

# Le langage forme-t-il une condition nécessaire de la rationalité?

JOËLLE PROUST C.N.R.S., Institut Jean-Nicod

Dans ce petit livre dense et enlevé, Ronald de Sousa réfléchit sur la manière dont l'évolution permet d'éclairer les particularités déroutantes de la rationalité humaine. La question est complexe, mais le travail interdisciplinaire qui est effectué depuis une quinzaine d'années permet de la poser dans des termes nouveaux. Ronald de Sousa pour sa part aborde le problème en relevant ce qui ressemble à un paradoxe : «la rationalité tant vantée de l'être humain est surtout fondée sur notre capacité de nous montrer irrationnels» (2004, p. 3). Il y a deux façons, remarque-t-il, d'être hors rationalité : soit en n'en relevant pas, soit en l'enfreignant. Comprendre l'évolution de la rationalité permet de distinguer l'arationalité de l'irrationalité. Elle permet aussi de dissoudre le paradoxe.

Quel est le type de manquement qui donne lieu à l'irrationalité? La réponse de l'auteur est qu'il résulte du *design* de la rationalité, plutôt que d'une faillite caractérisée de la raison sous l'empire d'une force extérieure à elle (passion, imagination, etc.). Ce qu'on sait de l'évolution permet en effet d'exclure qu'ait été sélectionné un système unique d'arbitrage rationnel indifférent à son application. Des «formes» d'arbitrage (des algorithmes, soit des pratiques inférentielles automatisées) ont plutôt émergé à la faveur d'un bricolage modulaire, en fonction de besoins apparus successivement au cours de la phylogénèse. D'où les diverses formes de la rationalité humaine. La rationalité «stratégique» optimise les procédures, étant donné des buts. La rationalité «axiologique» détermine les normes de pertinence des buts : elle s'occupe de fixer les utilités. La rationalité «épistémique» s'intéresse à la justification des actions par des raisons.

*Dialogue XLVI* (2007), 1-00

© 2007 Canadian Philosophical Association /Association canadienne de philosophie

## 2 Dialogue

L'irrationalité peut se produire aux trois niveaux : elle traduit un conflit entre les modules. Ce conflit peut-il trouver un arbitrage non-modulaire au niveau «épistémique»? Telle est la première question cruciale que soulève l'ouvrage et qu'à mon sens il résout avec brio — en répondant affirmativement — à la lumière de considérations évolutionnistes.

C'est la première opposition qui me retiendra dans ce qui suit, celle qui oppose le rationnel à l'arationnel, et donc définit le territoire de la rationalité, (où les infractions prennent sens). Le repérage de ce territoire trouve ici aussi des éléments de réponse pour Ronald de Sousa, dans la sélection naturelle — en analysant ses limites. L'argumentation est la suivante.

- 1) La sélection naturelle est productrice de finalité et de normativité.
- 2) La capacité représentationnelle a été sélectionnée relativement tôt dans la phylogénèse : il y a des précurseurs de la rationalité dans (ce qu'il nomme) «l'infra-humain».
- 3) Mais une des distinctions capitales pour la rationalité est celle qui oppose la simple représentation des buts et la représentation des possibilités alternatives. La première caractérise la simple finalité biologique, tandis que la seconde permet à l'individu de faire valoir ses propres valeurs.
- 4) L'expression langagière est essentielle à la représentation des possibilités alternatives et des valeurs individuelles.

### 1. Langage, universalité, particularité, spécificité

Les deux premiers points mériteraient discussion, mais, pour simplifier le débat, je supposerai bien fondées la définition étiologique de la fonction et la conception de la normativité qui est censée en découler. Le troisième est *la* thèse stratégique, qui ancre l'essence du rationnel dans une capacité proprement humaine. Les deux formes de représentation invoquées s'opposent selon Ronald de Sousa :

- 1) par leur faculté inégale de penser l'inexistant;
- 2) par la manière dont elles représentent la généralité. Au premier sens, *général* s'oppose à *spécifique*, au second, à *particulier* (2004, p. 72).

Le critère de l'inexistence est emprunté à Chisholm (1957) commentant Brentano (1874) dans le cadre d'un plaidoyer pour des «critères gramma-

ticaux» de l'intentionnalité<sup>1</sup>. Dans l'interprétation de Chisholm, ce critère pose que, pour qu'un être quelconque soit pourvu de la capacité représentationnelle, il doit pouvoir se représenter des entités inexistantes, c'est-à-dire former des représentations éventuellement vides. Selon Chisholm il est essentiel à l'intentionnalité que puisse être *enfreint* le principe de la généralisation d'un jugement existentiel (principe valide dans les contextes non-intensionnels) : dire que Paul croit que le Père Noël passe par la cheminée n'implique pas que le Père Noël existe en fait, alors que dire que deux corps sont soumis à la gravité implique que les corps existent.

Le critère de généralité, développé par Strawson, pose que, pour tout système doué d'intentionnalité, il n'existe pas de limitation a priori entre les combinaisons sémantiques possibles autorisées par la syntaxe : si je peux penser f(a), je dois pouvoir penser g(a) ou f(b). Ou, comme le commente Ronald de Sousa, «la pensée est capable d'envisager [...] des objets en général». Précisons ici que le principe de généralité tel qu'on vient de le formuler s'établit une fois que l'on a introduit un autre principe, le *principe de l'objectivité*, en vertu duquel un être pensant doit pouvoir réidentifier des individus numériquement identiques, ce qui suppose davantage que de reconnaître des propriétés sortales, aussi spécifiques qu'elles soient.

Ronald de Sousa observe que l'animal est capable de représenter le spécifique, non le particulier, en vertu d'au moins deux limitations : il ne peut pas penser l'inexistant, et manque aussi de ce fait au principe de généralité ; il ne peut pas penser l'identité numérique de l'individu, et par conséquent il ne peut pas faire référence au particulier «en tant que tel» (2004, p. 73). En d'autres termes, les animaux sont dépourvus de ce que Kripke (1982) appelle des «désignateurs rigides» : des symboles langagiers ou mentaux qui font référence à un individu numériquement identique, indépendamment de ses propriétés apparentes.

Ces deux critères sont-ils en mesure d'opérer la rupture que Ronald de Sousa cherche à effectuer entre êtres dotés ou non d'intentionnalité — précondition de la rationalité? Prenons d'abord «l'inexistence intentionnelle». Notons que ce critère n'est Brentanien que par l'un des contresens les plus persistants de l'histoire de la traduction philosophique. Les latinistes savent qu'*inexistere* (ou *inexistere*) n'est nullement privatif, et veut dire «exister dans». En faisant de l'inexistence la propriété des états mentaux, Brentano entend souligner le fait que les phénomènes mentaux sont «les phénomènes qui contiennent un objet intentionnellement à l'intérieur d'eux-mêmes»<sup>2</sup>.

Au-delà de la fidélité à Brentano, le critère de l'inexistence de Chisholm-de Sousa met la barrière de l'intentionnalité trop haut<sup>3</sup>. Il est vrai qu'une représentation reste une représentation qu'elle soit vraie ou fausse, satisfaite ou non, ait ou non un référent dans la réalité, et qu'un mécanisme de révision (l'équivalent pratique de la reconnaissance de l'inadéquation d'une représentation) est nécessaire à tout système intentionnel. Mais

#### 4 Dialogue

cette condition n'équivaut pas à requérir d'un tel système qu'il puisse explicitement reconnaître son erreur, *c'est-à-dire s'attribuer à lui-même une croyance reconnue comme sans objet ou comme fausse*. Pour comprendre ce qu'est un concept vide, ou une représentation sans instantiation, comme «licorne» ou «Père Noël», il faut posséder une «théorie de l'esprit», c'est-à-dire comprendre comment soi-même ou un autre peuvent former des représentations fausses ou vides. Cette capacité est uniquement maîtrisée par l'être humain adulte<sup>4</sup>. Le critère de la compréhension des représentations vides a un autre inconvénient pour qui cherche à comprendre les formes minimales de l'intentionnalité : il suppose que la métareprésentation fasse son apparition en même temps que la capacité représentationnelle, ce qui contrevient aux hypothèses du développement graduel, «machiavellien», de l'intelligence sociale.

Quant au deuxième critère, qui porte sur l'objectivité, Ronald de Sousa rappelle qu'on ne peut pas passer d'une caractérisation conceptuelle aussi précise qu'elle soit à la détermination d'un individu numériquement identique. Il voit là une impossibilité logique, qui traduit une propriété profonde de l'intentionnalité, celle de pouvoir faire référence à un individu «nu», un «bare particular», qui reste numériquement unique même s'il dispose d'un double qualitatif (pourvu des mêmes propriétés). Cette capacité référentielle, selon l'auteur, est inséparable «de la maîtrise d'un langage suffisamment riche pour pouvoir distinguer un nom propre d'un nom commun» (2004, p. 74). Le lecteur est tenté d'objecter que l'animal social peut discerner les individus qui composent son groupe. Mais de Sousa sort alors une nouvelle carte : si les chauve-souris peuvent distinguer les individus coopératifs du groupe, elles ne peuvent pas «établir une distinction entre un individu qui porterait une parfaite ressemblance à tel compagnon et ce compagnon lui-même [...]. Seul un acte de métacognition permettrait d'établir une distinction notionnelle entre la situation où il reconnaît un congénère et la situation où il croit le reconnaître à tort» (2004, p. 82-83).

De nouveau ici, le naturaliste aura du mal à admettre une condition aussi forte, dans laquelle on ajoute à l'objectivité, condition nécessaire de l'intentionnalité, un mode de reconnaissance qui a été sélectionné ultérieurement (nous allons y revenir). Ronald de Sousa a raison de soutenir qu'un animal ne peut avoir l'équivalent de conditions de correction de ses pensées que s'il peut réidentifier un individu comme le même. Mais pourquoi exclure que l'animal non-langagier puisse le faire? Il existe en effet des mécanismes *non-langagiers* de correction du recueil de l'information qui ont la fonction de catégoriser des proto-objets. Ces mécanismes perceptifs confèrent aux animaux qui en sont pourvus (entre autres : serpents, oiseaux, mammifères) l'aptitude de faire référence à des objets représentés comme stables indépendamment de leurs propriétés<sup>5</sup>. Qu'une telle chose soit possible est confirmé par les travaux de Ballard et de Pylyshyn, qui

montrent que la référence directe à des objets effectuée par les indexicaux langagiers peut également être obtenue par ce qu'on appelle des «indexicaux visuels»<sup>6</sup>. Ces différents mécanismes référentiels non-langagiers suggèrent d'élargir le domaine du représentationnel en y admettant les contenus non-conceptuels — par ailleurs requis pour une pleine compréhension du figuratif et de l'analogique dans la pensée.

C'est une chose pourtant de faire référence «*de re*» à un individu, autre chose de se représenter que l'individu pourrait, en dépit des apparences, être un autre que celui que l'on croit reconnaître. Cette nouvelle capacité suppose l'intervention d'un découplage entre des plans de référence (réel, potentiel, contrefactuel, fictionnel)<sup>7</sup>. Ce découplage paraît de nouveau n'être pleinement accessible qu'au sujet humain adulte, même si le jeu de faire semblant et la tromperie tactique sont des formes de découplage «implicite» présentes chez l'animal non-humain.

Pour conclure sur ce point, Ronald de Sousa défend l'idée qu'un être intentionnel doit pouvoir disposer :

- 1) d'un langage armé d'indexicaux et de noms propres;
- 2) d'une capacité métareprésentationnelle.

À mon sens, l'éthologie et la psychologie cognitive de la vision permettent de rejeter la première condition ; la deuxième condition ne répond pas à ce que l'on sait de l'évolution de l'intentionnalité : la capacité représentationnelle a dû exister avant que se développe une capacité métareprésentationnelle en vertu des pressions machiavelliennes instaurées par la première.

## **2. De l'intentionnalité à la rationalité : de la simple représentation des buts à la représentation des possibilités alternatives?**

Un système intentionnel est-il automatiquement rationnel? Pour Ronald de Sousa, la réponse est clairement négative. La rationalité suppose une capacité de réflexion langagière sur soi et sur ses valeurs qui est inaccessible aux animaux non-humains. L'inconvénient de cette conception est qu'elle ne permet pas de penser l'évolution de la rationalité<sup>8</sup>. Dans la mesure où la représentation des buts permet un contrôle flexible du comportement, et implique que soient effectués des compromis entre l'«exploration» du domaine et son «exploitation», il semble justifié de considérer qu'une dimension essentielle de la rationalité stratégique est déjà présente chez les mammifères et les oiseaux, pour ne citer qu'eux. Les théoriciens de l'apprentissage animal en donnent maint exemple<sup>9</sup>.

La rationalité stratégique ouvre la voie à de nouvelles pressions sélectives pour qu'émerge la rationalité métacognitive grâce à laquelle l'animal peut contrôler ses comportements sur la base de la qualité information-

## 6 *Dialogue*

nelle de ses propres états cognitifs<sup>10</sup>. Des animaux de diverses espèces, grands primates, singes, dauphins et mêmes oiseaux, s'avèrent capables d'évaluations qu'on croyait propres à l'homme : ils peuvent apprécier leurs propres chances de réussir à une tâche, et choisir entre diverses modalités d'action celle qui est la plus appropriée dans un contexte donné. C'est là une forme de métacognition qui semble ne pas requérir de capacité métareprésentationnelle; en effet, les mêmes animaux échouent aux tâches de théorie de l'esprit<sup>11</sup>. Faut-il suivre Ronald de Sousa en considérant qu'il s'agit là de formes non rationnelles (arationnelles) de comportement parce que non-réflexives? Ou bien lui fausser compagnie, pour explorer les atouts de formes implicites, procédurales, non langagières, de réflexivité?

On ne peut répondre à cette question principielle qu'en s'appuyant sur des raisons de méthode. Pour l'auteur, la rationalité ne devient une question déterminée que si l'agent expose verbalement ses raisons. Tandis que chez l'animal non langagier, «les paramètres sont impossibles à démêler» (2004, p. 106). La difficulté de méthode est réelle, mais justifie-t-elle l'abandon de toute tentative de «démêlage des paramètres»? Le débat philosophique entre les conceptions réalistes et attributives des états mentaux n'est pas clos; cependant les naturalistes peuvent estimer que le pessimisme méthodologique est réfuté par le progrès des recherches comparatives. Il est parfaitement possible d'étudier les états épistémiques et motivationnels des animaux et des bébés, en testant ce qu'ils apprennent d'une situation donnée ou en examinant comment ils délibèrent devant des alternatives. Le langage est-il quant à lui un outil attributif méthodologiquement fiable? L'une des leçons que l'on tire de l'ouvrage de Ronald de Sousa est qu'on ne peut pas croire un agent sur parole quant à ses motifs, lesquels sont très largement influencés par ses émotions du moment. C'est certainement plutôt dans le registre exécutif que dans le registre représentationnel que le langage a joué le rôle essentiel pour l'évolution de la rationalité : le langage fournit une mémoire de travail compacte, combinable, et extériorisable. Il permet de disposer d'une information considérable pour guider la décision; il permet de découpler les plans de référence. Il permet enfin la communication des représentations et des motivations d'un agent à l'autre, c'est-à-dire la culture, qui devient un creuset de valeurs partagées. La rationalité ne naît pas avec le langage : mais avec lui, elle se plie aux contraintes de la justification verbale, contraintes qui ont plus à avoir avec les nécessités de la coordination sociale qu'avec la causalité de l'action. L'une des manières de juger de l'impact du langage sur la rationalité humaine serait d'intégrer plus systématiquement à l'enquête sur l'évolution de la rationalité les données de la physiologie cérébrale, de la psychologie comparative et de l'anthropologie cognitive. Mais avoir posé le problème n'est pas l'un des moindres mérites de l'ouvrage de Ronnie de Sousa.

## Notes

- 1 Selon les partisans de la conception «grammaticale» «ou linguistique» de l'intentionnalité, celle-ci présuppose que des conditions portant sur l'usage logique des termes utilisés pour attribuer l'intentionnalité, usage caractérisé par leur «intensionnalité».
- 2 On trouvera une explication de texte détaillée dans McAlister (1974) et dans Jacob (2004). Cf. aussi Dupuy, (1994).
- 3 Jacob (2004) apporte d'autres arguments en ce sens.
- 4 Peut-être aussi dans certains contextes par les chimpanzés: Tomasello *et al.* (2003).
- 5 Voir sur ce point le rôle de la recalibration multimodale dans Proust (1997) et (1999).
- 6 Sur l'hypothèse de l'indexage visuel, cf. Ballard (1997), Pylyshyn (2001).
- 7 On appelle *découplage* le processus qui permet de former des représentations alternatives d'une même situation perçue, imaginée ou conçue. Cf. Proust, (2002a, 2002b, 2003), Sterelny (2003).
- 8 Godfrey-Smith (1996).
- 9 Cf. Gallistel (1990).
- 10 Proust (2006).
- 11 Cf. Smith *et al.* (2003).

## Références bibliographiques

- Ballard, D. H., M. M. Hayhoe, P. K. Pook, et R. P. N. Rao  
1997 «Deictic Codes for the Embodiment of Cognition», *Behavioral and Brain Sciences*, vol. 20, n° 4, p. 723-757.
- Brentano, F.  
1874 *Psychologie vom Empirischen Standpunkt*, 3 vol., Leipzig, Felix Meiner Verlag.
- Call, J., et M. Tomasello  
1999 «A Non-Verbal Theory of Mind Test: The Performance of Children and Apes», *Child Development*, vol. 70, p. 381-395.
- Chisholm, R.M.  
1957 *Perceiving : A Philosophical Study*. Ithaca [NY], Cornell University Press.
- De Sousa, R.  
1984 «Teleology and the Great Shift», *The Journal of Philosophy*, vol. 81, p. 647-653.  
2004 *Évolution et rationalité*, Paris, Presses universitaires de France.
- Dupuy, J. P.  
1994 *Aux origines des sciences cognitives*, Paris, Éditions La Découverte.
- Gallistel, R. C.  
1990 *The Organization of Learning*, Cambridge [MA], MIT Press.

8 *Dialogue*

- Godfrey-Smith, P.  
1996 *Complexity and the Function of Mind in Nature*, Cambridge [MA], Cambridge University Press.
- Hampton, R. R.  
2001 «Rhesus Monkeys Know When They Remember», *Proceedings of the National Academy of Sciences USA*, vol. 98, p. 5359-5362.
- Jacob, P.  
2004 *L'intentionnalité*, Paris, Odile Jacob.
- Kripke, S. A.  
1982 *La logique des noms propres*, Paris, Minuit.
- McAlister, L.  
1974 «Chisholm and Brentano on Intentionality», *Review of Metaphysics*, vol. 28, n° 2, p. 328-338.
- Proust, J.  
1997 *Comment l'esprit vient aux bêtes*. Paris, Gallimard.  
1999 «Mind, Space and Objectivity in Non-Human Animals», *Erkenntnis*, vol. 511, p. 41-58.  
2002a «Imitation et agentivité», dans J. Nadel et J. Decety, dir., *Imiter pour découvrir l'humain*, Paris, Presses universitaires de France  
2002b «Can "Radical" Theories of Simulation Explain Mental Concept Acquisition?», dans J. Dokic and J. Proust, dir., *Simulation and Knowledge of Action*, Amsterdam, John Benjamins, p. 201-228.  
2003 *Les animaux pensent-ils?* Paris, Bayard.  
2006 «Rationality and Metacognition in Non-Human Animals», dans S. Hurley et M. Nudds, dir., *Rational Animals?*, Oxford, Oxford University Press.
- Pylyshyn, Z.  
2001 «Visual Indexes, Preconceptual Objects, and Situated Vision», *Cognition*, vol. 80, p. 127-158.
- Smith, J. D., W. E. Shields, et D. A. Washburn  
2003 «The Comparative Psychology of Uncertainty Monitoring and Metacognition», *Behavioral and Brain Sciences*, vol. 26, n° 3, p. 317- 373.
- Sterelny, K.  
2003 *Thought in a Hostile World : The Evolution of Human Cognition*, Oxford, Blackwell.
- Tomasello, M., J. Call, et B. Hare  
2003 «Chimpanzees Understand Psychological States — The Question Is Which Ones and to What Extent», *Trends in Cognitive Science*, vol. 74, p. 153-156.