

Gradhiva

Revue d'anthropologie et d'histoire des arts

15 | 2012

Robots étrangement humains

Aux origines de l'anthropomorphisme

Intersubjectivité et théorie de l'esprit

The origins of anthropomorphism. Intersubjectivity and theory of mind

Gabriella Airenti

Traducteur : Matteo Severi



Édition électronique

URL : <http://journals.openedition.org/gradhiva/2314>

DOI : 10.4000/gradhiva.2314

ISSN : 1760-849X

Éditeur

Musée du quai Branly Jacques Chirac

Édition imprimée

Date de publication : 16 mai 2012

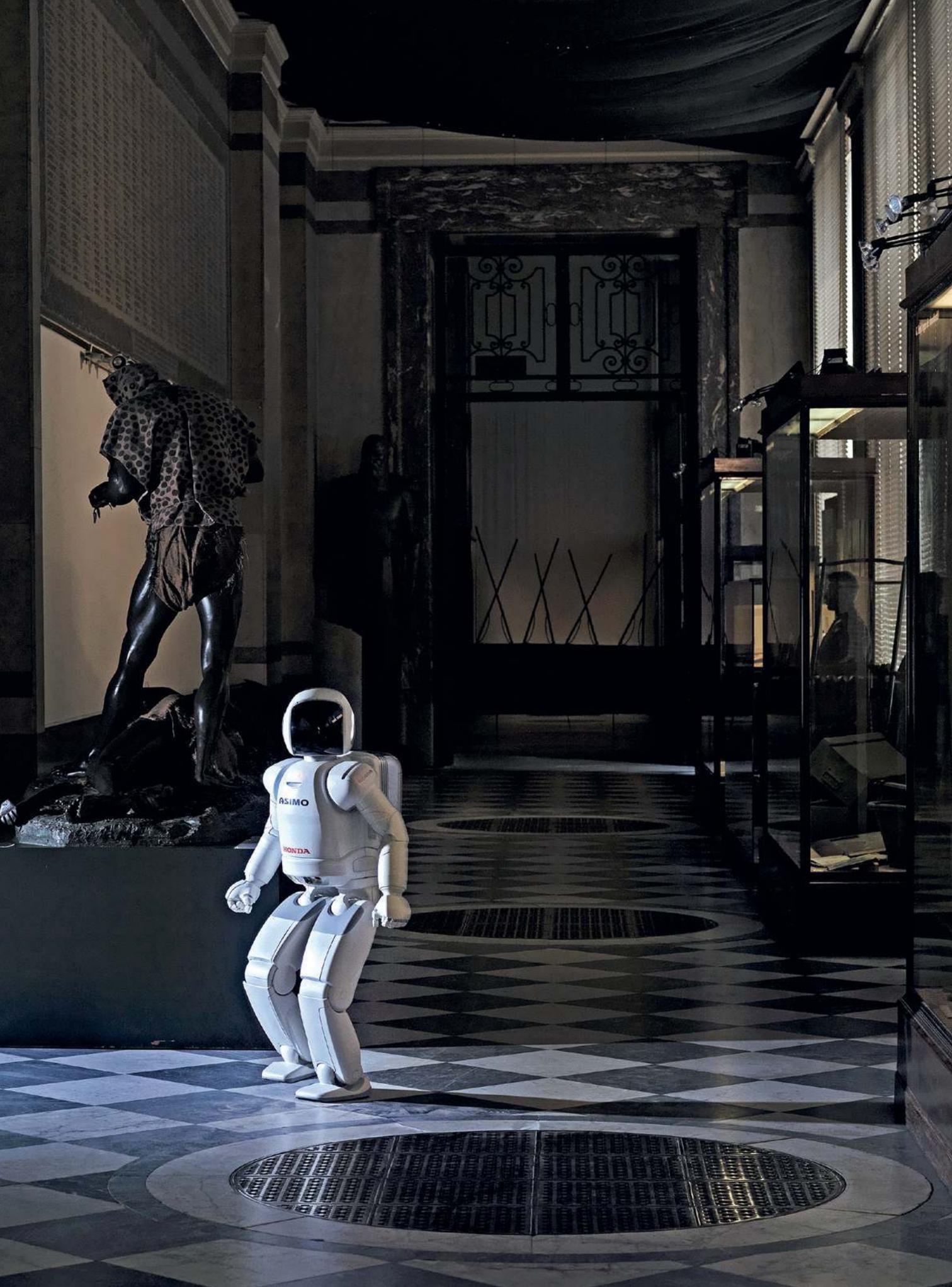
Pagination : 34-53

ISBN : 978-2-35744-047-0

ISSN : 0764-8928

Référence électronique

Gabriella Airenti, « Aux origines de l'anthropomorphisme », *Gradhiva* [En ligne], 15 | 2012, mis en ligne le 16 mai 2015, consulté le 30 avril 2019. URL : <http://journals.openedition.org/gradhiva/2314> ; DOI : 10.4000/gradhiva.2314



Aux origines de l'anthropomorphisme

Intersubjectivité et théorie de l'esprit

Gabriella Airenti

Fig. 1 Le robot Asimo de Honda
© Photo Vincent Fournier -
vincent@vincentfournier.co.uk

• • •

1. Dans certaines sociétés, les relations que les humains établissent avec les animaux avec qui ils partagent leur habitat sont l'expression directe de phénomènes d'anthropomorphisme. Les animaux sont des interlocuteurs avec lesquels on établit un lien d'empathie, et l'une des modalités fondamentales pour la manifestation de ce lien est l'imitation. Une analyse des relations entre les hommes et les animaux qui va dans ce sens a été faite par Florence Brunois (2005) auprès des Kasua de Nouvelle-Guinée. Des formes complexes d'imitation comme moyen d'établir une relation avec un interlocuteur animal ont été observées dans une tout autre région, chez les Quechua de l'Amazonie péruvienne (Gutierrez Choquevilca 2010).

Dès la première enfance, les humains attribuent des traits de comportement anthropomorphe – tels des intentions, des perceptions et des sentiments – aux animaux, aux artefacts et même aux phénomènes naturels. Pourquoi cette attitude est-elle si répandue? Nous nous proposons ici d'analyser les bases psychologiques de l'anthropomorphisme quotidien, et dans cette perspective nous pensons qu'il est utile d'étudier l'émergence des premières formes de la pensée anthropomorphique chez le jeune enfant pour retrouver ainsi les racines de ses manifestations dans la pensée adulte. Il semble en effet que les êtres humains tendent à avoir une pensée dichotomique en la matière. Ils comprennent les êtres doués de mouvement de deux façons : soit l'objet en question est vu comme l'expression d'un dessein prédéterminé – qu'il soit biologique ou mécanique – et est alors assimilé à une machine; soit il semble doté, au moins en apparence, d'une vie indépendante, et se voit attribuer des caractéristiques humaines et une vie mentale propre. Il est remarquable que cela se produise également lorsque nous interprétons, par exemple, la vie animale¹. Dans nos interactions quotidiennes avec les animaux, nous avons en effet tendance à leur attribuer des émotions et des états mentaux typiquement humains, mais la question se pose également dans le domaine scientifique, en particulier dans l'étude des primates. Une grande partie du débat entre éthologues vise aujourd'hui à établir si les primates sont capables de formes de pensée et de socialisation semblables à celles des humains ou si, au contraire, leur comportement est déterminé par quelques schèmes simples orientant strictement leur vie sociale. Ce n'est que rarement que l'on prend en considération une troisième possibilité, à savoir que les primates ne s'adapteraient pas aux changements de situation de la même manière que les humains (Servais 2007); les animaux auraient donc un esprit complexe, mais différent de celui des humains.

La question de l'anthropomorphisme peut aussi être traitée autrement, par l'analyse des relations qui s'instaurent entre les êtres humains et les artefacts avec lesquels ils interagissent. Cet aspect émerge avec une force particulière dans le fameux travail de Masahiro Mori sur l'*uncanny valley* (dont la traduction est publiée dans ce numéro). Dans cet essai, Mori a étudié le degré de plus ou moins grande familiarité qui transparaît dans les réactions des individus face aux robots. Comment un robot doit-il être conçu pour être accepté au mieux par les humains ? La réponse de Mori est qu'une machine qui se présente en tant que telle ne provoque aucun sentiment de familiarité. Seuls les robots qui reproduisent certaines caractéristiques humaines, sans toutefois être des répliques parfaites de l'être humain, sont en mesure de susciter ce sentiment. Enfin, Mori considère que l'imitation peut avoir un effet paradoxal : les robots qui imitent le plus fidèlement les caractéristiques humaines provoquent une perte complète de familiarité. Dans ce cas, le robot sera perçu comme un objet monstrueux. En effet, aussi perfectionnée soit-elle, l'imitation des caractéristiques humaines sera très difficilement parfaite : un petit décalage dans la vitesse du mouvement d'un des muscles du visage suffit à transformer un sourire en une grimace artificielle et troublante. Toujours selon Mori, un bon modèle de familiarité serait fourni par les marionnettes du théâtre traditionnel japonais, qui ressemblent assez peu à de véritables êtres humains, mais qui, vues en mouvement sur scène, font illusion.

Ce point de vue n'a pas fait l'unanimité parmi les spécialistes, d'autant que le modèle de l'*uncanny valley* est plus une piste de travail qu'une hypothèse prouvée scientifiquement. En réalité, les observations de Mori intéressent aussi les chercheurs non directement impliqués dans la robotique, et ce pour deux raisons. La première est qu'elles posent la question de l'anthropomorphisme en termes relationnels. Ce choix s'oppose à ceux qui, réactualisant les questionnements classiques de l'intelligence artificielle, affirment que pour construire un robot efficace il faut se demander « qu'est-ce qu'un humain ? » (Kahn *et al.* 2007). En effet, un robot est un artefact comme n'importe quel autre, et en tant que tel il n'a pas besoin de constituer un bon modèle d'être humain. Ce qui compte, c'est que le robot est censé interagir avec des êtres humains, donc la question pertinente est plutôt : quels sont les facteurs pouvant déterminer chez les humains l'exercice de la projection anthropomorphique, et ainsi susciter de la familiarité envers ces artefacts ?

Les observations de Mori sont également intéressantes parce qu'elles montrent que l'imitation de l'apparence humaine est un faux problème. La reproduction quasi parfaite des caractéristiques humaines n'invite pas nécessairement les êtres humains à l'interaction. Une ressemblance excessive de l'artefact à l'être humain semble bloquer la projection de caractéristiques mentales humaines et en souligner au contraire l'aspect mécanique et artificiel. Le fait que l'on puisse virtuellement attribuer des caractéristiques anthropomorphes à n'importe quel objet montre que la vraisemblance n'est pas le critère fondamental. L'anthropomorphisme est un type de *relation* que l'on établit avec un animal ou un artefact, et c'est sur cette relation qu'il faut se concentrer si on veut comprendre un tel phénomène. Je vais donc examiner ici l'attribution de caractéristiques anthropomorphiques à des êtres inanimés à partir de l'étude des relations que les très jeunes enfants établissent avec des objets, animés et inanimés.



Fig. 2 Douchan Stanimirovitch, *Pantin imitant un homme habillé élégamment assis dans un fauteuil tenant un verre à la main*, 1955-1960
© Archives Alinari, Florence, Dist. RMN/Douchan Stanimirovitch.

L'anthropomorphisme chez l'enfant

Dans la littérature psychologique classique, on établissait une différence fondamentale entre la pensée de l'enfant et celle de l'adulte. Pour Jean Piaget, les enfants, jusqu'à l'âge de sept ans environ, ont du mal à distinguer entre ce qui relève du domaine physique et du domaine mental. Penser et rêver sont des actions comme les autres. La pensée est produite par la voix. L'enfant est à la fois égocentrique – dans la mesure où il comprend le monde en fonction de ses propres désirs, besoins, sensations, perceptions – et animiste, puisqu'il attribue aux phénomènes physiques des états mentaux. Entre ces deux modes, il n'y a pas de contradiction, car dans un cas comme dans l'autre il s'agit du fruit de la confusion conceptuelle de l'enfant : « ... faute d'une prise de conscience de la subjectivité de la pensée, de l'intentionnalité, de l'effort, etc., ces éléments intérieurs sont attribués à n'importe quel tableau externe susceptible de correspondre aux mouvements et à l'activité propres, par une analogie immédiate et non pas conceptuelle. » (Piaget 1978 [1945] : 268)

C'est précisément l'absence de compréhension réflexive de la dimension mentale de l'expérience qui permet au jeune enfant d'attribuer des motivations propres à un quelconque objet en mouvement. Seule l'acquisition de capacités cognitives avancées l'amènera plus tard à sortir de ces formes de pensée primitives. L'adulte serait-il immunisé contre ces formes de cognition ? Normalement oui. Selon Piaget, l'anxiété ou des sentiments très forts de peur ou de désir sont seuls capables de ramener l'adulte à des formes de pensée magique ou à des comportements superstitieux (Piaget 1926).

Cette vision de l'enfant a été largement remise en cause par la recherche récente. Nous savons maintenant qu'en situation expérimentale, à partir de l'âge de trois ans, les enfants sont en mesure de distinguer entre des actions physiques et des actions mentales. Quand on demandait aux enfants s'ils pouvaient toucher ou voir un objet physique tel qu'une paire de ciseaux ou une bicyclette, la réponse était positive, tandis qu'elle était négative lorsque la même ques-



Fig. 3 Un robot interagissant avec un enfant autiste, Hatfield, UK, 2011. AP Photo/Alastair Grant.

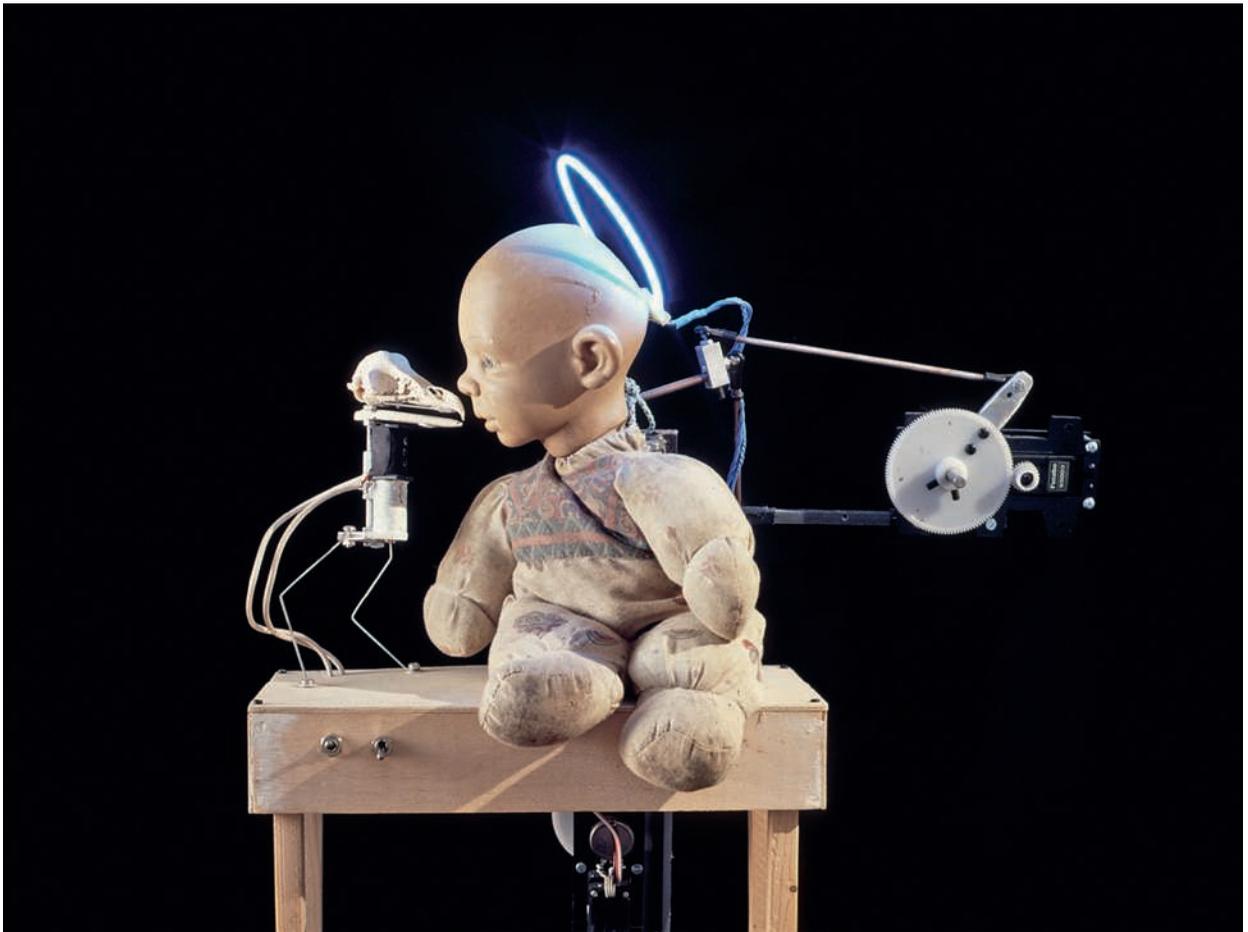


Fig. 4 Christiaan Zwanikken, *Little Franciscus*, 2004 © Christiaan Zwanikken.

tion, rapportée à des objets mentaux tels qu'une bicyclette imaginée, pensée ou rêvée, leur était posée. Cette capacité de distinction continuait de se manifester même lorsqu'on introduisait des *close impostors*, c'est-à-dire des objets ou des sensations dont la présence physique n'est pas immédiatement évidente, comme les ombres, la fumée, le son ou la douleur. Les précisions que les enfants donnaient pour motiver leurs réponses montraient qu'ils reconnaissaient bien la nature physique de ces objets, et ne les considéraient pas comme des objets mentaux au même titre que les rêves ou les pensées (Wellman 1990). À partir de ces conclusions, Henry Wellman a pu réinterpréter les résultats des expériences de Piaget. On se souviendra que les enfants étudiés par Piaget – qui leur demandait par exemple d'expliquer ce qu'est un rêve – disaient que les rêves sont de la fumée. Ils utilisaient donc un langage propre aux phénomènes physiques pour décrire des phénomènes mentaux. Piaget en déduisait que les enfants ne faisaient pas la distinction entre ce qui relève du physique d'une part et du mental de l'autre. Or, selon Wellman, cette conclusion est fallacieuse : en décrivant ainsi les phénomènes mentaux, un enfant se comporte en fait exactement comme le ferait un adulte qui dirait avoir vu quelque chose en rêve. Les

résultats de Piaget ne permettent donc pas de conclure que les enfants par lui interrogés avaient des difficultés à distinguer entre des phénomènes de nature ontologique différente (physique ou mentale), mais plutôt qu'ils avaient du mal à définir dans un langage approprié des concepts particulièrement complexes.

Les résultats que nous venons de décrire concernent des enfants âgés de trois ans. La recherche actuelle est allée au-delà, en consacrant beaucoup d'efforts à l'étude des capacités cognitives des nouveau-nés. Nous disposons désormais de techniques qui nous aident à mieux connaître la perception des très jeunes enfants et des nouveau-nés, ce qui nous a permis de considérer séparément la capacité de l'enfant à opérer des distinctions entre différents aspects de son environnement – et à y réagir de façon différenciée –, et à verbaliser et penser ces distinctions. C'est en particulier le cas pour l'identification des objets et la distinction entre objets animés et inanimés.



Fig. 5 « Ensemble mannequin d'accouchement : le nouveau-né » réalisé par Angélique Marguerite du Coudray, XVIII^e siècle. Photo Jacques Petitcolas, musée Flaubert et d'histoire de la médecine, CHU de Rouen.

Des expériences menées sur des nouveau-nés ont montré que, dès trois à quatre mois, ceux-ci sont en mesure d'identifier des objets à travers quelques propriétés de base (délimitation, cohésion, rigidité, action par contact) qui se manifestent dans le mouvement (Spelke 1990). Pour les enfants très jeunes, dont on ne peut obtenir de réponse verbale, les psychologues du développement se fondent sur des indicateurs qui leur permettent de déceler les attentes de l'enfant. Dans une de ces expériences, on montrait à des nouveau-nés une baguette en mouvement. Quand celle-ci disparaissait momentanément derrière un écran, le nouveau-né s'attendait à la voir réapparaître identique à son apparence initiale, et s'étonnait de la voir finalement scindée en deux morceaux.

L'identification d'un objet physique en tant que tel est donc implicite dans les inférences du nouveau-né. La capacité de distinguer entre les objets animés et les objets inanimés est aussi très précoce. Prenons un exemple. On donnait à des enfants âgés de douze mois des jouets de deux sortes : certains avaient la forme d'animaux, d'autres, celle de véhicules. On observait ensuite l'ordre suivant lequel les enfants les touchaient. Ils ne procédaient pas dans un ordre aléatoire, mais touchaient d'abord l'un après l'autre tous les jouets appartenant à l'une des deux catégories (les animaux, par exemple), puis tous ceux appartenant à l'autre catégorie, montrant ainsi leur capacité à distinguer les deux catégories d'objets. Malgré les similitudes possibles entre des jouets représentant des objets animés et d'autres figurant des objets inanimés – un oiseau en plastique, par exemple, ressemblait bien plus à l'avion qu'au petit chien –, les enfants ne les regroupaient pas en fonction de leur aspect physique, mais en respectant la distinction entre mouvement biologique et mouvement mécanique (Mandler et Bauer 1988).

Les enfants sont donc capables de faire la différence entre deux formes de mouvement, l'une mécanique, l'autre biologique. David Premack (1990) a donc proposé l'hypothèse suivante, à savoir que la distinction entre objets animés et inanimés est liée à la perception de deux formes de mouvement différentes. Les enfants naîtraient ainsi dotés de deux propriétés spécifiques et innées : une propriété causale qui déterminerait la perception des objets ne pouvant se mouvoir par eux-mêmes, et une propriété intentionnelle dont dépendrait la perception du mouvement autonome des êtres biologiques. La présence d'un mouvement autonome conduirait donc l'enfant à attribuer une caractéristique mentale, c'est-à-dire une intentionnalité, à un objet.

La même problématique a orienté d'autres recherches, qui ont donné des résultats surprenants concernant l'anthropomorphisme. György Gergely et Gergely Csibra (2003) ont montré comment, chez les enfants âgés de douze mois, on peut déjà déceler la présence d'une pensée téléologique, qui s'exprime par l'attente d'un comportement rationnel, orienté vers un but, de la part des agents. Les enfants s'attendent à ce que les agents atteignent leur but de la façon la plus efficace possible dans des conditions données. Cette expérience est intéressante non seulement parce que l'étude portait sur de très jeunes enfants, mais aussi parce qu'ici les agents des actions interprétées par les enfants n'étaient pas des personnes, mais des boules colorées qui bougeaient sur un écran. Initialement, la boule A devait contourner un obstacle pour rejoindre la boule B. Ensuite, on présentait aux enfants deux situations différentes, qui avaient en commun que l'on avait retiré l'obstacle : dans la condition dite « compatible », la boule A avançait en ligne droite jusqu'à la boule B, tandis que dans la condition dite « incompatible » la boule A rejoignait la boule B par un chemin indirect, bien qu'aucun obstacle ne l'empêchât d'emprunter le chemin le plus court. Les enfants montraient

qu'ils trouvaient cette seconde situation « irrationnelle », inattendue. On peut en conclure que les enfants âgés d'un an s'attendent à ce que les agents aient des intentions et poursuivent leurs objectifs de façon rationnelle. Il est particulièrement significatif ici que l'expérience n'ait donné aucune indication pouvant faire penser que les agents avaient des caractéristiques humaines : il s'agissait de deux boules de couleurs et de dimensions différentes qui bougeaient sur un écran.

Une expérience similaire a abouti à des résultats encore plus surprenants. Dans ce cas, les agents de l'expérience étaient des petits blocs de bois colorés (en forme de carré, de cercle ou de triangle), pourvus de deux yeux schématiques, qui se déplaçaient sur une colline. Le bloc A « essayait » de remonter la colline, et les deux autres pouvaient intervenir de deux façons : l'action du bloc B consistait à pousser le bloc A vers le haut de la colline, « l'aidant » ainsi à atteindre son objectif, tandis que l'action du bloc C consistait à repousser le bloc A vers le bas de la colline, l'« empêchant » ainsi d'arriver jusqu'en haut. Ces situations ont été présentées à des enfants de six et dix mois. Les premiers s'étonnaient de voir le bloc A se rapprocher du bloc C, celui qui l'avait gêné, et quand on demandait aux seconds de choisir un des deux blocs, ils montraient une nette préférence pour le bloc B, le « gentil », plutôt que pour le bloc C, le « méchant ». En enlevant les yeux des blocs de bois et en supprimant le mouvement autonome des agents, on n'obtenait pas les mêmes résultats. La thèse des auteurs est que les enfants témoignent leur préférence pour les agents qui réalisent des actions sociales positives, et que ce principe serait une base innée de la cognition morale (Hamlin, Wynn et Bloom 2007). Quoi qu'il en soit, cette expérience montre que les enfants de six mois attribuent naturellement des comportements sociaux spécifiques aux humains, ainsi qu'à des petits blocs de bois simplement dotés d'yeux.

Les expériences que nous avons présentées sont très riches en informations concernant les fondements cognitifs de l'anthropomorphisme. Elles nous montrent tout d'abord que la réalité est loin de correspondre à la confusion censée caractériser les très jeunes enfants selon Piaget. Nous apprenons que l'enfant, même très jeune, est prédisposé à opérer des distinctions précises sur des aspects complexes de la cognition sociale. La perception de différentes formes de mouvement, réel ou imputé, comme dans l'étude des modes de catégorisation des jouets, l'amène à distinguer les êtres animés des êtres inanimés, et par là même les êtres intentionnels des êtres non intentionnels. Nous pouvons comparer ces résultats, issus des réponses spontanées de très jeunes enfants, à des réponses plus élaborées obtenues par les chercheurs auprès d'enfants plus âgés, qui, pour reprendre Annette Karmiloff-Smith (1992), ont appris à formuler des explications au sujet de leur propre comportement. Considérons l'expérience suivante : on montrait à des enfants de trois à quatre ans des photographies représentant des objets inhabituels, comme des animaux inconnus, des statues bizarres et d'autres objets difficilement reconnaissables. On leur demandait ensuite si chacun de ces objets était capable ou non de monter ou descendre une colline de façon autonome. Dans leurs réponses, les enfants ne tenaient pas compte de la ressemblance (par exemple entre deux images très semblables, l'une d'un objet et l'autre d'un animal), mais avaient plutôt recours à la distinction entre deux types de mouvement, autonome et non autonome. Pour justifier leurs choix, ils pouvaient parfois aller contre l'évidence, en affirmant par exemple voir les pattes d'un animal qui n'étaient pas sur la photographie ; ou, au contraire, en expliquant qu'une statue ne pouvait pas se mouvoir puisqu'elle n'avait pas de pattes, alors que, sur l'image, elle en avait (Massey et Gelman 1988).



Fig. 6 Elfoïd P1, téléphone humanoïde développé par l'université d'Osaka et ATR Hiroshi Ishiguro Laboratory. Photo Emiko Ogawa © Photo Hiroshi Ishiguro, Osaka University, ATR.

Une première conclusion s'impose : il semble que les êtres humains, dès la première enfance, soient prédisposés à considérer qu'une certaine forme de mouvement (réel ou imputé) implique une intentionnalité. De plus, le fait d'attribuer une intentionnalité comporte un jugement. Ainsi, une action qui va dans le sens des objectifs de l'agent est bonne, et de la même manière une action allant à l'encontre de ses objectifs est mauvaise.

Jusqu'ici, nous avons évoqué un phénomène de l'ordre de la perception. L'attribution d'une intentionnalité s'étend à toutes les entités dotées d'une capacité de mouvement autonome. Que se passe-t-il quand il n'y a aucun mouvement apparent ? Comment les enfants attribuent-ils, en ce cas, une intentionnalité à l'artefact ? Examinons encore quelques études expérimentales. Dans l'une d'elles, on demandait à des enfants de dix-huit mois d'imiter les actions d'une personne ou d'un robot. Toutes les actions qu'on leur demandait d'imiter étaient inachevées. Les enfants n'imitaient que celles tentées par une personne, jamais par un robot (Meltzoff 1995). Ainsi, quand l'expérimentateur ne parvenait pas à achever une action, les enfants cherchaient à comprendre son intention, ce qu'il avait voulu faire, et quand on leur demandait d'imiter son action, ils la portaient à terme. À l'inverse, ils ne se posaient aucune question sur la finalité de l'action inachevée



Fig. 7 Hiroshi Ishiguro posant avec Elfoïd P1, téléphone humanoïde, et Telenoïd, université d'Osaka/NTT Docomo/Qualcomm. Avec l'aimable autorisation d' eSeL.at (Lorenz Seidler), <http://esel.at>.

du robot. Un phénomène similaire de non-attribution d'intentionnalité a été observé dans le comportement d'enfants de dix mois à qui l'on demandait d'imiter des actions simples effectuées soit par une expérimentatrice, soit par Big Dog, une grande marionnette en forme de chien caractérisée par une expression de visage neutre/amicale, et animée par un humain caché à l'intérieur. Les enfants imitaient les actions dans les deux cas quand il s'agissait d'actions achevées, mais quand il s'agissait d'imiter des actions incomplètes, seules celles de l'expérimentatrice étaient imitées (Legerstee et Markova 2008).

D'autres études ont examiné la distinction entre objets animés et inanimés chez les très jeunes enfants, en situation de communication. Par exemple, on a montré que les enfants âgés de douze mois utilisaient le pointage proto-déclaratif quand ils s'adressaient à des êtres humains pour attirer leur attention. Ils n'en faisaient pas autant quand ils s'adressaient à des robots en forme de petit chien, qui répondaient en ouvrant et en fermant les yeux de façon contingente. Les enfants âgés de six mois s'attendent à ce que les êtres humains communiquent avec d'autres êtres humains et à ce qu'ils manipulent des objets inanimés. Ils s'étonnent de voir un expérimentateur parler à une personne cachée derrière un rideau, pour découvrir ensuite qu'il s'agissait d'un balai. Dès quatre mois, quand une personne disparaît derrière un paravent, ils vocalisent pour attirer son attention, tandis que lorsque l'on fait disparaître un jouet derrière le paravent, ils tendent les mains comme pour le saisir (Legerstee 2005).

Ainsi, d'un côté, les enfants, même très jeunes, sont capables d'identifier les objets, de distinguer entre les êtres animés et inanimés, et de leur attribuer une intentionnalité de façon différenciée, de l'autre, ils étendent dans certains cas l'attribution d'intentionnalité aux objets inanimés. Ce n'est donc pas la « confusion » dont parlait Piaget qui les porte à attribuer une intentionnalité à des objets inanimés, mais c'est bien autre chose; nous pouvons aussi en conclure que les êtres humains sont dotés d'une pensée téléologique du réel qui les conduit à interpréter les actions d'autrui comme des actions intentionnelles. Dans certaines circonstances, cette attribution est étendue aux objets. Mais comment expliquer qu'une telle attribution, si fondamentale, n'advienne pas dans certains cas? Nous avons vu qu'une des hypothèses avancées pour expliquer ce phénomène est que les enfants sont en mesure de distinguer différentes formes de mouvement. Ainsi, si des boules bougent sur un écran selon des modalités non mécaniques, l'enfant leur attribue une intentionnalité, tandis qu'un chien mécanique, pour habile qu'il soit, n'est pas un partenaire adéquat. Mais le type de mouvement suffit-il à l'enfant pour établir cette distinction? Dans le cas de Big Dog, c'était un être humain – même s'il était caché – qui réalisait les mouvements qu'on demandait aux enfants d'imiter, et pourtant l'intentionnalité qu'ils lui attribuaient n'était pas équivalente à celle qu'ils attribuaient à l'expérimentatrice.

Dans des situations proches de celles de la vie quotidienne, le comportement de l'enfant apparaît plus complexe qu'il ne l'était dans les expériences décrites plus haut. Celles-ci nous montrent que la pensée téléologique semble bien être une caractéristique fondamentale de la pensée humaine, mais elles ne suffisent pas à expliquer le comportement de l'enfant dans des situations plus complexes. Pour clarifier cette question, on peut avancer l'hypothèse que les distinctions que l'enfant opère ne dépendent pas des caractéristiques des objets en présence, mais de la situation d'interaction qui s'établit entre lui et eux. Il s'agit donc de passer de la simple analyse des caractéristiques des objets avec lesquels les enfants interagissent à l'analyse de la situation d'interaction.

Anthropomorphisme et interaction

Les travaux que nous avons vus auparavant nous ont montré comment l'enfant est en mesure d'agir en opérant des distinctions fondamentales bien avant de pouvoir les conceptualiser de façon explicite, et même d'en parler. Comment ce schéma de base, qui oriente la pensée téléologique, évolue-t-il ? Revenons à l'anthropomorphisme chez l'adulte : pourquoi sommes-nous aussi enclins à attribuer des caractéristiques mentales aux objets inanimés ? Comment fonctionne cette attribution très fréquente de caractéristiques mentales humaines aux artefacts ?

Commençons par une remarque : cette attribution de facultés humaines à l'objet, qu'elle s'exprime en paroles ou en actes, ne peut se produire que dans le cadre d'une interaction. On peut difficilement imaginer en effet que quelqu'un *pense* qu'une machine à café est un être humain, alors qu'on n'est pas surpris de voir une personne *traiter* une machine à café comme si elle avait affaire à un être intentionnel. Nous pouvons très bien dire à la machine à café : « Allez, fais vite, je suis en retard ! » Ou à notre voiture : « S'il te plaît, ne t'arrête pas maintenant ! » ou bien « Où t'es-tu cachée ? Je suis pourtant sûr de t'avoir garée ici hier soir. » Mais cette interaction verbale ne suppose pas une croyance stable concernant l'ontologie des objets. La caractéristique fondamentale de l'anthropomorphisme est de s'exprimer sous la forme d'un dialogue. Personne ne pourrait sérieusement : « Ma voiture est une folle qui se cache quand je la cherche. » Mais on peut facilement imaginer que cette phrase trouve sa place dans une *interaction*. Ainsi, c'est bien l'aspect relationnel qui prime dans l'anthropomorphisme. Remarquons par ailleurs que ce dernier ne concerne pas que les intentions, mais aussi les sentiments et les émotions.

On peut en conclure que l'anthropomorphisme possède deux caractéristiques fondamentales :

1. En tant qu'expression d'une pensée téléologique de base, il constitue une manière de se représenter les non-humains en les assimilant à des humains, leur attribuant des perceptions, des croyances, des intentions et des émotions.
2. Il se manifeste essentiellement dans les interactions.

Nous pouvons donc formuler l'hypothèse que la projection anthropomorphique ne dépend pas seulement de l'attribution d'humanité à telle ou telle caractéristique d'un animal ou d'un objet mais qu'il faut, pour que la projection opère, que ces caractéristiques se manifestent dans une situation d'interaction spécifique, qui prenne la forme d'un dialogue. Pour éclaircir cet aspect, nous aurons recours à un exemple qui remonte à la préhistoire de l'intelligence artificielle. En 1966, Joseph Weizenbaum se proposa de mettre au point un programme capable d'interagir avec l'utilisateur. Il eut l'idée de donner à l'ordinateur le rôle d'un psychologue rogersien. Le programme avait un fonctionnement extrêmement simple, avec un nombre limité de réactions possibles : poser des questions simples (« Parlez-moi de votre père », par exemple), reprendre les derniers mots de son interlocuteur, relancer celui-ci (« Parlez-moi un peu plus de... »), etc. On aurait pu croire que l'utilisateur humain, au bout d'une interaction de quelques minutes, allait facilement s'apercevoir du caractère préprogrammé des interventions de l'ordinateur et perdre ainsi l'illusion de se trouver face à un interlocuteur doté d'états mentaux humains. Le concepteur de ce programme, nommé ELIZA, fut très étonné de constater qu'au contraire les utilisateurs humains entretenaient de longues conversations avec l'ordinateur, auquel il leur arrivait souvent de confier des éléments intimes de leur vie.



Fig. 8 Série *Robots are real* : les mains de Zeno © Timothy Archibald.

Les individus qui interagissaient avec ELIZA pensaient-ils réellement avoir affaire à un psychothérapeute ? Sûrement pas. Dans la plupart des cas, il s'agissait d'étudiants ou de chercheurs qui travaillaient dans le même laboratoire que son créateur et savaient donc qu'ils conversaient avec un ordinateur. Et pourtant, dans la situation d'interaction, ils se comportaient comme s'ils avaient affaire à un interlocuteur réel. Il semble que ce soit, d'une certaine façon, la situation d'interaction en tant que telle qui nous conduit à considérer l'autre en interlocuteur valable. La forme paraît avoir une force bien plus importante que le fond et donner du sens au contenu même quand il est très pauvre. En l'occurrence, il est significatif que la situation d'interaction ait été extrêmement simple.

Une fois reconnue l'importance de la forme « dialogue » dans le déclenchement de l'anthropomorphisme, revenons aux enfants, ou plutôt aux nouveau-nés, et voyons comment se manifestent leurs premières interactions. Les êtres humains interagissent les uns avec les autres dès la naissance, et dès lors expriment donc une intersubjectivité (Trevarthen 1979). Dès les premiers mois, ils engagent des formes primitives de dialogue avec les adultes, formes que Mary Catherine Bateson (1979) définissait comme des « proto-conversations ». C'est en cherchant à comprendre pourquoi il nous semble naturel de nommer « dialogues » ou « conversations » les échanges extrêmement simples qui ont lieu entre un nouveau-né et un adulte que nous pourrions mieux en comprendre les caractéristiques fondamentales et en donner une définition plus précise. De toute évidence, il manque à ces conversations ce qui fait habituellement la matière d'une conversation, à savoir le langage. Les nouveau-nés ne répondent que par un répertoire de sons et de gestes très limité : sourires, gazouillis, grimaces. Et pourtant, il ne nous semble pas insensé de penser qu'avec les adultes s'instaure une certaine forme de dialogue. Ce n'est donc pas le contenu qui donne sa définition au dialogue, mais bien la forme de l'interaction (Airenti 2010). De ce point de vue, la caractéristique fondamentale d'un dialogue est que les locuteurs parlent à tour de rôle. Un dialogue prévoit que les interlocuteurs fassent chacun leur tour des gestes adressés à l'interlocuteur. Ce schéma de l'alternance est coercitif en tant qu'il donne sens à tout ce qui se passe au sein du dialogue, qu'il s'agisse de mots, de gestes ou de sons, comme dans le cas du nouveau-né. Il faut noter également que l'adulte, qui attribue consciemment du sens à toutes les manifestations de l'enfant, reconnaît ainsi qu'il y a bien dialogue, mais l'enfant aussi le reconnaît. Ceci a été démontré par les psychologues du développement grâce au paradigme de la *still face* (Tronick *et al.* 1978). Si on demande à la mère d'un nouveau-né âgé de six semaines d'interrompre une situation de dialogue en s'immobilisant quelques secondes et en passant d'une expression concernée à une autre, complètement neutre, l'enfant se montre très perturbé, et cherche par tous les moyens à attirer l'attention de sa mère et à la ramener dans l'échange.

Les êtres humains ont une tendance innée à interagir avec leurs semblables en utilisant la forme du dialogue, dont la caractéristique principale est d'être un format qui conduit à attribuer du sens à toute forme de contenu qui y est inclus, qu'il soit simple ou complexe. Ce format se met en place dès les premières interactions que connaît l'enfant, et donc bien avant qu'il soit en mesure d'attribuer des caractéristiques mentales à son interlocuteur.

Mais n'y a-t-il aucun lien entre ce que fait l'enfant et ce que fait l'adulte ? Les conversations adultes se caractérisent par le format de l'interaction, soit l'alternance du tour de parole, mais aussi par la pertinence de ce qui est dit. Chacun d'entre nous s'attend à ce que l'autre participe au dialogue en apportant une

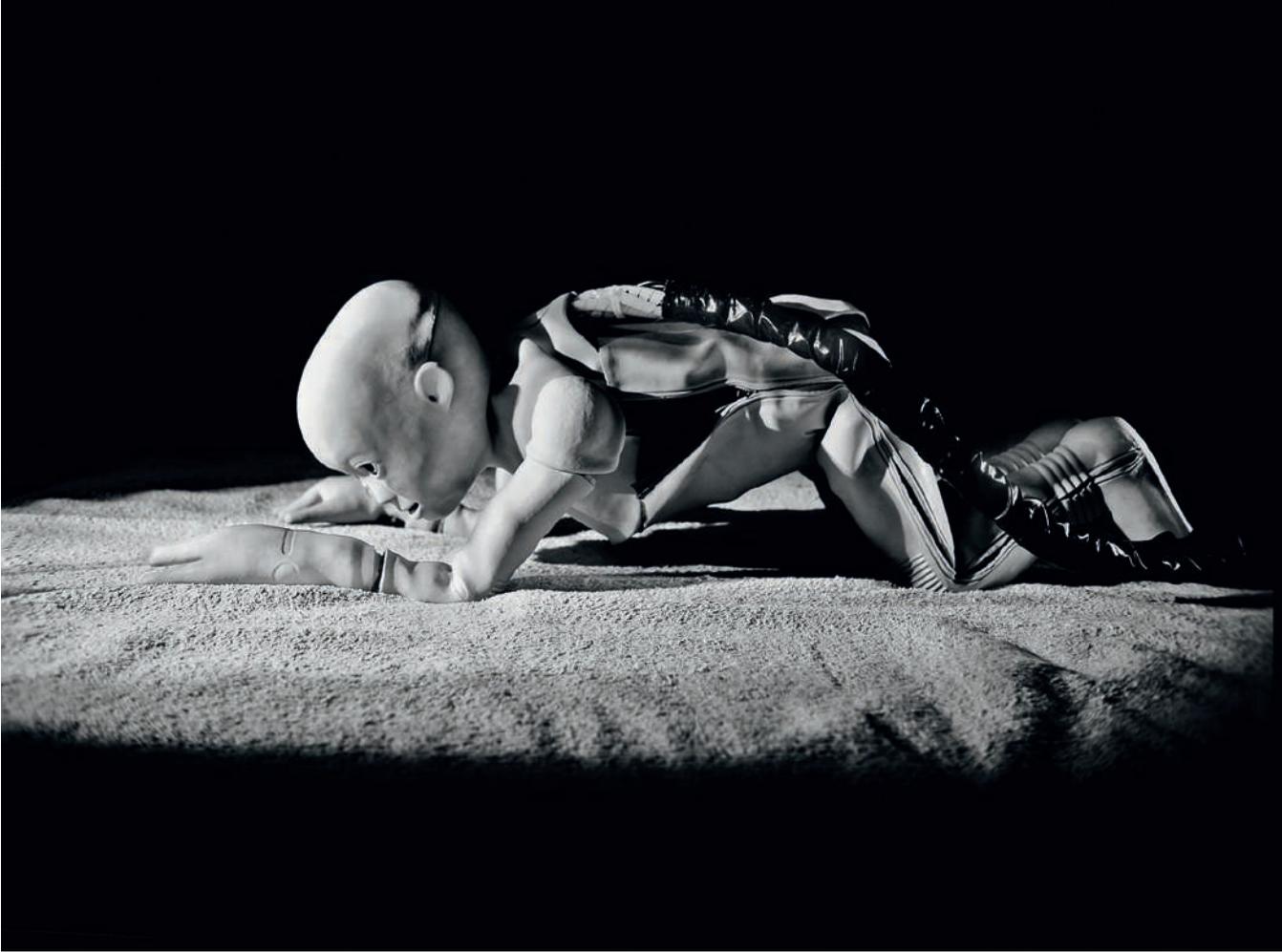


Fig. 9 CB2, les stades de développement des enfants servent de plus en plus de modèle en robotique
© Max Aguilera-Hellweg.

contribution pertinente (Grice 1975). Comment les nouveau-nés, qui ne se servent pas des mots, peuvent-ils intervenir dans un dialogue de façon pertinente? Et comment peuvent-ils reconnaître que leur interlocuteur est intervenu de façon pertinente? Les adultes et les enfants utilisent pour cela une capacité dont ils sont dotés dès la naissance : l'imitation. Plus de trente ans de recherches, désormais, ont montré que les nouveau-nés sont capables d'imiter des expressions faciales quelques heures seulement après la naissance (Meltzoff et Moore 1977). Les dialogues entre un adulte et un enfant âgé de quelques mois font largement appel à l'imitation : l'enfant imite l'adulte, et l'adulte imite l'enfant. C'est à travers l'imitation réciproque, ou plus précisément à travers les variations autour d'une imitation, que se construit la pertinence. La mère répète un comportement de l'enfant en l'exagérant, ou bien lui répond par un comportement proche de celui du nouveau-né mais différent, ou encore opposé. C'est parce qu'ils interviennent dans le cadre d'un dialogue que ces comportements acquièrent une pertinence. Nous pouvons donc conclure que les êtres humains sont prédisposés à interagir entre eux selon un format bien précis, à savoir le dialogue, qui prévoit la circulation de la parole (ou des gestes) en alternance, la réciprocité des actions et la pertinence des interventions.

En étudiant de très jeunes enfants, nous avons identifié deux processus cognitifs qui apparaissent primordiaux pour comprendre la tendance essentielle qui pousse les êtres humains à l'anthropomorphisme, à savoir la pensée téléologique et la dimension du dialogue. Nous pouvons désormais donner de l'anthropomorphisme une définition relationnelle : l'anthropomorphisme signifie mettre un

objet ou un animal dans la position d'interlocuteur au sein d'un dialogue. Le fait que ces procédés cognitifs soient déjà présents chez les nouveau-nés montre qu'il s'agit de structures cognitives fondamentales, qui n'ont besoin que de très peu de contenu pour se manifester. Ces procédés sont à la base de toutes les interactions, y compris celles des adultes. C'est ainsi que nous pouvons comprendre certaines formes d'interaction comme l'effet ELIZA (dont il était question plus haut). Même si le contenu de l'interaction est pauvre, l'effet ne l'est pas, car l'interaction entre l'ordinateur et l'utilisateur se fait sous la forme d'un dialogue, et c'est ce format qui lui donne son sens.

Ces considérations nous permettent de montrer le développement de l'anthropomorphisme, depuis celui du jeune enfant jusqu'à celui de l'adulte. En effet, que les enfants participent à des dialogues dès les premiers mois de leur existence ne signifie nullement qu'il n'y a pas de changements, même majeurs, avec l'âge dans la capacité et la propension au dialogue. Les enfants les plus jeunes participent à des dialogues sans être en mesure d'analyser leurs propres états mentaux ni ceux des autres. Avec l'âge, ils passent de la pensée téléologique simple à l'attribution différenciée d'émotions, de désirs, de croyances à leurs interlocuteurs, et apprennent ainsi à entrer en empathie avec eux. Les contenus qui entrent dans le format de base deviennent donc de plus en plus complexes. C'est aussi le cas quand l'interlocuteur n'est pas une personne réelle, mais un partenaire imaginaire.

Une des étapes du développement de l'enfant réunit tous les aspects que nous avons traités jusqu'ici, il s'agit du jeu de fiction. L'enfant que nous avons vu participer à des dialogues avec les adultes dès les premiers mois commence autour de dix-huit mois à intégrer à ses jeux des objets inanimés. Ces nouveaux jeux de fiction l'amènent à construire des situations dans lesquelles la vie réelle et le monde imaginaire sont intimement liés. Les deux aspects que nous avons abordés plus haut, à savoir la pensée téléologique et l'aptitude au dialogue, sont appliqués à des objets de la vie quotidienne, qui prennent ainsi des caractéristiques imaginaires. L'enfant ne confond pas la vie réelle et la situation imaginaire, et il peut passer de l'une à l'autre sans difficulté. Comme le remarquait Karl Bühler (1999 [1930]), un enfant peut jeter au feu un bout de bois qu'il traitait jusque-là comme un bébé. Dans le jeu de fiction, il expérimente des relations, explore des sentiments, des émotions, des états mentaux. L'enfant qui fait semblant de prendre le bout de bois pour un bébé parle avec lui comme il s'imaginerait qu'une mère le ferait. Dans ce cas précis, traiter un objet comme s'il s'agissait d'une personne est un exercice qui permet à l'enfant d'affiner ses capacités de compréhension de soi-même et des autres (Harris 2000).

Revenons donc à la question de la ressemblance et de la familiarité. Un bout de bois ne ressemble en rien à un bébé. La seule caractéristique qui en fait un candidat possible pour tenir ce rôle est qu'on peut le tenir dans ses bras. Il n'est pas capable de se mouvoir de façon autonome et ne peut pas répondre quand on lui parle. Et pourtant, l'enfant peut établir avec lui un rapport de familiarité. Pourquoi? Parce qu'il établit avec l'objet une relation de dialogue qui implique le déclenchement d'une pensée téléologique de base. Le même enfant trouvera beaucoup moins familier un petit robot en forme de chien au comportement beaucoup plus sophistiqué, mais avec lequel il est impossible d'établir une interaction verbale imaginaire. En ce cas, au lieu de l'attirer, le robot pourrait même l'effrayer. Toute familiarité est donc liée à la possibilité de placer un objet inanimé dans la position d'interlocuteur dans un dialogue².

• • •
2. Une situation dans laquelle l'aspect iconique joue un rôle constitutif dans la définition pragmatique d'un interlocuteur a été analysée, dans une perspective analogue, par Carlo Severi (2009).

Conclusions

Nous avons cherché ici à analyser certaines caractéristiques cognitives fondamentales des êtres humains, qui peuvent nous aider à mieux comprendre le phénomène de l'anthropomorphisme. Il semblerait que deux caractéristiques qui se manifestent déjà chez les très jeunes enfants soient à l'origine de cette tendance. La première est que les humains ont une vision téléologique de l'action. Ils s'attendent ainsi à ce qu'un acteur poursuive ses objectifs de façon rationnelle. La seconde est que les êtres humains ont une tendance très précoce à établir des relations avec leurs semblables en ayant recours à la forme du dialogue. Un enfant peut entrer dans une relation de dialogue bien avant d'avoir acquis la capacité de mettre en œuvre des modèles complexes pour se représenter ses propres états mentaux et ceux des autres. Dans les relations qu'il entretient avec les adultes de son entourage, il développe sa propre capacité de compréhension de l'esprit, des émotions, des sentiments, des désirs, des croyances. Dès la première enfance, il exerce ses capacités relationnelles dans le jeu de fiction, en utilisant des objets auxquels il attribue des caractéristiques humaines. Ainsi il met cet objet dans la position d'interlocuteur dans le cadre d'un dialogue possible. Les adultes ne s'adonnent plus au jeu de fiction, mais dans certains cas ils peuvent assigner des intentions, des émotions ou des croyances à des objets dans le cadre d'une situation d'interaction. Il est vrai que nous attribuons des caractéristiques anthropomorphiques à des animaux, des artefacts, ou même des objets du quotidien. Mais ces caractéristiques ont un seul point commun : elles sont toujours de type relationnel. L'anthropomorphisme est lié au besoin de construire des relations avec les autres. Quand nous mettons un objet dans la position d'interlocuteur, nous lui supposons les mêmes caractéristiques mentales que celles des êtres humains. Réciproquement, la personne, l'animal ou l'objet avec qui nous entrons en relation ne sont pensables pour nous que comme des sujets dotés d'un esprit présentant les caractéristiques que nous attribuons à l'esprit humain. Les aspects relationnel et communicatif, qui dépendent du mode de l'interaction, et l'attribution de caractéristiques mentales, qui dépend de l'exercice d'une pensée téléologique, sont indissociables. Ainsi, du point de vue de la psychologie de la cognition, l'anthropomorphisme apparaît comme une forme précoce de pensée téléologique, inhérente au dialogue.

Centre de sciences cognitives,
Département de psychologie, université de Turin
gabriella.arenti@unito.it

Traduction de l'italien par Matteo Severi

Bibliographie

AIRENTI, Gabriella

2010 « Is a naturalistic theory of communication possible? », *Cognitive Systems Research* 11 : 165-180.

BATESON, Mary Catherine

1979 « The epigenesis of conversational interaction: a personal account of research development », in Margaret Bullowa (éd.), *Before speech. The beginning of human communication*. Cambridge, Cambridge University Press : 63-67.

BRUNOIS, Florence

2005 « Man or animal: who copies who? Interspecific empathy and imitation among the Kasua of New Guinea », in Alessandro Ortelli, Gherardo Ortalli et Glauco Sanga (éd.), *Animal Names*. Venise, Istituto Veneto di Scienze, Lettere ed Arti : 369-381.

BÜHLER, Karl

1999 [1930] *The Mental Development of the child*. Londres, Routledge.

GERGELY, György et CSIBRA, Gergely

2003 « Teleological reasoning in infancy: the naive theory of rational action », *TRENDS in Cognitive Sciences* 7 : 287-292.

GRICE, Paul H.

1975 « Logic and conversation », in Peter Cole et Jerry L. Morgan (éd.), *Syntax and semantics*, vol. III : *Speech Acts*. New York, Academic Press : 41-58.

GUTIERREZ CHOQUEVILCA, Andréa-Luz

2010 « Imaginaire acoustique et apprentissage d'une ontologie animiste », *Ateliers d'anthropologie* 34, <http://ateliers.revues.org/8553>.

HAMLIN, J. Kiley, WYNN, Karen et BLOOM, Paul

2007 « Social evaluation by preverbal infants », *Nature* 450 : 557-560.

HARRIS, Paul L.

2000 *The Work of the imagination*. Oxford, Basil Blackwell.

KAHN, Peter H. Jr, ISHIGURO, Hiroshi, FRIEDMAN, Batya, KANDA, Takayuki, FREIER, Nathan G., SEVERSON, Rachel L. et MILLER, Jessica

2007 « What is a human? Toward psychological benchmarks in the field of human-robot interaction », *Interaction Studies* 8(3) : 363-390.

KARMILOFF-SMITH, Annette

1992 *Beyond modularity: a developmental perspective on cognitive science*. Cambridge, MIT Press.

LEGERSTEE, Maria

2005 *Infants' sense of people: precursors to a theory of mind*. Cambridge, Cambridge University Press.

LEGERSTEE, Maria et MARKOVA, Gabriela

2008 « Variations in 10-month-old infant imitation of people and things », *Infant Behavior and Development* 31 : 81-91.

MANDLER, Jean M. et BAUER, Patricia J.

1988 « The cradle of categorization. Is the basic level basic? », *Cognitive Development* 3 : 247-264.

MASSEY, Christine M. et GELMAN, Rochel

1988 « Preschooler's ability to decide whether a photographed unfamiliar object can move itself », *Developmental Psychology* 24 : 307-317.

MELTZOFF, Andrew

1995 « Understanding the intentions of others: re-enactment of intended acts by 18-month-old children », *Developmental Psychology* 31 : 838-850.

MELTZOFF, Andrew et MOORE, Keith M.

1977 « Imitation of facial and manual gestures by human neonates », *Science* 198 : 75-78.

MORI, Masahiro

1970 « The uncanny valley », *Energy* 7(4) : 33-35.

PIAGET, Jean

1926 *La Représentation du monde chez l'enfant*. Paris, Alcan.

1978 [1945] *La Formation du symbole chez l'enfant : imitation, jeu et rêve, image et représentation*. Neuchâtel, Delachaux et Niestlé ; version en ligne sur le site de la fondation Jean Piaget, <http://www.fondationjeanpiaget.ch>.

PREMACK, David

1990 « The infant's theory of self-propelled objects », *Cognition* 36 : 1-16.

SERVAIS, Véronique

2007 « The report and the command: the case for the relational view in the study of communication », *Intellectica* 2-3 : 85-104.

SEVERI, Carlo

2009 « La parole prêtée. Comment parlent les images », *Cahiers d'anthropologie sociale* 5 : 11-41.

SPELKE, Elizabeth S.

1990 « Principles of object perception », *Cognitive Science* 14 : 29-56.

TREVARTHEN, Colwyn

1979 « Communication and cooperation in early infancy: a description of primary intersubjectivity », in Margaret Bullowa (éd.), *Before speech*. Cambridge, Cambridge University Press : 321-347

TRONICK, Edward, ALS, Heidelise, ADAMSON, Lauren, WISE, Susan et BRAZELTON, Berry T.

1978 « The infant's response to entrapment between contradictory messages in face-to-face interaction », *Journal of the American Academy of Child Psychiatry* 17 : 1-13.

WEIZENBAUM, Joseph

1966 « ELIZA – a computer program for the study of natural language communication between man and machine », *Communications of the ACM* 9 : 36-45.

WELLMAN, Henry M.

1990 *The Child's Theory of mind*. Cambridge, MIT Press.

Résumé / Abstract

Gabriella Airenti, *Aux origines de l'anthropomorphisme. Intersubjectivité et théorie de l'esprit* – Dès la première enfance, les humains attribuent des traits de comportement anthropomorphe, tels des intentions, des perceptions et même des sentiments, aux artefacts. Pourquoi cette attitude est-elle si répandue ? Dans le champ de la psychologie, on a longtemps attribué ces comportements à l'exercice d'une forme primitive de la pensée, et notamment à l'incapacité prêtée aux jeunes enfants de distinguer entre le niveau physique et le niveau mental de la réalité. On admettait aussi que cette incapacité, propre aux enfants, pouvait, dans des situations affectivement chargées ou difficiles, affecter le comportement des adultes. Des recherches récentes ont montré, toutefois, que les très jeunes enfants sont parfaitement en mesure de distinguer entre ce qui relève du physique et ce qui relève du mental. Nous proposons donc dans cet article l'hypothèse qu'à l'origine de l'anthropomorphisme l'on ne trouve pas une modalité spécifique de la pensée, mais plutôt des modalités d'interaction précocement apprises. Plusieurs études consacrées au développement social de l'enfant montrent en effet que ce qu'on appelle « anthropomorphisme » est une extension aux non-humains de modalités interactionnelles propres au dialogue entre humains.

Gabriella Airenti, *The origins of anthropomorphism. Intersubjectivity and theory of mind* – Since a very young age humans attribute anthropomorphic features like intentions, perceptions and even feelings to artifacts. Why is this behavior so common? In the past developmental psychologists have considered this attitude as a primitive form of thinking. They maintained that young children were unable to distinguish between physical and mental events. Adults would adopt a similar attitude only when facing affectively charged or difficult situations. However, recent research has shown that also young children make distinctions between the “physical” and the “mental”. In this article I argue that anthropomorphism is not a primitive form of thinking but has its origins in interaction modalities acquired during infancy. A number of studies devoted to social development show that what we call “anthropomorphism” is an extension to non-humans of forms of interactions typical of human dialogue.