, vol. 105, no 1, p. 55-79.
- Brandom, Robert. 1979. « Freedom and Constraint by Norms ». *American Philosophical Quarterly*, vol. 16, no 3, p. 187-196.
- , 1983. « Asserting ». *Noûs*, vol. 17, no 4, p. 637-650.
- , 1997a. « Replies ». *Philosophy and Phenomenological Research*, vol. 57, no 1, p. 189-204.
- , 1997b. « Review: Précis of Making It Explicit ». *Philosophy and Phenomenological Research*, vol. 57, no 1, p. 153-156.
- , 2000a. « Facts, Norms, and Normative Facts: A reply to Habermas ». *European Journal of Philosophy*, vol. 8, no 3, p. 356-374.
- , 2000b. « Quelques thèmes pragmatistes dans l'idéalisme de Hegel: Négociation et administration dans la conception hégélienne de la structure et du contenu des normes conceptuelles ». *Philosophiques*, vol. 27, no 2, p. 231-261.
- , 2002. *Tales of the Mighty Dead. Historical Essays in the Metaphysics of Intentionality*. Cambridge, Mass.: Harvard University Press.
- , 2009. *L'articulation des raisons: une introduction à l'inférentialisme*. Paris: Cerf.
- , 2011. *Rendre explicite: raisonnement, représentation et engagement discursif (I, II)*. Paris: Cerf.

- Burge, Tyler. 1979. « Individualism and the mental ». *Midwest Studies in Philosophy*, vol. 4, p. 73-122.
- 1992. « Philosophy of Language and Mind: 1950-1990 ». *The Philosophical Review*, vol. 101, no 1, p. 3-51.
- Byrne, Alex. 1996. « On misinterpreting Kripke's Wittgenstein ». *Philosophy and Phenomenological Research*, vol. 56, no 2, p. 339-343.
- Carnap, Rudolf. 1932. « The Elimination of Metaphysics Through Logical Analysis of Language ». *Erkenntnis*, vol. II, p. 60-81.
- 1939. *Foundations of Logic and Mathematics*. Chicago: The University of Chicago Press.
- 1963. « Intellectual Autobiography ». In *The Philosophy of Rudolf Carnap*, P. A. Schilpp (éd.). La Salle, Illinois: Open Court (The Library of Living Philosophers, 11).
- 1985. « Le dépassement de la métaphysique par l'analyse du langage ». In *Manifeste du Cercle de Vienne et autres écrits*, A. Soulez (éd.). Paris: Presses universitaires de France.
- Cavell, Stanley. 1962. « The Availability of Wittgenstein's Later Philosophy ». *The Philosophical Review*, vol. 71, no 1, p. 67-93.
- 1996. *Les voix de la raison. Wittgenstein, le scepticisme, la moralité et la tragédie*. Paris: Seuil.
- 2009. *Dire et vouloir dire*. Paris: Cerf.
- Chastain, Charles. 1975. « Reference and Context ». In *Minnesota Studies in the Philosophy of Science Vol. VII*, K. Gunderson (éd.). Minneapolis: University of Minnesota Press.
- Chauviré, Christiane. 2004. *Le moment anthropologique de Wittgenstein*. Paris: Kimé.
- Chomsky, Noam. 1986. *Knowledge of Language: its nature, origin and use*. New York: Praeger.
- Churchland, Paul M. 1981. « Eliminative Materialism and the Propositional Attitudes ». *The Journal of Philosophy*, vol. 78, no 2, p. 67-90.
- Corblin, Francis. 1985. « Remarques sur la notion d'anaphore ». *Revue québécoise de linguistique*, vol. 15, no 1, p. 173-195
- Davidson, Donald. 1974. « Replies to David Lewis and W. V. Quine ». *Synthese*, vol. 27, no 3/4, p. 345-349.

- , 1980. « Toward a unified theory of meaning and action ». *Grazer Philosophische Studien*, vol. 11, p. 1-12.
- , 1985. « A new basis for decision theory ». *Theory and Decision*, vol. 18, no 1, p. 87-98.
- , 1989. « The Conditions of Thought ». In *The Mind of Donald Davidson*, J. Brandl et W. L. Gombocz (éd.). Amsterdam: Rodopi.
- , 1990a. « The Structure and Content of Truth ». *Journal of Philosophy*, vol. 87, no 6, p. 279-328.
- , 1990b. « Representation and Interpretation ». In *Modelling the Mind*, K. A. Mohyeldin et al. (éd.). Oxford: Clarendon Press.
- , 1991a. « Les animaux rationnels ». In *Paradoxes de l'irrationalité*. Combas: L'Éclat.
- , 1991b. « Qu'est-ce qui est présent à l'esprit ? ». In *Lieux et transformations de la philosophie, Les Cahiers de Paris VIII*. Paris: Presses Universitaires de Vincennes.
- , 1992a. « Thinking causes ». In *Mental Causation*, J. Heil et A. R. Mele (éd.). Oxford: Oxford University Press.
- , 1992b. « En quel sens un langage est-il social ? ». In *Wittgenstein et la philosophie d'aujourd'hui*, A. Soulez et J. Sébestik (dir.). Paris: Klincksieck.
- , 1993a. *Actions et événements*. Paris: PUF.
- , 1993b. *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1967. « Vérité et signification ». In *Enquête sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1968. « Dire que ». In *Enquête sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1970. « Sémantique pour les langues naturelles ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1973a. « L'interprétation radicale ». In *Enquête sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1973b. « Pour défendre la Convention T ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.

- , 1993b, 1974a. « La croyance et le fondement de la signification ». In *Enquêtes sur la vérité et la signification*. Nîmes: J. Chambon.
- , 1993b, 1974b. « Sur l'idée même de schème conceptuel ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1975. « Pensée et discours ». In *Enquêtes sur la vérité et la signification*. Nîmes: J. Chambon.
- , 1993b, 1976. « Réponse à Foster ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1977a. « La méthode de la vérité en métaphysique ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1977b. « La réalité sans la référence ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1979a. « Modes et actes de langage ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1979b. « La citation ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1993b, 1979c. « L'inscrutabilité de la référence ». In *Enquêtes sur la vérité et l'interprétation*. Nîmes: J. Chambon.
- , 1996. « The folly of trying to define truth ». *The Journal of Philosophy*, vol. 93, no 6, p. 263-278.
- , 2001a. « Externalisms ». In *Interpreting Davidson*, P. Kotatko, P. Pagin et G. Segal (éd.). Stanford: CSLI.
- , 2001b, 1983. « A Coherence Theory of Truth and Knowledge ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2001b, 1984. « The First Person Authority ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2001b, 1987. « Knowing One's Own Mind ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2001b, 1988. « The Myth of the Subjective ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2001b, 1990. « Epistemology Externalized ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.

- , 2001b, 1991. « Three Varieties of Knowledge ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2001b, 1992. « The Second Person ». In *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2003. « Indéterminisme et anti-réalisme ». In *Cent ans de philosophie américaine*, J.-P. Cometti et Cl. Tiercelin (dir. publ.). Pau: Presses de l'Université de Pau.
- , 2004, 1995. « The Problem of Objectivity ». In *Problems of Rationality*. Oxford: Oxford University Press.
- , 2004, 2001. « What Thought Requires ». In *Problems of Rationality*. Oxford: Oxford University Press.
- , 2005, 1986. « A Nice Derangement of Epitaphs ». In *Truth, Language and History*. Oxford: Clarendon Press.
- , 2005, 1993. « Method and Metaphysics ». In *Truth, Language and History*. Oxford: Clarendon Press.
- , 2005, 1994. « The Social Aspect of Language ». In *Truth, Language and History*. Oxford: Clarendon Press.
- Davies, David. 2009. « L' 'empirisme transcendantal' de McDowell ». *Philosophiques*, vol. 36, no 1, p. 195-204.
- Dennett, Daniel. 1971. « Intentional systems ». *Journal of Philosophy*, vol. 68, no 4, p. 87-106.
- , 1987. *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Diamond, Cora. 1989. « Rules: Looking in the Right Place ». In *Wittgenstein: Attention to Particulars*, D. Z. Philipps et P. Winch (éds.). New York: St Martin Press.
- Domenach, Elise. 2003/2. « Naturalisme et scepticisme chez Cavell, McDowell et Strawson ». *Revue de Métaphysique Et de Morale*, vol. 38, p. 159-179.
- Duhem, Pierre. 1914. *La théorie physique, son objet, sa structure*. Paris: Chevalier & Rivière.
- Dummett, Michael. 1973. « Original Sinn ». In *Frege: Philosophy of Language*. Londres: Duckworth.
- , 1978. *Truth and Other Enigmas*. Londres: Duckworth.
- , 1982. *The Interpretation of Frege's Philosophy*. Londres: Duckworth.

- , 1986. « A Nice Derangement of Epitaphs: Some comments on Davidson and Hacking ». In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, E. Lepore (éd.). Oxford: Blackwell.
- , 1991. *The Logical Basis of Metaphysics*. Cambridge, Mass.: Harvard University Press.
- , 1993. « What is a theory of meaning ?(1) ». In *The Seas of Language*. Oxford: Clarendon Press, p.1-13.
- Engel, Pascal. 1981. « Davidson on perspective ». *Critique*, vol. 37, no 409-410, p. 578-594.
- , 1989. « Interprétation, charité et mentalité prélogique ». *Revue philosophique*, vol. 4, p. 543-558.
- , 1994. *Davidson et la philosophie du langage*. Paris: PUF.
- Esfeld, Michael. 2001. *Holism in Philosophy of Mind and Philosophy of Physics*. Dordrecht: Kluwer Academic Publishers.
- Evans, Gareth, et John McDowell. 1976. *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press.
- Evnine, Simon. 1991. *Donald Davidson*. Oxford: Polity Press.
- Falk, Barrie. 1994. « Doing What One Meant to Do ». *Synthese*, vol. 98, no 3, p. 379-399.
- Fennell, John. 2000. « Davidson on meaning normativity: Public or social ». *European Journal of Philosophy*, vol. 8, no 2, p. 139-154.
- Feyerabend, Paul. 1987. *Adieu la raison*. Paris: Seuil.
- Field, Hartry. 1989. *Realism, Mathematics and Modality*. Oxford: Blackwell.
- Fine, Arthur. 1984. « And Not Anti-Realism Either ». *Noûs*, vol. 18, no 1, p. 51-65.
- Fodor, Jerry A. 1990. *A Theory of Content and Other Essays*. Cambridge, Mass.: MIT Press.
- Fodor, Jerry A., et Ernest Lepore. 1992. *Holism: A Shopper's Guide*. Oxford: Blackwell.
- Follesdal, Dagfinn. 1975. « Meaning and Experience ». In *Mind and Language. Wolfson Lectures 1974*, S. Guttenplan (éd.). Oxford: Clarendon Press.
- , 1995. « In What sense is Language Public ? ». In *On Quine*, P. Leonardi et M. Santambrogio (éd.). Cambridge: Cambridge University Press.

- Forbes, Graeme. 1983. « Scepticism and Semantic Knowledge ». *Proceedings of the Aristotelian Society*, vol. 84, p. 223-237.
- Foster, J. A. 1976. « Meaning and Truth theory ». In *Truth Meaning and Essays in Semantics*, G. Evans et J. McDowell. Oxford: Clarendon Press, p.1-32.
- Frege, Gottlob. 1971. « Sens et dénotation ». In *Écrits logiques et philosophiques*. Paris: Seuil.
- Friedman, Michael. 1997. « Philosophical Naturalism ». *Proceedings and Addresses of the American Philosophical Association*, vol. 71, no 2, p. 5 - 21.
- , 2002. « Kant, Kuhn, and the rationality of science ». *Philosophy of Science*, vol. 69, no 2, p. 171-190.
- Gadamer, H. G. 1996. *Vérité et méthode*. Paris: Seuil.
- Gibbard, Allan. 1996. « Review Essays: Thought, Norms, and Discursive Practice: Commentary on Robert Brandom, Making It Explicit ». *Philosophy and Phenomenological Research*, vol. 56, no 3, p. 699-717.
- Gibson, Roger. 1988. *Enlightened Empiricism*. Gainesville: University Press of Florida.
- Ginsborg, Hannah. 2006. « Reasons for Belief ». *Philosophy and Phenomenological Research*, vol. 72, no 2, p. 286-318.
- Glüer, Kathrin. 2006. « Triangulation ». In *The Oxford Handbook of Philosophy of Language*, E. Lepore et B. Smith (éd.). Oxford: Oxford University Press.
- Gochet, Paul. 1978. *Quine en perspective: essai de philosophie comparée*. Paris: Flammarion.
- Goldfarb, Warren. 1985. « Kripke on Wittgenstein on Rules ». *The Journal of Philosophy*, vol. 82, no 9, p. 47-488.
- Goodman, Nelson. 1984. *Faits, fictions et predictions*. Paris: Minuit.
- Grandy, Richard. 1973. « Reference, Meaning, and Belief ». *The Journal of Philosophy*, vol. 70, no 14, p. 439-452.
- Gupta, A. 1993. « A Critique of Deflationism ». *Philosophical Topics*, vol. 21, no 1, p. 57-81.
- Habermas, Jürgen. 1978. *Raison et légitimité: problèmes de légitimation dans le capitalisme avancé*. Paris: Payot.

- , 1979. « What is Universal Pragmatics ? ». In *Communication and the Evolution of Society*. Boston: Beacon Press.
- , 1987a. *Logique des sciences sociales*. Paris: PUF.
- , 1987b. *Théorie de l'agir communicationnel, Vol.1*. Paris: Fayard.
- , 1988. *Le discours philosophique de la modernité*. Paris: Gallimard.
- , 1992a. *De l'éthique de la discussion*. Paris: Flammarion.
- , 1992b. « Towards a critique of the theory of meaning ». In *Postmetaphysical Thinking*. Cambridge: Mass: MIT Press.
- , 1993. *La pensée postmétaphysique*. Paris: Armand Colin.
- , 1995. *Sociologie et théorie du langage*. Paris: Armand Colin.
- , 1997. *Droit et démocratie: entre faits et normes*. Paris: Flammarion.
- , 1998. *On the Pragmatics of Communication*. Cambridge, Mass.: MIT Press.
- , 1999. *Morale et Communication*. Paris: Champs-Flammarion.
- , 2000. « From Kant to Hegel: On Robert Brandom's Pragmatic Philosophy of Language ». *European Journal of Philosophy*, vol. 8, no 3, p. 322-355.
- , 2001. *Vérité et justification*. Paris: Gallimard.
- , 2003. *L'éthique de la discussion et la question de la vérité*. Paris: Grasset.
- Hacking, Ian. 1986. « The Parody of Conversation ». In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, E. Lepore (éd.). Oxford: Blackwell.
- , 1993. « On Kripke's and Goodman's Uses of 'Grue' ». *Philosophy*, vol. 68, no 265, p. 269-295.
- Hare, R.M. 1952. *The Language of Morals*. Oxford: Oxford University Press.
- Haugeland, John. 1998. *Having Thought: Essays in the Metaphysics of Mind*. Cambridge: Harvard University Press.
- Heal, Jane. 1989. *Fact and meaning : Quine and Wittgenstein on philosophy of language*. Oxford: Basil Blackwell.

- Heath, Joseph. 1998. « What is a Validity Claim ? ». *Philosophy and Social Criticism*, vol. 24, no 4, p. 23-41.
- , 2001. « Bandom et les sources de la normativité ». *Philosophiques*, vol. 28, no 1, p. 27-46.
- Hegel, G. W. F. 1941. *Phénoménologie de l'Esprit*. Paris: Aubier-Montaigne.
- , 2003. *La raison dans l'histoire*. Paris: 10/18.
- Horwich, Paul. 1990a. *Truth*. Oxford: Basil Blackwell.
- , 1990b. « Wittgenstein and Kripke on the Nature of Meaning ». *Mind & Language*, vol. 5, no 2, p. 105-121.
- , 1995. « Meaning, Use and Truth: On Whether a Use-Theory of Meaning Is Precluded by the Requirement that Whatever Constitutes the Meaning of a Predicate Be Capable of Determining the Set of Things of Which the Predicate is True and to Which It Ought to be Applied ». *Mind*, vol. 104, no 414, p. 355-368.
- , 1998. *Meaning*. Oxford: Oxford University Press.
- , 2005. *Reflections on Meaning*: Oxford University Press.
- Hoy, D. C., et T. McCarthy. 1994. *Critical Theory*. Oxford: Blackwell.
- Hume, David. 1995. *Traité de la nature humaine - Livre I et appendice*. Paris: Garnier Flammarion.
- , 2001. *Traité de la nature humaine, Livre III*. Paris: Garnier Flammarion.
- Johnston, Mark. 1993. « Objectivity Refigured: Pragmatism without Verificationism ». In *Reality, Representation, and Projection*, J. Haldane et W. Crispin (éd.). New York: Oxford University Press.
- Kant, Emmanuel. 1944. *Critique de la raison pure*. Paris: PUF.
- , 1974. *Prolégomènes à toute métaphysique future*. Paris: Vrin.
- Kenny, A. J. P. 1963. *Action, Emotion And Will*. London: Routledge & Kegan Paul.
- Kibble, Rodger. 2006. « Speech acts, commitment and multi-agent communication ». *Computational & Mathematical Organization Theory*, vol. 12, no 2, p. 127-145.
- Kremer, Michael. 2000. « Wilson on Kripke's Wittgenstein ». *Philosophy and Phenomenological Research*, vol. 60, no 3, p. 571-584.

- Kripke, Saul. 1982. *La logique des noms propres*. Paris: Minuit.
- , 1996. *Règles et langage privé*. Paris: Seuil.
- Kuhn, T. S. 1983. *La structure des révolutions scientifiques*. Paris: Flammarion.
- , 1990. *La tension essentielle: tradition et changement dans les sciences*. Paris: Gallimard.
- Ladrière, Jean. 1967. « Les limites de la formalisation ». In *Encyclopédie de la pleïade*. Paris: Gallimard.
- Lance, M. N., et J. O'Leary-Hawthorne. 1997. *The Grammar of Meaning*. Cambridge: Cambridge University Press.
- Laugier, Sandra. 2001/3. « Où se trouvent les règles ? ». *Archives de philosophie*, vol. 64, p. 505-524.
- Laurier, Daniel. 1983. « Tarski, Davidson et la signification ». *Dialogue*, vol. 22, no 4, p. 595-620.
- , 1994. « Holismes ». In *Lire Davidson, interprétation et holisme*, P. Engel (dir. publ.). Combas: L'Éclat.
- , 2002. *L'esprit et la nature*. Montréal: Les Presses de l'Université de Montréal.
- , 2004. « La publicité et l'interdépendance du langage et de la pensée ». *Dialogue*, vol. 43, no 2, p. 281-316.
- Lesniewski, Stanislaw. 1929. « Grundzüge eines neuen Systems der Grundlagen der Mathematik ». *Fundamenta Mathematica*, vol. 14, p. 1-81.
- Levin, Michael. 2007. « Bundling Hume with Kripkenstein ». *Synthese*, vol. 155, p. 35-64.
- Lewis, David. 1974. « Radical Interpretation ». *Synthese*, vol. 27, no 3/4, p. 331-344.
- , 1979. « Scorekeeping in a language game ». *Journal of Philosophical Logic*, vol. 8, no 1, p. 339-359.
- Lorenz, Kuno. 1981. « Dialogical Logic ». In *Dictionary of Logic as Applied in the Study of Language*, W. Marciszewski (ed.). The Hague: Martinus Nijhoff.
- Lorenzen, P., et K. Lorenz. 1978. *Dialogische Logik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Lukes, Steven. 1982. « Relativism in Its Place ». In *Rationality and Relativism*, Martin Hollis et Steven Lukes (éd.). Cambridge Mass.: MIT Press, p. 261-305.

- MacFarlane, John. 2010. « Pragmatism and Inferentialism ». In *Reading Brandom: On Making It Explicit*, B. Weiss et J. Wanderer (éd.). London: Routledge.
- 2011. « What is assertion? ». In *Assertion*, J. Brown et H. Cappelen (éd.). Oxford: Oxford university Press.
- Mackie, John L. 1977. *Inventing Right and Wrong*. Harmondsworth: Penguin.
- Malcolm, Norman. 1966. *Ludwig Wittgenstein: A Memoir*. Oxford: Oxford University Press.
- Marion, Mathieu. 2012. « Game Semantics and The Manifestation Thesis ». In *The Realism-Antirealism Debate in the Age of Alternative Logics*, S. Rahman, G. Primiero et M. Marion (éd.). Netherlands: Springer
- Martin, C. B., et J. Heil. 1998. « Rules and Powers ». *Noûs*, vol. 32, p. 283-312.
- McCarthy, Thomas. 1989. « Contra Relativism: A Thought Experiment ». *Zeitschrift für philosophische Forschung*, vol. 43, no 2, p. 318-330.
- 1991. *Ideals and Illusions : On Reconstruction and Deconstruction in Contemporary Critical Theory*. Cambridge, Mass: MIT Press.
- 1992a. « Ironie privée et décence publique ». In *Lire Rorty. Le pragmatisme et ses conséquences*. Paris: L'Éclat.
- 1992b. « Une réponse à la réplique de Rorty: la théorie ironiste comme vocation ». In *Lire Rorty. Le pragmatisme et ses conséquences*. Paris: Édition de l'Éclat.
- 1996. « Philosophy and critical theory: A reply to Richard Rorty and Seyla Benhabib ». *Constellations*, vol. 3, no 1, p. 95-103.
- McDowell, John. 1981. « Anti-Realism and the Epistemology of Understanding ». In *Meaning and Understanding*, H. Parrett et J. Bouveresse (éd.). Berlin: de Gruyter, p. 225-248.
- 1984. « Wittgenstein on following a Rule ». *Synthese*, vol. 58, no 3, p. 325-363.
- 1987. « In Defense of modesty ». In *Michael Dummett, Contributions to Philosophy*, B. Taylor (éd.). Dordrecht: Martinus Nijhoff.
- 1992. « Meaning and Intentionality in Wittgenstein's Later Philosophy ». *Midwest Studies in Philosophy*, vol. 17, no 1, p. 40-52.
- 1997. « Brandom on Representation and Inference ». *Philosophy and Phenomenological Research*, vol. 57, no 1, p. 157-162.

- , 1998. *Mind Value and Reality*. Cambridge: Harvard University Press.
- , 1999. « Experiencing the World ». In *Reason and Nature. Lecture and Colloquium in Münster*, M. Willaschek (éd.). Münster: LIT-Verlag.
- , 2000. « Responses ». In *Reason and Nature. Lecture and Colloquium in Münster*, M. Willaschek (éd.). Münster: LIT-Verlag.
- , 2001/3. « Non-cognitivism et règles ». *Archives de philosophie*, vol. 64, p. 457-477.
- , 2002. « Responses ». In *Reading McDowell. On Mind and World*, N. H. Smith (éd.). Londres et New York: Routledge.
- , 2003. « Subjective, Intersubjective, Objective ». *Philosophy and Phenomenological Research*, vol. 67, no 3, p. 675-681.
- , 2007. *L'esprit et le monde*. Paris: Vrin.
- McGinn, Colin. 1983. *The Subjective View: Secondary Qualities and Indexical Thoughts*. Oxford: Oxford University Press.
- , 1984. *Wittgenstein on Meaning*. Oxford: Basil Blackwell.
- Mead, G. H. 1963. *L'esprit, le soi et la société*. Paris: PUF.
- Miller, Alexander. 2007. *Philosophy of Language*, Seconde Édition. London & New York: Routledge.
- Moreau, Joseph. 1939. *La construction de l'idéalisme platonicien*. Paris: Boivin et Cie.
- Mulligan, K., P. Simons et B. Smith. 1984. « Truth makers ». *Philosophy and Phenomenological Research*, vol. 44, no 1, p. 287-321.
- Mumford, Stephen. 1998. *Dispositions*. Oxford: Oxford University Press.
- Nagel, Thomas. 1993. *Du point de vue de nulle part*. Combas: L'Éclat.
- O'Shea, James R. 2007. *Wilfrid Sellars: Naturalism with a Normative Turn*. Cambridge: Polity Press.
- Pataut, Fabrice. 1991. « Quelle logique une sémantique anti-réaliste peut-elle espérer justifier ? ». *L'âge de la science: lectures philosophiques*, vol. 4, p. 121-151.
- Piaget, Jean. 1970. *L'épistémologie génétique*. Paris: PUF.
- Platon. 1998. *Euthyphron, Lachès*. Paris: Garnier-Flammarion.

- Platts, Mark. 1980. *Reference, Truth and Reality: Essays on the Philosophy of Language*. Londres: Routledge & Kegan Paul.
- Popper, Karl. 1998. « Sur la théorie de l'esprit objectif ». In *La connaissance objective*. Paris: Champs-Flammarion.
- Putnam, Hilary. 1974. « The Refutation of Conventionalism ». *Noûs*, vol. 8, no 1, p. 25-40.
- , 1975. « The Meaning of 'Meaning' ». In *Mind, Language and Reality: Philosophical Papers, vol. 2*. Cambridge: Cambridge University Press, p. 215-272.
- , 1983. *Realism and Reason, Philosophical papers, vol.3*. Cambridge: Cambridge University Press.
- , 1984. *Raison, vérité et histoire*. Paris: Minuit.
- , 1985. « Why reason can't be naturalized ». In *Realism and Reason*. Cambridge: Cambridge University Press, p.3-24.
- , 1990. *Représentation et réalité*. Paris: Gallimard.
- , 1994. « On truth ». In *Words and Life*, J. Conant (éd.): Harvard University Press, p. 315-329.
- Quine, W. V. O. 1962. « Le mythe de la signification ». In *La philosophie analytique*. Paris: Éditions de Minuit.
- , 1965. *Mathematical Logic*. New York: Harper Torchbooks.
- , 1969. « Reply to Strawson ». In *Words and Objections, essays on the work of W. V. O. Quine*, D. Davidson et J. Hintikka (éd.). Dordrecht, Boston: Reidel Pub. Co.
- , 1970. « Grades of Theoreticity ». In *Experience and Theory*, L. Foster et J. W Swanson (dir.). Cambridge (Mass.): University of Massachussets.
- , 1973. *The Roots of Reference*. La salle, Illinois: Open Court.
- , 1975. *La philosophie de la logique*. Paris: Aubier.
- , 1976a. « Carnap and Logical Truth ». In *The Ways of Paradox and Other Essays*. Cambridge, Mass.: Harvard University Press.
- , 1976b. « Truth by Convention ». In *The Ways of Paradox and Other Essays*. Cambridge, Mass.: Harvard University Press.
- , 1976c. « Necessary Truth ». In *The Ways of Paradox and Other Essays*. Cambridge, Mass.: Harvard University Press.

- , 1977a. *Le mot et la chose*. Paris: Champs-Flammarion.
- , 1977b. *Relativité de l'ontologie et autres essais*. Paris: Aubier Montaigne.
- , 1981. *Theories and Things*. Cambridge (Mass.): Harvard University Press.
- , 1986. « Reply to Hellman ». In *The Philosophy of W.V. Quine*, P. A. Schilpp et L. E. Hahn (éd.). La Salle, Illinois: Open Court.
- , 1992. *Quiddités. Dictionnaire philosophique par intermittence*. Paris: Seuil.
- , 1992a. « Vérité ». In *Quiddités: dictionnaire philosophique par intermittence*. Paris: Seuil.
- , 1993. *La poursuite de la vérité*. Paris: Seuil.
- , 1995. *From Stimulus to Science*. Cambridge (Mass.): Harvard University Press.
- , 2003. « Deux dogmes de l'empirisme ». In *Du point de vue logique*, S. Laugier (éd.). Paris: Vrin.
- Ramberg, Bjørn T. 1989. *Donald Davidson's Philosophy of Language: An Introduction*. Oxford: Blackwell.
- Raz, Joseph. 1999. « Explaining Normativity: On Rationality and the Justification of Reason ». *Ratio*, vol. 12, no 4, p. 354-379.
- Read, R. J., et A. M. Crary. 2000. *The New Wittgenstein*. New York: Routledge.
- Renaut, Alain. 1997. *Kant aujourd'hui*. Paris: Aubier.
- Rivenc, François. 2004. « Contre la déflation de la vérité ». *Dialectica*, vol. 58, no 4, p. 517-528.
- , 2008. *Lecture de Quine*. Coll. «Cahiers de logique et d'épistémologie»: College publications.
- Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
- , 1986. « Pragmatism, Davidson and Truth ». In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, E. Lepore (éd.). Oxford: Blackwell.
- , 1990a. *L'homme spéculaire*. Paris: Seuil.
- , 1990b. *Science et solidarité - la vérité sans le pouvoir*. Paris: Éditions de l'éclat.

- , 1992. « Vérité et liberté: Réponse à Thomas McCarthy ». In *Lire Rorty*. Paris: L'Éclat.
- , 1993. « Putnam and the Relativist Menace ». *The Journal of Philosophy*, vol. 90, no 9, p. 443-461.
- , 1994a. « Solidarité ou objectivité ? ». In *Objectivisme, relativisme et vérité*. Paris: PUF.
- , 1994b. « Are Assertions Claims to Universal Validity ». *Deutsche Zeitschrift fuer Philosophie*, vol. 42, p. 988-995.
- , 1995a. *L'espoir au lieu du savoir*. Paris: Albin Michel.
- , 1995b. « Is Truth a Goal of Inquiry? Davidson Vs. Wright ». *The Philosophical Quarterly*, vol. 45, no 180, p. 281-300.
- , 1997a. « What Do You Do When They Call You a 'Relativist'? ». *Philosophy and Phenomenological Research*, vol. 57, no 1, p. 173-177.
- , 1997b. *Contingence, ironie et solidarité*. Paris: Armand Colin.
- , 1998a. *Truth and Progress: Philosophical papers, Vol. 3*. Cambridge: Cambridge University Press.
- , 1998b. « Robert Brandom on Social Practices and Representations ». In *Truth and Progress: Philosophical Papers, Vol. 3*. Cambridge: Cambridge University Press.
- , 2000a. « Universality and Truth ». In *Rorty and his critics*. Malden: Blackwell Publishers.
- , 2000b. « Response to Jürgen Habermas ». In *Rorty and his Critics*, R. Brandom (éd.). Oxford: Blackwell.
- Rouilhan, Philippe. 1984. « Le menteur. Sur la théorie de la vérité de Tarski ». *Le temps de la réflexion*, vol. 5, p. 271-291.
- , 1988. *Frege, Les paradoxes de la représentation*. Paris: Minuit.
- , 1998. « Tarski et l'universalité de la logique ». In *Le formalisme en question*, F. Nef et D. Vernant (éd.). Paris: Vrin.
- Sainsbury, Richard M. 1987. *Paradoxes*. Cambridge: Cambridge University Press.

- Scharp, Kevin. 2003. « Communication and Content: Circumstances and Consequences of the Habermas-Brandt Debate ». *International Journal of Philosophical Studies*, vol. 11, no 1, p. 43-61.
- , 2005. « Scorekeeping in a Defective Language Game ». *Pragmatics and Cognition*, vol. 13, no 1, p. 203-226.
- Searle, John. 1983. *L'intentionnalité: essai de philosophie des états mentaux*. Paris: Minuit.
- Searle, John. 1998. *La construction de la réalité sociale*. Paris: Gallimard.
- Sellars, Wilfrid. 1953. « Inference and Meaning ». *Mind*, vol. 62, no 247, p. 313-338.
- , 1956. *Empirisme et philosophie de l'esprit*. Combas: L'Éclat.
- , 1963. « Philosophy and the scientific image of man ». In *Science, Perception, and Reality*, R. Colodny (éd.). Atascadero, California: Ridgeview Publishing Company.
- , 1980a. *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*. Reseda, California: Ridgeview Publishing Company.
- , 1980b. « Language, Rules and Behavior ». In *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, J. Sicha (éd.). Reseda, California: Ridgeview Publishing Company.
- Seymour, Michel. 1994. *Pensée, langage et communauté: une perspective anti-individualiste*. Paris: Bellarmin Vrin.
- , 2000. « La philosophie de la logique ». In *Précis de philosophie analytique*, P. Engel (dir. publ.). Paris: PUF.
- Shanker, Stuart. 1987. *Wittgenstein and the Turning-Point in the Philosophy of Mathematics*. London: Croom Helm.
- Smith, M., D. Lewis et M. Johnston. 1989. « Dispositional Theories of Value ». *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 63, p. 89-174.
- Smolin, L. 2007. *Rien ne va plus en physique !* Paris: Dunod.
- Soames, Scott. 1997. « Skepticism about meaning: Indeterminacy, normativity and the rule-following paradox ». *Canadian Journal of Philosophy, Supplementary Volume*, vol. 23, p. 211-249.
- , 1998a. « Facts, Truth Conditions, and the Skeptical Solution to the Rule-Following Paradox ». *Noûs*, vol. 32, p. 313-348.

- , 1998b. « Skepticism about Meaning, Indeterminacy, Normativity, and the Rule-Following Paradox ». *Canadian Journal of Philosophy, Supplementary*, vol. 23, p. 211-250.
- , 1999. *Understanding Truth*. Oxford: Oxford University Press.
- Strawson, Galen. 1994. *Mental Reality*. Cambridge, Mass.: Mit Press.
- Stroud, Barry. 1965. « Wittgenstein and Logical Necessity ». *The Philosophical Review*, vol. 74, no 4, p. 504-518.
- Strydom, Piet. 2006. « Intersubjectivity – interactionist or discursive? Reflections on Habermas' critique of Brandom ». *Philosophy & Social Criticism*, vol. 32, no 2, p. 155-172.
- Swindal, James. 2007. « Can a discursive pragmatism guarantee objectivity ? : Habermas and Brandom on the correctness of norms ». *Philosophy and Social Criticism*, vol. 33, no 1, p. 113-126.
- Tarski, A., et R. L. Vaught. 1957. « Arithmetical extensions of relational systems ». *Compositio mathematica*, vol. 13, p. 81-102.
- Tarski, Alfred. 1935, 1983. *Logic, Semantics, Metamathematics, 1923-1938*. Indianapolis: Hackett Publishing Company.
- , 1944. « The Semantic Conception of Truth: and the Foundations of Semantics ». *Philosophy and Phenomenological Research*, vol. 4, no 3, p. 341-376.
- , 1972. *Logique, sémantique, métamathématique (I), 1923-1944*. Paris: A. Colin.
- Taylor, Charles. 1998. *Hegel et la société moderne*. Paris: Cerf et Presses de l'université Laval.
- Velleman, J. David., et Paul A. Boghossian. 1989. « Colour as a Secondary Quality ». *Mind, New Series*, vol. 98, no 389, p. 81-103.
- Wiggins, David. 1998. « A sensible subjectivism ». In *Needs Values and Truth*. Oxford: Clarendon Press.
- Williams, Meredith. 1991. « Blind obedience: Rules, Community and the Individual ». In *Meaning Scepticism*, Klaus Puhl (éd.). New York: De Gruyter.
- , 1999. « Wittgenstein on Representation, Privileged Objects and Private Language ». In *Wittgenstein, Mind and Meaning: Towards a Social Conception of Mind*. London: Routledge et Kegan Paul.

- , 2000. « Wittgenstein and Davidson on the sociality of language ». *Journal for the Theory of Social Behaviour*, vol. 30, no 3, p. 299-318.
- Williams, Michael. 1980. « Coherence, Justification, and Truth ». *The Review of Metaphysics*, vol. 34, no 2, p. 243-272.
- , 1984. « Review of Reason, Truth and History ». *Journal of Philosophy*, vol. 80, p. 257-261.
- , 1986. « Do we (epistemologists) need a theory of truth? ». *Philosophical Topics*, vol. 14, no 1, p. 223-242.
- , 1991. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*. Oxford: Blackwell.
- , 1996. « Exorcism and Enchantment ». *The Philosophical Quarterly*, vol. 46, no 182, p. 99-109.
- , 1999. « Meaning and Deflationary Truth ». *The Journal of Philosophy*, vol. 96, no 11, p. 545-564.
- Wilson, George M. 1994. « Kripke on Wittgenstein and Normativity ». *Midwest Studies in Philosophy*, vol. 19, no 1, p. 366-390.
- , 1998. « Semantic Realism and Kripke's Wittgenstein ». *Philosophy and Phenomenological Research*, vol. 58, no 1, p. 99-122.
- Wittgenstein, Ludwig. 1975. *Remarques philosophiques*. Paris: Gallimard.
- , 1976. *De la certitude*. Paris: Gallimard.
- , 1980. *Grammaire philosophique*. Paris: Gallimard.
- , 1983. *Remarques sur les fondements des mathématiques*. Paris: Gallimard.
- , 1994. *Remarques sur la philosophie de la psychologie (I) et (II)*. Mauvezin: Trans. Europ Reppress.
- , 2002. *Remarques mêlées*. Paris: Garnier Flammarion.
- , 2005. *Recherches philosophiques*. Paris: Gallimard.
- Wolenski, Jan. 2009. « The Rise and Development of logical semantics in Poland ». In *The Golden Age of Polish Philosophy*, S. Lapointe, J. Wolenski, M. Marion et W. Miskiewicz (éd.). Dordrecht: Springer.

- Wright, Crispin. 1980. *Wittgenstein on the Foundations of Mathematics*. Londres: Duckworth.
- , 1981. « Rule-following, objectivity and the theory of meaning ». In *Wittgenstein: To Follow A Rule*, S. H. Holtzman et C. M. Leich (éd.): Routledge.
- , 1984. « Kripke's Account of the Argument Against Private Language ». *The Journal of Philosophy*, vol. 81, no 12, p. 759-778.
- , 1988. « The Inaugural Address: Moral Values, Projection and Secondary Qualities ». *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 62, p. 1-26.
- , 1992. *Truth and Objectivity*. Cambridge Mass: Harvard University Press.
- , 1993. *Realism, Meaning and Truth*. Oxford: Blackwell.
- , 2001. *Rails to Infinity*. Cambridge Mass: Harvard University Press.
- , 2007. « Rule-following without reasons: Wittgenstein's quietism and the constitutive question ». *Ratio*, vol. 20, no 4, p. 481-502.
- Zalabardo, Jose. 1997. « Kripke's Normativity Argument ». *Canadian Journal of Philosophy*, vol. 27, no 4, p. 467-488.