

L'interprétation ordinaire, entre simulation et méta-représentation

Jérôme Dokic

Volume 32, Number 1, Spring 2005

Questions d'interprétation

URI: <https://id.erudit.org/iderudit/011061ar>

DOI: <https://doi.org/10.7202/011061ar>

[See table of contents](#)

Publisher(s)

Société de philosophie du Québec

ISSN

0316-2923 (print)

1492-1391 (digital)

[Explore this journal](#)

Cite this article

Dokic, J. (2005). L'interprétation ordinaire, entre simulation et méta-représentation. *Philosophiques*, 32(1), 19–37.
<https://doi.org/10.7202/011061ar>

Article abstract

In this essay, I examine some aspects of the debate between a perceptual model of communication, according to which testimony is a source of knowledge about the communicated fact, and an inferential model of communication, according to which testimony requires from the hearer an inference from the used signs, the speaker's mental states, and other features of the context. From a reflection on the nature of the capacity for metarepresentation, and its dependence on the capacities of social perception and simulation, I show that the ascription of even complex mental states to the speaker does not preclude communication from functioning on a mode in which the hearer directly inherits the information transmitted by the speaker. Understanding is grounded on the natural tendency to accept what is intelligibly presented to us as true.

L'interprétation ordinaire, entre simulation et méta-représentation*

JÉRÔME DOKIC

École des hautes études en sciences sociales et Institut Jean-Nicod, Paris
dokic@ehess.fr

RÉSUMÉ. — Cet essai porte sur quelques aspects de l'opposition entre un modèle perceptif de la communication, selon lequel le témoignage est une source de connaissance directe du fait témoigné, et un modèle inférentiel de la communication, selon lequel le témoignage implique de la part du récepteur une inférence à partir des signes utilisés, des états mentaux de l'émetteur et du reste du contexte. À partir d'une réflexion sur la nature de la capacité de méta-représentation, et sur sa dépendance à l'égard des capacités de perception sociale et de simulation, je montre que l'attribution d'états mentaux souvent complexes à l'émetteur n'empêche pas le fonctionnement d'un mode primitif de communication dans lequel le récepteur hérite directement de l'information transmise par l'émetteur. La compréhension, aussi réflexive soit-elle, reste enracinée dans la tendance naturelle à accepter ce que l'émetteur cherche à nous faire savoir.

ABSTRACT. — In this essay, I examine some aspects of the debate between a perceptual model of communication, according to which testimony is a source of knowledge about the communicated fact, and an inferential model of communication, according to which testimony requires from the hearer an inference from the used signs, the speaker's mental states, and other features of the context. From a reflection on the nature of the capacity for metarepresentation, and its dependence on the capacities of social perception and simulation, I show that the ascription of even complex mental states to the speaker does not preclude communication from functioning on a mode in which the hearer directly inherits the information transmitted by the speaker. Understanding is grounded on the natural tendency to accept what is intelligibly presented to us as true.

1. Introduction : deux modèles de la communication

Une théorie de l'interprétation des propos d'autrui, dans un acte de communication linguistique ordinaire, peut être conçue selon deux modèles antagonistes.

D'un côté, la phénoménologie de la communication semble encourager un *modèle perceptif de la communication*¹. La communication peut

* Je remercie Martin Montminy et deux évaluateurs anonymes pour leurs remarques utiles sur une version plus ancienne de cet essai.

1. T. Reid, *Essais sur les facultés intellectuelles de l'homme* (1785), dans R. E. Beanblossom et K. Lehrer, dir., *Inquiry and Essays*, Indiana, 1983; C. A. J. Coady, *Testimony. A philosophical Study*, Oxford, Clarendon Press, 1992; T. Burge, « Content preservation », *Philosophical Review*, 102, 1993, p. 457-488; J. McDowell, « Meaning, Communication, and Knowledge » (1980) et « Knowledge by Hearsay » (1993), repris dans J. McDowell, *Meaning, Knowledge, and Reality*, Cambridge (Mass.), Harvard University Press, 1998; F. Récanati, « Does Linguistic Communication Rest on Inference? », *Mind and Language*, 17 (1-2), 2002, p. 105-126.

prendre des formes multiples, mais le témoignage, conçu comme la transmission d'une information d'un émetteur à un récepteur, est l'une de ses formes fondamentales². On peut parler de ce que l'on ne saurait percevoir; le contenu du témoignage n'est donc pas nécessairement un contenu de perception possible. Mais, comme la perception, le témoignage est une source de connaissance. Dans un contexte normal, lorsque quelqu'un m'annonce qu'il pleut en regardant par la fenêtre, alors que je n'ai pas constaté moi-même le temps qu'il fait, j'acquiesce le savoir qu'il pleut de manière directe, non inférentielle. De même que l'expérience perceptive est normalement accompagnée de la croyance correspondante (je vois qu'il pleut, et je le crois), la compréhension des propos de l'émetteur est normalement liée à l'assentiment (l'émetteur me dit qu'il pleut, et je le crois). Certes, je suis parfois victime d'une illusion, mais je suis naturellement enclin à tenir pour vrai le contenu conceptuel de mon expérience perceptive. De même, il arrive que je me méfie de mon interlocuteur et que je réfléchisse à ses intentions avant d'accepter ou non ce qu'il me dit. Mais, dans tous les cas, je suis au moins naturellement enclin à donner mon assentiment au message que l'émetteur cherche à me transmettre.

De l'autre côté, force est de reconnaître que la communication repose le plus souvent sur une théorie de l'esprit. Par « théorie de l'esprit », j'entends la capacité de mentalisation qui nous permet d'attribuer à autrui et à nous-mêmes des propriétés psychologiques, y compris, éventuellement, des attitudes propositionnelles comme la croyance, le désir et l'intention³. L'interprétation d'un acte de langage s'accompagne d'une représentation souvent complexe des attitudes propositionnelles de l'émetteur. Le récepteur sait ce que l'émetteur a dit lorsqu'il est capable de répondre à des questions du type « Pourquoi mon interlocuteur a-t-il utilisé ces mots ? », « Qu'est-ce qu'il a voulu dire par là ? », etc. Certains auteurs considèrent que la dépendance de la communication par rapport à la théorie de l'esprit nous oblige à renoncer au modèle perceptif au profit d'un *modèle inférentiel de la communication*⁴. Selon ce modèle, le message que l'émetteur cherche à transmettre est *inféré* par le récepteur sur la base de

2. Conformément à l'usage philosophique, j'utilise le terme « témoignage » dans un sens large, détaché du contexte juridique formel, pour désigner une situation de communication linguistique dans laquelle l'émetteur *fait savoir* quelque chose au récepteur.

3. Même si je continuerai à utiliser le terme « théorie de l'esprit », désormais traditionnel, l'appellation « capacité de mentalisation » est préférable dans la mesure où elle ne préjuge pas du caractère théorique de la capacité en question. Selon la théorie de la simulation mentale, par exemple (cf. section 6 ci-dessous), la capacité de mentalisation est essentiellement pratique.

4. D. Sperber, « Metarepresentations In An Evolutionary Perspective », in D. Sperber, dir., *Metarepresentations: A Multidisciplinary Perspective*, Oxford, Oxford University Press, 2000, p. 117-137; A. Bezuidenhout, « Is Verbal Communication a Purely Preservative Process ? », *Philosophical Review*, 107, 1998, p. 261-288; G. Origgi et D. Sperber, « Evolution, communication and the proper function of language », in P. Carruthers et A. Chamberlain, dir., *Evolution and the Human Mind: Language, Modularity and Social Cognition*, Cambridge, Cambridge University Press, 2000.

prémisses concernant les signes utilisés, les croyances, désirs et intentions de l'émetteur, et tout autre trait pertinent du contexte. Dans la mesure où l'inférence en question est rarement déductive, on ne peut plus concevoir la compréhension comme une source de connaissance directe relative à la vérité de ce qui est dit. La compréhension du message de l'émetteur est détachée de l'assentiment que le récepteur est susceptible de lui donner. Un tel assentiment peut être le produit d'une nouvelle inférence, sur la base de prémisses concernant la fiabilité de l'émetteur, et le degré de confiance que le récepteur est prêt à lui accorder.

Mon but dans cet article n'est pas de trancher entre ces deux modèles. Le débat sur la nature de la communication linguistique est fort complexe, et je ne saurais prétendre en traiter ici tous les aspects. Je vais néanmoins essayer d'apporter quelques éléments qui, à mon avis, plaident en faveur du modèle perceptif. Ces éléments concernent ce que Tyler Burge appelle le « principe d'acceptation » (*acceptance principle*), à savoir la thèse selon laquelle toute personne est *prima facie* légitimée à accepter ce qui lui est présenté de manière intelligible comme vrai. Comme Burge, mais en invoquant des arguments différents, je tenterai de montrer que notre tendance à adhérer à un contenu propositionnel se manifeste déjà sur le plan de la *compréhension* (ou de ce que Frege appelle la « saisie de la pensée »). Par conséquent, l'adhésion spontanée à ce que l'émetteur nous dit peut avoir, dans certains contextes, une garantie épistémique indépendante d'une inférence partant de prémisses sur les états mentaux de l'émetteur. Cette conséquence me semble être une pierre dans le jardin du modèle inférentiel de la communication.

2. Les scénarios canoniques

Dans ce qui suit, je vais m'intéresser à une version spécifique du modèle perceptif. Cette version se place dans un cadre philosophique plus général, qui concerne les rapports entre le langage et la théorie de l'esprit. Certains auteurs considèrent qu'une théorie de l'esprit ne peut être maîtrisée que par un organisme qui possède déjà la faculté de langage⁵, et donc que l'évolution biologique de cette faculté ne dépend pas du développement préalable de la capacité de mentalisation⁶. Dans la même veine, des travaux en psychologie du développement font l'hypothèse que la maîtrise de la *syntaxe* des compléments propositionnels (qui impliquent en français la conjonction de subordination « que », comme dans « Pierre croit que Marie le trompe ») est une condition antérieure nécessaire de la compréhension de la notion de *croyance fausse*, au cœur de notre théorie de l'esprit⁷.

5. D. Davidson, *Enquêtes sur la vérité et l'interprétation*, traduction française Pascal Engel, Nîmes, Éditions Chambon, 1993.

6. D. Dennett, *La conscience expliquée*, traduction française Pascal Engel, Paris, Odile Jacob, 1993.

7. Voir aussi, par exemple, J. de Villiers, « Language and theory of mind: what are the developmental relationships? », in S. Baron-Cohen, H. Tager-Flusberg, et D. J. Cohen, dir., *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*, Oxford, Oxford University Press, 2000, p. 83-123.

D'autres auteurs maintiennent au contraire que la communication au sens propre du terme, qui inclut mais ne se limite pas à la communication linguistique, exige qu'une théorie de l'esprit soit déjà en place. La thèse de la dépendance de la communication par rapport à la capacité de mentalisation remonte au moins aux travaux de Paul Grice, et c'est un thème central de la théorie de la pertinence développée par Dan Sperber et Deirdre Wilson⁸. Selon ces auteurs, la communication repose, du point de vue du récepteur, sur l'attribution à l'émetteur d'intentions complexes, en particulier l'intention méta-représentationnelle de rendre mutuellement manifeste à tous les protagonistes de l'acte de communication que l'émetteur a une intention informative spécifique. Dans cette perspective, on peut admettre que le langage et la théorie de l'esprit ont subi une évolution parallèle⁹, mais l'avènement du premier dépend unilatéralement de l'existence de la seconde, qui a évolué à la suite d'autres pressions subies au cours de l'évolution¹⁰.

A priori, l'opposition entre un modèle perceptif et un modèle inférentiel de la communication est indépendante de la question de savoir si la faculté de langage dépend d'une théorie de l'esprit ou vice-versa. On peut admettre que le langage dépend d'une théorie de l'esprit, mais que celle-ci n'est pas exercée dans la communication comme les partisans du modèle inférentiel le prétendent. De même, on peut défendre la thèse selon laquelle la théorie de l'esprit dépend du langage, mais dès lors qu'elle est en place, l'image du témoignage comme source de connaissance non inférentielle ne peut plus être maintenue. Si je mentionne ici la question de la priorité du langage et de la théorie de l'esprit dans l'ordre de l'explication, c'est que la version la plus convaincante du modèle perceptif exploite une analogie entre la communication humaine et la communication animale¹¹. Selon cette version, des mécanismes fondamentaux qui sous-tendent la communication humaine sous-tendent également la communication animale. Dans la mesure où ces mécanismes ne dépendent pas d'une capacité sophistiquée de mentalisation, la communication humaine comporte une dimension fondamentale qui précède l'exercice de notre théorie ordinaire de l'esprit.

Les mécanismes supposés être communs à la communication humaine et à la communication animale concernent des transferts directs de croyances et de désirs (ou plus généralement de motivations) de l'émetteur au récepteur.

8. P. Grice, *Studies in the Way of Words*, Boston, Harvard University Press, 1989; D. Sperber et D. Wilson, *La pertinence*, Paris, Minuit, 1989.

9. G. Origgi, et D. Sperber, « Evolution, communication and the proper function of language ».

10. Je pense ici aux travaux sur le développement de l'intelligence sociale ou « machiavélique » : voir aussi, par exemple, A. Whiten et R. W. Byrne, dir., *Machiavellian Intelligence II*, Cambridge, Cambridge University Press, 1997.

11. Cette analogie est particulièrement claire dans J. McDowell, « Meaning, Communication, and Knowledge ».

Considérons les deux scénarios suivants, que j'appellerai « scénarios canoniques » :

- I. *Extension de la perception.* Une créature crie ou gesticule en réaction naturelle ou spontanée à quelque fait (positif ou négatif) et une seconde créature, entendant les cris ou voyant les gestes de la première, réagit à son tour *comme si* elle faisait elle-même l'expérience directe de ce fait. Il y a ici un *transfert naturel de croyance* : le « récepteur » en vient à croire quelque chose initialement cru par l'« émetteur ».
- II. *Extension de l'action.* Une créature exprime par ses cris et gestes son désir d'atteindre quelque objet, provoquant ainsi une autre créature à venir à son secours. Il y a ici un *transfert naturel de désir* : le « récepteur » en vient à désirer l'action même désirée par l'« émetteur », par exemple amener quelque objet à portée de main de ce dernier.

Des versions plus ou moins rudimentaires de ces scénarios se réalisent chez de nombreuses espèces non humaines (telles que l'abeille, certains oiseaux et les singes vervets). Par ailleurs, les transferts de croyances ou de désirs aux sens décrits par I et par II ne semblent pas dépendre d'une théorie de l'esprit sophistiquée. Il faut intervenir un « système informationnel » qui repose avant tout sur des relations causales fiables¹². Les deux scénarios impliquent un flux d'information, au sens naturel, non intentionnel du terme, du monde à l'émetteur, puis de l'émetteur au récepteur. Dans le scénario I, la réponse de l'émetteur véhicule une information sur un fait donné, dont le récepteur hérite par sa propre réponse différenciée. Dans le scénario II, la réponse de l'émetteur véhicule une information sur la désidérabilité relative d'une certaine action, dont le récepteur hérite en désirant lui-même cette action. La question est de savoir si ces scénarios ont quoi que ce soit à voir avec la communication humaine ordinaire.

3. La spécificité de la communication humaine

Dans une série d'essais récents, Dan Sperber a apporté une réponse négative à cette question. Il s'est prononcé en faveur de l'hypothèse d'une discontinuité entre les scénarios canoniques et la communication humaine. À propos de l'idée que les scénarios canoniques impliquent une forme de proto-langage à partir duquel la communication telle que nous la connaissons a pu se développer, il écrit :

Ou bien la communication linguistique ancestrale, quoiqu'elle fût à de nombreux égards plus simple, était fondée sur le même type de mécanisme communicationnel que la communication linguistique moderne. Dans ce cas, elle présupposait la capacité de méta-représentation, et ne pouvait donc pas la précéder. Ou bien la communication ancestrale linguistique était strictement une

12. Sur cette notion de système informationnel et sa relation avec le témoignage, cf. G. Evans, *The Varieties of Reference*, Oxford, Clarendon Press, 1982, chap. 5, surtout p. 123.

affaire de codage/décodage, et alors il n'y a aucune raison de supposer que nos ancêtres avaient les ressources nécessaires pour prendre conscience du caractère représentationnel de leurs signaux, pas plus que les abeilles ou les singes vervets¹³.

Selon Sperber, dès lors que la communication repose sur une théorie méta-représentationnelle de l'esprit, elle prend une forme radicalement différente de celle des transferts de croyances et de désirs dans les scénarios canoniques. Dans la théorie de la pertinence, le message n'est pas directement perçu mais inféré à partir des intentions de communication du locuteur, elles-mêmes dérivées de prémisses concernant les signes utilisés et le reste du contexte. Les interactions proto-linguistiques des scénarios canoniques — qui étendent les capacités perceptives et agentives des sujets — ne survivent pas au développement de la communication proprement dite (ou alors seulement en tant que vestiges). Le véritable précurseur de la communication n'est pas un proto-langage fondé sur des signes naturels, mais la capacité de mentalisation. Il n'y a aucun lien privilégié entre le proto-langage et la théorie de l'esprit qui pourrait expliquer de manière plausible l'évolution ou le développement du premier à la seconde.

L'objection de Sperber comporte plusieurs éléments spécifiques distincts. Premièrement, Sperber et les théoriciens de la pertinence prétendent que le modèle perceptif de la communication, du moins en tant qu'il fait jouer un rôle central aux scénarios canoniques, a partie liée avec la conception de la communication comme un code. Selon cette conception, le message communiqué est globalement codé dans le signe, plus ou moins indépendamment du contexte. Je laisserai cet argument de côté dans cet essai, car je me rallie à la critique qu'en a faite François Récanati¹⁴. Le modèle perceptif de la communication peut et doit rejeter la conception de la communication comme un processus de simple codage (du côté de l'émetteur) et décodage (du côté du récepteur). Le message communiqué dépend au moins autant du contexte extra-linguistique que du signe utilisé.

Un second élément de l'objection de Sperber concerne la *transparence* des signes utilisés dans les scénarios canoniques. Pour qu'un transfert de croyance ou de désir ait lieu, il suffit que l'émetteur et le récepteur se tiennent l'un l'autre pour de simples intermédiaires causaux. Or il semble qu'une condition nécessaire de la communication soit que l'émetteur et le récepteur considèrent certains événements *comme* des signes; dans les termes de Sperber, ils doivent prendre conscience de leur « caractère représentationnel ». Une telle prise de conscience requiert une théorie de l'esprit, seule habilitée à donner un sens à l'idée que l'émetteur et le récepteur utilisent sciemment des signes. Cet argument, destiné à mettre en évidence le fossé qui sépare les scénarios canoniques des processus ordinaires de communication, sera l'objet de la section 4 ci-dessous. Dans la section 5, j'examinerai une variante de cet

13. D. Sperber, « Metarepresentations In An Evolutionary Perspective », p. 4.

14. F. Récanati, « Does Linguistic Communication Rest on Inference? ».

argument, selon laquelle les scénarios canoniques impliquent des signes naturels alors que la communication humaine implique des signes conventionnels ou institutionnels, qui peuvent être utilisés librement, indépendamment des exigences du contexte présent. En réponse à cet argument, je présenterai les grandes lignes d'une solution de continuité entre les signes naturels et les signes d'institution.

Un troisième aspect, plus général, de l'objection de Sperber concerne la capacité de méta-représentation engagée dans la communication humaine. Il faut admettre que cette capacité n'est pas seulement exercée pour prendre conscience du caractère représentationnel des signes utilisés. Le récepteur d'un acte de communication doit tenir compte des intentions complexes de l'émetteur, qui ne sont pas toujours compatibles avec la transmission de la vérité. Comme le rappellent Origgi et Sperber, un problème majeur de la théorie de la communication est d'expliquer comment les bénéfices de la communication peuvent excéder le coût d'éventuelles fraudes, mensonges ou manipulations :

Dans le cas de la communication humaine, l'explication de la manière dont le coût d'une possible tromperie [*deception*] est maîtrisé implique le fait que la compréhension et l'acceptation [*acceptance*] sont deux étapes distinctes dans le processus global¹⁵.

En général, la communication humaine est hautement réflexive ou méta-représentationnelle¹⁶. Ce trait l'oppose incontestablement à la communication animale. Un partisan du modèle perceptif de la communication doit expliquer la réflexivité de la communication humaine sans souscrire au point de vue inférentialiste selon lequel la compréhension et l'adhésion sont deux étapes distinctes d'un processus normal de communication. Une telle explication sera esquissée dans les sections 7 et 8.

4. La perception sociale

Considérons le problème de la transparence des signes impliqués dans les scénarios canoniques. Dans tout acte de communication digne de ce nom, les signes doivent être compris comme tels par les intéressés. Cette compréhension est psychologique et fait donc intervenir quelque théorie de l'esprit.

La question est de savoir si cette théorie de l'esprit est nécessairement méta-représentationnelle, c'est-à-dire si elle porte explicitement sur des représentations mentales. La réponse est négative si l'on peut donner un sens à l'opacité de signes en restant sur le terrain de la perception. En effet, les représentations perceptives sont toujours des représentations de premier ordre. On ne perçoit jamais des représentations mentales comme telles. Les paradigmes de représentations mentales (ou d'états mentaux qui impliquent essentiellement

15. D. Origgi, et D. Sperber, « Evolution, communication and the proper function of language », p. 161-162.

16. Dans le cadre de cet essai, j'utilise « réflexif » dans le même sens que « méta-représentationnel ».

des représentations mentales) sont les croyances et les désirs, conçus comme attitudes propositionnelles¹⁷. Or le concept de croyance ou de désir, tel qu'il figure dans notre psychologie ordinaire, n'est pas un concept d'observation. On peut observer autrui avoir mal, ou avoir peur, mais on ne peut pas l'observer croire ou désirer quoi que ce soit¹⁸. Les croyances et les désirs sont attribués sur la base du comportement observable de manière holiste : on ne peut pas déchiffrer une croyance à partir d'un comportement observable, mais seulement à partir d'un ensemble ouvert de comportements. Par suite, si une théorie de l'esprit peut être mobilisée sur le plan perceptif, elle ne sera pas méta-représentationnelle, mais au contraire pré-réflexive.

Les recherches psychologiques sur l'évolution et le développement de la théorie de l'esprit ont identifié plusieurs « précurseurs » de la capacité méta-représentationnelle de mentalisation. L'attention partagée, les comportements de pointage proto-impératif et proto-déclaratif, la compréhension des intentions en action, l'imitation et le comportement empathique précèdent apparemment la réflexion telle qu'elle s'exerce dans l'attribution d'attitudes propositionnelles comme la croyance ou le désir¹⁹.

Nos capacités pré-réflexives de mentalisation ont une composante perceptive importante. On a décrit la capacité de détecter visuellement les postures et les attitudes corporelles d'autrui (qui semble avoir son fondement causal dans le cortex temporal) comme une capacité de perception sociale, ou une « fenêtre ouverte sur l'esprit d'autrui²⁰ ». De même, on peut établir par observation qu'une personne est dans une relation épistémique directe à un fait, par exemple qu'elle voit qu'il pleut. Une telle relation épistémique est directe au sens où elle n'est pas médiatisée par une représentation mentale. Sa représentation peut rester de premier ordre : contrairement à la représentation d'une croyance, elle ne mobilise pas essentiellement la capacité de méta-représentation. Ainsi, certains psychologues du développement font l'hypothèse (anticipée par Wittgenstein dans *De la certitude*) que l'enfant commence à comprendre ce qu'est le savoir avant de comprendre ce qu'est la croyance²¹.

17. mais pas forcément comme relations à des propositions; cf. section 7 ci-dessous.

18. Le cas du désir est plus complexe : on ne peut pas observer le désir en tant qu'attitude propositionnelle, mais on peut observer le désir en tant qu'émotion. J'en dis un peu plus sur la distinction entre les états mentaux observables et les autres dans J. Dovic, « The Sense of Ownership: An Analogy between Sensation and Action », in N. Eilan. et J. Roessler, dir., *Agency and Self-Awareness. Issues in Philosophy and Psychology*, Oxford, Clarendon Press, 2003, p. 321-344.

19. Voir par exemple les essais réunis dans Baron-Cohen, S., *Understanding Other Minds*. Pour une théorie pré-réflexive de l'attention partagée et de son rôle dans la communication, cf. I. Brinck., « Attention and the evolution of intentional communication », *Pragmatics and Cognition*, 9.2, 2001, p. 259-277.

20. T. Allison, A. Puce et G. McCarthy, « Social perception from visual cues: role of the STS region », *Trends in Cognitive Science* 4:7, 2000, p. 267-278.

21. Cf. K. Bartsch et H. M. Wellman, *Children Talk About the Mind*, Oxford, Oxford University Press, 1995, et J. Perner, *Understanding the Representational Mind*, Cambridge (Mass.), MIT Press, 1991.

À nouveau, la saisie du concept de croyance est plus complexe que celle du concept de savoir ou d'information comme relation directe au monde, car elle suppose la maîtrise d'une théorie méta-représentationnelle de l'esprit, que l'enfant n'acquiert qu'à trois ou quatre ans.

Dans les versions purement causales des scénarios canoniques, le contenu de la perception du récepteur à qui l'on apprend que *p* est principalement « *p* ». Autrement dit, le contenu de la perception du récepteur s'apparente à celui d'une perception directe du fait que *p*. La médiation de l'émetteur dans la transmission de l'information est totalement transparente. Par contre, le récepteur qui dispose de capacités perceptives de mentalisation peut faire l'expérience d'une situation plus complexe : le contenu de sa perception a à peu près la forme suivante : « L'émetteur sait et me fait savoir que *p* ». Le récepteur ne perçoit pas seulement le fait que *p* ; il perçoit le fait qu'un émetteur cherche à lui révéler le fait que *p*. Le scénario reste canonique, mais n'est pas purement causal, puisque le récepteur prend conscience de l'émetteur comme un acteur social cherchant à communiquer. À nouveau, cette prise de conscience ne repose pas nécessairement sur l'attribution à l'émetteur de représentations mentales comme la croyance ou le désir²².

On peut ainsi espérer rendre compte de la phénoménologie de la compréhension, qui est à la fois perceptive et psychologiquement engagée. Certes, la « théorie de l'esprit » mobilisée au niveau perceptif est assez rudimentaire. Elle permet de repérer des situations psychologiques relationnelles, qui impliquent des objets intentionnels concrets, mais elle est incapable de rendre compte des cas de « méreprésentation », où le sujet n'est précisément pas en relation cognitive avec un objet réel ou avec un fait. *A fortiori*, elle est loin de pouvoir fonder la conception d'une croyance fausse, qui a partie liée avec une théorie méta-représentationnelle de l'esprit.

5. La simulation comme proto-langage opératoire

Comme nous l'avons vu, les scénarios canoniques impliquent des *signes naturels*. Dans certains contextes, lorsqu'un émetteur affirme qu'il pleut en regardant par la fenêtre, il signifie naturellement la pluie locale. L'occurrence du signe « Il pleut » co-varie de manière fiable avec la pluie locale, en vertu d'une chaîne causale qui la relie à celle-ci. Or les signes naturels ne sont pas des signes à *proprement parler*. L'une des caractéristiques essentielles des signes

22. J. McDowell, *Meaning, Knowledge, and Reality*, p. 433. F. Récanati, « Does Linguistic Communication Rest on Inference? », affirme que lorsque nous comprenons une occurrence du démonstratif complexe « ce gars », nous ne devons pas penser au référent *comme* étant la personne à laquelle l'émetteur a l'intention de faire référence. S'il y a une représentation des intentions du locuteur (dans un contexte normal), elle reste sub-personnelle plutôt que consciente. Si mes remarques ci-dessus sont correctes, le défenseur du modèle perceptif n'est pas obligé de reléguer au niveau sub-personnel les représentations psychologiques qui sous-tendent la compréhension. Au contraire, au moins certaines d'entre elles participent de la phénoménologie ordinaire de la compréhension.

impliqués dans la communication humaine est qu'ils peuvent être convoqués librement, indépendamment du contexte présent.

On peut répondre à cette objection en essayant de mettre en évidence la *continuité* qui existe entre les signes naturels et les signes non naturels ou institutionnels. Au XVIII^e siècle, dans l'*Essai sur l'origine des connaissances humaines*, Condillac a donné une description de la transition des signes naturels aux signes d'institution dont certains éléments restent à mon avis valables :

Plus ils se familiarisèrent avec ces signes, plus ils furent en état de se les rappeler à leur gré. Leur mémoire commença à avoir quelque exercice ; ils purent disposer eux-mêmes de leur imagination, et ils parvinrent insensiblement à faire avec réflexion ce qu'ils n'avoient fait que par instinct. D'abord tous deux se firent une habitude de connaître, à ces signes, les sentiments que l'autre éprouvait dans le moment ; ensuite ils s'en servirent pour se communiquer les sentiments qu'ils avoient éprouvés. Celui, par exemple, qui voyait un lieu où il avait été effrayé, imitait les cris et les mouvements qui étoient les signes de la frayeur, pour avertir l'autre de ne pas s'exposer au danger qu'il avait couru²³.

L'*imitation* ou la *simulation* est invoquée ici comme un mécanisme cognitif permettant un premier découplage du signe par rapport au contexte présent. Nous sommes capables d'imiter une action relationnelle — dans l'exemple de Condillac, l'homme fait semblant d'avoir peur de quelque chose — en l'absence de son *relatum* (par exemple, un lion rencontré dans le passé à l'endroit où se trouvent les protagonistes de l'acte de communication). Selon une conception plausible de la simulation, la simulation d'une action ou d'une réponse spontanée implique des mécanismes cognitifs également mobilisés dans l'action ou la réponse originale. L'homme qui fait semblant d'avoir peur d'un lion se trouve dans un état épistémique *analogue* (mais non identique) à celui dans lequel il se trouvait au moment de la rencontre réelle avec le lion²⁴. On peut faire ici deux remarques à propos de la suggestion de Condillac.

Premièrement, elle met en évidence la continuité entre les signes naturels et les signes d'institution. Plus précisément, elle explique comment certains signes naturels peuvent être également considérés comme des signes d'institution, et repris en dehors de leur contexte d'origine. Supposons que l'homme décrit par Condillac a initialement répondu de manière spontanée à la présence d'un lion en produisant un cri caractéristique. Si sa réponse co-varie de manière fiable avec la présence d'un lion, elle a à peu près la même signification que notre énoncé « Il y a ici un lion ».

Notre homme est aussi capable d'imiter sa réponse spontanée à la présence d'un lion ; il lui suffit de faire semblant d'être actuellement en présence

23. Condillac, *Essai sur l'origine des connaissances humaines*, Paris, Vrin, 2002, II, 1, ch. 1, § 3.

24. A. Goldman, « In defense of the simulation theory », *Mind and Language*, 7, 1992, p. 104-119 ; A. Goldman, « Simulation and Mental Concepts », in J. Dokic. et J. Proust, dir., *Simulation and Knowledge of Action*, Amsterdam/Philadelphia, John Benjamins, 2002.

de l'animal, dans un contexte qu'il choisit librement. Notons que l'homme n'imité pas sa réponse à un fait indépendamment identifié, mais la *reproduit*. En définitive, seul le fait est explicitement représenté dans la simulation, dont le processus n'implique nullement la capacité de méta-représentation. Lorsque l'homme imite sa rencontre passée avec un lion, il réactive des mécanismes cognitifs mobilisés lors de celle-ci. Il utilise un signe, dont la signification s'approche de notre énoncé « Il y a ici un lion », en dehors de son contexte d'origine. Il fait ainsi un premier pas dans la transformation du signe naturel en signe d'institution, dont l'occurrence est indépendante de la présence actuelle d'un lion.

Ma seconde remarque est que la suggestion Condillac est compatible avec le modèle perceptif de la communication. On pourrait objecter que ce modèle vaut au mieux pour les signes naturels, mais qu'il est obsolète pour les signes d'institution. Si les signes peuvent être utilisés librement dans le cadre d'une affirmation, c'est que leurs occurrences cessent d'être des réponses naturelles à des faits donnés, et donc ne peuvent plus en être les *révélateurs*. Cette objection repose sur un malentendu. Seules les occurrences de certains signes utilisés tels quels, dans le cadre d'une affirmation, peuvent prétendre constituer des réponses naturelles à des faits donnés. Si les signes sont utilisés comme parties sémantiques propres d'une affirmation plus complexe, par exemple dans la portée d'un opérateur de phrases, aucune garantie n'existe que leurs occurrences révèlent des faits donnés. Dans certains contextes, l'émetteur qui annonce qu'il pleut au moyen des mots « Il pleut » signifie naturellement la pluie locale, mais ces mots peuvent être également utilisés en l'absence de la pluie, par exemple, dans l'énoncé « S'il pleut, je prendrai mon parapluie ». Cet énoncé est sémantiquement complexe, et contient l'opérateur de phrases « si... alors... ».

Dans l'exemple de Condillac, le message transmis par l'émetteur a à peu près la même signification que notre énoncé « Il y avait ici un lion ». Il présente déjà une certaine complexité sémantique, puisqu'il contient comme partie propre un autre signe, à peu près synonyme de « Il y a ici un lion ». Comme nous l'avons vu, il s'agit du signe utilisé *dans la portée de la simulation*. On peut considérer que notre homme imite la présence d'un lion d'une manière spécifique, qui communique au récepteur le fait qu'un lion s'est trouvé, *dans le passé*, à l'endroit indiqué par le récepteur. En imitant la présence d'un lion de cette manière, l'émetteur utilise un signe qui se rapproche de l'opérateur temporel intensionnel « Il était le cas que », tel qu'il a été décrit par Arthur Prior. Le message communiqué a alors la forme « Il était le cas que (il y a ici un lion) » (ou « Dans le passé (il y a ici un lion) »). Il vaut la peine de noter que Prior considérerait cette forme comme l'analyse sémantique correcte de notre énoncé « Il y avait ici un lion » :

[Une] expression qui permet de construire une phrase à partir d'une autre phrase est un adverbe ou une expression adverbiale, comme « ne ... pas » ou « Ce n'est pas le cas que », ou « prétendument » ou « Il est prétendu que », ou

« Il est possible que ». [...] Je voudrais émettre l'hypothèse selon laquelle mettre un verbe au temps passé ou futur revient exactement au même qu'ajouter un adverbe à la phrase. « J'étais en train de déjeuner » est lié à « Je suis en train de déjeuner » de la même manière que « Je suis prétendument en train de déjeuner » lui est lié. C'est un accident historique que nous formons le temps passé en modifiant le temps présent, par exemple en substituant « étais » à « suis », plutôt qu'en lui accrochant un adverbe. Dans un langage rationalisé qui comporte des constructions uniformes pour des fonctions similaires, nous formerions le temps passé en préfixant à une phrase l'expression « Il était le cas que » [...] et le temps futur en préfixant « Il sera le cas que »²⁵.

La suggestion de Condillac est compatible avec le modèle perceptif de la communication dans la mesure où, si l'occurrence du signe qui correspond à « Il y a ici un lion » ne révèle pas la présence actuelle d'un lion dans l'environnement, l'occurrence du signe « Il y avait ici un lion » peut être considérée comme une réponse spontanée au fait qu'il y avait ici un lion. Par suite, cette réponse se transmet au récepteur, qui met en œuvre les mécanismes cognitifs, quels qu'ils soient, qui seraient activés s'il constatait lui-même la présence passée d'un lion, par exemple s'il percevait sur le sol des traces fraîches du félin.

Si les remarques qui précèdent sont sur la bonne voie, le niveau de communication réalisé dans l'exemple de Condillac implique déjà un langage pourvu d'une syntaxe rudimentaire. Ce langage implique des opérateurs intensionnels de phrases, tels que « Il était le cas que » ou « Il sera le cas que » (si l'homme décrit par Condillac est aussi capable de transmettre des messages relatant des situations futures, par exemple le fait qu'il y aura bientôt un lion ici). À l'analyse sémantique que Prior offre de ces opérateurs s'ajoute la thèse selon laquelle leur utilisation repose sur la simulation, capable de rendre présente à l'esprit une situation passée ou future. On peut rapprocher cette thèse de la théorie de la « deixis orientée par l'imagination » développée par Karl Bühler :

Si j'entends « Il pleut » sans autre introduction, je le considère comme une évaluation du temps qu'il fait dans la situation du discours ; il pleut là où se trouve le locuteur au moment présent [...] L'addition de « sur le lac de Constance » est un pas libérateur : « Il pleut sur le lac de Constance » ; cette formulation étendue peut être produite n'importe où ; son sens est dans une large mesure libéré de [...] la situation du discours²⁶.

Au niveau de langage qui m'intéresse ici, les opérateurs spatio-temporels, du type « sur le lac de Constance » ou « dans le passé », sont compris comme des *invitations à la simulation*, c'est-à-dire des invitations à rendre présentes, par

25. A. N. Prior, *Papers on Time and Tense*, dir., par P. Hasle, P. Øhrstrøm, T. Braüner, et J. Copeland, Oxford, Oxford University Press, 2003, p. 12-13.

26. K. Bühler, *Theory of Language* (1934), traduction anglaise D. F. Goodwin, Amsterdam, John Benjamins, 1990, p. 426, cité par F. Récanati, *Oratio Obliqua, Oratio Recta. An Essay on Metarepresentation*, Cambridge (Mass.), MIT Press, 2000, p. 89, qui contient un développement original de la théorie de Bühler.

l'imagination ou le souvenir, des situations délocalisées ou passées, ou simplement possibles. On peut ainsi émettre la spéculation selon laquelle l'évolution de la capacité de simulation est étroitement liée à la capacité proto-syntaxique d'utiliser un langage opératoire, sans préjuger ici de la question de savoir quelle capacité est prioritaire sur l'autre dans l'ordre de l'explication.

6. Les limites de la simulation mentale

L'introduction de la notion de simulation nous permet de faire un pas, des scénarios canoniques vers la communication linguistique ordinaire. Selon l'argument de la section précédente, elle rend possible l'émergence d'un proto-langage articulé. La maîtrise de ce langage, dont la syntaxe rudimentaire fait intervenir des opérateurs de phrases, ne semble pas requérir une théorie méta-représentationnelle de l'esprit. En effet, la capacité de simulation qui sous-tend l'utilisation de ces opérateurs ne manipule que des représentations de premier ordre. Dans la mesure où, comme je l'ai concédé aux partisans du modèle inférentiel de la communication, la communication linguistique repose essentiellement sur une théorie méta-représentationnelle de l'esprit, le fossé qui sépare un proto-langage opératoire du langage tel que nous le connaissons reste à combler.

Selon une certaine conception des rapports entre la simulation et la théorie de l'esprit, ce fossé est moins important qu'il ne paraît. On peut arguer que la capacité de simulation permet de comprendre certaines situations mentales inaccessibles au niveau purement perceptif. Par exemple, elle permet de faire des prédictions limitées sur le comportement d'un sujet qui croit faussement qu'il fait beau alors qu'il pleut : le sujet agit simplement *comme* s'il pleuvait. Une telle situation n'est pas accessible au niveau perceptif ; par hypothèse, le sujet n'est pas dans une relation observable avec le fait qu'il fait beau. Certains auteurs vont plus loin et défendent la thèse selon laquelle la capacité de simulation est la voie royale vers notre théorie ordinaire de l'esprit. Robert Gordon prétend que la capacité d'exprimer une conception étrangère du monde sur le mode du « comme si » est l'élément central sur lequel peut se fonder l'introduction ultérieure de concepts psychologiques, comme celui de croyance. Lorsque je simule le monde tel qu'autrui l'a trouvé, par exemple lorsque j'imagine qu'il fasse beau alors que je crois qu'il pleut, j'attribue au moins implicitement à autrui la croyance qu'il fait beau. Gordon considère que notre théorie de l'esprit, dans la mesure où elle est fondée sur la capacité de simulation, relève davantage d'une capacité pratique que d'une capacité théorique. Par conséquent, elle n'est pas aussi réflexive qu'elle ne paraît. Lorsque j'attribue à autrui la croyance qu'il fait beau en simulant le monde tel qu'il l'a trouvé, je ne fais pas normalement référence à la croyance d'autrui en tant que représentation mentale. Mon attribution n'est pas méta-représentationnelle ; au contraire, elle est entièrement tournée vers le monde, réel ou simulé²⁷.

27. R. Gordon, « Simulation Without Introspection or Inference from Me to You », in M. Davies et T. Stone, dir., *Mental Simulation*, Oxford, Blackwell, 1995 ; R. Gordon, « "Radical" »

Malheureusement, il n'est pas évident que la capacité de simulation soit une condition suffisante de la capacité d'attribuer, même implicitement, des croyances à soi-même et à autrui. On peut douter de l'existence d'un chemin de la simulation à la simulation mentale. Considérons le scénario suivant, emprunté au célèbre *test de la croyance fausse*²⁸. Soit un personnage, Maxi, en face de deux boîtes, bleue et rouge. Maxi range une barre de chocolat dans la boîte bleue, puis quitte la pièce. En son absence, quelqu'un déplace la barre de chocolat dans la boîte rouge. Maxi fait son retour dans la pièce. Où va-t-il chercher la barre de chocolat ? Selon la théorie de la simulation mentale, je peux répondre correctement à cette question, et donc attribuer implicitement une croyance fausse à Maxi, avant tout parce que je suis capable de me mettre mentalement à sa place, et de construire par l'imagination un modèle mental du monde tel qu'il l'a trouvé.

Certes, le sujet capable d'anticiper le comportement d'une autre personne ayant des croyances fausses, ou plus précisément incompatibles avec les siennes propres, ne maîtrise pas pour autant le concept de croyance. Comme Gordon le reconnaît lui-même, un tel sujet « n'est pas encore en mesure de comprendre que *ses propres croyances présentes* peuvent être elle-mêmes en désaccord avec les faits ». En effet, il « ne parviendra pas à une telle compréhension aussi longtemps qu'il déterminera quelles sont ses propres croyances présentes en se demandant quels sont les faits ». Selon Gordon, c'est précisément par un enchâssement multiple de modèles mentaux que le sujet parvient à la compréhension requise :

Pour faire apparaître ses propres croyances présentes comme différenciables des faits, [le sujet] devra simuler quelqu'un d'autre pour qui les faits sont différents — ou, plus généralement, adopter une perspective à partir de laquelle les faits sont différents, que cette perspective soit occupée ou non par une personne réelle — et, à partir de cette perspective étrangère, *simuler lui-même* [*simulate herself*]²⁹.

L'ordre de priorité établi par la philosophie cartésienne est ici comme inversé. La troisième personne est prioritaire par rapport à la première personne au sens où je dois passer par une représentation d'autrui pour accéder à mes propres croyances, conçues comme indépendantes des faits.

Gordon considère que le développement d'une certaine forme de *mémoire* peut constituer une autre manière de parvenir à la compréhension que notre perspective actuelle sur les choses pourrait être erronée. Pour reprendre les termes d'une autre expérience célèbre, supposons que l'on me

Simulationism », in P. Carruthers et P. K. Smith, dir., *Theories of theories of mind*, Cambridge, Cambridge University Press, 1996.

28. H. Wimmer et J. Perner, « Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception », *Cognition*, 13, 1983, p. 103-128.

29 R. Gordon, « Simulation Without Introspection or Inference from Me to You », p. 62.

présente un objet que je prends d'abord pour une pierre³⁰. Après avoir manipulé l'objet, je change d'avis, et reconnais qu'il s'agit en fait d'une éponge. Je peux avoir conservé le souvenir apparent d'une pierre, et parvenir ainsi à comprendre que ma perspective passée sur les choses était erronée. Or selon Gordon, si je simule une perspective future sur ma perspective présente, je peux comprendre, de manière analogue, que celle-ci pourrait être erronée — même si, par définition, je ne peux pas le *croire*.

À y regarder de plus près, toutefois, la méthode qui consiste à enchâsser des modèles mentaux ou des perspectives différentes sur le monde ne permet pas de fonder à elle seule la compréhension selon laquelle nos propres croyances, ou du moins certaines d'entre elles, pourraient être fausses. Cette méthode donne au mieux un sens à la possibilité que, selon une autre personne, j'ai tort; elle ne saurait donner un sens à la possibilité que cette personne a raison, et donc que j'ai *effectivement* tort. De même, la méthode donne au mieux un sens à la possibilité que selon mon moi futur, mon moi présent a tort; elle ne saurait donner un sens à la possibilité que mon futur moi a raison, et donc que j'ai *actuellement* tort.

Martin Montminy m'a objecté que la méthode décrite par Gordon permet de comprendre que *si* autrui ou mon moi futur a raison, alors j'ai tort, et, de là, comprendre la possibilité que j'aie tort. Je prétends au contraire que l'on ne peut pas comprendre qu'autrui ou mon moi futur a raison en restant sur le terrain de la simulation. La théorie de la simulation mentale n'explique pas comment je peux dépasser la dichotomie entre le monde tel que je l'ai trouvé, dont la représentation ne doit rien à la simulation, et le monde tel qu'autrui ou moi-même pourrait le trouver, dont la représentation implique la simulation. Alors que je conçois le premier monde comme une réalité, je conçois le second comme une simple possibilité, incapable de se réaliser (du moins lorsqu'il est en désaccord avec « mon » monde). Je peux imaginer quelqu'un contester la réalité du monde tel que je l'ai trouvé, mais je ne peux pas saisir la possibilité que l'une de mes croyances soit fausse, *indépendamment de l'opinion de qui que ce soit*.

7. Une ontologie d'états mentaux

Il ressort des remarques précédentes que la capacité de simulation ne peut pas se substituer à une théorie méta-représentationnelle de l'esprit, essentielle à toute communication digne de ce nom. En particulier, elle ne permet pas de fonder la maîtrise de la notion de croyance fausse, et donc de donner un sens à certaines possibilités auxquelles tout locuteur adulte doit être sensible, comme la possibilité que la croyance que l'émetteur cherche à lui transmettre est fausse. Il reste donc à montrer que l'acquisition d'une théorie de l'esprit capable de donner un sens à ces possibilités est compatible avec le modèle perceptif de la

30 J. H. Flavell, E. R. Flavell, et F. L. Green, « Development of the appearance-reality distinction », *Cognitive Psychology*, 15, 1983, p. 95-120.

communication, notamment avec l'idée d'un transfert naturel de croyance ou de désir.

La simulation permet de maintenir simultanément à l'esprit deux modèles mentaux incompatibles, par exemple le modèle du monde tel que je l'ai trouvé, dans lequel il pleut, et le modèle du monde tel qu'autrui pourrait le trouver, dans lequel il fait beau. Je crois qu'il pleut, et j'imagine qu'il fasse beau (par exemple, pour en tirer des conséquences additionnelles sur le monde dans lequel autrui croit être). Intuitivement, la maîtrise de la notion de croyance fausse requiert davantage que la construction de deux modèles de ce genre. Elle requiert la construction d'un modèle plus complexe, dans lequel la croyance coexiste avec un fait qui la rend fausse. Dans un tel modèle, la croyance est *explicitement* représentée dans un monde possible où elle est fausse. Or, comme les partisans de la théorie de la simulation mentale le revendiquent eux-mêmes, les modèles mentaux que la capacité de simulation permet de manipuler ne font aucune référence explicite aux croyances des sujets.

Notre théorie méta-représentationnelle de l'esprit fait apparaître les croyances comme des *objets* (au sens large, frégéen du terme) relativement indépendants de ce qui les rend vrais ou faux dans le monde réel. Ainsi, je peux parler de la croyance de Marie, tout en ignorant si elle est vraie ou fausse. Par suite, nous sommes capables d'identifier la même croyance à travers deux mondes (ou deux ensembles de mondes) possibles : un monde (ou un ensemble de mondes) dans lequel elle est vraie, et un monde (ou un ensemble de mondes) dans lequel elle est fausse. L'expression d'une telle théorie de l'esprit suppose un langage dont la syntaxe est plus riche que celle d'un langage purement opératoire. Apparemment, seule l'utilisation de la modalité *de re* permet d'identifier une croyance *à travers les mondes*, et ainsi d'exprimer l'idée que la même croyance peut être vraie dans un monde et fausse dans un autre :

(1) Ma croyance que *p* est vraie, mais il est possible qu'elle soit fausse.

L'énoncé (1) est *de re* relativement à ma croyance, au sens où l'anaphore « elle », qui apparaît dans la portée d'un opérateur modal (« il est possible que »), a pour antécédent la description « ma croyance que *p* », qui figure hors de la portée de cet opérateur. Quiconque comprend cet énoncé et le pose comme vrai est capable de donner un sens à l'idée que la *même* croyance, vraie dans le monde réel, est fausse dans quelque monde possible. La compréhension de cet énoncé va au-delà de la capacité de simulation, et dépend de la maîtrise d'une théorie méta-représentationnelle de l'esprit. Contrairement à la capacité de simulation, cette théorie dépend en effet d'une ontologie explicite d'états mentaux, comme les croyances. Une telle ontologie nous permet de concevoir la fausseté possible de certaines de nos croyances présentes³¹.

31. Je ne prétends pas ici qu'une ontologie qui réifie les états mentaux soit la seule manière de dépasser la dichotomie, mentionnée vers la fin de la section précédente, entre la réalité du monde tel que je l'ai trouvé et les mondes imaginaires tels qu'autrui ou moi-même pourrait les trouver. Je crois par contre que cette ontologie est celle de notre psychologie ordinaire.

Quelques points de clarification sont de rigueur. Premièrement, la thèse envisagée ici, selon laquelle notre théorie méta-représentationnelle de l'esprit suppose une ontologie d'états mentaux comme des objets, est indépendante de la question de savoir si ces objets doivent être tenus pour des objets concrets, tels des états cérébraux, ou pour des objets abstraits, tels des nombres. Par exemple, Helen Steward a argué que la distinction entre type et occurrence, qui s'applique surtout aux objets concrets, fait défaut dans le cas des croyances (de même qu'elle fait défaut dans le cas des nombres)³². Peut-être notre théorie méta-représentationnelle de l'esprit est-elle neutre en ce qui concerne la nature ontologique des croyances.

En second lieu, la thèse envisagée ici laisse ouverte la possibilité d'analyser plus avant la croyance, par exemple, comme une véritable relation à une *proposition*. Je ne suis pas persuadé qu'une telle analyse soit correcte, mais elle à première vue compatible avec la notion d'une croyance dont l'existence ou l'exemplification est relativement indépendante de ce qui la rend vraie ou fausse dans le monde.

8. Notre préférence naturelle pour la vérité

La capacité de méta-représentation rend possible une théorie de l'esprit essentiellement différente de celle qui est accessible aux capacités qui, comme la perception et la simulation, ne manipulent que des représentations de premier ordre. Elle permet l'attribution d'attitudes propositionnelles comme la croyance et le désir, et avec elle l'apparition de contraintes de rationalité plus complexes³³. Toutefois, il n'est pas évident que ces nouvelles possibilités de mentalisation rendent caduc le système informationnel sur lequel repose, selon le modèle perceptif, la communication linguistique ordinaire. En particulier, elles ne nous obligent pas à considérer la compréhension et l'adhésion comme deux *étapes distinctes* du processus de compréhension.

Certes, seule la capacité de méta-représentation permet d'identifier une représentation mentale comme telle, en la posant comme un objet relativement indépendant de ce qui la rend vraie ou fausse. Mais l'identification n'est pas la compréhension : on peut identifier une représentation sans la comprendre. Par exemple, je peux parler de la croyance que Marie vient d'exprimer, alors que je n'ai pas compris ce qu'elle a dit. L'intuition des philosophes « naturalistes », comme Condillac, Reid et plus tard Wittgenstein, est que la compréhension a essentiellement son origine dans l'expression. C'est parce que j'ai la disposition naturelle de répondre à des faits que je suis capable de

32. H. Steward, *The Ontology of Mind. Events, Processes, and States*, Oxford, Clarendon Press, 1997. Par exemple, on parle de la croyance que *p*, pour une proposition *p* donnée, et non pas des croyances que *p*. On parle aussi de deux personnes qui *partagent* une croyance.

33. Je pense ici aux différentes formes de *holisme* qui pèsent sur l'attribution des croyances et des désirs, et qui ne concernent pas les états mentaux relationnels directement observables dans le comportement.

les représenter³⁴. Lorsque la capacité de simulation est en place, ces réponses peuvent être différées, mais elles restent des réactions spontanées à une sollicitation réelle ou imaginée de l'environnement. De manière générale, la *compréhension* d'une représentation propositionnelle repose sur la simulation d'un *fait*. Comprendre une représentation, c'est faire comme si elle était *vraie*. Plus précisément, c'est se mettre dans un état épistémique analogue à celui dans lequel on serait si on *savait* que *p*.

En particulier, on établit qu'autrui croit que *p* essentiellement en simulant le monde tel qu'il l'a trouvé. Si les croyances sont *conçues* comme des représentations, elles ne sont pas directement *données* comme telles. Ce à quoi nous avons directement accès sont, au mieux, des relations épistémiques aux faits. Autrement dit, l'ontologie de l'interprétation introduit les représentations mentales comme des objets, mais son épistémologie reste simulationniste. C'est pourquoi le développement de la capacité de méta-représentation permet de *réinterpréter* les transactions du système informationnel comme impliquant des croyances et des désirs soumis à des exigences spécifiques de rationalité, mais il ne supprime pas le système lui-même, sur lequel repose toujours notre saisie du contenu de ces attitudes³⁵.

Le lien constitutif entre la compréhension et la simulation nous permet ainsi d'envisager un nouvel argument en faveur de la thèse qui lie la compréhension à l'adhésion spontanée ou *prima facie*, telle que la formule ici Burge :

Ce n'est pas seulement la rationalité d'une source qui témoigne d'un lien *prima facie* avec la vérité. Le contenu même d'un message intelligible présenté comme vrai le fait également. Car le contenu a une dépendance constitutive, en première instance, à l'égard de liens structurés (*patterned*) avec un sujet (*subject matter*). Dans des circonstances normales, ces liens garantissent une base de présentation de pensées vraies. Ainsi, le fait que les présentations ont un contenu doit avoir une origine dans le fait de présenter les choses correctement. La rationalité *prima facie* de la source renforce un lien *prima facie* avec la vérité déjà présent dans l'existence *prima facie* du contenu présenté³⁶.

Cette thèse n'implique pas logiquement le modèle perceptif de la communication. Mais elle sous-tend peut-être ce que Reid appelle le *principe de véracité* (qui concerne la propension de l'émetteur à dire la vérité) et le *principe de crédulité* (qui concerne la propension du récepteur à croire autrui).

34. Ici, la nature inclut notre « seconde nature », au sens de J. McDowell, *Mind and World*, 2^e édition (1^{re} édition 1994), Cambridge (Mass.), Harvard UP, 1996. Ce que l'on veut dire par là est, entre autres, que si les réponses spontanées des protagonistes des scénarios canoniques restent effectivement liées à l'action, celles d'un sujet conceptuel peuvent donner lieu à des états épistémiques détachés du contexte pragmatique actuel.

35. Sur la dépendance de la saisie du contenu des attitudes propositionnelles par rapport à la simulation, cf. aussi les travaux de Jane Heal, par exemple J. Heal, « Simulation, Theory, and Content », in P. Carruthers et P. K. Smith, dir., *Theories of Theories of Mind*.

36. T. Burge, « Content Preservation », p. 471.

Normalement, l'affirmation est l'expression spontanée d'un fait. Lorsque l'émetteur affirme que p , il est dans un état épistémique analogue à celui dans lequel il serait s'il savait que p . La réponse épistémique spontanée de l'émetteur au fait que p est alors transférée au récepteur. Même si celui-ci n'a pas les moyens, *indépendamment du témoignage de l'émetteur*, d'entrer en relation cognitive avec le fait que p , sa compréhension du message le place dans un état épistémique au moins analogue à celui dans lequel il serait s'il savait que p — c'est-à-dire un état dans lequel il peut *compter* sur le fait que p .

J'ajoute à cette description sommaire d'un acte de communication assertorique que la simulation dépend d'une base de réponses spontanées *non simulées* à des faits. Le flux de l'information du monde à l'émetteur, et de celui-ci au récepteur, ne peut pas être toujours simulé. Dans un contexte normal, il n'y a aucun obstacle à ce que la communication procède selon les termes du modèle perceptif et donc qu'il y ait réellement un transfert d'information. À nouveau, lorsque nous avons des raisons de mettre en doute la fiabilité de notre interlocuteur (sa bonne foi ou son intégrité épistémique), nous résistons à l'adhésion, mais lorsque nous n'avons aucune raison de ce genre, celle-ci l'emporte naturellement.

9. Conclusion

Dans cet essai, j'ai essayé de montrer que l'interprétation ordinaire des propos d'autrui engage au moins trois niveaux de représentation : le niveau perceptif, le niveau de la simulation et le niveau de la méta-représentation. Ces niveaux correspondent respectivement aux signes naturels, aux signes d'institution d'un langage opératoire et aux signes d'institution d'un langage syntaxiquement plus riche qui permet la construction de compléments propositionnels quantifiables. Selon la conception que j'ai développée, une analyse adéquate de la nature de la méta-représentation montre que celle-ci ne rivalise pas avec l'existence de transferts naturels de croyances ou de désirs de l'émetteur au récepteur d'un acte de communication assertorique. L'enracinement de la compréhension dans la perception et la simulation accrédite la thèse selon laquelle comprendre un message propositionnel, c'est déjà être enclin à y donner son assentiment. La communication est à la fois hautement réflexive et spontanée. Si cette thèse ne suffit pas à réfuter le modèle inférentiel de la communication, il est probable qu'elle soit appelée à jouer un rôle crucial dans la défense du modèle perceptif.