

> Charles Lenay

Schèmes interactionnistes de la cognition sociale.



> #Numéro 1
> Suppléance perceptive et cognition sociale
> Traductions
> CRED - Cognitive Research and Enactive Design (Costech-UTC)
> Enaction - > Philosophie - > Sciences cognitives

Citer cet article

Lenay, Charles. "Schèmes interactionnistes de la cognition sociale..", 18 juin 2017, *Cahiers Costech*, numéro 1.
URL <https://www.costech.utc.fr/CahiersCostech/spip.php?article32>

Résumé

Faut-il expliquer les interactions sociales par des capacités cognitives individuelles ou doit-on au contraire expliquer ces capacités par les interactions sociales ? Pour répondre nous nous plaçons dans une perspective pragmatiste, sensible au fait que les individus sociaux sont toujours déjà engagés dans des interactions plus ou moins réciproques, dépendant de leur corps et du contexte concret dans lequel ils se situent. Dans ce cas, l'activité cognitive ne peut se comprendre comme détachée du milieu social, précédant des relations qu'elle rendrait possibles. Mais si l'on veut que les interactions interindividuelles participent à la constitution de phénomènes de cognition sociale, il faut disposer de schèmes explicatifs permettant de rendre compte de compétences comme la discrimination des êtres intentionnels, la perception des intentions, ou l'imitation des expressions faciales. Nous contribuons ici à la recherche de tels schèmes en proposant un paradigme expérimental minimaliste dans lequel on peut observer la genèse ou le

fonctionnement de telles compétences. Ces situations d'observation très simples devraient alors permettre une discussion précise des mécanismes en jeu en proposant une façon de démêler ce qui explique et ce qui est expliqué dans les relations entre compétences individuelles et dynamique sociale. Dans le cadre d'une approche énaïve on pourra ainsi montrer le fonctionnement de l'engagement corporel dans l'organisation collective des interactions et dans la reconfiguration des structures individuelles

Auteur(s)



Charles Lenay, Professeur de sciences cognitives et de philosophie des sciences, ex-directeur du COSTECH . Habilité à diriger les recherches en Philosophie (17^e) et en Histoire des sciences (72^e), il consacre l'essentiel de ses recherches aux technologies cognitives : comment les outils participent à l'activité cognitive : raisonnement, mémorisation, perception, interaction,...

Plan

- 1 - Introduction : une difficulté des approches interactionnistes de la cognition sociale
- 2 - Paradigme expérimental minimaliste
- 3 - Croisements perceptifs et reconnaissance d'autrui
 - 3.1 - Dispositif expérimental du croisement perceptif
 - 3.2 - Nouvelle expérience, tâche collaborative.
 - 3.3 - Expérience de discrimination sonore
- 4 - Perception des intentions d'autrui
- 5 - Imitation des expressions faciales
- 6 - Discussion
 - 6.1 - Sur la méthode : schématisation et minimalisme
 - 6.2 - Approche sociale de la cognition sociale
 - 6.2.1 - Compétences initiales et dynamique d'interaction
 - 6.2.2 - Liaison de la dynamique sociale avec des propriétés déterminées du corps-objet du partenaire
 - 6.2.3 - Apprentissage associatif en situation d'interaction
 - 6.2.4 - Nouvelles compétences de cognition sociale
 - 6.3 - Distinction corps percevant / corps-objet
 - 6.4 - Enaction et technique
- 7 - Conclusion



Cliquer sur l'image pour l'agrandir

1 - Introduction : une difficulté des approches interactionnistes de la cognition sociale

Les recherches en sciences cognitives sont animées par un débat général entre des approches qui fondent leurs explications sur les seules compétences individuelles internes (soit par un système de représentation symboliques, soit par des structures et activités neuronales) et une multitude d'approches qui, dans une perspective pragmatiste, cherchent à prendre en compte les conditions concrètes, corporelles, techniques et sociales, des phénomènes cognitifs. Ce débat prend une forme plus vive encore quand la recherche porte sur la nature et les mécanismes de la « cognition sociale », c'est-à-dire les compétences comme la reconnaissance d'autrui, la compréhension de ses intentions, l'imitation de ses comportements et de ses expressions faciales, l'attention conjointe, l'empathie, l'apprentissage, le langage, etc. Faut-il partir de ces compétences individuelles pour expliquer les phénomènes sociaux ou faut-il partir des interactions sociales pour expliquer ces compétences individuelles ?

D'un côté, dans une forme d'individualisme méthodologique classique en psychologie, en philosophie de l'esprit ou en neurosciences, on considère que seules les compétences portées par l'organisme constituent la cognition à l'origine des interactions sociales. Que ce soit par inférence sur des représentations mentales (Theory Theory) ou par simulation interne des comportements observés (Simulation Theory), il s'agit de comprendre comment chacun peut deviner les intentions cachées d'autrui. En effet, pour ces approches, le comportement des sujets

intentionnels doit s'expliquer par des états mentaux internes cachés pour l'observateur extérieur. Dès lors, la base des relations sociales réside dans le « mind reading » c'est-à-dire l'art de deviner les états mentaux d'autrui (Baron-Cohen 1997). Ce serait seulement ensuite que pourraient se réaliser une attention conjointe et la construction d'un monde proprement social (Tomasello et al. 2005).

De l'autre côté, il s'agit plutôt de prendre en compte les structures sociales dans l'explication de la cognition individuelle. En effet, dans ce cadre pragmatiste, les approches, dites 4E, posent que la cognition est Embodied, Embedded, Extended et Enactive, c'est-à-dire qu'elle ne peut être expliquée sans prendre en compte les conditions concrètes des actions humaines. On met alors l'accent sur trois constats liés :

- i) Il y a toujours des conditions matérielles, corporelles et techniques, des relations sociales. Les interactions se déroulent dans un contexte concrètement structuré définissant les possibilités de percevoir, d'agir et d'interagir des différents acteurs.
- ii) Les relations sociales sont généralement réciproques et nous engageant dans une dynamique d'interaction. La situation d'observateur extérieur désengagé est secondaire et doit plutôt être artificiellement construite.
- iii) Les activités cognitives ou perceptives se réalisent en général par un engagement corporel dans un couplage sensorimoteur avec l'environnement. Elles sont donc potentiellement perceptibles à même l'activité (et non pas à déduire dans une reconstruction interne par inférence ou simulation) (Gallagher 2008).

Dans ce cadre, on peut défendre que l'environnement social n'est pas seulement un contexte pour l'activité individuelle mais que les interactions interindividuelles sont des conditions nécessaires à la réalisation d'une cognition sociale ou même, dans une version plus forte encore, que ces interactions sont constitutives de cette cognition (De Jaegher et al. 2010). On trouve un tel rôle central accordé aux interactions interpersonnelles dans les sciences du développement (Trevarthen 1979 ; 1993 ; Reddy 2008), les neurosciences sociales (Dumas 2011, 2014, Schilbach et al. 2013), l'approche des systèmes dynamiques (*dynamic systems approach to cognition*) (Thelen & Smith 1996), et certaines approches énaactives (De Jaegher et al. 2007, Froese 2016).

Déjà, dans la perspective pragmatiste développée par John Dewey la

cognition et la perception ne doivent pas être séparées des possibilités d'agir. Ce qu'il appelait le « reflex arc » se réalise suivant différents niveaux de « sensori-motor coordination ». Il écrit par exemple : “we begin in perception not with a sensory stimulus, but with a sensorimotor coordination ... it is the movement which is primary, and the sensation which is secondary, the movement of body, head and eye muscles determining the quality of what is experienced” (Dewey 1896 : 358). De même dans les conceptions de la perception active et en particulier pour les approches énaclives, la perception ne résulte pas d'un traitement interne des informations sensorielles, mais est produite dans la dynamique du couplage entre l'agent autonome et son environnement (Varela 1991). Par cette activité, il fait advenir, ou « énacler », les contenus de son expérience qui correspondent aux invariants de la dynamique sensorimotrice (O'Regan & Noë 2001). Ezequiel Di Paolo et Hanne De Jaegher (2016) parlent donc de “sense making” réalisé dans la régulation du couplage entre un système adaptatif autonome et son environnement. La perception d'un objet correspond à la maîtrise d'un couplage sensorimoteur avec cet objet (et non pas à une représentation mentale distincte de l'objet représenté). Le contenu intentionnel d'une perception est construit dans ce couplage. Si maintenant, on s'intéresse à une perception sociale comme la reconnaissance d'autrui, de ses intentions ou de ses expressions, celle-ci devrait s'expliquer par une forme de couplage avec autrui, ses intentions et ses expressions. Son contenu intentionnel devrait donc se construire dans un couplage social.

Si l'on admet une telle approche sensorimotrice de l'activité cognitive, on comprend que l'accroche entre les activités de couplage de plusieurs organismes puisse former des structures relationnelles dont le comportement et l'histoire dépassent ce que chacun y apporte (Di Paolo & De Jaegher 2016). Shaun Gallagher propose d'appeler « Interaction forte : le couplage des engagements mutuels co-régulés entre au moins deux agents autonomes, et où la co-régulation et le couplage s'affectent l'un l'autre, constituant une organisation autonome dans le domaine des dynamiques relationnelles. »¹. Cette coordination fait émerger un nouveau domaine d'objets, d'actions et de significations pour les sujets qui s'y engagent. On peut alors définir une « construction de sens collective » (*participatory sense-making*) comme « une coordination des activités intentionnelles dans l'interaction, où le processus de construction de sens individuel est affecté et où un nouveau domaine de construction sociale du sens peut se produire tel qu'il ne soit pas accessible pour chaque individu à lui seul. »² Les interactions sociales

jouent un rôle constitutif dans la mesure où la dynamique d'interaction produit par elle-même la performance de cognition sociale, comme par exemple la danse d'un couple, un geste d'échange, un travail collaboratif ou l'attention du nouveau-né pour sa mère lors d'une protoconversation.

Dans une telle approche, les phénomènes de cognition sociale peuvent s'expliquer en mobilisant le concept « d'intercorporéité », c'est-à-dire une forme de communauté concrète interindividuelle dans les processus situés d'interaction, communauté préalable à la constitution des intentions individuelles (Merleau-Ponty 1968, Gallagher 2013). Par exemple, les travaux des sciences du développement tendent à montrer l'existence d'une « Intersubjectivité primitive » (*Primary intersubjectivity*) dès les premières années de l'existence, c'est-à-dire une compréhension pragmatique mutuelle qui se réalise dans la dynamique des interactions corporelles et émotionnelles de la protoconversation entre l'enfant et l'adulte qui en prend soin (Trevarthen 1979, 1993, Trevarthen & Reddy 2007). Les actions et perceptions des différents organismes s'harmonisent dans un « cross-modal sensorimotor system » (Gallagher et Meltzoff 1996) et par ajustement mutuel réalisent une « common bodily intentionality that is shared across the perceiving subject » (Gallagher 2013 : 61). Cette intercorporéité de la *primary intersubjectivity* serait le substrat à partir duquel pourra se constituer une « *secondary intersubjectivity* » (à partir de un an) expliquant par exemple les phénomènes d'attention conjointe puis les *Communicatives and narrative competencies* (de 2 à 4 ans).

Du point de vue des neurosciences, les « Systèmes de Neurones Miroirs » (*Mirror Neuron System, MNS*), seraient le mécanisme qui explique la formation de cette intercorporéité. En effet, ces structures neuronales s'activent aussi bien quand le sujet exécute une action, ou quand il perçoit cette action exécutée par autrui. Elles permettent de comprendre les phénomènes d'imitation, en particulier l'imitation inconsciente des mouvements, postures et expressions (*chameleon effect*). De même, ces structures semblent impliquées dans notre capacité à partager des émotions à travers les expressions faciales et les postures corporelles (Gallese 2009). Dans ce cadre interactionniste, le système miroir neuronal permettrait donc de réaliser le substrat sub-personnel de l'intercorporéité en expliquant l'engagement des organismes dans des ajustements comportementaux, avant toute prise de conscience ou même distinction entre soi et autrui (Gallese 2001).³ Dans ce cas, les structures neuronales ne prétendent pas supporter des représentations

ou même des simulations internes, mais servent plutôt à expliquer un jeu d'interactions pratiques et perceptives interindividuelles.⁴

Le travail présenté ici se place dans une telle perspective interactionniste et vise à résoudre une difficulté générale que rencontrent ces approches. Cette difficulté concerne les compétences initiales que l'on accorde aux individus pour expliquer leur engagement dans les interactions sociales. Sont-elles des compétences de cognition sociale nécessaires au fonctionnement de ces interactions ou sont-elles les produits de ce fonctionnement ? Et si elles sont nécessaires, ne sont-elles pas alors aussi suffisantes ? Ou, si elles sont produites par les interactions, alors par quel mécanisme et pour quel rôle ?

Dans les approches interactionnistes qui ont été rapidement présentées ci-dessus, on accorde en fait d'importantes compétences initiales aux organismes individuels pour expliquer leur participation aux interactions. Avec la « primary intersubjectivity », on admet dès le départ comme « innate or early developing sensory-motor capacities », les compétences de reconnaissance du visage d'autrui, de perception de ses intentions et de participation aux jeux d'imitation et de régulation émotionnelle. Toutes ces compétences que l'on appelle de « cognition sociale » sont données au niveau individuel avant les interactions qu'elles permettent.

Mais si ces compétences initiales dédiées à la cognition sociale sont nécessaires à la mise en place des interactions, ne peuvent-elles pas alors aussi suffire à leur explication ? Ne peuvent-elles pas être interprétées sous une forme ou une autre d'individualisme internaliste ? Il suffira de s'appuyer sur ces compétences initiales pour la construction d'un système de représentation ou de simulation interne personnel, sans que les interactions interindividuelles ne jouent de réel rôle explicatif. C'est d'ailleurs bien ce que l'on observe avec les interprétations divergentes des systèmes de neurones miroirs (MNS). En effet, ces structures neuronales peuvent fort bien être mobilisées dans un cadre internaliste pour défendre la « Simulation Theory » : un « mind reading » passerait par une phase interne, indépendante de l'environnement, dans laquelle les intentions d'autrui seraient devinées en reconnaissant ce que l'on vivrait si l'on effectuait les comportements que l'on perçoit (Rizzolatti et al., 2002). On devrait alors admettre que la signification de ces compétences sociales puisse être comme produite par des structures individuelles internes.

Si au contraire, on veut maintenir une approche interactionniste, pragmatiste ou énaïve, la signification sociale de la cognition devrait plutôt se déployer dans les relations interindividuelles elles-mêmes. Mais si la signification des structures individuelles mobilisées dans les interactions ne se réalise que dans ces relations externes, on risque fort de tomber dans une dangereuse circularité : les compétences qui participent aux interactions sociales devraient être produites par ces interactions elles-mêmes ! Si l'on ne veut pas se donner d'avance des compétences individuelles de cognition sociale, comment peuvent être interprétées et expliquées les structures individuelles dont on observe le rôle dans les interactions interindividuelles ?

Cette difficulté se retrouve dans les diverses approches qui refusent l'individualisme méthodologique. On peut comprendre que dans le « strong interactionism » les capacités de résolution de tâches sociales puissent s'expliquer par la dynamique d'interaction présente entre les différents sujets. Mais les compétences individuelles mobilisées sont supposées déjà établies puisque ce sont les conditions des interactions. Si l'intercorporité est expliquée à partir de structures internes préalables, c'est bien là toute l'explication. L'*explanans* est individuel et le social est seulement l'*explanandum*.

Dans le cas du participatory sense making, cette difficulté prend la forme d'une tension entre autonomie des organismes et autonomie du collectif (De Jaegher and Di Paolo 2007). On admet que la dynamique d'interaction peut, en tant que telle, rendre compte d'une activité sociale située pour chaque participant. Mais si l'on accorde une forme d'autonomie à cette dynamique collective avec des propriétés et une histoire propres, cela signifie précisément qu'elle évolue suffisamment indépendamment des individus qui y participent. Dans la mesure où les compétences individuelles ne sont pas affectées, les structures émergentes de la dynamique d'interaction n'existent pas pour l'individu. Comment rendre compte de l'existence d'une signification sociale dans l'activité de sense-making individuelle ? Il semble que l'on soit face à l'alternative : soit la dynamique d'interaction modifie de façon hétéronome l'activité sensorimotrice de chaque organisme, soit elle reste d'un ordre distinct des activités que chacun poursuit pour lui-même.

L'objectif ici est de démêler ces difficultés pour défendre une approche interactionniste, en un sens réellement fort, dans laquelle le social participe à la modification ou la formation des activités cognitives

individuelles. Pour cela, il faut proposer des schèmes explicatifs qui permettent de rendre compte de la façon dont les compétences individuelles pourraient, au moins en partie, être produites à partir de la dynamique sociale⁵. Un enchevêtrement complexe entre collectif et individuel doit être affronté. Pour cela je propose d'articuler trois moments.

1. En toute rigueur, si l'on veut proposer un schème explicatif pour le développement des compétences individuelles de cognition sociale, il est *nécessaire* de se donner des conditions initiales dans lesquelles elles sont *suffisamment absentes*. Il faut donc pouvoir définir tout d'abord des compétences individuelles initiales *non sociales* permettant seulement un engagement dans un couplage avec le monde (par exemple un couplage perceptif).
2. Puis ensuite, sur la base de ces compétences initiales, il faut expliquer comment l'accroche de ces couplages dans une dynamique d'interactions interindividuelles relativement autonome permet l'apparition de propriétés collectives proprement sociales.
3. Et finalement, il faut encore montrer comment par l'appropriation individuelle des propriétés émergentes de la dynamique collective, l'on peut rendre compte de la constitution des compétences individuelles de cognition sociale. Les propriétés collectives doivent ainsi servir à expliquer la formation des compétences qui permettent un engagement dans un couplage dont le sens est social et qui permettront des interactions interindividuelles plus riches.

Pour réaliser cette tâche, nous proposons une méthode expérimentale minimaliste qui permet de clarifier aussi précisément que possible les articulations entre composantes individuelles et collectives de cette dynamique d'interaction, leur ordre causal, et leurs relations logiques d'*explanans* et d'*explanandum*.

Ces travaux s'appuient sur le paradigme expérimental du croisement perceptif et ont déjà été utilisés et repris dans le cadre de diverses discussions théoriques et philosophiques⁶. Il s'agit ici de les présenter à nouveau, sous une forme systématique, pour pouvoir en dégager les principes généraux et les schèmes explicatifs qu'ils permettent de proposer. On verra en particulier l'importance de l'engagement corporel dans les dynamiques d'interaction perceptive pour l'émergence de formes collectives de cognition sociale ; le sens de la distinction entre corps-percevant et corps-objet pour comprendre les processus d'imitation et plus généralement l'appropriation individuelle des dynamiques sociales ; ainsi que le rôle d'un lien entre propriétés déterminées du corps-objet et dynamique sociale pour la construction de

structures associatives individuelles.

Dans le chapitre qui suit, nous allons décrire l'origine et les principes de cette méthode expérimentale (ch.2). Puis dans une série de trois chapitres nous présenterons successivement les situations expérimentales permettant de dégager des schèmes explicatifs possibles de la reconnaissance d'autrui (ch.3), de la perception de ses intentions (ch.4) et de l'imitation de ses expressions faciales (ch.5). Dans la discussion (ch.6), nous reviendrons sur la méthodologie utilisée et nous tenterons de dégager quelques-uns des principes mobilisés par ces modes d'explication.

2 - Paradigme expérimental minimaliste

L'arrière-plan de la méthode minimaliste que nous allons présenter est le programme général d'une recherche visant à comprendre le rôle constituant de la technique pour les activités humaines (Stiegler 1998). Dans le cadre des approches embodied, embedded and extended, la prise en compte des dimensions techniques – outils, machines, environnements – est une façon de comprendre la participation de la culture et de l'histoire sociale dans la constitution des activités cognitives individuelles ou collectives. Là encore, il existe des versions plus ou moins fortes, depuis la conception de l'environnement technique comme simple contexte ou mode de déploiement pour une activité cognitive qui reste essentiellement interne ; jusque l'idée radicale d'un rôle constitutif des technologies cognitives qui rendraient possibles des opérations cognitives ou perceptives qui n'existeraient pas sans elles (Leroi-Gourhan 1964, Goody 1977, Clark 1998, 2004).

Pour étudier systématiquement la façon dont les diverses médiations techniques supportent des domaines d'activité spécifiques, on peut d'abord travailler sur le sens de la saisie d'un outil. Ici une approche énaïve est particulièrement utile. En effet, si la cognition et la perception se réalisent dans le couplage entre l'agent autonome et son environnement, le domaine des opérations possibles dépendra des invariants sensorimoteurs accessibles et donc des possibilités d'agir et de sentir de l'agent (Varela 1991). Un outil saisi et utilisé réalise un changement de ces répertoires d'actions et de sensations, et donne accès à un champ d'opérations possibles spécifiques. Ici, la « constitutivité technique des opérations cognitives » ne signifie pas que ces opérations pourraient se réaliser de façon purement externe sur des

supports artificiels, mais plutôt que les techniques jouent un rôle constitutif par leur participation au couplage entre les organismes et leur environnement. L'étude expérimentale de ce couplage consiste à utiliser la médiation technique pour contrôler systématiquement les entrées sensorielles, les possibilités d'action et observer les opérations cognitives ou perceptives qui deviennent possibles.

Nous avons d'abord travaillé en simplifiant au maximum les médiations techniques pour dégager les conditions minimales nécessaires pour la perception de la localisation spatiale (Lenay et al. 1997, O'Regan & Noe 2001, Lenay & Steiner 2010) et pour la reconnaissance de formes (Lenay et al. 2003). Ici, il est seulement utile de présenter brièvement le dispositif technique qui a servi aux expériences sur la reconnaissance de formes dans un espace bidimensionnel puisqu'il servira aussi de base à nos expériences de cognition sociale.



Le système *TACTOS* que nous avons développé permet une sorte de toucher des formes numériques sur l'écran d'ordinateur. Le curseur que l'utilisateur contrôle avec sa souris, est maintenant remplacé par un unique champ récepteur. Quand ce champ récepteur croise les pixels colorés d'une forme, une stimulation tactile en tout ou rien est délivrée sous l'index de la main libre. Même dans ce cas, les utilisateurs (qu'ils soient des personnes aveugles ou voyantes les yeux bandés) peuvent apprendre à localiser et reconnaître des formes simples. La perception est nécessairement active puisqu'il n'y a aucune spatialité intrinsèque des stimuli sensoriels. Elle se réalise donc essentiellement à travers une trajectoire perceptive qui peut être facilement enregistrée, analysée et modélisée (Stewart and Gapenne 2004). Nous disons que par ses conditions hautement restrictives le dispositif force un déploiement spatial et temporel de l'activité perceptive.



Fig. 1A : Système TACTOS.

La rencontre entre une matrice de 9 champs récepteurs et une forme contrôle l'activation des 9 picots de cellules braille. La forme à l'écran est ainsi perçue de façon tactile.

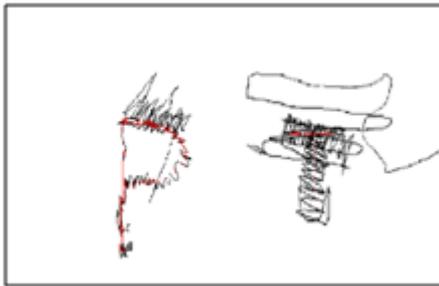


Fig. 1B : Exemples de trajectoires perceptives.

Nous utilisons un seul champ récepteur et une stimulation tactile en tout ou rien. Une zone rouge indique la réception par le sujet reçoit d'un stimulus tactile. On voit le déploiement spatial et temporel de l'activité perceptive quand un sujet reconnaît un P ou un T.

Nous pouvons alors appliquer cette méthode minimaliste à l'étude de la cognition sociale. En effet, dès lors que le dispositif Tactos permet une sorte de "toucher" des formes numériques à l'écran, il devient évident que cet espace peut être partagé entre plusieurs utilisateurs. Mais pour qu'une interaction perceptive soit possible il faut ajouter une chose essentielle : le « corps-percevant » dans l'espace d'interaction, ce qui correspond ici au champ récepteur, doit être lié à un « corps-objet », c'est-à-dire une forme pouvant être perçue par les autres participants. On peut alors construire des situations d'interaction interindividuelle minimales, en simplifiant autant que possible la complexité du contexte d'interaction partagé et les répertoires d'action et de sensation des participants. Ainsi le fonctionnement du couplage collectif et les capacités de cognition sociale qu'il réalise, sont clairement observables. Nous avons appliqué cette méthode pour étudier les phénomènes de la reconnaissance d'autrui, de la perception de ses intentions, et de l'imitation de ses expressions faciales. On espère alors que les mécanismes dégagés dans ces conditions artificiellement simplifiées pourront ensuite servir de schèmes explicatifs pour rendre compte des situations infiniment plus complexes des interactions sociales ordinaires.

Notons que travailler sur la cognition sociale à travers des médiations techniques permet de proposer une définition simple de ce que l'on entend par interaction sociale. Puisque les participants sont à distance, les conditions de leur interaction doivent être complètement spécifiées, sans ambiguïté, par les informations transmises pour leurs actions et perceptions. Ainsi, ils partagent un même espace d'interaction si les actions d'un participant animent en même temps un corps-objet perceptible dans l'espace d'action de l'autre participant et réciproquement. Pour une telle définition, simplement technique, il peut y avoir des formes d'interaction interindividuelles même si les sujets ne sont pas engagés dans une activité explicitement coordonnée. Cela permet d'éviter les difficultés et ambiguïtés qui résultent inévitablement des définitions cognitivistes comme celle proposée par Schoenherr et Westra : « Quand deux ou plus êtres humains conscients affectent mutuellement et en connaissance de cause les actions des uns des autres, ils sont engagés dans une interaction sociale »⁷. Introduire dès la définition de l'interaction sociale les notions de conscience et de connaissance interdit d'avance toute possibilité de comprendre comment les interactions peuvent rendre compte de cette conscience ou de ces connaissances sociales.

3 - Croisements perceptifs et reconnaissance d'autrui

Cette étude de la reconnaissance d'autrui est à présenter quelque peu en détail puisqu'elle a déjà été l'objet de nombreuses discussions (Auvray and Rohde 2012) et a conduit à des expériences et modélisations complémentaires.

Classiquement, dans le contexte de la philosophie de l'esprit et du paradigme représentationnel en sciences cognitives, on considère que le problème de la reconnaissance d'un autre sujet revient à la question de l'adoption d'une « attitude intentionnelle ». Il s'agit alors de déterminer les critères et mécanismes utilisés par le sujet pour décider de traiter les objets perçus, ou bien comme de simples choses obéissant à une causalité mécanique, ou bien comme des agents intentionnels qui agissent en fonction de buts et de représentations internes. De façon cohérente, les études expérimentales menées dans ce cadre établissent une séparation stricte entre les sujets observés et la scène qui est observée. Comme derrière un miroir sans teint, l'observateur est dans la position très artificielle d'un voyeur observant un potentiel agent intentionnel.

Si l'on adopte plutôt une perspective interactionniste, il est clair que les organismes vivants sont toujours déjà engagés dans des relations interpersonnelles. Par conséquent, la question de la reconnaissance d'autrui doit être fondamentalement mutuelle. La reconnaissance d'un agent intentionnel doit se produire durant une interaction où l'agent perçu peut réciproquement reconnaître l'observateur comme étant lui-même un agent intentionnel. Nous avons donc voulu étudier ces situations où deux activités perceptives se rencontrent, comme c'est par exemple le cas dans le toucher mutuel, le croisement de regard, ou la protoconversation entre la mère et le nouveau-né. Nous désignons ces situations par les termes de « croisement perceptif » (Lenay et al 2006, Auvray et al. 2009). Notre hypothèse était que le croisement perceptif permettait la reconnaissance mutuelle des deux sujets, non pas par une opération cognitive d'attribution d'intentionnalité mais comme une expérience directe, l'expérience de « voir autrui autrui me voir » (*seeing the other seeing me* (Gallagher 2014, p.3)).

3.1 - Dispositif expérimental du croisement perceptif

Pour étudier cette situation, nous avons utilisé le paradigme expérimental minimaliste que nous venons de présenter de façon à mettre en place un croisement perceptif élémentaire. Si champ récepteur et corps-objet coïncident et se déplacent en même temps, quand le champ récepteur d'un participant rencontre le corps-objet d'un autre participant, le champ récepteur de ce second participant rencontre le corps-objet du premier. Il y a croisement perceptif.

Pour permettre une analyse précise des dynamiques de perception et d'interaction l'espace d'action des participants est réduit à un espace unidimensionnel partagé (mouvements gauche-droite de leur champ récepteur commandés par la souris d'ordinateur) et le répertoire des retours sensoriels est réduit à une stimulation en tout ou rien (un stimulus tactile délivré sous un doigt de la main libre quand le champ récepteur passe au-dessus d'un pixel noir). Les participants ont les yeux bandés et ne peuvent interagir que via le dispositif. Pour chaque participant trois objets sont présents.

- Le corps-objet de l'autre participant (qui recouvre exactement son champ récepteur : 4 pixels de largeur). Quand deux participants occupent une même position, chacun reçoit une stimulation tactile. C'est le "croisement perceptif".
- Un objet fixe (4 pixels de largeur). L'objet fixe pour le participant 1 est invisible pour le participant 2 et est placé à une position différente.
- Un objet mobile (4 pixels de largeur) que l'on appelle « leurre mobile ». Pour assurer que le leurre mobile ait la même richesse de mouvements que le corps-objet du partenaire nous l'attachons par un lien virtuel rigide à ce corps-objet. Le leurre mobile reproduit donc, à une distance constante, exactement les mouvements réalisés par le partenaire.

Les participants savent qu'il y a ces trois objets mais ils ne savent pas que l'objet mobile est attaché au champ récepteur du partenaire. Ils ont pour consigne de cliquer quand ils pensent être en train de rencontrer leur partenaire.

Cette situation expérimentale permet de tester notre hypothèse théorique : les sujets seraient capables de distinguer le corps-objet du partenaire par le seul fait qu'il correspond à un champ récepteur sensible et animé par une activité perceptive, même s'il a objectivement la même forme et les mêmes mouvements que le leurre mobile ?

Les résultats sont riches et apparemment contradictoires (Auvray et al. 2009). D'une part, c'est un succès : les résultats pour l'ensemble des participants et des sessions montrent qu'une majorité de clics (63%) se produisent quand les deux partenaires se trouvent effectivement à une même position, en situation de croisement perceptif (cf. Fig.3). Ces résultats semblent montrer que les participants réussissent à distinguer le champ récepteur de leur partenaire, ce qui est surprenant puisque, par construction, le leurre mobile a le même comportement objectif que ce champ récepteur. Comme indiqué plus haut, la limitation des entrées sensorielles force un déploiement spatial et temporel des activités perceptives qui peuvent être facilement enregistrées et analysées. L'intérêt de cette méthode minimaliste est que l'on peut connaître tout ce qui s'est produit pour les sujets au long de leur interaction, c'est-à-dire toutes les actions effectuées et données sensorielles reçues.

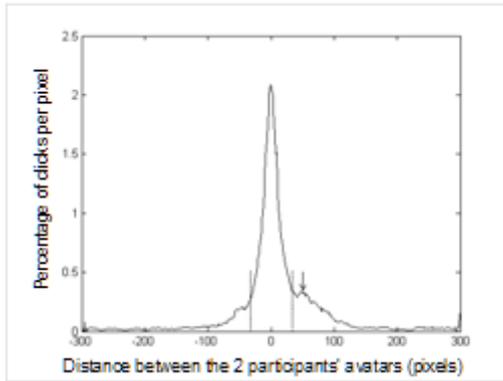


Fig.3. distribution des fréquences

Si l'on regarde la distribution des fréquences de clics en fonction de la distance entre les champs récepteurs des deux participants, on observe un pic très net pour zéro, quand les partenaires sont effectivement face-à-face, et un pic de plus faible amplitude à 50 pixels correspondant aux clics sur le leurre mobile.

Or, l'analyse des stimulations reçues par chaque participant révèle que derrière ce succès collectif se cache un échec individuel : 52% des stimulations reçues proviennent d'un croisement perceptif, 33% proviennent de rencontres avec l'objet fixe et seulement 15% des stimulations proviennent du leurre mobile. Si maintenant, on calcule le rapport entre ces stimuli sensoriels et les clics effectués on trouve 0.33 pour l'objet fixe, 1.26 pour le croisement perceptif et 1.51 pour le leurre mobile (Tab 1).

	Receptor field		Mobile lure		Fixed object	
Percentage of clicks	65.9 %	± 3.9	23.0 %	± 10.4	11.0 %	± 8.9
Percentage of stimuli	52.2 %	± 15.2	15.2 %	± 6.2	32.7 %	± 11.8
Ratio clicks / stimuli	1.26		1.51		0.33	

Tab 1. Ratio entre clics et stimuli en fonction des causes de ces stimuli.

Avec le rapport 0.33 il est clair que l'objet fixe est assez bien reconnu puisque les sujets cliquent peu sur lui alors qu'ils le croisent fréquemment. Mais les rapports entre clics et stimuli montrent qu'en général chaque participant ne semble pas distinguer entre les stimuli dus au croisement perceptif et ceux dus au leurre mobile (1.26 vs. 1.51). La différence des fréquences de clics entre leurre mobile et partenaire (23% vs. 66%) peut être entièrement expliquée par les stratégies de déplacement des sujets qui sont telles que les rencontres avec le leurre

mobile sont beaucoup moins fréquentes que les rencontres avec le partenaire (15% vs. 52%). La probabilité de cliquer à la suite d'une stimulation due à la rencontre avec autrui n'est pas significativement différente de celle de cliquer après une rencontre avec le leurre mobile (elle est même moins bonne). Le succès collectif ne peut être expliqué par une capacité *individuelle* de reconnaître un autre sujet au moyen d'une sensation particulière. Le succès des participants s'explique entièrement par la capacité du *système* supra-individuel qu'ils forment ensemble à se stabiliser en configuration de face-à-face, et non pas parce que chacun aurait reconnu dans le pattern des stimuli un indice permettant de différencier la présence du partenaire de celle du leurre mobile. La seule différence réside dans l'interaction elle-même. Dans cette première expérience où l'objectif est de reconnaître la présence d'un autre sujet, les participants échouent alors que l'action collective réussit, ce qui nous permet de savoir que la dynamique collective s'établit indépendamment d'une telle reconnaissance.

Il est facile de comprendre comment les trajectoires perceptives des deux sujets s'accrochent l'une à l'autre au cours du croisement perceptif (Lenay et al. 2011) et cela a aussi été confirmé par diverses formes de modélisation robotique (Di Paolo et al. 2008, Iizuka H. and Di Paolo E. 2007). Nous y reviendrons dans la discussion théorique.

Ici, l'interaction sociale est constitutive de la reconnaissance d'autrui dans la mesure où la tâche correspondante (cliquer sur le partenaire) est collectivement résolue indépendamment des jugements individuels. Dès lors, la dynamique d'interaction facilite la tâche individuelle de discrimination d'autrui en amenant les participants face à face (Di Paolo 2016). Cependant, cette solution collective ne semble pas accessible à chaque participant puisque l'on n'observe pas de capacité à associer la présence de l'autre sujet à la suite de telle ou telle stimulation sensorielle (sa probabilité de cliquer est la même que ce soit à la suite d'une stimulation due au partenaire ou due au leurre mobile) (Froese and Di Paolo 2011, Michael and Overgaard 2012)⁸. Cette compétence collective ne permet donc pas à chaque individu de s'engager dans des activités ou des relations plus complexes qui seraient basées sur la reconnaissance d'autrui.

Si l'on veut maintenant rendre compte de la formation d'une reconnaissance individuelle de la présence du partenaire, il faut envisager une seconde étape. Mais notons bien que nous avons là un bon point de

départ. Si l'on veut proposer un schème explicatif pour le développement des compétences individuelles de cognition sociale, il est nécessaire de se donner des conditions initiales dans lesquelles elles sont suffisamment absentes alors que les propriétés sociales sont présentes. Notre hypothèse est que c'est seulement dans un second temps, par l'appropriation individuelle des propriétés émergentes de la dynamique collective, que l'on peut rendre compte de la constitution des compétences individuelles de cognition sociale. Pour le comprendre, deux nouvelles expériences ont été réalisées.

3.2 - Nouvelle expérience, tâche collaborative.

Tom Froese a repris l'expérience du croisement perceptif en y apportant seulement deux subtiles différences (Froese et al. 2014) : il place les paires de sujets dans l'esprit d'une tâche collaborative (ils se considèrent comme une équipe qui doit accumuler plus de bons clics que les autres équipes) et surtout il demande à chaque sujet de ne cliquer qu'une fois pour chaque session d'une minute, ce qui renforce cette attention collaborative. Ces changements de consignes sont alors suffisants pour que l'on obtienne 88,5% de bons clics sur le partenaire, 8,6% sur le leurre mobile et 2,9% sur l'objet fixe. Avec 73% de stimulations dues au champ récepteur du partenaire, 13% dues au leurre mobile et 14% dues à l'objet fixe, ces résultats démontrent que, même au niveau individuel, la probabilité de cliquer est plus grande à la suite d'une stimulation due au champ récepteur du partenaire qu'à la suite d'une stimulation due au leurre mobile.

L'examen des trajectoires perceptives croisées montre alors que les sujets réussissent la tâche parce qu'ils mettent en place une dynamique d'interaction plus riche. En effet, la grande différence par rapport à la première expérience est que maintenant l'objectif partagé avec le partenaire est d'écarter les « faux-amis » (les périodes de rencontre avec le leurre qui pourraient tromper parce que par malchance il semblerait s'attacher à soi). C'est seulement à la suite d'une interaction suffisamment longue qu'ils prennent le risque du clic. Pour cela ils doivent déterminer la position d'une singularité sensorielle ayant la plus grande chance de correspondre au partenaire.

De façon très intéressante, les trajectoires montrent l'émergence d'une organisation générale suivant un jeu de « turn taking » : un des deux participants s'immobilise quelques instants, s'offrant comme objet

localisable pour la perception de son partenaire (recevant passivement des stimuli), puis il reprend ses mouvements de perception active alors que maintenant son partenaire s'immobilise pour s'offrir comme objet. Ainsi entrerait en jeu une sorte de dialogue ou de lutte entre sujets tentant de s'objectiver mutuellement, et n'y réussissant que par une alternance des rôles de sujet et d'objet (Sartre 1943). Nous reviendrons sur ce schème explicatif dans la discussion.

Voyons d'abord comment on peut atteindre le même résultat avec une autre expérimentation où l'on distingue clairement la composante déterminable d'autrui et la composante dynamique de l'interaction.

3.3 - Expérience de discrimination sonore

Dans cette variante on reprend encore le protocole expérimental initial du croisement perceptif, mais cette fois, il n'y plus de stimulation tactile : les trois objets (objet fixe, leurre mobile, champ récepteur du partenaire) sont associés à trois sons différents faciles à distinguer (Lenay & Stewart 2012). L'association des trois sons aux trois objets change de façon aléatoire à chaque session expérimentale. La tâche est de déterminer à chaque session quel son est associé au champ récepteur du partenaire. Dans ces conditions les sujets réussissent quasi parfaitement la tâche : 85% de bonne identification du partenaire dès la 4^e session de deux minutes.

L'examen des trajectoires perceptives permet d'expliquer ce résultat : comme dans la première expérience, il y a un attracteur des activités perceptives des deux partenaires. Mais maintenant, les retours sensoriels peuvent être différenciés. Il est alors facile pour les sujets de reconnaître qu'ils se retrouvent bien en face du partenaire. En effet, pour le sujet le croisement perceptif est caractérisé par des stimulations sensorielles fréquentes combinées avec une position indéterminée, contrairement à l'objet fixe qui peut donner des stimulations sensorielles fréquentes mais dans une position bien déterminée, et contrairement à l'objet mobile qui ne donne que des stimulations sensorielles peu fréquentes. Le croisement perceptif définit une propriété qui peut ainsi être liée à un caractère différentiel du corps-objet reconnu par ailleurs.

Nous reviendrons plus loin sur ce schème explicatif. Voyons auparavant comment la même méthode permet d'aborder d'autres aspects de la cognition sociale.

4 - Perception des intentions d'autrui

De nombreux travaux de neurosciences comme de psychologie expérimentale visent à mettre en évidence et expliquer la perception des intentions d'autrui (Meltzoff, 2005, Pacherie 2005). Les structures neuronales comme le *Mirror Neuron System* (MNS), qui sont activées de façon spécifique aussi bien lors de la perception d'un comportement finalisé donné que lors de l'exécution de ce même comportement, peuvent servir à cette explication : l'intention du mouvement perçu serait comprise par l'intention que l'observateur aurait s'il réalisait ce mouvement (Gallese et al. 2004). Ainsi ces structures neuronales donneraient d'avance à chaque individu la capacité d'une compréhension directe, sans raisonnement explicite, des intentions d'autrui. Ces structures sont clairement porteuses d'une signification sociale mais cette signification est-elle en dernier ressort une construction interne où est-elle le produit des interactions ? Dans le premier cas on peut revenir à une conception internaliste de la cognition et les interactions sociales concrètes ne seront que les conséquences de ces compétences individuelles internes. Dans le second cas, ce devrait plutôt être les interactions sociales qui expliquent ces structures internes. Pour trouver les moyens d'envisager cette seconde possibilité, on peut proposer une nouvelle expérience minimaliste.

Dans cette expérience, les participants partagent un espace bidimensionnel. Chaque participant est doté d'un unique champ récepteur couplé à un seul stimulateur tactile. Le corps-objet est un cercle attaché au champ récepteur (cf. Fig 4). Les participants déplacent leur champ récepteur (avec le corps objet) avec leur souris d'ordinateur. Comme pour le toucher, la perception est parfaitement réciproque : je ne peux toucher une autre personne sans être touché par elle. Il n'y a ni objet fixe, ni leurre mobile. L'un des participants a une mission très simple que l'autre doit deviner « aller vers le haut à droite » ; « à gauche » ; etc. Le suiveur essaie de trouver son guide et de rester autant que possible en contact avec lui. Le guide doit être collaboratif en essayant d'aider son partenaire à deviner ses intentions. Ils n'ont pas d'autres moyens d'interaction que le dispositif.

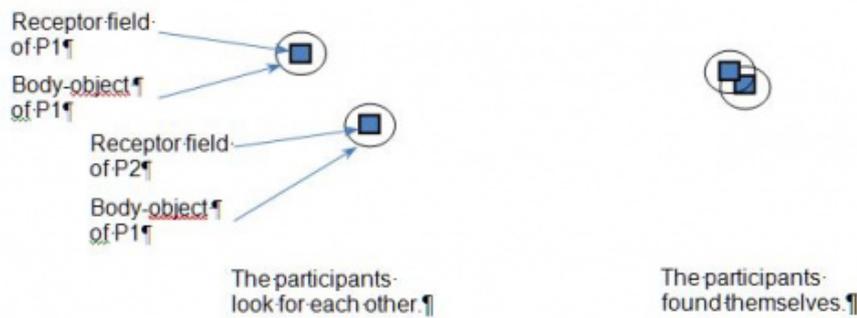


Fig. 4 : un espace bidimensionnel

Les participants P1 et P2 partagent un espace bidimensionnel dans lequel ils déplacent leur champ récepteur (avec leur corps-objet). Quand le champ récepteur d'un participant rencontre un pixel coloré, il reçoit un stimulus tactile sous un doigt de sa main libre. Les participants doivent se trouver et se suivre.

La plupart des participants réussissent la tâche qui pourtant n'est pas facile puisqu'il n'y a aucune information spatiale dans la donnée sensorielle : le champ récepteur n'est qu'une petite surface (ce n'est pas une rétine qui pourrait indiquer localement un mouvement). A chaque instant le suiveur n'a pas d'autres moyens pour deviner la direction du mouvement du guide que de quitter le contact (par exemple par un petit balayage autour de lui), et s'engager dans une direction en espérant pouvoir le retrouver. Progressivement, avec prudence et attention, le suiveur réussit à anticiper la direction du mouvement que lui indique le guide. Les conditions de restrictions techniques extrêmes de cette situation d'interaction produisent une équivalence entre la perception du geste d'autrui et celle de son propre geste. Il n'y a pas une perception des mouvements du partenaire qui précéderait son imitation ou sa simulation, mais accroche de l'activité perceptive de l'observateur aux mouvements du guide.

Ajoutons que ces conditions d'interaction minimalistes sont telles que la perception des mouvements du guide ne peut être que la perception des intentions de ces mouvements, puisqu'à chaque instant le suiveur doit parier sur une direction. Il y a une compréhension des intentions d'autrui par partage de ces intentions (Merleau-Ponty, 1945 ; Wilkerson, 1999). Ceci se voit très clairement si l'on propose la nouvelle mission paradoxale au guide qui consiste à suivre son partenaire (Fig.5). La dynamique d'interaction de cette situation où les participants se suivent mutuellement fait émerger des directions collectives qui semblent finalisées bien qu'aucune orientation intentionnelle ne préexistait.



Fig.5. dynamique d'interaction

On a séparé les trajectoires perceptives des deux sujets en interaction. Ici le guide (trajectoire de droite) a reçu la consigne paradoxale de suivre son suiveur. A travers la dynamique d'interaction il y a émergence d'une intention partagée. Les participants s'inventent ensemble les orientations d'un parcours, ici aller à gauche puis en haut à droite.

On examinera plus loin dans la discussion générale si cette situation limite peut éclairer la question du statut des compétences de perception des intentions. Auparavant, voyons une troisième expérience à propos de l'imitation des expressions faciales (Lenay & Stewart 2012, Froese et al. 2012).

5 - Imitation des expressions faciales

Comment expliquer la compétence initiale d'imitation des gestes et expression par le nouveau-né ? Comment comprendre que dans les premiers jours après sa naissance il puisse établir une relation d'équivalence entre les mouvements qu'il observe et les données motrices et proprioceptives de ses propres actions. Pour résoudre ce problème de correspondance (*correspondence problem* (Brass and Heyes, 2005)), il semble difficile de s'appuyer sur un apprentissage par le nouveau-né, surtout quand il s'agit des actions comme ses propres mouvements faciaux qu'il ne peut voir lui-même (*opaque actions*). Meltzoff et Moore (1977) supposent donc l'existence de compétences individuelles innées, un « Système actif d'appariement intermodal » (*Active Intermodal Matching system, AIM*) qui réalise une représentation supra-modale des actions corporelles qui sont vues ou réalisées. De même façon, on peut aussi supposer l'existence d'un Système neuronal miroir inné qui réaliserait un câblage entre les expressions faciales perçues et les expressions produites (Rizzolatti et al., 2002).

On a vu que même dans un cadre interactionniste, on admettait que cette imitation néonatale était une compétence initiale permettant l'apparition d'une « intersubjectivité primaire » (*primary intersubjectivity*), c'est-à-dire des processus très précoces qui engagent le nouveau-né dans des interactions dyadiques avec l'adulte (Trevarthen, 1979). L'intégration intermodale de la vision et de la proprioception permet au nouveau-né de

construire un sens pragmatique des expressions d'autrui par les réponses qu'elle provoque (Gallagher and Meltzoff, 1996). Elle permet de comprendre l'advenue d'une dynamique d'interaction réciproque qui produit différentes formes de synchronisation dans des séquences qui échappent en partie au contrôle de l'un ou l'autre participant.

Cependant, là encore, si l'on accepte un « système actif d'appariement intermodal » ou un « Système de neurones miroir » innés, la perspective interactionniste est fragilisée. La cognition sociale n'est pas d'abord expliquée par les interactions interindividuelles mais par le fonctionnement de structures internes. La question reste de comprendre l'origine ou la genèse de ces structures.

Le chemin que nous proposons d'explorer ici consiste à chercher les conditions pour l'apparition de « phénomènes mimétiques » au cœur des interactions perceptives, c'est-à-dire de mettre en place une situation où ce n'est pas l'imitation qui explique les interactions mais la dynamique d'interaction qui produit l'imitation. Nous avons donc repris notre expérience du croisement perceptif dans un espace unidimensionnel, mais cette fois, les participants peuvent modifier la forme qu'ils présentent à leur partenaire. En accord avec notre approche minimaliste, nous avons choisi la modification du corps la plus simple possible : la distance relative entre le corps-objet et le champ récepteur (cf. Fig.6). Le champ récepteur n'est plus directement perceptible par le partenaire. Seul est perceptible le corps-objet attaché à ce champ récepteur. La distance entre son champ récepteur et son corps-objet peut être activement modifiée par le participant. En cliquant sur le bouton droit ou gauche de sa souris d'ordinateur, il peut déplacer son corps-objet vers la droite ou vers la gauche relativement à son champ récepteur (deux pixels pour chaque clic). Avec leur souris d'ordinateur les participants déplacent leur champ récepteur dans un espace unidimensionnel partagé et cherchent à se percevoir mutuellement. Cependant, ils ne peuvent connaître la position de leur corps-objet relativement à leur champ récepteur.

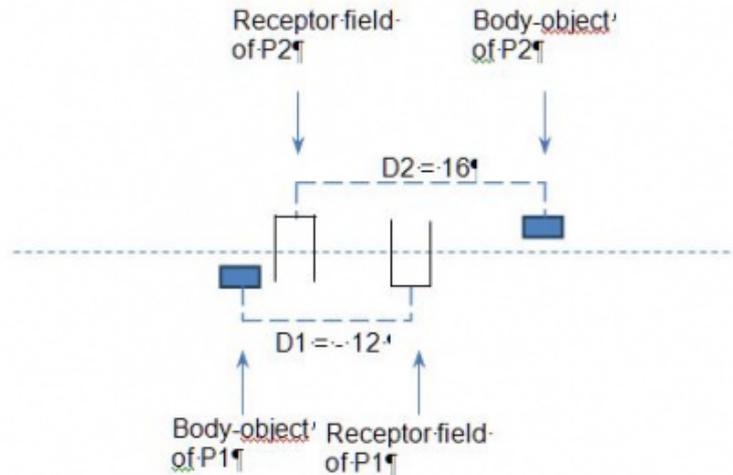


Fig.6. perception mutuelle

Les participants partagent un espace unidimensionnel (un tore) dans lequel ils déplacent leur champ récepteur en même temps que leur corps-objet (qui lui est lié à une distance D). Quand le champ récepteur d'un participant rencontre le corps-objet de l'autre participant il reçoit un stimulus tactile sous un doigt de sa main libre. Ici $D2 > D1$; si les participants recherchent un croisement perceptif ils seront entraînés dans une dérive collective vers la droite.

Nous appelons $D1$ la position relative du corps-objet du participant 1 par rapport à son champ récepteur, et $D2$ pour le participant 2. Quand $D1 + D2 = 0$, si le champ récepteur du participant 1 est en face du corps-objet du participant 2, alors le champ récepteur du participant 2 est aussi en face du corps-objet du participant 1. Dans cette configuration il ne devrait pas y avoir de difficulté à réaliser et maintenir un croisement perceptif dans la mesure où les deux partenaires peuvent se percevoir mutuellement en même temps. Au contraire, si $D1 + D2 < 0$, le croisement perceptif devrait être déséquilibré, chaque participant devant se déplacer vers la gauche pour retrouver son partenaire. De même, si $D1 + D2 > 0$, le croisement perceptif devrait dériver vers la droite. Au cours de leur interaction les participants doivent sentir cette dérive et utiliser les clics de la souris d'ordinateur pour stabiliser leur croisement perceptif.

L'objectif donné aux participants est de stabiliser leur croisement perceptif dans une position suffisamment fixe. Ils savent que leurs clics sur les boutons droit ou gauche de leur souris d'ordinateur peuvent les aider mais ils ne connaissent pas la nature de ces actions. On observe clairement une convergence vers une situation où $D1 + D2 = 0$, situation que l'on peut assimiler à une imitation. En 3 minutes, le déséquilibre $D1 + D2$ est réduit à moins de 30% de sa valeur initiale. Même si les participants ne savent pas la position relative de leur corps-objet, aussi

bien au commencement qu'à la fin, leur clics agissent sur une variable commune, la distance relative $D1+D2$, qui détermine l'équilibre des interactions perceptives. En amenant cette variable à zéro, ils produisent une stabilisation du croisement perceptif qui, du point de vue d'un observateur extérieur, correspond à une ressemblance en miroir des images qu'ils présentent à leur partenaire. Il faut noter qu'il y a une infinité de solutions possibles à la ressemblance entre les deux participants, c'est-à-dire de couples de valeurs vérifiant l'égalité $D1+D2=0$, et effectivement, la variance des solutions croit au cours des séquences d'interaction (Lenay & Stewart 2012).

L'analyse des résultats expérimentaux indique que les participants réussissent à accorder leurs corps en réponse à la stabilité du processus d'interaction parce qu'ils sont capables de percevoir la dérive systématique dans le flux des interactions produite par le déséquilibre dans les configurations corporelles. Cette dérive ne peut être réduite aux actions d'un des participants. Au contraire, elle émerge de leur interaction. L'imitation est donc rendue possible par une propriété collective du processus d'interaction en tant que totalité. Elle s'explique parce que l'ajustement de la synchronisation du croisement perceptif est réalisé par des actions sur la variable spatiale correspondant aux configurations corporelles des participants.

Dans cette expérience, « l'imitation » est le résultat et non la cause. Elle n'est pas réalisée par la comparaison entre la perception externe de la configuration corporelle du partenaire et la perception ou la connaissance interne de ses propres configurations corporelles. Mais dans un second temps, cette imitation *de facto* devrait permettre un tel apprentissage. Nous allons y revenir.

6 - Discussion

6.1 - Sur la méthode : schématisation et minimalisme

La radicalité du minimalisme des expériences que nous avons présentées permet d'adopter une approche externaliste de la perception. Comme indiqué plus haut, les activités perceptives sont externalisées sous forme de trajectoires qui peuvent être aisément enregistrées, ce qui permet une analyse complète des relations sensorimotrices. On peut d'ailleurs, inversement, étudier l'effet d'un relâchement du minimalisme par augmentation du nombre de champs récepteurs. Si, plutôt qu'un

unique champ récepteur, l'utilisateur commande les déplacements d'une rétine composée d'une *matrice* de champs récepteurs (par exemple 4 x 4 champs récepteurs), il y a un parallélisme spatial entre les stimuli reçus par l'utilisateur à un instant donné et la surface couverte par cette rétine dans l'environnement virtuel. Si l'on étudie la façon dont les trajectoires perceptives se modifient à mesure qu'augmente ce parallélisme, on observe une économie de mouvements et donc une forme d'internalisation progressive des processus perceptifs (Sribunruangrit 2004). Le parallélisme des entrées sensorielles permet une relative séparation entre perception et action externe, même si des mouvements actifs de la matrice semblent toujours nécessaires.

Dans le cas des interactions sociales, il est particulièrement révélateur de partir d'une situation maximale externalisée en n'accordant aux participants qu'un bit d'information à chaque instant et en contraignant leur espace d'action à une ou deux dimensions de déplacements en translation. Une telle dynamique déployée sous nos yeux semble réaliser les conditions empiriques pour une discussion théorique en réduisant au maximum les ambiguïtés conceptuelles sur les composantes individuelles ou collectives de la cognition humaine.

Il y a bien sûr un écart énorme entre ce type de situation expérimentale artificielle et les conditions normales de réalisation des activités humaines. Les espaces d'action et de sensation sont appauvris au maximum, il n'y a pas de structure sociale générale préexistante et, pire encore, les interactions ne concernent à chaque fois que deux personnes. On peut cependant faire deux remarques.

1. La première est que notre objectif est seulement de proposer des façons de comprendre les phénomènes, ce que nous appelons des schèmes explicatifs. Il ne s'agit pas de résoudre les questions de la cognition sociale humaine dans toute leur complexité, mais seulement de proposer des schèmes qui pourront servir de modèle pour des explications adaptées aux situations réelles. Notre hypothèse est qu'une dynamique d'interaction identifiée dans des conditions minimalistes pour expliquer une capacité de cognition sociale (permettant telles performances dans la situation d'observation) peut être retrouvée dans les conditions complexes des situations naturelles pour expliquer des performances cognitives similaires. C'est alors une question purement empirique de vérifier si ce schème explicatif fonctionne encore et s'il propose une explication suffisamment parcimonieuse des phénomènes.

2. La seconde remarque est qu'une telle approche explicative n'est pas réductionniste. Il ne s'agit pas de ramener un phénomène observé à des causes plus simples et différentes qui en expliqueraient seulement l'apparence, comme quand on prétend réduire les phénomènes cognitifs à une dynamique neuronale. Au contraire, il s'agit pour chaque situation expérimentale de construire un monde qui fonctionne, même s'il s'agit d'un monde très simple. A chaque fois il s'agit de donner l'occasion de la constitution d'un monde propre, un "Umwelt", spécifique qui correspond au système général des phénomènes accessibles pour les participants. La réussite est la viabilité d'une éaction et non pas l'adéquation d'une représentation. C'est une validation technologique, à la façon de ce que Francisco Varela appelle une « validation par construction » dans les travaux sur la vie artificielle (Varela 1997). Ici, même si les modes d'interaction et les structures de l'environnement sont aussi simples que possible, ils permettent tout de même de comprendre l'éaction d'un monde partagé entre les sujets en interaction. En même temps, ce type de situation expérimentale est suffisamment simple pour que l'examen des activités croisées des différents sujets permette l'abduction d'hypothèses minimales sur leurs stratégies d'action individuelles.

6.2 - Approche sociale de la cognition sociale

On a remarqué en introduction que pour dépasser les difficultés du « mind reading », on devait commencer par reconnaître que la plupart des activités intentionnelles étaient directement perceptibles à même les comportements, sans avoir besoin d'inférence sur des représentations mentales. Cependant, cela ne suffit pas.

Si les mécanismes de cette perception étaient eux-mêmes essentiellement internes, on reviendrait au problème des intentions cachées. C'est ce que l'on risque en admettant au départ les compétences individuelles de reconnaissance du visage, de reconnaissance des intentions et d'imitation des expressions (Gallagher 2013, 61). Il faut donc ajouter que l'activité perceptive elle-même doit être suffisamment incarnée et exprimée dans le comportement. C'est ce qui

se réalise dans les conditions minimalistes que nous avons adoptées.

Mais c'est encore insuffisant. Si l'on veut que l'activité perceptive individuelle ait un sens social, il faut que cette activité elle-même puisse être configurée à travers la dynamique d'interaction interindividuelle. C'est du moins une des spécificités des schèmes explicatifs proposés ici.

Les trois expériences que nous avons présentées correspondent à trois schèmes explicatifs pour la reconnaissance d'autrui, la perception de ses intentions, et l'imitation des expressions faciales. L'intérêt de ces schèmes explicatifs est que ces phénomènes de cognition sociale ne sont pas expliqués par des capacités individuelles qui les précéderaient. Au contraire, c'est la dynamique des interactions sociales qui explique l'émergence de ces phénomènes et, possiblement, secondairement, la reconfiguration des capacités individuelles.

Les compétences individuelles initialement mobilisées dans ces situations n'ont pas d'avance une signification sociale. Elles consistent simplement en le maintien de l'autonomie individuelle par la recherche des invariants sensorimoteurs d'un couplage stable. L'ordre logique de ces schèmes explicatifs consiste à partir d'interactions sociales *de facto* résultant du couplage de ces simples capacités individuelles à travers une structure minimale de l'environnement physique et social déjà là (ch. 6.2.1). On observe alors l'émergence d'une organisation des interactions réalisant concrètement le phénomène de cognition sociale qui nous intéresse (reconnaissance d'autrui, perception de ses intentions, imitation de ses expressions). Puis, la capacité individuelle à se saisir des résultats de la dynamique collective peut se réaliser par une liaison entre cette dynamique et des propriétés perceptibles du corps-objet du partenaire (ch. 6.2.2), liaison qui donne alors les conditions d'un apprentissage associatif (ch. 6.2.3). On peut ainsi comprendre la stabilisation de nouvelles structures individuelles qui permettront de nouveaux types d'engagements sociaux. Nous allons détailler ces trois étapes pour nos trois situations de cognition sociale avec leur conséquence générale (ch. 6.2.4.), puis nous examinerons les conditions générales d'application de cette approche de la cognition sociale (ch. 6.3 et 6.4).

6.2.1 - Compétences initiales et dynamique d'interaction

Pour la **reconnaissance d'autrui**, la tâche est collectivement résolue même si chacun est incapable de reconnaître la présence d'autrui

derrière une stimulation sensorielle particulière. Les compétences initiales nécessaires n'ont rien de social. Elles correspondent simplement à la capacité de rechercher un couplage stable avec une forme perçue, c'est-à-dire ici déterminer la position spatiale d'un objet par la maîtrise d'une loi sensorimotrice réversible, un balayage par aller-retour autour d'une variation de stimuli sensoriels. La stratégie très générale de chaque sujet consiste donc à inverser le mouvement de son champ récepteur à la suite d'un événement sensoriel. Ainsi, s'il croise son partenaire il reviendra sur ses pas alors que celui-ci fera de même. Les deux champs récepteurs devront alors entrer dans une sorte de danse (chacun oscillant l'un autour de l'autre). Cela peut être décrit comme constituant un attracteur de la dynamique d'interaction collective, un attracteur qui n'est pas un point spatialement situé, mais une région qui peut se déplacer. C'est là une condition suffisante pour la formation d'une accroche des deux dynamiques perceptives qui tendra pour chacun à augmenter la probabilité de la présence de son partenaire. La dynamique d'interaction émergente du croisement perceptif explique la résolution collective de la tâche (cliquer sur le corps-objet du partenaire).

Pour la **perception des intentions** les compétences individuelles initiales nécessaires consistent en la possession d'un savoir-faire permettant la poursuite d'une cible dans un espace bidimensionnel. Le couplage des activités perceptives individuelles produit *de facto* un mouvement intentionnel partagé.

Pour **l'imitation des expressions faciales**, les compétences initiales nécessaires sont à peine plus complexes. Il s'agit d'être capable de maintenir un couplage perceptif et de percevoir sa dérive pour déterminer des actions de correction. Une forme très simple d'imitation résulte directement de la dynamique d'interaction, même quand les participants ignorent ce qu'ils donnent à percevoir à leur partenaire. Ils ne procèdent pas par comparaison entre une forme ou un comportement perçu et la forme ou le comportement produit. L'imitation est produite dans la recherche d'un équilibre du croisement perceptif réalisé par le double ajustement spatial des liens entre champs récepteurs et corps-objet.

Ce dispositif expérimental peut sembler trop simplifié et trop artificiel pour que ces observations puissent être appliquées au phénomène de l'imitation néonatale. Cependant, si l'on utilise ce schème explicatif pour décrire les jeux d'imitation dans la proto-conversation, on comprend que le nouveau-né soit capable de reconnaître un accord dans l'interaction

sans avoir besoin de déterminer au préalable si ses expressions faciales reproduisent correctement celles de l'adulte (Reddy, 2003 ; Trevarthen & Reddy, 2007).

6.2.2 - Liaison de la dynamique sociale avec des propriétés déterminées du corps-objet du partenaire

Pour la **reconnaissance d'autrui**, nous avons vu dans deux expériences complémentaires que la capacité individuelle de reconnaissance d'autrui nécessitait une détermination suffisante du corps-objet du partenaire. Soit (ch. 3.2) parce qu'un « turn taking » permet à chacun de se présenter tour à tour comme objet dans une position suffisamment déterminée pour pouvoir être désigné à la fin de l'essai, soit (ch. 3.3.) par des traits différentiels sonores des corps-objets. Dans les deux cas, les sujets peuvent reconnaître et désigner autrui en liant ce corps-objet déterminé à la dynamique d'interaction du croisement perceptif. L'application de ce schème explicatif à la situation de proto-conversation entre la mère et l'enfant autorise à assimiler cette forme complexe d'interaction à un type de croisements perceptifs et permet de comprendre que son absence (comme dans le cas d'une interaction avec un enregistrement) puisse être reconnue et vécue négativement par l'enfant (Trevarthen 1979, 1993). Dans cette situation naturelle complexe, il y a bien sûr pour l'enfant une multitude d'éléments additionnels qu'il perçoit de façon déterminée, en particulier le visage de sa mère et les autres objets dans l'environnement. La dynamique du croisement perceptif est liée avec la perception cooccurrence de ce visage.

Remarquons que cette capacité ne correspond pas aux effets d'un « détecteur de contingence » individuel, mais simplement à l'association entre les dynamiques du croisement perceptif et un indice (les propriétés sonores des objets). La dynamique collective émergente peut être ressaisie par le sujet en associant l'indétermination de position du partenaire avec une propriété sensorielle suffisamment déterminable. Dans cette situation encore, il semble que le croisement perceptif joue un rôle constitutif pour la perception d'autrui dans la mesure où ce qu'est la perception d'autrui, dans les limites de cette expérience minimaliste, est en fait la présence de la dynamique particulière du croisement perceptif liée à la propriété perçue.

Pour la **perception des intentions**, la dynamique d'interaction lie directement les mouvements des deux partenaires. La reconnaissance

du geste du partenaire est liée à la reconnaissance du geste produit pour le percevoir.

Pour l'**imitation des expressions faciales**, on a vu que l'accord entre les partenaires pouvait se faire sur une imitation *de facto* qui avait une infinité de solutions. Si l'on veut que les individus puissent s'approprier l'imitation d'une expression particulière il faudrait, outre les conditions actuelles de l'expérience, leur donner aussi un accès perceptif à la forme particulière du « visage » du partenaire (ici la grandeur du lien entre champ récepteur et corps-objet). Dans les situations naturelles, cela signifie lier la dynamique d'interaction réalisant une imitation *de facto*, à la perception d'une expression faciale particulière.

6.2.3 - Apprentissage associatif en situation d'interaction

La liaison établie par la situation sociale entre d'une part, l'organisation émergente de la dynamique d'interaction et d'autre part, des propriétés déterminées du corps-objet du partenaire, devrait ensuite permettre un apprentissage associatif individuel de cette corrélation.

Pour la **reconnaissance d'autrui**, la stabilisation des interactions dans le croisement perceptif permet sa liaison avec la position du corps-objet du partenaire ou avec un indice sonore qui peuvent prendre le sens d'une rencontre avec autrui. Dès lors est possible un apprentissage individuel de l'association entre ces propriétés du corps-objet et le comportement de croisement perceptif. Cela signifierait, en situation naturelle, l'apprentissage individuel de la liaison entre le stimulus que représente un visage et la dynamique d'interaction perceptive. Cette nouvelle compétence individuelle de reconnaissance de la présence d'autrui dans un visage perçu trouverait ainsi son origine et son contenu intentionnel dans l'interaction sociale qu'elle permet.

Pour la **perception des intentions**, on suppose généralement que des structures internes comme le MNS permettent de comprendre dans le mouvement perçu, un mouvement équivalent que l'on pourrait produire. Or, dans l'expérience très simple que nous avons proposée, ce sont les activités perceptives elles-mêmes qui sont dynamiquement liées. Ceci donne une piste pour expliquer l'origine de ces structures internes. Le geste perçu est associé à un geste effectué parce qu'il est perçu en étant effectué. Les arrangements neuronaux qui commandent la réalisation d'un geste spécifique sont en même temps mobilisés lors de la

perception de ce geste réalisé par autrui. Dans la mesure où l'interaction stabilise un geste donné, on peut comprendre que puissent se construire des structures qui associent les sensations multimodales concomitantes reçues, qu'elles soient d'origine extéroceptive (le geste d'autrui), ou proprioceptive (son propre geste) et ceci devrait rester vrai quand, avec un parallélisme des entrées sensorielles (rétine), une séparation tend à s'établir entre l'activité perceptive et les mouvements réalisés ou observés. On a donc là une piste pour expliquer, par le sens fonctionnel des interactions, la formation de structures internes de type MNS qui permettront ensuite de reconnaître l'intention d'un mouvement observé sans avoir besoin de l'effectuer. Cette compétence nouvelle trouve son origine et son contenu intentionnel dans l'engagement qu'elle permet dans des interactions avec la personne observée.

Pour l'**imitation des expressions faciales**, là aussi, la dynamique collective du croisement perceptif fait advenir une situation d'interaction sensorimotrice qui est suffisamment stable pour servir de base pour un apprentissage associatif. Nous avons vu que dans la situation naturelle de la protoconversation, la perception d'un accord précède une conscience explicite de ce à propos de quoi il y a accord (Reddy 2008). L'imitation résulte spontanément de la régulation mutuelle de la dynamique collective d'interaction, et c'est cette interaction sociale qui donne au nouveau-né la motivation et le moyen pour lier sa perception d'autrui avec ses sensations. Ce serait seulement plus tard que l'enfant découvrirait que ce qu'il fait est en fait une imitation. Et dès lors, sur la base d'un accord dans le croisement perceptif, il pourra supposer que ses propres expressions faciales, qu'il ne peut voir, ressemblent à celle de son partenaire qu'il voit.

Ce qui est ici particulièrement intéressant, c'est qu'il s'agit d'une association entre le visage du partenaire et une dynamique d'interaction qui est socialement signifiante. La logique classique de l'imitation est inversée. C'est une imitation *de facto* résultant de la dynamique collective, qui procure le moyen de lier l'image perçue à des sensations proprioceptives. L'existence de structures comme les neurones miroirs pour les expressions faciales peut être expliquée par une telle situation sociale qui permet une association entre différents flux synchronisés d'inputs sensoriels multimodaux et proprioceptifs, input sensoriels provenant à la fois du comportement de l'autre sujet et des propres actions du sujet⁹.

Remarquons que la modification du lien entre champ récepteur et corps-objet sont des modifications d'une inscription corporelle des stratégies d'action (Maillet et al. 2008). En effet, la distance entre champ récepteur et corps-objet définit la façon dont chacun intervient dans la dynamique d'interaction et la dérive collective qui s'en suit. Suivant ce schème explicatif, une expression faciale correspond à une configuration de la relation entre corps percevant et corps-objet, configuration qui peut être associée à une forme donnée du visage pour chaque type d'engagement interactionnel. Dans le cadre d'une conception relationnelle des émotions (Dumouchel 1995, Damiano 2015) un apprentissage du sens émotionnel des expressions devient ainsi concevable. La forme du visage du partenaire est en même temps associée à une structure relationnelle signifiante particulière (par exemple un moment de joie et d'approche ou un moment de dégoût et de recul), il pourra donc y avoir association entre cette structure sociale et cette expression faciale.

6.2.4 - Nouvelles compétences de cognition sociale

Dans sa généralité, ce type de schème explicatif ne décide pas en faveur d'une conception héréditaire (innéiste) ou environnementale de la cognition sociale humaine. Mais il milite clairement en faveur d'une approche interactionniste, et donc une approche pleinement sociale de la cognition sociale. L'apprentissage perceptif que nous supposons consiste à profiter de configurations stables de la dynamique d'interaction pour favoriser la formation de structures internes (neuronales) associant les flux d'actions et de sensation qui résultent en même temps de cette dynamique d'interaction interindividuelle et du couplage avec des éléments déterminables de la situation (propriétés du corps-objet du partenaire). L'association entre cette dynamique socialement signifiante et une propriété déterminée peut aussi bien être le résultat d'un apprentissage perceptif associatif individuel, que de la sélection de caractères héréditaires qui réalisent cette association. Le point logique crucial ici est que les structures neuronales individuelles qui participent à cette association puissent être en partie le résultat et non la cause première des dynamiques d'interaction interindividuelle.

En cohérence avec l'approche éactive adoptée ici, nous ne demandons pas que ces structures internes soient des représentations « tenant lieu » des situations qui ont servi à leur apprentissage, comme si elles pouvaient se substituer à la présence actuelle des interactions sociales. Ces structures sont plutôt à comprendre comme les supports de

stratégies d'action stabilisant ces situations et permettant l'engagement dans des formes de couplage social plus riches. Par exemple, l'association entre la forme déterminée d'un visage et la dynamique du croisement perceptif permettra ensuite à un individu de s'appuyer sur la seule perception d'un visage pour s'engager dans une interaction en y supposant d'avance l'existence d'une présence perceptive qui pourra lui répondre. De même, pariant d'avance sur une intentionnalité dans un mouvement observé chacun pourra s'engager dans d'autres mouvements ; ou, percevant une expression faciale, il pourra répondre par une autre expression.

En même temps, la liaison entre des propriétés perceptivement déterminées et la signification sociale de dynamiques d'interaction, permet de comprendre que l'on puisse s'engager dans un couplage avec des traces ou images en trouvant un sens social à ce couplage, même en l'absence d'interaction actuelle avec autrui. C'est là une voie possible pour rendre compte dans un cadre éactif des capacités d'imaginer, raisonner et planifier. Il suffit de comprendre ces activités, non pas comme la manipulation de représentations mentales, mais comme le résultat d'un jeu de couplage avec des traces de la présence sociale.

6.3 - Distinction corps percevant / corps-objet

Au long de ces explications, il a été nécessaire de distinguer corps percevant et corps perçu. Cette distinction s'imposait d'abord du point de vue des médiations techniques. En effet, pour qu'une interaction interindividuelle soit possible il faut définir pour chaque individu, d'une part ses champs récepteurs, qui permettent la réception de stimuli sensoriels lorsque son corps-percevant rencontre des objets, et d'autre part son corps-objet, qui peut affecter autrui. Il faut alors que ces deux aspects du corps soient liés, soit de façon constante (ch. 3 et 4), soit de façon variable (ch. 5).

D'une certaine façon, cette distinction correspond à la distinction phénoménologique entre le corps propre et le corps pour autrui (Merleau-Ponty 1945), et il ne s'agit surtout pas de les opposer comme dans un dualisme du corps et de l'esprit. Ici, dans un cadre pragmatiste, il s'agit seulement d'une distinction entre le corps tel qu'il permet de percevoir et le corps tel qu'il entre dans les interactions sociales, c'est-à-dire tel qu'il affecte les autres et peut être perçu.

L'individu ne perçoit pas le corps qu'il donne à percevoir aux autres, il ne perçoit pas son corps-objet. Cela permet de caractériser une différence concrète de points de vue situés, distincts les uns des autres. Il est constitutif du social que nous ne percevons jamais exactement les mêmes choses (sinon, on ne pourrait pas se percevoir mutuellement). Dans le cas du visage, on distinguera donc le visage comme organe de perception et le visage comme organe d'interaction. Une distinction qui fonctionne sans être directement accessible par l'individu. Son rôle est proprement social.

Le corps-objet ne participe pas à l'activité perceptive individuelle initiale, mais il participe *indirectement* à la perception sociale puisqu'il est nécessaire à l'établissement des dynamiques d'interaction qui permettent à chacun de donner un sens social à son activité. Cette distinction est fonctionnellement nécessaire à la formation des interactions sociales que nous avons examinées. Dans les différentes expériences, les croisements perceptifs correspondent à une dynamique qui échappe à chaque participant parce qu'elle dépend d'actions perceptives portant sur des corps-objets que les participants ne perçoivent pas eux-mêmes et qui sont pourtant liés à leur activité perceptive. C'est la composante radicalement indéterminable du corps-objet d'autrui. Cependant, par ailleurs, on a vu que l'appropriation individuelle d'une organisation de la dynamique d'interaction passait par des moments où chacun pouvait suffisamment déterminer des propriétés du corps-objet de son partenaire pour ensuite les lier à cette organisation. Ainsi, la distinction entre le corps-percevant et le corps-objet rebondit au niveau du corps-objet comme une distinction entre le corps-objet qui sert à entrer dans une dynamique d'interaction (dans laquelle il ne peut être précisément déterminé) et le corps-objet qui porte des propriétés déterminables. Cette tension qui se maintient semble bien constitutive de l'expérience du visage d'autrui : à la fois quelque chose qui résiste à sa détermination et pourtant qui persiste à être présent et que l'on peut partiellement déterminer.¹⁰

6.4 - Enaction et technique

Le collectif, la dynamique d'interaction qui échappe à chaque individu, peut en retour doter cet individu de nouveaux pouvoirs d'agir, d'interagir et donc de capacités de sense-making qui n'existaient pas avant cet engagement social (Di Paolo & De Jaegher 2016). Le social transforme le champ de l'autonomie individuelle en lui donnant accès à un domaine

d'action et de sens nouveau. Mais dire que la dynamique des interactions collective façonne les structures individuelles de la cognition sociale, n'est-ce pas admettre un déterminisme social qui remettrait en question l'autonomie individuelle ? Ce n'est pas le cas si l'on comprend ce retour du social sur l'individu comme un mécanisme d'appropriation individuelle qui peut bien sûr tout aussi bien être source d'aliénation que de pouvoir d'action (empowerment).

Pour comprendre cette rétroaction constituante du collectif sur l'individu, on peut prendre le modèle de l'outil. Socialement fabriqué, il est donné à l'individu comme nouveau pouvoir de percevoir, d'agir et d'interagir. Le social donne à chacun un environnement technique et langagier constitutif des activités individuelles. Or l'outil reçu pose d'abord le problème de son appropriation. Il force l'apprentissage de stratégies spécifiques, c'est-à-dire par exemple de configurations neuronales permettant la maîtrise de son usage.

Pour les mécanismes de cognition sociale que nous avons présentés, le corps-objet joue le rôle de l'outil. Il est donné par autrui (puisqu'il n'est pas directement perceptible individuellement) et il sert à entrer dans la dynamique des interactions sociales. Le corps-objet joue un rôle constituant non pas directement, comme corps percevant, mais indirectement, comme corps constituant des interactions qui donnent leur sens aux différentes formes de la cognition sociale. Les étapes d'appropriation que nous avons décrites correspondent à l'apprentissage individuel de l'usage de ce premier « outil » des interactions sociales puisqu'il sert à reconnaître autrui, à percevoir ses intentions ou à imiter ses expressions.

Dans une perspective énaïve, le corps-objet ne peut directement faire partie du sense-making individuel. Il n'a de sens que par les interactions. Si l'on veut comprendre la cognition sociale, il faut donc admettre que la clôture opérationnelle de l'autonomie individuelle soit reconfigurée : l'énaïe (le « sense-making ») se réalise dans un couplage *outillé* par ce corps-objet, couplage avec les formes signifiantes du social et en premier lieu, autrui, ses intentions, ses émotions et toutes les formes d'organisation collective (langues, normes, valeurs). Dès lors, comme l'outil que l'on conserve et que l'on peut emmener avec soi, chacun conserve son corps-objet en l'absence de l'environnement social qui l'en a doté. Comme les autres traces de la présence sociale, le corps-objet, premier support de la réflexivité, sert à rejouer, même dans la solitude, le

sens des interactions interindividuelles.

7 - Conclusion

Nous avons proposé trois schèmes explicatifs pour rendre compte dans un cadre interactionniste des trois premières formes de la cognition sociale : la reconnaissance d'autrui, la perception de ses intentions et l'imitation de ses expressions faciales. Pour cela, nous avons proposé une méthode minimaliste qui permet de démêler les différentes composantes des dynamiques d'interaction interindividuelles. Et nous avons montré que des compétences initiales non sociales permettent l'émergence d'organisations collectives socialement signifiantes, des situations d'interaction qui donnent l'opportunité de la formation de structures individuelles associant des propriétés déterminées des corps-objets et ces situations d'interaction.

Dans les trois cas, l'intercorporéité, comprise comme dynamique d'interaction corporelle interindividuelle, est *l'explanans* et non pas *l'explanandum* des structures et activités cérébrales spécifiques. Plutôt que de rendre compte des capacités de cognition sociale humaines à l'aide de structures neuronales, il s'agissait de montrer comment des dynamiques d'interaction qui font sens dans le monde social humain peuvent rendre compte de telles structures. Dans ces trois schèmes, l'explication des activités sociales est basée sur le sens de ces activités. Elle n'est basée ni sur des représentations méta-cognitives internes, ni sur des structures infrapersonnelles. Elle est elle-même sociale. C'est maintenant une question empirique de vérifier si de tels schèmes explicatifs établis dans des conditions minimalistes sont des outils utiles pour rendre compte des phénomènes dans les situations plus complexes de la vie sociale ordinaire.

Remerciements

L'auteur remercie la Région Hauts-de-France et le Fond Européen de Développement Régional (FEDER) 2014/2020 pour le financement de cette recherche.

Bibliographie

- *Auvray, Malika, Charles Lenay, and John Stewart. 2009. "Perceptual interactions in a minimalist virtual environment". *New ideas in psychology* 27 (1) : 32–47.
- *Auvray, Malika, and Marieke Rohde. 2012. "Perceptual crossing : the simplest online paradigm". *Frontiers in human neuroscience* 6.
- Baron-Cohen, Simon. 1997. "How to build a baby that can read minds : Cognitive mechanisms in mindreading". *The maladapted mind*, 207–239.
- Brass, Marcel, and Cecilia Heyes. 2005. "Imitation : is cognitive neuroscience solving the correspondence problem ?" *Trends in Cognitive Sciences* 9 (10) : 489–495.
- Clark, Andy. 1998. *Being there : Putting brain, body, and world together again*. The MIT Press.
- Clark, Andy. 2004. *Natural-Born Cyborgs : Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.
- Damiano, Luisa, Paul Dumouchel, and Hagen Lehmann. 2015. "Towards human–robot affective co-evolution overcoming oppositions in constructing emotions and empathy". *International Journal of Social Robotics* 7 (1) : 7–18.
- *De Jaegher, Hanne, Ezequiel Di Paolo, and Shaun Gallagher. 2010. "Can social interaction constitute social cognition ?" *Trends in cognitive sciences* 14 (10) : 441–447.
- *De Jaegher, Hanne, and Ezequiel Di Paolo. 2007. "Participatory sense-making". *Phenomenology and the Cognitive Sciences* 6 (4) : 485–507.
- *Deschamps, Loïc, Charles Lenay, Katia Rovira, Gabrielle Le Bihan, and Dominique Aubert. 2016. "Joint Perception of a Shared Object : A Minimalist Perceptual Crossing Experiment". *Frontiers in psychology* 7.
- Dewey, John. 1896. "The reflex arc concept in psychology." *Psychological review* 3 (4) : 357.
- *Di Paolo, Ezequiel, Marieke Rohde, and Hiroyuki Iizuka. 2008. "Sensitivity to social contingency or stability of interaction ? Modelling the dynamics of perceptual crossing". *New ideas in psychology* 26 (2) : 278–294.
- *Di Paolo, Ezequiel, and Hanne De Jaegher. 2016. "Neither individualistic, nor interactionist". *Embodiment, Enaction, and Culture*.
- Di Paolo, Ezequiel, and Hanne De Jaegher. 2012. "The interactive brain hypothesis". *Frontiers in human neuroscience* 6.
- Dumas, Guillaume. 2011. "Towards a two-body neuroscience". *Communicative & Integrative Biology* 4 (3) : 349.
- Dumas, Guillaume, Julien Laroche, and Alexandre Lehmann. 2014. "Your body, my body, our coupling moves our bodies". *Frontiers in human neuroscience* 8.

- Dumouchel, Paul. 1995. *Emotions : essai sur le corps et le social*. Institut Edition Synthelabo.
- Froese, Tom, and Ezequiel Di Paolo. 2011. "The enactive approach : theoretical sketches from cell to society". *Pragmatics & Cognition* 19 (1) : 1–36.
- *Froese, Tom, Charles Lenay, and Takashi Ikegami. 2012. "Imitation by social interaction ? Analysis of a minimal agent-based model of the correspondence problem". *Front. Hum. Neurosci* 6 : 202.
- Froese, Tom, Hiroyuki Iizuka, and Takashi Ikegami. 2014. "Embodied Social Interaction Constitutes Social Cognition in Pairs of Humans : A Minimalist Virtual Reality Experiment". *Scientific Reports* 4 (janvier).
- Froese, Tom. 2016. "Interactivity Should Aim to Extend, Not Reject, the Conceptual Foundations of Enaction". Édité par Tom Froese, Sebastjan Vörös, and Alexander Riegler. *Constructivist Foundations* 11 (2) : 247–49.
- Gallagher, Shaun, and Andrew N. Meltzoff. 1996. "The earliest sense of self and others : Merleau-Ponty and recent developmental studies". *Philosophical psychology* 9 (2) : 211–233.
- Gallagher, Shaun. 2008. "Direct perception in the intersubjective context". *Consciousness and Cognition* 17 (2) : 535–543.
- Gallagher, Shaun. 2008. "Intersubjectivity in perception". *Continental Philosophy Review* 41 (2) : 163–178.
- Gallagher, Shaun. 2013. "When the problem of intersubjectivity becomes the solution". *The infant mind : Origins of the social brain*, Guilford Press, Toronto, 48–74.
- Gallagher, Shaun. 2014. "In your face : transcendence in embodied interaction". *Frontiers in human neuroscience* 8.
- Gallese, Vittorio. 2001. "The 'shared manifold' hypothesis. From mirror neurons to empathy". *Journal of consciousness studies* 8 (5–6) : 33–50.
- Gallese, Vittorio, Christian Keysers, and Giacomo Rizzolatti. 2004. "A unifying view of the basis of social cognition". *Trends in cognitive sciences* 8 (9) : 396–403.
- Gallese, Vittorio. 2007. "Before and below 'theory of mind' : embodied simulation and the neural correlates of social cognition". *Philosophical Transactions of the Royal Society B : Biological Sciences* 362 (1480) : 659–669.
- Gallese, Vittorio. 2009. "The two sides of mimesis : Girard's mimetic theory, embodied simulation and social identification". *Journal of Consciousness Studies* 16 (4) : 21–44.
- Gallese, Vittorio, and Valentina Cuccio. 2015. "The paradigmatic body". *Open MIND*, 1–23.
- Goody, Jack. 1977. *The domestication of the savage mind*. Cambridge University Press.
- *Hanneton, Sylvain, Olivier Gapenne, Christelle Genouel, Charles Lenay, and Catherine Marque. 1999. "Dynamics of shape recognition through a minimal visuo-tactile sensory substitution interface". In *Third international conference on cognitive and neural systems*.

Boston, USA.

- *Iizuka, Hiroyuki, and Ezequiel Di Paolo. 2007. "Minimal Agency Detection of Embodied Agents". In *Advances in Artificial Life*, édité par Fernando Almeida e Costa, Luis Mateus Rocha, Ernesto Costa, Inman Harvey, and António Coutinho, 485-94. Lecture Notes in Computer Science 4648. Springer Berlin Heidelberg.
- *Lenay, Charles, Stéphane Canu, and Pierre Villon. 1997. "Technology and perception : the contribution of sensory substitution systems". In *Cognitive Technology, International Conference on*, 44-53. Los Alamitos, CA, USA : IEEE Computer Society.
- *Lenay, Charles, Olivier Gapenne, Sylvain Hanneton, Catherine Marque, and Christelle Genouëlle. 2003. "Sensory substitution : Limits and perspectives". In *Touching for Knowing, Cognitive psychology of haptic manual perception*, édité par Y. Hatwell, A. Streri, and E. Gentaz, 275-292. Amsterdam/Philadelphia : John Benjamins Publishing Company.
- *Lenay, Charles, Malika Auvray, François D. Sebbah, and John Stewart. 2006. "Perception of an intentional subject : An enactive approach". In *Third International Conference on Enactive Interface*, 37-38. Montpellier, France.
- Lenay, Charles, and Pierre Steiner. 2010. "Beyond the internalism/externalism debate : the constitution of the space of perception". *Consciousness and Cognition*, no 19 : 938-52.
- *Lenay, Charles, John Stewart, Marieke Rohde, and Amal Ali Amar. 2011. "You never fail to surprise me : the hallmark of the Other : Experimental study and simulations of perceptual crossing". *Interaction Studies* 12 (3) : 373-396.
- *Lenay, Charles and John Stewart. 2012. "Minimalist approach to perceptual interactions". *Frontiers in Human Neuroscience* 6.
- Leroi-Gourhan, André. 1993. *Gesture and speech*. MIT Press.
- Levinas, Emmanuel. 1979. *Totality and infinity : An essay on exteriority*. Vol. 1. Springer.
- *Maillet, Barthelemy, Charles Lenay, and Anne Guenand. 2008. "Designing for interpersonal tactile interaction over distance". In *HFT2008*, 391-98. Kuala Lumpur.
- Meltzoff, Andrew N., and M.Keith Moore. 1977. "Imitation of facial and manual gestures by human neonates". *Science* 198 (4312) : 75.
- Meltzoff, Andrew N. 2005. "Imitation and other minds : The "like me" hypothesis". *Perspectives on imitation : From neuroscience to social science* 2 : 55-77.
- Merleau-Ponty, Maurice. 1996. *Phenomenology of perception*. Motilal Banarsidass Publishes.
- Merleau-Ponty, Maurice. 1968. *The visible and the invisible : followed by working notes*. Northwestern University Press.
- *Michael, John, and Søren Overgaard. 2012. "Interaction and social cognition : A comment on Auvray et al.'s perceptual crossing paradigm". *New Ideas in Psychology* 30 (3) : 296-299.

- *O'Regan, J. Kevin, and Alva Noë. 2001. "A sensorimotor account of vision and visual consciousness". *Behavioral and brain sciences* 24 (5) : 939–972.
- Pacherie, Elisabeth. 2005. "Perceiving intentions." In *A Explicação da Interpretação Humana*, edited by J. Sàágua, 401–414. Edições Colibri.
- Reddy, Vasudevi. 2008. *How infants know minds*. Harvard Univ P^r.
- Rizzolatti, Giacomo, Luciano Fadiga, Leonardo Fogassi, and Vittorio Gallese. 2002. "From mirror neurons to imitation : Facts and speculations". *The imitative mind : Development, evolution, and brain bases* 6 : 247.
- Sartre, Jean-Paul. 2012. *Being and nothingness*. Open Road Media.
- Schilbach, Leonhard, Bert Timmermans, Vasudevi Reddy, Alan Costall, Gary Bente, Tobias Schlicht, and Kai Vogeley. 2013. "Toward a second-person neuroscience". *Behavioral and Brain Sciences* 36 (04) : 393–414.
- *Sribunruangrit, Nitiphan, Catherine K. Marque, Charles Lenay, Sylvain Hannequin, Olivier Gapenne, and Clotilde Vanhoutte. 2004. "Speed-accuracy tradeoff during performance of a tracking task without visual feedback". *Neural Systems and Rehabilitation Engineering, IEEE Transactions* 12 (1) : 131–139.
- Stiegler, Bernard. 1998. *Technics and time : The fault of epimetheus*. Vol. 1. Stanford University Press.
- *Stewart, John, and Olivier Gapenne. 2004. "Reciprocal modelling of active perception of 2-D forms in a simple tactile-vision substitution system". *Minds and Machines* 14 (3) : 309–330.
- Thelen, Esther, and Linda B. Smith. 1996. *A dynamic systems approach to the development of cognition and action*. MIT press.
- Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. 2005. "Understanding and sharing intentions : The origins of cultural cognition". *Behavioral and brain sciences* 28 (5) : 675–690.
- Trevarthen, Colwyn. 1979. "Communication and cooperation in early infancy : A description of primary intersubjectivity". *Before speech : The beginning of interpersonal communication* 1 : 530–571.
- Trevarthen, Colwyn. 1993. "The self born in intersubjectivity : The psychology of an infant communicating." In *The perceived self : Ecological and interpersonal sources of self-knowledge*, 121–173. New York USA : Cambridge University Press.
- Trevarthen, Colwyn, and Vasudevi Reddy. 2007. "Consciousness in infants". *The Blackwell companion to consciousness*, 37–57.
- Varela, Francisco J. (1997) "Patterns of life : Intertwining identity and cognition". *Brain and cognition* 34, no 1 : 72–87.
- Varela, Francisco J., Evan Thompson, and Eleanor Rosch. 1991. *The embodied mind :*

Cognitive science and human experience. MIT press.

Wilkerson, William S. 1999. "From bodily motions to bodily intentions : The perception of bodily activity". *Philosophical Psychology* 12 (1) : 61–77.

1 « Strong Interaction : a mutually engaged coregulated coupling between at least two autonomous agents where the coregulation and the coupling mutually affect each other, constituting a self-sustaining organization in the domain of relational dynamics. » (Gallagher 2013 : 59)

2 « the coordination of intentional activity in interaction, whereby individual sense-making processes are affected and new domains of social sense-making can be generated that were not available to each individual on her own » (De Jaegher & Di Paolo, 2007 : 497)

3 « It appears therefore that there is a we-centric dimension in the experience of a given emotional / affective state, and that it is underpinned by the activity of a common neural substrate. » (Gallese 2009 : 30)

4 « Mirror activation, on this interpretation, is not the initiation of simulation ; it is part of an enactive intersubjective perception of what the other is doing » (Gallagher 2013 : 64)

5 On retrouve un objectif de même ordre dans la « l'hypothèse des cerveaux en interaction (Interactive brain hypothesis) proposée par Di Paolo et De Jaegher (2012), mais comme on va le voir notre proposition est sensiblement différente.

6 Les articles notés d'une étoile dans la bibliographie utilisent explicitement les résultats d'une telle approche minimaliste que ce soit pour la perception spatiale ou la cognition sociale.

7 « When two or more conscious human beings mutually and knowingly affect one another's actions, they are engaged in a social interaction. » (Schoenherr et Westra 2017 : 7)

8 "While the behavior of the participants is, unbeknownst to them, guided by the global dynamics of the interaction process to an appropriate solution to the given task, their individual sense-making remains qualitatively unaffected with respect to its solitary point of reference" (Froese and Di Paolo 2011, 20).

9 "Combined with evidence for their plasticity, this suggests that mirror neurons may develop as a result of the agent's skilful involvement in social interaction rather than being the wellspring of capacities for social understanding." (De Jaegher et al. 2010, 7)

10 C'est là, nous semble-t-il un répondeur possible de la description du visage telle que proposée par Emmanuel Levinas dans *Totalité et infini*. En effet, Levinas décrit l'épreuve d'autrui comme mise en question de l'intentionnalité spontanée ; ce qui peut se comprendre comme trouble de la perception : autrui est ce qui ne se laisse pas phénoménaliser, à même la phénoménalité (Lenay & Sebbah 2016). Notons que cette

lecture de Lévinas est assez différente de celle proposée par Gallagher (2014) qui vise plutôt à trouver la transcendance de l'altérité dans l'interaction elle-même.

+

- Cet article inédit est largement inspiré d'un texte à paraître en langue anglaise.
- Site présentation de TACTOS : <http://www.intertact.net>
- Vidéo courte de présentation de TACTOS par Charles Lenay (2:40) : https://webtv.utc.fr/watch_video.php?v=M1OBD289W59S
- Vidéo de présentation de TACTOS par Charles Lenay en séminaire de recherche (34:13) : http://www.utc.fr/~wprecip/seminaire/costech/01-Lenay/Pres/co/publi_Video.html#segment_WAiafnoloZdVh9VEyNtc6f1