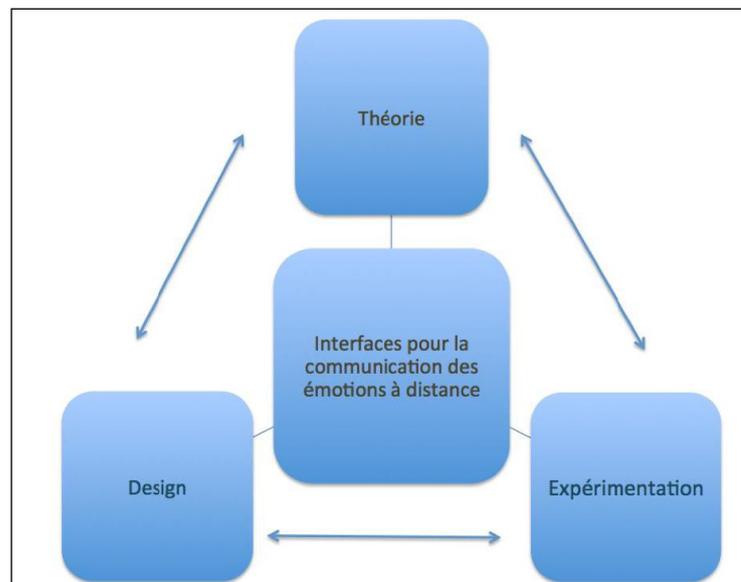


Par **Gabrielle LE BIHAN**

*Design d'interaction pour un « contact » à distance :
suppléance perceptive du toucher et médiatisation de
l'attention conjointe*

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC



Soutenue le 04 décembre 2014

Spécialité : Technologies Cognitives, Management de
l'Innovation et Systèmes Complexes

D2159

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

ÉCOLE DOCTORALE « SCIENCES POUR L'INGÉNIEUR »

Thèse

Présentée pour l'obtention du grade de Docteur de l'UTC

en spécialité

« Technologies Cognitives, Management de l'Innovation et Systèmes Complexes »

Gabrielle LE BIHAN

DESIGN D'INTERACTION POUR UN
« CONTACT » À DISTANCE

**Suppléance perceptive du toucher et médiatisation de
l'attention conjointe**

Sous la direction de Charles LENAY

Soutenue publiquement le 4 décembre 2014 devant le jury composé de :

Mme Françoise DARSES – Rapporteur – Institut de Recherche Biomédicale des Armées

M. Pascal SALEMBIER – Rapporteur – Université de Technologie de Troyes

Mme Beth DAWSON – Examineur – American University of Paris, Orange Entreprises PME

M. Pierre LÉVY – Examineur – Eindhoven University of Technology

Mme Anne GUÉNAND – Examineur – Université de Technologie de Compiègne

M. Charles LENAY – Directeur – Université de Technologie de Compiègne

Remerciements

Ce travail de thèse n'aurait pu aboutir – et n'aurait même été entrepris ! – sans l'aide, la bienveillance et le soutien de nombreuses personnes, qui m'ont entourée au cours de mon doctorat. Je me tourne ici vers elles avec gratitude, afin de leur adresser ces quelques mots de remerciement, qui ne font qu'esquisser ma reconnaissance à leur endroit. Je ne peux citer tout le monde, mais je remercie de tout cœur tous ceux qui ont été là au cours de cette thèse : merci à tous.

Tout d'abord, je remercie vivement Charles Lenay pour sa passion, ses encouragements, sa guidance et ses remarques. Merci pour ces années d'encadrement et de soutien ; je ressors transformée de cette expérience, et je sais à quel point tout aurait été différent sans le plaisir que j'ai eu à travailler ensemble. Merci de m'avoir donné le goût de la recherche !

Je remercie bien sûr chaleureusement le jury et les rapporteurs, pour le temps consacré à la lecture et à l'évaluation de ce travail, pour leurs remarques bienveillantes et leur retour enrichissant.

J'adresse un remerciement général à tous les membres de Costech et plus particulièrement du Cred, pour leur présence, leurs encouragements, leurs remarques et pour tous les débats enflammés du laboratoire.

Loïc Deschamps a notamment été de ceux qui m'ont convaincue de me lancer dans la grande aventure de la thèse, aventure dans laquelle il m'a soutenue, conseillée et aidée ; précieuse a été sa collaboration sur public/privé – merci encore d'avoir levé le voile sur quelques mystères d'excellence !

Je remercie Hugues Choplin pour son encadrement ; nos points étaient toujours des moments agréables, qui me forçaient à prendre du recul sur mon travail et me permettaient une vision différente. Ces rendez-vous ont été un véritable moteur.

Grand merci à Dominique Aubert, pour son accueil chaleureux au sein du groupe, ainsi que pour ses précieux conseils, pour le développement de public/privé et pour les discussions variées !

Je remercie Jérôme Mara pour son aide autour du développement de Touch Through et des expériences associées ; un excellent travail, mené avec rapidité et efficacité !

Matthieu Tixier m'a lui aussi apporté une grande aide, notamment par ses conseils, ses encouragements et son calme. Merci d'avoir si bien mené la barque d'ITOIP !

Je remercie Christine Devaux, Jocelyne Morel et Sylvie Carlier pour leur très efficace aide, leur rapidité et leur gentillesse.

Je remercie aussi bien sûr Olivier Gapenne, qui a été mon encadrant de master, mon responsable de groupe puis mon directeur d'école doctorale ! Ses encouragements et ses remarques ont été d'une grande aide.

Je remercie Nicolas Esposito, Pierre Steiner et Véronique Havelange pour leur gentillesse et pour m'avoir donné l'opportunité d'assurer quelques-uns de leurs cours.

Je remercie Bruno Majot, Sophie Pellegrin et Thao Lang, pour leur aide à m'engager sur la voie de la recherche, et pour m'avoir fourni des outils au long de ce voyage – valables aussi pour le suivant !

Merci aussi à Yamazaki-sensei, qui m'a accueillie dans son université japonaise au cours de mon stage de master et m'a permis de rencontrer le monde de la recherche.

Je remercie bien sûr ma famille et mes amis, pour leur soutien et leur intérêt ! La terrible question de la date de la soutenance a enfin trouvé sa réponse ; je sais que vous attendiez ce moment avec impatience... Merci à toutes les personnes qui ont pu venir en ce jour spécial, et à toutes celles qui ont pensé à moi.

Maman, Papa, merci pour tout (et pour le reste aussi !). Même si pour moi ce n'est pas médecine, je suis fière de marcher dans vos pas de docteurs. Vous m'avez toujours soutenue, encouragée, poussée ; je mesure la chance que j'ai de vous avoir pour parents, et vous suis redevable au-delà des mots.

Coco, ma petite Côtelette, merci d'avoir été là le plus longtemps possible – presque jusqu'au bout. Tu me manques atrocement. Je t'entends encore m'encourager et me poser les questions auxquelles je ne songe pas. Merci – pour cela, et pour tout le reste.

Coti, merci pour ton soutien ; chère consœur doctorante, je sais que tu me comprends ! Vivement ta propre soutenance – un événement « plus important qu'un mariage », ce sont tes propres mots !

Christine, Philippe, Dorine, Antoine, Elena, Alexandre, Canelle : merci pour votre accueil, votre gentillesse et votre générosité, qui me touchent profondément.

Quentin et Élodie, merci pour vos petites histoires de vie, et pour votre présence au loin.

Les Amibes : merci de votre soutien indéfectible !

La bande du lycée : merci d'être encore là après toutes ces années.

Merci enfin à celui qui partage ma vie, Nicolas. Ton soutien et le bonheur que j'ai à passer mon quotidien avec toi m'ont portée durant ces années ; tout aurait été infiniment plus difficile sans toi. Merci de m'aider à changer et de me faire réfléchir. Je sais bien que je te reconnaîtrai au toucher entre mille ; merci pour la magie de ta main dans la mienne.

Résumé

Notre recherche se concentre sur la conception d'interfaces pour la communication émotionnelle à distance. Notre travail s'inscrit notamment dans le cadre de l'élaboration d'espaces numériques d'interactions tactiles, rendus possibles par la connexion en réseau de dispositifs de suppléance perceptive. Nous adoptons une approche éactive et interactionniste de la cognition sociale. Nous utilisons de plus la méthodologie minimaliste, qui offre une opportunité originale d'étudier la constitution des rencontres interpersonnelles au sein d'espaces mettant en place des couplages sensori-moteurs inédits. Nous déroulons deux méthodologies, en articulant les domaines de la recherche fondamentale, de l'étude expérimentale et du design.

Dans un premier temps, nous partons de la théorie autour du thème du toucher, avant de proposer une solution d'interface : l'application smartphone Touch Through. Elle permet une suppléance perceptive du toucher en se basant sur des propriétés fonctionnelles de ce sens. Touch Through sert ensuite à une expérimentation basée autour d'un test de Turing minimaliste, afin de tester la possibilité d'un sentiment de présence à distance avec notre application et de la mettre entre les mains d'utilisateurs.

Dans un second temps, nous adressons la question de l'attention conjointe avant de mettre en place une étude expérimentale autour de ce thème (tâche de distinction entre des objets privés et des objets communs). Forts des résultats, nous proposons alors un prototype d'interface pour le réseau social facebook, en mettant en avant des possibilités d'attention conjointe au sein du site et en offrant une interaction inédite autour des posts.

Mots-clés : Contact, Interactions à distance, Émotions, Toucher, Interactions Tactiles, Suppléance perceptive, Design d'interaction, Attention conjointe, Dynamique d'interaction, Croisements perceptifs, Minimalisme

Interaction design for a remote “contact”

Abstract

Our research is aimed at designing interfaces for the remote communication of emotions. In particular our work is part of the development of tactile interaction digital spaces. These are made possible by a network connection between perceptual supplementation devices. We consider an enactive and interactionist approach of social cognition. In addition, we use a minimalist methodology. This offers an original opportunity to study the very constitution of interpersonal encounters within spaces where unprecedented sensorimotor couplings happen. We work through two methodologies, each articulating the fields of fundamental research, experimental study and design.

In a first phase, we start from theory around the theme of touch, before proposing an interface solution – the smartphone application ‘Touch Through’. This application allows a perceptual supplementation of touch using its functional properties. ‘Touch Through’ is then used in a experimental study based on a minimalist Turing test. This experiment is about testing the possibility of a presence in absence feeling through our application. It was also an opportunity to have users handling it.

In the second phase we address the question of joint attention. After a study of this question, we design an experiment where we propose a task of discrimination between private and common objects. Based on the outcome of this experiment, we then propose an interface prototype for the social network ‘facebook’. With this prototype interface, the facebook website is enhanced with possibilities of joint attention. In addition, facebook users can take part in an original interaction around posts.

Keywords: Contact, Remote interactions, Emotions, Touch, Tactile Interactions, Perceptual supplementation, Interaction design, Joint Attention, Interaction Dynamics, Perceptual crossings, Minimalism

Laboratoire COSTECH – Connaissance, Organisation et Systèmes TECHniques
Equipe d'accueil du ministère EA 2223
Centre Pierre Guillaumat
B.P.60319
60203 COMPIÈGNE Cedex

Sommaire

0	Introduction	13
0.1	Environnement et démarche de recherche : objet d'étude théorique et conception d'interfaces	14
0.2	ITOIP – Interactions Tactiles pour l'OrientatIon, l'Information et la Présence.....	16
0.3	La communication médiatisée : le cas de la visiophonie	17
0.4	Motivations et enjeux	29
0.4.1	Enjeu sociétal : le besoin de contact à distance	29
0.4.2	Enjeu et méthode de conception : design d'interfaces	31
0.4.3	Enjeu scientifique : un niveau fondamental	32
1	Cadre théorique et méthodologique.....	35
1.1	Éléments théoriques.....	35
1.1.1	De la cognition	35
1.1.1.1	Énaction et approche sensori-motrice de la perception	35
1.1.1.1.1	Présentation de l'énaction	35
1.1.1.1.2	L'outil et sa saisie dans un cadre éactif.....	38
1.1.1.1.3	Le croisement perceptif.....	44
1.1.1.2	Les interactions : supports de la cognition sociale	49
1.1.1.3	De l'impact des émotions	52
1.1.1.3.1	La difficile question des émotions	52
1.1.1.3.1.1	Difficulté de définition	52
1.1.1.3.1.2	Difficulté de mesure	56
1.1.1.3.2	Émotions et raison.....	57
1.1.1.3.3	Les émotions en interaction.....	58
1.1.2	La suppléance perceptive comme cadre de recherche.....	63
1.1.2.1	Présentation.....	63
1.1.2.1.1	La substitution sensorielle.....	63
1.1.2.1.2	Ou plutôt : la suppléance perceptive	66
1.1.2.2	Tactos.....	68
1.1.2.2.1	Présentation	68
1.1.2.2.2	État d'avancement et résultats.....	74
1.1.2.2.2.1	Analyses d'usage	74
1.1.2.2.2.2	Supports pédagogiques numériques	75
1.1.2.2.2.3	Émoticônes tactiles	77

1.1.2.2.2.4	Perspective : Tactos mobile	78
1.1.2.2.2.5	De l'individuel au collectif : Intertact.....	79
1.1.3	Communication et interaction à distance	84
1.1.3.1	La médiatisation pour le croisement perceptif.....	84
1.1.3.2	Sentiment de présence, connectedness, relatedness, social awareness,... : la présence à distance.....	87
1.1.3.3	La communication des émotions à distance	91
1.2	Méthodologie	93
1.2.1	Méthodologie globale : au croisement des disciplines.....	93
1.2.2	Méthodologies design	93
1.2.2.1	Conception et appropriation des interfaces.....	94
1.2.2.2	Concevoir pour l'interaction.....	96
1.2.2.3	Conception centrée utilisateurs.....	98
1.2.3	Méthodologie expérimentale.....	100
1.2.4	Le minimalisme comme cadre de travail	101
1.2.5	L'approche phénoménologique.....	103
1.3	Conclusion : l'articulation Théorie-Design-Expérimentation comme enjeu de recherche.....	105
2	La présence à distance : du théorique vers l'expérimentation en passant par le design 107	
2.1	Introduction.....	107
2.2	Contexte théorique.....	107
2.2.1	Interactions interpersonnelles mobilisant le toucher	107
2.2.1.1	Sens du toucher.....	107
2.2.1.1.1	Propriétés.....	107
2.2.1.1.1.1	Le toucher au cours du développement	107
2.2.1.1.1.2	Le toucher : rencontres physiques et conscience de soi	112
2.2.1.1.1.3	Fonctionnement	113
2.2.1.1.2	Caractéristiques du toucher	115
2.2.1.2	Communication par le toucher.....	116
2.2.1.2.1	Le toucher : un canal de communication spécifique.....	116
2.2.1.2.2	Le toucher comme sens du contact et du touchant.....	119
2.2.1.3	Médiation des interactions haptiques.....	122
2.2.1.3.1	État de l'art des dispositifs pour la communication émotionnelle à distance	

2.2.1.3.1.1	Systèmes de médiation haptique.....	125
2.2.1.3.1.2	Autres systèmes pour la communication émotionnelle	146
2.2.1.3.2	Le toucher dans les environnements virtuels partagés	153
2.2.1.3.3	Synthèse	154
2.2.2	Suppléance perceptive du toucher.....	156
2.3	Design d'interface : Touch Through.....	157
2.3.1	L'application Touch Through : introduction.....	157
2.3.2	L'application Touch Through : description	158
2.3.3	Premiers retours utilisateurs	164
2.3.4	L'application Touch Through : enjeux.....	166
2.4	Étude expérimentale : toucher à distance avec Touch Through (Distal vs Proximal)	167
2.4.1	Population.....	169
2.4.2	Matériel	169
2.4.3	Procédure.....	170
2.4.4	Hypothèses opérationnelles.....	172
2.4.5	Résultats	173
2.4.5.1	Analyses des réponses individuelles.....	173
2.4.5.2	Analyses des réponses collectives	182
2.4.5.3	Analyse des trajectoires	194
1.1.1.1.1	Interaction des sujets avec les deux types d'interactant.....	195
2.4.5.3.1	Interaction entre les sujets en fonction de la qualité des réponses	208
2.4.5.3.1.1	Au niveau individuel.....	208
2.4.5.3.1.2	Au niveau collectif.....	211
2.4.6	Discussion	218
2.4.7	Conclusion.....	223
2.5	L'articulation Théorie-Design-Expérimentation : premiers résultats.....	226
3	L'attention partagée : du théorique vers le design en passant par l'expérimentation. 228	
3.1	Introduction.....	228
3.2	Attention conjointe, objet partagé : contexte théorique.....	231
3.2.1	Attention conjointe : définitions et enjeux	231
3.2.2	L'attention conjointe comme coordination perceptive.....	235
3.2.3	Les étapes de développement de l'attention conjointe.....	240
3.2.4	Inscription dans le mouvement et dans le contexte.....	243

3.3 Étude expérimentale objets communs et objets privés : une expérience minimaliste mettant en jeu l'attention conjointe	249
3.3.1 Population.....	249
3.3.2 Matériel	249
3.3.3 Procédure.....	251
3.3.4 Hypothèses opérationnelles.....	255
3.3.5 Résultats	256
3.3.5.1 Analyses des réponses individuelles.....	256
3.3.5.2 Analyses des réponses collectives	260
3.3.5.3 Analyses des trajectoires.....	265
3.3.5.3.1 Interaction des sujets avec les deux types d'objets	266
3.3.5.3.2 Interaction entre les sujets autour des deux types d'objets	270
3.3.5.3.3 Interaction entre les sujets en fonction de la qualité des clics.....	274
3.3.6 Discussion	276
3.3.7 Conclusion.....	284
3.4 L'attention conjointe dans facebook : expliciter le public et le privé	287
3.4.1 Facebook et le déficit en attention conjointe ; qu'est-ce qui est privé, qu'est-ce qui est public ?	287
3.4.2 Proposition de design	290
3.4.2.1 La distinction public/privé des éléments de facebook.....	290
3.4.2.2 L'attention conjointe en coïncidence temporelle sur facebook	295
3.4.3 Retours utilisateurs	300
3.4.4 Conclusion.....	301
3.5 L'articulation Théorie-Design-Expérimentation : une deuxième approche	303
4 Perspectives et conclusion générale	305
4.1 Perspectives	305
4.1.1 Une expérience préliminaire sur un « langage » tactile	305
4.1.2 Population.....	308
4.1.3 Matériel	308
4.1.4 Procédure.....	309
4.1.5 Résultats	311
4.1.6 Remarques générales tirées de nos observations.....	316
4.1.7 Pistes futures	316
4.2 Conclusion générale.....	317
Bibliographie	321

ANNEXES	345
ANNEXE A : Note concernant l'analyse des trajectoires perceptives.....	345
ANNEXE B : Exemples de trajectoires perceptives (Turing avec Touch Through)	348
Exemple 1 : Interaction entre deux partenaires (condition Distale)	348
Exemple 2 : Interaction entre deux partenaires (condition Proximale Visuelle)	348
Exemple 3 : Interaction entre deux partenaires (condition Proximale Vibratoire).....	349
Exemple 4 : Interaction entre un sujet humain et un robot (reconnaissance)	350
Exemple 5 : Interaction entre un sujet humain et un robot (non-reconnaissance).....	350
ANNEXE C : Exemples de trajectoires perceptives (Commun/Privé)	352
Exemple 1 : Clic Mutuel en condition « Mono » (stratégie orientée objets)	352
Exemple 2 : Clic Mutuel en condition « Mono » (stratégie orientée sujet)	353
Exemple 3 : Clic Mutuel en condition « Différenciée » (stratégie orientée objets)	354
Exemple 4 : Clic Mutuel en condition « Différenciée » (stratégie orientée sujet)	355
Exemple 5 : Clic Mutuel en condition « Parallélisme » (stratégie orientée objets).....	356
Exemple 6 : Clic Mutuel en condition « Parallélisme » (stratégie orientée sujet).....	357
ANNEXE D : Exemples de trajectoires perceptives (expérience préliminaire sur un	
« langage » tactile).....	358
Exemple 1 : binôme 1	358
Exemple 2 : binôme 2	358
Exemple 3 : binôme 3	359
Exemple 4 : binôme 4	360

0 Introduction

Notre travail se place dans le cadre de l'élaboration de dispositifs pour la communication (et plus particulièrement la communication émotionnelle) à distance. On sait bien l'importance de la communication dans la vie humaine, et on constate que des trésors d'ingéniosité ont été développés pour pouvoir communiquer à grande distance, depuis les signaux de fumée jusqu'à internet. Nous considérons qu'une communication à distance a lieu entre deux individus ne se trouvant pas dans le même espace physique (c'est-à-dire, dans notre optique, ne se trouvant plus à portée ni de la voix, ni de la vue – ni donc encore moins de la main...). Les dispositifs de communication à distance ont pour but de permettre à des individus éloignés d'échanger non seulement des messages, mais aussi simplement des signes de présence. Nous nous intéressons plus précisément à la dimension émotionnelle de ce type d'échanges, et c'est autour de cette thématique que s'articule notre objectif de conception.

Afin de traiter cette question, nous allons dans la première partie de ce travail présenter notre cadre théorique et méthodologique. Nous adoptons en effet une approche éactive, incarnée et située de la cognition sociale, afin de placer le processus même de l'interaction au cœur de la compréhension interpersonnelle. Nous allons traiter de la question des émotions, non seulement dans le cadre de la cognition, mais aussi dans le cadre de l'interaction. Nous allons présenter le cadre théorique de la suppléance perceptive, qui permettra de traiter notre objet dans une perspective originale. Puis nous traiterons de l'interaction à distance et des limites posées par ce type particulier de communication. Nous exposerons aussi notre méthodologie, qui articule des méthodes de travail diverses, liées au monde du design mais aussi à celui de la psychologie expérimentale, et complétées par les approches minimaliste et phénoménologique.

Dans la seconde partie, nous allons explorer notre cadre de travail par l'étude de la présence à distance. Nous allons exposer la thématique d'entrée que nous adoptons pour ce thème, à savoir le sens du toucher et son rôle dans la communication des émotions. De là, en nous appuyant sur notre cadre théorique, nous allons présenter la première interface que nous avons conçue pour répondre à notre question de recherche : il s'agit d'une application smartphone, appelée Touch Through, qui est basée sur la suppléance perceptive du toucher. Enfin, nous allons présenter l'étude expérimentale que nous avons réalisée en utilisant Touch Through. Notre expérimentation permet d'étudier le concept de toucher à distance que nous proposons ainsi que l'expérience ressentie par les utilisateurs, qui se sont retrouvés confrontés à une tâche du type test de Turing.

Enfin, dans la troisième partie de ce travail, nous opérerons un développement autour de la thématique de l'attention conjointe, qui est centrale dans le cadre d'une interaction entre plusieurs individus autour d'un objet ou événement commun. Nous avons dans un premier temps

conçu une étude expérimentale autour de cette question, en faisant appel au dispositif de suppléance perceptive Tactos développé dans notre laboratoire. Nous proposons aux sujets une tâche de discrimination entre un objet commun perçu par les deux sujets et deux objets privés, qui ne peuvent respectivement être perçus que par un seul des deux sujets. Suite à l'analyse des résultats de cette étude, nous exposerons notre deuxième solution de design, à savoir un prototype d'interface ayant pour but d'explicitier le mécanisme de l'attention conjointe dans le réseau social en ligne facebook.

Mais tout d'abord, nous allons dans cette introduction présenter notre environnement de recherche, puis dans un second temps exposer le projet ITOIP (Interactions Tactiles pour l'OrientatIon, l'InformatIon et la Présence) à partir duquel notre travail se développe. Enfin, à partir d'un rapide rappel des questions que posent les différentes médiations de la communication, en particulier la visiophonie, nous exposerons nos hypothèses et les enjeux qui s'articulent autour de notre travail.

0.1 Environnement et démarche de recherche : objet d'étude théorique et conception d'interfaces

Le groupe CRED

Les travaux de cette thèse interdisciplinaire ont été réalisés au sein du groupe de recherche CRED (Cognitive Research and Enaction Design) du laboratoire COSTECH (Connaissance, Organisation et Systèmes TECHniques) de l'Université de Technologie de Compiègne (UTC).

Le groupe CRED pratique une recherche technologique articulant sciences humaines, design technologique et réflexion philosophique. Le CRED est donc une équipe profondément interdisciplinaire (et ce depuis sa création en 1998), qui organise ses projets autour d'une exigence à la fois :

- scientifique et philosophique dans le domaine des recherches sur la cognition,
- technologique notamment à travers le design de dispositifs dits de suppléance perceptive,
- sociétale en étant particulièrement attentif aux questions d'accessibilité, d'éthique et d'esthétique partagées.

Le CRED s'intéresse notamment à l'étude des phénomènes perceptifs, en particulier dans le cas de leur prothétisation (lorsqu'une interface médiatise la perception). Les travaux fondamentaux menés peuvent adopter différentes formes : expérimentations, réflexions philosophiques, modélisations,... Ils concernent des sujets variés, allant de la genèse de nouvelles modalités perceptives (comme par exemple celle permise par le dispositif Tactos), à la

description de l'expérience vécue, en passant par l'apprentissage perceptif. Dans le cadre de ces recherches a été développé un dispositif de perception du monde numérique mettant en œuvre la modalité tactile ; il s'agit du système Tactos. Ce dispositif est dans la continuité des travaux en conception de systèmes haptiques (ou tactiles) permettant l'exploration de formes graphiques (Sribunruangrit, 2004) et de l'ergonomie de ces dispositifs, ainsi que de leur possible intégration par des collégiens dans un contexte d'usage scolaire pour l'apprentissage de la géométrie (Ali-Ammar, 2005).

La recherche-action : une démarche particulière

La réalisation de cette thèse et des travaux associés s'inscrit dans une logique de recherche-action. Les interactions interpersonnelles haptiques, notamment médiatisées, sont au cœur du questionnement de recherche exposé dans ce manuscrit. Il s'agit en effet plus largement de se questionner sur la transmission des émotions à distance ; nous envisageons notamment la piste d'un « toucher à distance » comme étant une réponse adéquate à cette problématique. Dans ce cadre, nous avons réalisé une interface sous la forme d'une application pour smartphone, Touch Through. Le fonctionnement de cette application et sa prise en main par les usagers nous ont permis d'approfondir notre réflexion. Cette démarche de conception s'est vue accompagnée de la construction de nouvelles pratiques et de nouvelles connaissances, notamment grâce à la réalisation d'études de psychologie expérimentale.

Les problèmes empiriques auxquels nous cherchons une solution amènent invariablement des questions théoriques. De même, certaines questions théoriques sont à la source d'expérimentations empiriques pour y répondre. L'action, considérée à travers la mise en place de dispositifs spécifiques, se trouve être à la fois objectif et support de recherche. Cependant, articuler une démarche de conception et une démarche de recherche ne va pas de soi, et nécessite questionnement, recul et précautions. Comme le souligne Bruce Archer, « l'activité du praticien ne peut en effet être considéré comme participant de la recherche qu'à la condition qu'elle s'accorde aux critères de la recherche. Elle doit donc avoir pour but l'émergence de nouveaux savoirs, être menée de manière systématique et être exprimée sans ambiguïté. Ses données et ses méthodes doivent faire preuve d'un caractère transparent, et son résultat en terme de savoirs doit être transmissible. Mais comme toute recherche-action, la recherche à travers l'action du praticien doit être reconnue comme étant très probablement non-objective et presque certainement spécifique à une situation. » (Archer, 1995, p. 13, traduction personnelle.)¹

¹ « Practitioner activity can count as research if, and only if, it accords with the criteria of research. It must be knowledge directed, systematically conducted, unambiguously expressed. Its data and methods must be transparent and its knowledge-outcome transmissible. But like all

Les connaissances retirées de ce type de recherche sont plus souvent des « règles conditionnelles » que des « règles générales » (Hummels, 2005). Pour notre part, nous visons par ce travail à répondre à certaines des questions que se posent les designers devant concevoir un dispositif de communication à distance.

Du point de vue du design, tout l'enjeu est d'offrir des solutions d'interfaces portant des modes d'interaction émergents et qui puissent peu à peu intégrer la sphère des usages. La démarche que nous proposons ici insiste sur le besoin de prendre en compte deux dualités qui se répondent, qui sont toujours existantes, mais qui se retrouvent cristallisées lorsque l'on considère les interactions médiatisées. Il s'agit d'une part de la dualité entre interface *tenue* et interface *déposée* ; la première est constituante pour le sujet de l'expérience perceptive, notamment du contact avec autrui (en d'autres termes l'interface est utilisée, et devient transparente, ouvrant de nouvelles possibilités d'agir et de percevoir) tandis que la seconde est constituée par le sujet (en d'autres termes c'est l'interface lorsqu'elle est perçue mais pas utilisée). D'autre part, il s'agit de la dualité entre le corps percevant et le corps-image ; le premier permet au sujet de percevoir les autres sujets et l'environnement, le second permet au sujet d'être perçu par les autres.

0.2 ITOIP – Interactions Tactiles pour l'Orientation, l'Information et la Présence

Le projet ITOIP (Interactions Tactiles pour l'Orientation, l'Information et la Présence) est un projet financé par la région Picardie, porté par le groupe du CRED du laboratoire COSTECH de l'UTC et dont les partenaires principaux sont l'ESAD (Ecole Supérieure d'Art et de Design d'Amiens), l'APICADEV (Association PICARde des Aveugles et DEFicients Visuels), le laboratoire Heudiasyc de l'UTC et le laboratoire Roberval de l'UTC. Le projet, qui s'étend de septembre 2011 à mai 2014, vise à proposer de nouvelles interfaces mobiles à destination des publics non et mal voyants afin de les aider au quotidien dans leurs déplacements, recherches d'informations et échanges médiatisés². La technologie utilisée est celle du dispositif Tactos et offre plusieurs possibilités. Le projet vise d'une part à proposer une version mobile de Tactos à coupler avec un smartphone et, d'autre part, à développer Intertact, qui est un site spécialement conçu pour être utilisé en interaction par le biais de Tactos.

Action Research, research through practitioner action must be recognised as very probably non-objective and almost certainly situation-specific. »

² J'ai notamment participé à la rédaction des actes d'une communication orale :

- Tixier, M., Lenay, C., Le Bihan, G., Gapenne, O. & Aubert, D. (2013). Designing Interactive Content with Blind Users for a Perceptual Supplementation System, Proceedings of TEI, Barcelona, Spain. February 9 - 13, 2013

Orientation

Cette facette du projet a pour objectif de rendre des cartes et des plans accessibles aux personnes en situation de handicap visuel. Chaque carte et chaque plan doivent être consultables simplement, d'où l'intérêt de les implémenter sur un smartphone.

Information

Equipé de la technologie Tactos, le téléphone permet l'exploration tactile des interfaces graphiques. Le dispositif reprend le rôle de la souris ou du touchpad pour la navigation, mais ajoute en plus un retour tactile.

Présence

Le principe de cette partie du projet est de permettre des interactions émotionnelles entre utilisateurs, notamment à travers des sensations tactiles. Ceci vise la capacité de manifester sa présence aux autres.

Le travail présenté dans cette thèse s'inscrit en particulier dans l'axe présence du projet, qui traite notamment les problématiques d'espace interpersonnel et d'interaction. Il se partage en plusieurs sections, dont celle à laquelle ce travail est affilié : communication émotionnelle et langage.

0.3 La communication médatisée : le cas de la visiophonie

Nous allons ici nous concentrer sur des concepts importants pour la communication à distance, dont une des techniques les plus importantes et les plus anciennes est l'écriture et la plus récente celle des dispositifs de communication numériques, et en particulier la visiophonie. Il ne s'agit pas ici de faire un historique exhaustif des systèmes de communication ; il est cependant important, à l'orée de notre travail, de rappeler la diversité des modes de communication à distance. Cette diversité des façons et des propriétés spécifiques des différents modes de communication à distance va nous permettre de définir l'espace des possibles dans lequel nous allons travailler. Ce qui nous intéresse plus spécifiquement, ce sont les supports techniques de la communication, qui sont nombreux, et leurs propriétés : signaux de fumée, écrit, morse, transfert de la voix, etc. En effet, la façon de communiquer, le support technique de la communication, n'est pas neutre vis-à-vis de ce qui est communiqué quand il s'agit d'un contenu émotionnel, alors qu'il est relativement neutre pour la transmission d'une l'information qui peut être codée de diverses façons. Certains supports semblent plus propices à l'échange émotionnel

et au sentiment de présence. Notre principe directeur est donc la non-neutralité des dispositifs de communication relativement à la nature de la rencontre et de sa qualité émotionnelle. Cette hypothèse préliminaire nous oblige à revenir sur les dispositifs existants dans l'histoire de la communication à distance, en nous arrêtant plus particulièrement sur celui de la visiophonie.

Tout d'abord, rappelons que les systèmes de communication à distance ont tout d'abord été basés sur un système de codage, que ce soit par le biais de signaux de fumée, de calculi ou encore plus tard de l'écriture. Jack Goody a posé l'importance de l'écriture dans notre pensée occidentale. L'écriture nous permet en effet d'acquérir et de stocker des connaissances, et de les véhiculer ensuite, tant dans le temps que dans l'espace. À ses débuts, l'écriture a notamment été utilisée comme moyen pour mémoriser des activités ou des quantités trop nombreuses ou complexes. Ainsi, l'écriture n'a certainement pas été d'abord inventée pour transcrire la parole, mais pour faire état d'échanges entre des personnes (Goody, 1979). Le fait qu'elle ait progressivement aussi permis la transcription de la parole est cependant capital pour la communication à distance : on peut inscrire des messages qui seront portés à quelqu'un qui se trouve dans un autre espace, qu'il s'agisse d'une autre rue, d'une autre ville, d'un autre pays ou d'un autre continent. Il faut cependant remarquer la réduction qu'opère le passage de la parole au texte. Toutes les modulations de la voix et toute la communication non verbale sont notamment perdues.

Les systèmes d'écriture sont devenus, au cours de leur développement par transformations successives, de plus en plus simples et abstraits, rendant ainsi possible une diffusion de plus en plus large de la pratique la lecture (ibid). En même temps, l'écriture introduit des différences dans la manière de se servir du langage et dans les processus cognitifs, et notamment elle « donne à la parole une forme permanente » (ibid, p.143). Les mots, une fois écrits, deviennent des objets durables, alors que simplement prononcés ils sont évanescents. Cela entraîne des modifications sur les communications qui sont considérables, tant dans le temps que dans l'espace. On peut ainsi envoyer un message à une personne très lointaine, ou voir un message perdurer dans le temps bien après que son auteur soit décédé. En effet, grâce au stockage du texte sur un support externe qui en garde la trace sans modification, l'écrit garantit son exacte répétition, ce qui est impossible dans le cadre d'une interaction en face-à-face. *Les paroles s'envolent, les écrits restent...*

L'évolution des supports de l'écriture a joué un rôle crucial dans la transformation des activités humaines. C'est ce que décrit par exemple I. Illich dans Du visible au lisible à propos de l'évolution du livre au XII^{ème} siècle. C'est en effet à cette période que le livre est passé de sa forme roulée en parchemin à celle du codex. I. Illich s'appuie sur L'art de lire de Hugues de Saint-Victor, chanoine qui présente dans cet écrit la façon dont était vécue la lecture à son époque. Les versions de la Bible étaient des copies faites à la main, illustrées de nombreuses

enluminures. Les mots y étaient tous attachés, et il n’existait alors ni ponctuation ni numérotation de page. Lire impliquait donc nécessairement une activité de marmottage : les mots écrits ne prenaient sens dans l’esprit que lorsqu’ils étaient prononcés (Illich, 1991).

Lorsque le livre sous forme de codex est apparu, la numérotation des pages s’est avérée possible. On a aussi pu, avec l’apparition du codex, se libérer une main : nul besoin de tenir un codex à deux mains, étant donné qu’il n’est pas besoin de le dérouler, mais simplement d’en tourner les pages (ibid). La prise de notes a ainsi été rendue possible. Avec le codex est aussi arrivé l’espace entre les mots, qui a permis la lecture silencieuse. Ainsi donc au XII^{ème} siècle s’est constituée la culture du livre et de la lecture que nous connaissons aujourd’hui.

Une autre révolution permise par l’apparition du codex est celle de l’imprimerie, survenue au milieu du XV^{ème} siècle (Chartier, 1994). Elle vient proposer une alternative à la copie manuscrite et permet la multiplication et la circulation des textes, ce qui a notamment permis le développement de la pratique de la lecture par un plus grand nombre de personnes.

On peut évidemment utiliser l’écrit pour la communication à distance ; il faut alors faire transiter le message écrit vers son destinataire. Le message sera exactement tel que l’expéditeur l’a écrit (sauf en cas de dégradation du support ou de falsification), mais en contrepartie il pourra plus facilement que dans une interaction en face à face être interprété d’une manière différente de celle qui était pensée à l’origine, un problème déjà souligné par Platon au IV^{ème} siècle avant J.-C. Ainsi, il écrit dans Phèdre : « C’est que l’écriture a, tout comme la peinture, un grave inconvénient. Les œuvres picturales paraissent comme vivantes ; mais si tu les interrogues, elles gardent un vénérable silence. Il en est de même des discours écrits. Tu croirais certes qu’ils parlent comme des personnes sensées ; mais, si tu veux leur demander de t’expliquer ce qu’ils disent, ils te répondent toujours la même chose. Une fois écrit, tout discours roule de tous côtés ; il tombe aussi bien chez ceux qui le comprennent que chez ceux pour lesquels il est sans intérêt ; il ne sait point à qui il faut parler, ni avec qui il est bon de se taire. S’il se voit méprisé ou injustement injurié, il a toujours besoin du secours de son père, car il n’est pas, par lui-même capable de se défendre ni de se secourir. » (Platon, 1922).

L’échange épistolaire entre deux personnes leur permet une sorte de conversation, dont les échanges sont espacés par le temps et les aléas du parcours qui les sépare. La lettre permet dans une certaine mesure la transmission des émotions ; elle permet d’inscrire ce que l’on ressent afin de pouvoir le laisser savoir à la personne à qui on la destine. Cependant, la phase d’explicitation pose un biais ; il n’est pas aisé de poser des mots sur ce que l’on ressent, et cela n’assurera pas que la personne qui lira ce qui est décrit comprendra le sentiment que l’on cherchait à lui communiquer. Le lyrisme et l’expression des émotions sont le moteur principal de la communication émotionnelle par le biais de l’écrit. Le genre du roman épistolaire apparu au XVII^{ème} siècle montre bien que l’écriture et la lecture de simples lettres peuvent donner un

sentiment de proximité. L'un des ressorts propres à ce genre est en effet lié à l'impression que le lecteur a de s'introduire dans l'intimité des personnages à leur insu (Boyer, 1982).

Avec l'arrivée des télécommunications, les messages ont pu être acheminés de manière bien plus rapide et leur support s'est radicalement transformé. Bien loin du cavalier qui devait chevaucher parfois des jours durant pour transmettre une lettre d'un endroit à un autre, on peut aujourd'hui communiquer avec quelqu'un se trouvant par exemple à des centaines ou même des milliers de kilomètres de manière quasi-immédiate par le biais d'un téléphone ou d'un ordinateur.

Les télécommunications ont tout d'abord fait appel au texte écrit codé (par l'utilisation du télégraphe), puis sont devenues une technique de l'oral au service de la communication interpersonnelle à distance (par l'utilisation du téléphone). Ces deux modalités ont été déclinées et parfois conjuguées (par exemple avec le logiciel Skype, on peut s'écrire des messages tout en ayant une conversation audio).

L'invention de l'ordinateur a eu lieu au moment de la seconde guerre mondiale, et son développement aboutit dans les années 70 à la création du courrier électronique. La création du protocole internet remonte à 1972 (Herrenschmidt, 1999). Les messages écrits (par exemple les mails ou les chats) sont aujourd'hui presque instantanément acheminés à l'autre bout du monde via Internet. Ce réseau permet aussi de communiquer la voix et la vidéo. Ces moyens de télécommunication accroissent la mobilité des personnes et renforcent la concentration des activités (Rallet, 2000). Ils ont non seulement changé l'échelle des communications interpersonnelles, mais ils ont aussi produit une forme de dématérialisation des supports, notamment de l'écrit. Non pas parce qu'il n'y aurait plus de support du tout, mais parce que l'accès à ce qui est inscrit se fait de manière indirecte : on ne peut ouvrir un disque dur et y voir les informations inscrites comme sur un livre, il faut passer par le biais d'un ordinateur pour avoir accès à ce qu'il contient. Les textes ont alors pu se détacher de la forme du livre (Chartier, 1994).

La transmission des messages écrits grâce aux télécommunications n'a de différence avec la correspondance traditionnelle que la vitesse, ce qui a cependant un grand impact ; le contenu d'une lettre qui va mettre des jours entiers à arriver à son destinataire n'est pas pensé ni écrit de la même façon qu'une phrase de chat ou qu'un SMS, qui transitera en quelques secondes et pourra aussi avoir une réponse quasi immédiate. Les questions d'accélération ne sont donc pas à négliger ; d'après Rosa Hartmut, elles sont cruciales et constituent l'expérience majeure de la modernité (Hartmut, 2010).

Les nouveaux moyens de communication remettent en cause les relations de proximité, au point que Alain Rallet et André Torre posent la question : être proche est-il encore nécessaire à l'heure d'internet ? (Rallet & Torre, 2007) Si la réponse est en partie oui, les réseaux de télécommunication permettent cependant d'agrandir l'échelle géographique des interactions.

Selon le moyen utilisé, le transfert des messages ou du canal audio (et/ou vidéo) peut se faire en moins de quelques secondes, ce qui permet une conversation pour ainsi dire en temps réel. Ainsi une interaction synchrone à distance devient possible. De plus, la richesse de la transmission audio permet de restaurer les composantes émotionnelles qui s'expriment dans la prosodie. De même, la transmission de la vidéo permet de restaurer la complexité des expressions faciales. Il semblerait donc que l'on puisse retrouver une situation équivalente au face à face.

La conversation audio-visuelle médiatisée permet donc une communication à distance complexe qui pourrait au premier abord sembler aussi efficace que l'interaction en face à face. Elle se révèle en fait assez radicalement différente de la situation d'interaction réelle, et ne permet pas, du fait de ses différences et notamment de sa fragilité, un sentiment de présence important. Développons quelque peu le cas de la visiophonie, qui nous semble révélateur des difficultés du « contact » distal. Voir en quoi la visiophonie diffère de l'interaction en face à face nous permettra de justifier les premières directions de notre recherche.

Élaborée dès les années 1930, présentée lors des expositions universelles de 1958 et 1967, la visiophonie utilise de nos jours deux interfaces principales :

- visiophonie par le biais de l'ordinateur (avec des logiciels de type Skype, en conversation avec ou sans chat,...),

- visiophonie par le biais du téléphone portable. On pensera notamment à l'iPhone 4 (et modèles suivants) qui permet au choix de se filmer ou de filmer ce que l'on a devant les yeux. Cette technologie fonctionne de façon correcte depuis 2003 environ.

Mais les premiers dispositifs étaient de type filaire, associés à un poste fixe comme les premiers téléphones. C'est ce type de dispositif qui a été mis en place de façon expérimentale dans la ville de Biarritz en 1984, où il n'a pas rencontré un franc succès (Kaminow, Li, & Willner, 2010).

Michel De Fornel, qui a travaillé sur le cas de Biarritz, s'est intéressé à l'interaction par visiophonie et a dégagé les caractéristiques spécifiques de ce type d'interaction (qui n'est ni un appel téléphonique, ni une conversation en face-à-face). Dans Contraintes systémiques et contraintes rituelles dans l'interaction visiophonique, il s'intéresse aux pratiques et usages liés aux technologies de communication (De Fornel, 1988). Il souhaite procéder à l'analyse conversationnelle des interactions sociales. Il expose les variations structurales entre les interactions médiatisées (en particulier par le visiphone, qui permet d'explorer le canal verbal et le canal non-verbal dans le cas d'une communication à distance) et les interactions en face-à-face ordinaires. Il présente les propriétés du cadre interactionnel de l'échange visiophonique, et se demande quels sont la place et le statut de l'objet technique, de l'instrument de communication, dans l'espace interactionnel. Il commence par exposer les contraintes systémiques et les contraintes rituelles (notions posées par E. Goffman) liées à la communication visiophonique.

« Les premières peuvent se définir comme « les conditions et les dispositifs » permettant de faciliter une transmission efficace de la parole entre les interlocuteurs » (Goffman, 1987, p. 19). Les secondes concernent « la façon dont chaque individu doit se conduire vis à vis de chacun des autres, afin de ne pas discréditer sa propre prétention tacite à la respectabilité, ni celle des autres à être des personnes dotées de valeur sociale et dont il convient de respecter les diverses formes de territorialité » (ibid). Les contraintes systémiques ont un poids particulièrement important lors des communications médiatisées : il y a les contraintes du système d'interaction entre les personnes et les contraintes découlant de l'objet technique. De nombreux problèmes pratiques apparaissent au cours d'une communication visiophonique, et beaucoup sont liés aux contraintes structurelles de l'objet technique (par exemple les modifications de taille, de voix, de distance à la caméra,...). Les relations des participants s'en voient bouleversées. Les contraintes structurelles sont la représentation écranique (le dispositif tend notamment à entraîner la fixité et la déviation du regard) et la nécessité pour les interactants d'intégrer des activités de gestion du dispositif durant la conversation. Un travail interactionnel est nécessaire de la part des interactants. Ils doivent élaborer un ensemble de consignes de positionnement interactionnel. C'est l'acquisition d'un savoir-faire lié à l'interaction qui permet des pratiques visiophoniques maîtrisées. La conversation visiophonique est en fait un système d'activités situées. « L'objet technique – l'instrument de communication – est donc aussi une forme d'action en ce qu'il contribue à la mise en forme de l'interaction. » (De Fornel, 1988, p. 41). L'image permet l'introduction d'une forme de co-présence des interactants dans la communication visiophonique. L'espace, le territoire commun où se joue l'interaction, n'est pas une zone physique, mais un espace formé par un chevauchement maintenu de manière continue entre les segments transactionnels des participants, c'est-à-dire que les participants ont une orientation physique coordonnée qui permet l'émergence de ce territoire commun. La disposition spatio-temporelle des participants les uns vis-à-vis des autres est essentielle. Un espace commun est créé de façon coopérative entre les participants qui vont ajuster leurs corps vis-à-vis de la caméra disposée chez eux. En visiophonie, cet espace transactionnel partagé n'est pas physique concret, mais physique abstrait, et donc fragile (« chacun des participants doit gérer la relation de son interaction avec l'environnement extérieur et la prise en charge collective n'est pas possible, ce qui entraîne une fragilité très forte de l'espace transactionnel partagé » (ibid, p.44)).

Dans « Alors tu me vois ? » Objet technique et cadre interactionnel dans la pratique visiophonique, Michel de Fornel approfondit la question des modifications des pratiques interactionnelles liées à l'utilisation de médias destinés à la communication interpersonnelle. Dans le cas de la visiophonie, il s'intéresse au problème du « positionnement interactionnel », qui est en fait la difficulté des participants à avoir leur image captée correctement par le dispositif (De Fornel, 1992). Confier le maintien du bon positionnement à l'utilisateur pose

problème car cela perturbe le déroulement correct de la conversation visiophonique. M. De Fornel souligne la difficulté de créer un espace interactionnel commun à partir de deux espaces physiques séparés. « Loin d'être transparent, l'objet technique s'interpose. Il oblige les participants de l'interaction à se soumettre à des contraintes techniques et introduit des distorsions importantes de la communication. » (ibid, p.114) L'auteur revient ensuite sur ce qu'est l'interaction en face-à-face, laquelle implique co-présence physique. Goffman définit ainsi la co-présence : « Les personnes doivent sentir qu'elles sont suffisamment proches pour que tout ce qu'elles fassent soit perçu, ce qui inclut leur réaction aux autres » (Goffman, 1963, p. 17). La disposition spatio-temporelle des participants les uns par rapport aux autres est essentielle. En effet, les personnes en interaction doivent créer une forme d'espace commun. L'image, malgré la différence d'environnement spatial et l'absence de co-présence physique, permet bien l'introduction d'un sentiment de co-présence des interactants ; c'est une forme de co-présence virtuelle qui émerge. Cependant, un positionnement correct doit être maintenu pendant toute la durée de l'interaction afin de permettre la création et le maintien du territoire interactionnel commun dans lequel la transaction communicative se déroule. Les interactants doivent si possible être centrés sur l'image, et ils ne doivent surtout pas sortir du cadre, car cela constitue une rupture immédiate et complète de l'engagement. Le visage occupe une place centrale dans la régulation de l'interaction ; les interactants doivent se positionner non seulement pour voir, mais aussi pour être vu, ce qui constitue une contrainte propre à la visiophonie. Il faut à la fois une bonne attitude de réception, et une bonne attitude de production.

Enfin, dans Le cadre interactionnel de l'échange visiophonique, Michel De Fornel s'intéresse plus particulièrement aux modes d'engagement mutuels utilisés par les participants au cours d'une interaction visiophonique. Il examine à nouveau les stratégies des interactants pour créer un cadre interactionnel. Dans les interactions focalisées, la disposition spatio-temporelle des interactants les uns par rapport aux autres est essentielle. Ils vont de façon coopérative se positionner et ajuster leurs corps de manière à créer un espace commun. L'interaction visiophonique est une interaction focalisée ; les participants y orientent leurs corps de façon à surmonter la différence d'environnement spatial. Ils parviennent ainsi à créer le territoire interactionnel commun permettant à la transaction communicative de se dérouler. Les participants sont alors en coprésence virtuelle ; la communication visiophonique entraîne une distorsion de la forme de la communication de type face à face. L'aspect de la personne avec laquelle on communique est en effet modifié du fait de la médiation de l'écran, ainsi que la distance par rapport à l'objectif de la caméra. La présence d'autrui perd son caractère sensible et tactile, le sentiment de présence physique est atténué. « Au visiophone, le corps des participants a une fonction structurante de l'interaction mais en relation avec un objet technique : par leur activité corporelle – la posture, le positionnement, l'orientation visuelle – la situation

visiophonique devient un contexte doté de sens, un domaine intersubjectif d'actions pratiques. » (De Fornel, 1994, p. 111) Les participants s'isolent de leur environnement immédiat – mais à la différence du face à face, chacun d'eux s'isole de son environnement respectif pour construire un environnement commun. Dans l'interaction visiophonique, il est très difficile de maintenir une égalité de participation, c'est-à-dire que contrairement au face à face, les participants ne peuvent contribuer à construire de façon égale leur interaction de manière à ce qu'elle soit plus ou moins isolée, plus ou moins séparée du reste de l'environnement concret (ibid). Les interactants n'ont pas les mêmes droits quant à l'ouverture ou à la fermeture face à toute intrusion extérieure dans une interaction visiophonique, car ils ne partagent pas le même espace physique concret. Ils ne peuvent donc y exercer un contrôle commun. Ils doivent intervenir de façon séparée s'ils veulent maintenir une frontière entre l'espace de leur interaction et l'environnement général dans laquelle il s'inscrit. L'espace transactionnel partagé est donc très fragile, étant donné qu'une prise en charge collective de la gestion de la relation de l'interaction avec l'environnement extérieur n'est pas possible et doit être effectuée par les interactants, chacun de son côté. La fragilité de l'espace interactionnel tient aussi au fait que les participants doivent maintenir un positionnement correct pendant toute la durée de l'interaction visiophonique. Mais ils doivent aussi aider leur interlocuteur à garder un positionnement correct, notamment tout ce qui concerne les sorties de cadre (qui constituent, comme cela a déjà été dit, une forme de rupture de l'engagement, étant donné que le visage occupe une place centrale dans la régulation de l'interaction). Le maintien d'un espace transactionnel commun ne peut être dissocié d'un contrôle visuel réciproque, lequel repose sur la possibilité d'échange de regards entre les deux participants. Il est nécessaire dans l'interaction visiophonique de se positionner tant pour voir que pour être vu, ce dernier point échappant à certains utilisateurs, qui sont persuadés d'être vus par leur partenaire lorsqu'ils le voient (ce qui est en effet le cas lors de l'interaction en face à face habituelle). Il est intéressant de noter que comme dans le face à face, l'échange de regards joue un rôle important dans la régulation des tours de parole. Les gestes permettent eux aussi aux participants d'organiser leur interaction et peuvent avoir une valeur communicative. La médiation de l'écran du visiophone tend à entraîner la fixité du regard du destinataire. « Au visiophone, le destinataire peut fixer continûment son interlocuteur, quel que soit le moment du discours, alors qu'une telle fixation intense provoquerait très rapidement de la gêne et de l'embarras en face à face. » (ibid, p. 117) Mais ce regard n'est pourtant pas intrusif, il est plutôt la marque de l'attention de l'interlocuteur. Les participants ont de plus le souci de « minimiser la possibilité d'apparition de problèmes sur le plan rituel, en particulier, dans les termes de Goffman, de problèmes pour la face » (ibid, p.120) (Goffman définit comme suit la face : « On peut définir le terme de face comme étant la valeur sociale positive qu'une personne revendique effectivement à travers la ligne d'action que les autres supposent qu'elle a adoptée au cours d'un

contact particulier. La face est une image du moi délinéée selon certains attributs sociaux approuvés, et néanmoins partageables, puisque, par exemple, on peut donner une bonne image de sa profession ou de sa confession en donnant une bonne image de soi. » (Goffman, 1974, p. 9)) La présentation de soi et la maîtrise de sa propre image sont centrales dans l'interaction visiophonique, qui nous donne le sentiment d'un état permanent de visibilité. La fragilité de la face qu'entraîne ici l'objet technique est liée à certains moments clés de l'interaction (notamment l'ouverture de l'échange). « De façon générale sauver sa face et ne pas la faire perdre à l'autre sont deux aspects inséparables du déroulement de l'interaction. » (De Fornel, 1994, p. 122) La conversation visiophonique possède un caractère « menaçant pour la face », ce qui se sent à travers la présence de nombreux compliments et évaluations positives, qui ont sans doute pour effet de rassurer l'interlocuteur sur l'image qu'il donne de lui-même. La dimension confirmative dans la méta-communication (la communication à propos de la réussite de cette communication) occupe une place très importante dans la visiophonie.

Diverses améliorations techniques (en particulier les logiciels comme Skype) ont permis le développement récent de la visiophonie en la rendant plus mobile (et sans surcoût). Son usage est soit privé (communication en famille ou entre amis, ou encore avec parfois de parfaits inconnus, comme sur le site chatroulette), soit professionnel (visioconférences, guichets visiophoniques, télé médecine, murs ou salles de téléprésence, téléenseignement,...)

S'il n'y a chez chacun qu'une seule caméra ou webcam, soit on regarde l'objectif (et donc le partenaire a l'impression qu'on le regarde), soit on regarde l'image du partenaire (qui alors n'a pas l'impression qu'on le regarde ; c'est l'effet « faux-jeton »), soit encore on regarde l'image que le partenaire voit de nous (qui peut d'ailleurs être incrustée dans une partie de l'écran). Mais est-il plus important de pouvoir agir sur ce que l'on voit ou sur ce que l'on donne à voir, sur ce que notre partenaire voit ? Il serait alors intéressant d'étudier la possibilité de dispositif de visiophonie où les interlocuteurs auraient la main sur le placement et les mouvements des caméras dans l'espace filmé chez l'interlocuteur. En quoi le déplacement de la caméra virtuelle pourra-t-elle améliorer la conversation visiophonique ? Il semble que cela permettrait de passer outre ce regard dévié et fixe (la fixité n'étant pas nécessairement négative, elle n'est pour autant pas la norme de l'interaction normale en face à face). Le rapport au visage s'en verra changé. Se pose aussi la question de l'oubli de la technique. Quel est l'impact du fait que l'image de la personne que l'on voit ne soit pas en taille réelle ? La téléprésence est un outil de visiophonie qui combine l'échelle 1 et la sensation de se regarder les yeux dans les yeux ; elle pourrait permettre de répondre à ces questions. D'autre part, permet-elle un oubli de la technique, de rendre le dispositif transparent ? Cela est-il seulement faisable ?

Dans Efficient Dense-Stereo and Novel-view Synthesis for Gaze Manipulation in One-to-one Teleconferencing, les auteurs proposent un nouvel algorithme permettant, à partir de deux

images captées de part et d'autre d'un écran, de synthétiser l'image d'une caméra virtuelle pouvant se trouver à n'importe quel endroit du moniteur afin de faciliter le contact visuel lors d'une interaction en visiophonie (Criminisi, Shotton, Blake, & Torr, 2003). Le but des auteurs est de réussir à placer la caméra virtuelle de manière à ce que l'utilisateur ait l'air de la regarder en face – ce qui constitue donc une solution supérieure à celle qui consisterait à avoir la caméra virtuelle sans cesse au centre de l'écran, ce qui obligerait l'utilisateur à toujours regarder à cet endroit-là. Les auteurs se proposent pour cela de placer la caméra virtuelle derrière la fenêtre où s'affiche la conversation visiophonique, et où s'affiche l'image de l'interlocuteur.

Des systèmes comme celui-ci et celui de la téléprésence peuvent donc recalculer l'image affichée chez l'interlocuteur quand on regarde son image sur notre écran, afin de lui donner l'impression qu'on le regarde lui. Le problème qui se pose alors est celui d'un croisement perceptif visuel qui s'établit de manière constante et en est dérangeant pour les utilisateurs (car on ne peut alors pas échapper au regard du partenaire). On a l'impression d'être dévisagé par un regard fixe. Dans C'est très touchant, Charles Lenay souligne l'importance du décalage et de l'alternance dynamique entre croisement de regard et évitement (Lenay, 2010) : ce qui est le plus intéressant n'est pas la situation où l'on est dans l'immersion perceptive directe, mais plutôt ce qui est sur les bords du croisement perceptif, ce qui est oblique, tangent.

Gaver et al ont quant à eux développé un système appelé The Virtual Window (la fenêtre virtuelle). Ce système s'appuie sur les mouvements de la tête ayant lieu dans un bureau pour contrôler les mouvements d'une caméra dans un bureau distant. Le résultat ressemble à une fenêtre permettant l'exploration de scènes lointaines, plus qu'à un écran plat montrant des images mouvantes. Les auteurs ont développé un prototype de leur dispositif et ont observé des utilisateurs ; leurs analyses suggèrent qu'il permet de surmonter les limites des configurations habituelles de l'espace virtuel. Il permet notamment un champ de vision élargi, une réduction des discontinuités visuelles, aide à la négociation mutuelle de l'orientation, fournit des informations sur la profondeur et permet de fournir une conscience de la caméra (Gaver, Smets, & Overbeeke, 1995).

Se pose aussi la question de l'utilisabilité et de l'acceptation par les utilisateurs, question notamment traitée par Dourish (1993). Il pose en effet que les espaces virtuels, en constante amélioration, permettent d'intégrer de l'audio et de la vidéo afin de fournir un environnement de communication riche, souvent en vue de permettre la collaboration. Mais il souligne d'autre part que cette connectivité soulève d'importantes questions concernant l'intimité, la protection et le contrôle ; il est essentiel de considérer ces questions pour tirer le meilleur parti des nouvelles technologies. Ainsi, Dourish a développé une infrastructure (RAVE) prenant en compte ces problèmes dans un environnement de travail. Un aspect-clé de ce travail est la relation existant entre deux facettes du système développé, à savoir d'une part les éléments technologiques

déterminant la façon dont le système va se comporter, et d'autre part les éléments sociaux déterminant les utilisations et les comportements acceptables (ibid).

La visiophonie a aussi permis le développement d'une palette d'outils numériques, comme les espaces virtuels partagés pour des réunions à distance, les éditeurs de texte partagés, les logiciels de dessins ou de modélisation partagés,... Ces outils illustrent bien l'idée que de nombreux moyens de communication sont mis en place non seulement pour *communiquer* à distance, mais pour *collaborer* à distance, donc plus dans un contexte de travail que d'échange émotionnel. Il s'agit du vaste domaine du CSCW (Computer Supported Cooperative Work). L'eScience est un terme aussi utilisé pour désigner la conception de systèmes visant à aider à la collaboration à distance de groupes de scientifiques (Fraser et al., 2006).

Comme exemple de système de collaboration à distance, on pourra citer les laboratoires de recherche Rank Xerox de Cambridge à EuroPARC, qui ont développé un espace numérique en vue de permettre le travail en coopération et la sociabilité informelle entre des individus se trouvant dans différentes parties d'un même bâtiment. Une infrastructure omniprésente retransmettant l'audio et la vidéo a été mise en place, permettant aux individus d'accéder les uns aux autres même en étant dans des pièces différentes. Heath et Luff rapportent, suite à des observations préliminaires du système en utilisation, que même si la technologie peut fournir un accès audio et vidéo en temps réel aux autres membres du personnel, elle transforme cependant le comportement visuel des utilisateurs, ce qui entraîne des conséquences sur la distribution et la réception de la parole. En particulier, ils notent que la communication interpersonnelle au travers de cet espace numérique entraîne certaines asymétries dans l'interaction. Ces asymétries modifient l'impact des modalités visuelles et vocales (Heath & Luff, 1991, 1992) ; on pourra rapprocher ce résultat des travaux de Michel De Fornel présentés un peu plus haut.

On notera que les recherches sur la collaboration à distance se focalisent non seulement sur la visiophonie, mais aussi sur la possibilité pour les interactants d'accéder et d'utiliser une palette de ressources utiles se trouvant à proximité, telles que des objets et des outils. On a donc pu assister à l'émergence de technologies conçues afin de permettre aux utilisateurs d'accéder aux environnements les uns des autres, ainsi qu'aux caractéristiques de ces environnements. Par exemple, Luff et al ont mis en place un système créant un espace numérique partagé riche, conçu afin de permettre aux participants de se référer et de pouvoir pointer des objets et outils se trouvant dans l'environnement distant de leurs interactants (Luff et al., 2003).

La visiophonie n'a pas été explicitement conçue pour prendre en compte les informations non-verbales, qui sont pourtant particulièrement expressives, notamment en ce qui concerne le côté émotionnel d'une interaction. Même si la visioconférence présente l'avantage de permettre les informations non-verbales, des études (voir Ames, Go, Kaye, & Spasojevic, 2010) montrent que la vidéo n'est traitée que comme une « caractéristique technique ». Le soin est laissé aux

utilisateurs de s'approprier cette composante, c'est-à-dire de la faire marcher en tant que pratique sociale. Une étude récente sur la visioconférence entre couples séparés par de longues distances menée par Neustaedter & Greenberg (2012) fait état de la richesse des pratiques inventées par les utilisateurs, et de ce qu'il en résulte en terme d'expérience. En d'autres termes, le téléphone et les systèmes de visiophonie sont conçues pour une fonction technique (à savoir transmettre le texte, le son et la vidéo), et c'est secondairement, à travers l'usage, qu'ils peuvent être porteurs d'une expérience pour les utilisateurs. L'émotion se transmet alors par le contenu ; mais l'émotion réside-t-elle seulement dans l'information ?

Ce qui est caractéristique dans la visiophonie, c'est que l'on contrôle (et que l'on doit contrôler) le point de vue que notre partenaire a sur nous, ce qui est contradictoire avec les conditions d'une communication classique, où au contraire ce que l'on présente à autrui nous échappe. Avec la visiophonie, il n'est pas facile de savoir si l'on est bien vu par son partenaire. Est-ce que je vois que je suis vu ? Quand est-ce que je vois quand je suis vu ? Lorsque l'on voit, on n'est pas pour autant sûr d'être vu. Nous soulignons ici qu'au contraire, lorsque l'on touche, on est sûr d'être touché ; la réciprocité est une caractéristique du sens du toucher. N'est-ce pas là le sens profond du contact, recherché dans toute communication émotionnelle ? Pour nous, il réside là un enjeu essentiel. Remarquons d'ailleurs les différences entre les modalités perceptives du sonore, du visuel et du tactile. Je ne peux pas toucher sans être touché ; lorsque je touche, je suis touché et bien sûr touchable. Au contraire, je peux voir sans être vu, mais je reste cependant toujours visible. Mais dans l'audition, je peux entendre sans être entendu, et sans être entendable. Le niveau de non-réciprocité des perceptions des différents sujets est donc encore plus grand en audition qu'en vision. On pourrait ainsi établir une classification des sens basée sur leurs propriétés relationnelles, et non sur leurs propriétés fonctionnelles. Ces distinctions ont un rôle-clé dans la conception des modes d'interaction.

Des signaux de fumée à skype, de très nombreux dispositifs de communication à distance de plus en plus performants ont été proposés. De l'utilisation de signes et symboles intermédiaires, communiqués d'une manière de plus en plus rapide et efficace, au transfert direct de la parole telle qu'elle est prononcée, se trouvant parfois accompagné de l'image des interactants ; voici ce que permettent les différents outils développés pour permettre la communication à distance. L'étude des systèmes de communication à distance, qui ont donc tout d'abord été asynchrones puis de plus en plus synchrones avant d'aboutir à la visiophonie, nous a permis en mettre en évidence les différentes limites existant dans la communication à distance. Ces limites impactent notamment le caractère émotionnel de ce type de communication, en le rendant plus difficilement accessible. Notre objectif est de créer des dispositifs d'interaction émotionnelle instaurant un contact, un espace de rencontre émotionnelle, qui ne souffrirait pas des dissociations que l'on observe dans les différents dispositifs que l'on vient de présenter.

Nous pouvons dégager un certain nombre d'hypothèses préliminaires par l'étude de ces dispositifs ayant été utilisés jusqu'ici. Par exemple, la visiophonie souffre de plusieurs défauts qui n'en font pas un outil aussi proche de l'interaction en face à face qu'on pourrait le croire, notamment sur le plan du transfert émotionnel. On constate que les interactants n'arrivent pas à créer facilement un espace commun par le biais de la visiophonie, et qu'ils doivent sans cesse réajuster leur position et s'assurer qu'ils sont vus. Chacun est et reste dans son propre espace, tandis qu'un véritable espace commun devrait se définir par l'interchangeabilité des points de vue (Schutz, 2005). C'est là le problème majeur de la visiophonie : les interactants doivent sans cesse construire l'espace d'interaction, en veillant à recevoir et à produire une image adéquate, ce qui casse l'immersion dans l'interaction elle-même et a un impact négatif sur le sentiment de présence. Il faut prendre garde à la fragilité de l'espace transactionnel partagé que l'on conçoit. De plus, la nécessité de vérifier que l'on est bien vu (et ce même si l'on voit bien son partenaire) va déjà à l'encontre d'un contact. En effet, le tactile implique une réciprocité immédiate : si je touche, alors je suis touché, et inversement. Il s'agit pour nous d'éviter les limites présentées dans les systèmes que nous allons proposer.

0.4 Motivations et enjeux

Les enjeux de cette recherche concernent des domaines divers ; il existe notamment un enjeu sociétal, un enjeu de conception et un enjeu scientifique. Notre travail concerne en effet à la fois la théorie à un niveau fondamental, le design (non seulement les interfaces conçues mais aussi le processus de conception), et l'expérimentation.

0.4.1 Enjeu sociétal : le besoin de contact à distance

Il convient tout d'abord de rappeler que dans la communication, ce qui est primordial n'est pas tant les messages échangés, la simple transmission d'information, mais le développement d'un sentiment de présence partagé, ce que l'on appelle « contact », même à distance. Le contact n'est pas seulement un besoin, mais est un élément nécessaire à notre survie. Nous communiquons, nous renouvelons sans cesse nos contacts, nous construisons autour de nous un réseau de relations, et ceci dans l'optique de pouvoir nous développer durablement au sein d'une communauté, en vue d'augmenter nos chances de survie et de reproduction (Bougnoux, 2001). Nous avons un besoin d'interaction physique, corporel, avec autrui, qui s'il n'est pas satisfait amène à notre périlclitation (voir par exemple le développement des enfants sauvages relaté dans Malson, 2002).

La mondialisation et la démocratisation des transports, notamment de l'avion, permettent une mobilité croissante, qui a parfois pour conséquence paradoxale la séparation physique des individus ; ainsi donc on peut se retrouver séparé de sa famille, de ses amis, de ses relations. Comment faire alors pour garder ce contact vital même à distance ? Les nouveaux moyens de communication, permis notamment grâce à internet, sont utilisés comme des possibilités de maintenir les liens des relations interpersonnelles. Les sites de réseaux sociaux, au succès fulgurant, sont un exemple de cette tendance. Des études concernant l'activité d'adolescents sur des sites de type réseaux sociaux montrent que l'activité en ligne des individus est principalement concentrée sur les relations interindividuelles (Boyd & Ellison, 2007; Desavoie, 2011).

Permettre le maintien du contact est un enjeu majeur. Selon le centre dédié aux études des relations à grande distance (Center for the Study of Long Distance Relationships), environ 3,6 millions d'habitants des Etats-Unis mariés ont fait l'expérience d'une relation à longue distance en 2006, et le taux de mariages à longue distance a augmenté de 30% entre 2000 et 2005. En 2005, le nombre estimé de couples à longue distance au niveau universitaire s'élevait entre 4 et 4,5 millions³. Les couples illustrent une situation particulièrement critique de séparation, d'autant qu'il arrive souvent que le couple ait des enfants. Cependant il y a bien d'autres cas de séparation pouvant survenir ; avec la famille plus étendue, avec les amis, avec les relations de voisinage et du travail,.... Maintenir le contact avec tous ces liens nécessite de pouvoir communiquer, mais pas uniquement. Il s'agit non seulement de proposer aux individus des moyens efficaces pour communiquer, mais aussi des expériences riches permettant l'échange émotionnel, afin de permettre le maintien des liens.

Le contact avec autrui peut être considéré comme une coïncidence spatiale entre deux sujets percevants. Nous entendons le terme « coïncidence » au sens de l'état d'événements qui se réalisent en même temps et au même endroit, et non dans le sens de hasard, concours de circonstance. Dans le monde physique, cette définition du contact comme coïncidence spatiale est facile à appréhender. Cependant, il n'est pas aussi simple de la considérer lorsqu'elle s'inscrit dans un système de communication. Il s'agit alors d'étendre l'idée d'une coïncidence des positions à un espace virtuel, partagé entre les individus, à un espace créé par les interfaces et le réseau de communication. Par exemple, lors d'une communication téléphonique, les interlocuteurs se rencontrent dans un espace sonore partagé. Dans le cas de la visiophonie, cet espace partagé est, comme on l'a vu, le résultat d'une difficile coordination des interlocuteurs en face de leurs caméras et écrans. Cependant, d'après Lenay (2010), l'idée du contact comme coïncidence spatiale doit être critiquée et analysée. Selon lui, la coïncidence spatiale stricte ne

³ <http://www.longdistancerelationships.net/faqs.htm>

permet pas de faire l'expérience du contact avec autrui. En effet, si nous occupons des positions exactement identiques, si nos perceptions sont confondues, alors nous ne pouvons en fait nous percevoir l'un l'autre, de même que nous ne pouvons percevoir ce avec quoi l'on perçoit (il est impossible de voir nos propres yeux sans recourir à un miroir). Pour illustrer ce point, Lenay prend l'exemple de l'écran de télévision : le point de la perception de ce qui passe à l'écran (donc celui de la caméra) est le même pour tous les spectateurs – spectateurs qui ne se perçoivent donc pas mutuellement. Pour percevoir autrui, et pour qu'il nous perçoive réciproquement, il est donc nécessaire que nous partagions un même espace, mais pas que nous soyons exactement à la même position : il faut que nous soyons face à face et non confondus en un même point. Nous devons pouvoir déplacer nos points de vue l'un par rapport à l'autre. L'hypothèse considérée ici est donc plutôt que pour pouvoir percevoir autrui, il faut pouvoir percevoir son activité perceptive.

Pour parvenir à maintenir le contact à distance, il s'agit en partie de permettre aux individus de se percevoir à distance, et donc de percevoir leurs activités perceptives réciproques. Nous allons pour cela privilégier deux pistes : d'une part celle du toucher à distance (aussi paradoxale que puisse sembler cette expression ; voir partie 2), et d'autre part celle de l'attention conjointe (voir partie 3).

0.4.2 Enjeu et méthode de conception : design d'interfaces

Notre enjeu de design d'interaction se situe à la fois au niveau du design d'interface et du design de la structure, du mode de fonctionnement de l'espace d'interaction. Notre objectif est de réaliser un « contact » entre utilisateurs à distance, que ce soit via des entrées sensorielles visuelles, auditives et/ou tactiles. Il s'agit de trouver les supports d'un sentiment de présence partagé et d'échange émotionnel.

Le développement d'interfaces proposant des modes d'interaction inédits n'est pas sans difficultés. Cela reste cependant une occasion de se confronter aux problèmes de méthodologie rencontrés par les designers lorsqu'ils développent un produit radicalement nouveau – ici pour l'interaction à distance. Par notre démarche, nous allons proposer des méthodes et outils à mettre en œuvre lors du développement de ce type de projets. Ces méthodes et outils s'appuient sur des bases théoriques, ainsi qu'une méthode expérimentale, et s'intègrent dans un processus de conception innovante. Nous cherchons à proposer des principes et méthodes pouvant être utiles pour les designers se trouvant confrontés aux difficultés propres à ce type de conception.

Les approches de conception centrées utilisateurs consistent en la mise en place d'outils et de méthodes permettant de considérer l'activité des utilisateurs (Norman, 1988). Elles prônent

l'implication de l'utilisateur au cours du développement ; l'utilisateur est considéré comme ayant une expertise relative de l'activité du produit, mais il est également alors vu comme un co-concepteur (Darses, 2004). Dans une optique de développement d'interfaces pour la communication à distance, la conception participative constitue un moyen d'améliorer le processus de conception, notamment du fait d'une meilleure expression des besoins dès l'amont. Comme on le verra, une situation d'innovation radicale ne peut s'appuyer sur une expérience ou un retour utilisateur préalable. Une des issues proposées ici sera de trouver l'inspiration à partir d'une recherche fondamentale autour des médiations techniques dans les interactions humaines. Notre travail participera ainsi à une recherche expérimentale sur les questions de cognition sociale comme celle de la reconnaissance d'autrui ou celle de l'attention conjointe afin de pouvoir atteindre son objectif de conception.

L'objet conçu et le processus de conception sont deux aspects qui doivent cohabiter dans une démarche de recherche en conception (Stephan, 2005). Pour pouvoir innover, il faut prendre en considération les fondements et les conditions de la conception, et non simplement appliquer de nouvelles techniques et développer de nouvelles technologies. Et de même, la recherche en conception ne peut se faire simplement à un niveau théorique ; les théories doivent se vérifier de manière empirique.

0.4.3 Enjeu scientifique : un niveau fondamental

Notre vie quotidienne se peuple toujours plus de nouvelles interfaces ; si le nombre de designers d'interaction, professionnels dédiés au développement de ces interfaces, augmente de manière parallèle, il n'existe pourtant pas de véritable discipline scientifique portant l'interaction outillée au centre de ses préoccupations. L'interaction entre individus est traitée par diverses sciences sociales (notamment la sociologie et la psychologie), mais aussi par les sciences cognitives, par le domaine de l'IHM et par l'ergonomie. Cependant l'interaction entre individus au travers de ces interfaces n'est pas étudiée de manière systématique.

Ainsi que le souligne Dan Saffer (2006, p. 4, traduction personnelle) : « Le design d'interaction est un art – un art appliqué, comme la fabrication de meubles ; ce n'est pas une science. Bien que d'excellentes pratiques aient émergé au cours des trente dernières années, cette discipline doit encore parvenir à mettre en place des règles strictes et rapides pouvant être prouvées par des méthodes scientifiques, et qui se voient vérifiées à tous les coups. »⁴

⁴ « Interaction design is an art an - applied art, like furniture making; it's not a science. Although best practices have emerged over the past three decades, this discipline has yet to arrive at hard and fast rules that can be proven via scientific methods and that are true in all instances. »

Il est difficile pour le design en général, et le design d'interaction en particulier, de s'imposer comme science, ou comme recherche. Cependant, il s'agit sans aucun doute d'une activité qui produit des connaissances nouvelles, ayant parfois un caractère fondamental (Schneider, 2005). En ce sens, nous positionnons notre activité de recherche comme étant en partie une activité de « recherche par le design » (Frens, 2006). Pour cela, nous allions un travail de conception, d'observation des usages et de l'expérience vécue, ainsi que des études de psychologie expérimentale plus traditionnelles.

Ainsi, notre recherche va consister en l'articulation de trois pôles : théorie, expérimentation, et design (voir figure 1). Nous allons donc appliquer une méthodologie novatrice, mais nous ne faisons pas pour autant une recherche méthodologique, puisque nous ne comparons pas notre démarche avec une autre. Nous allons valider par la pratique un certain type de méthodologie, original, qui articule étude expérimentale, théorie et design, tout en prenant en compte l'expérience utilisateur.

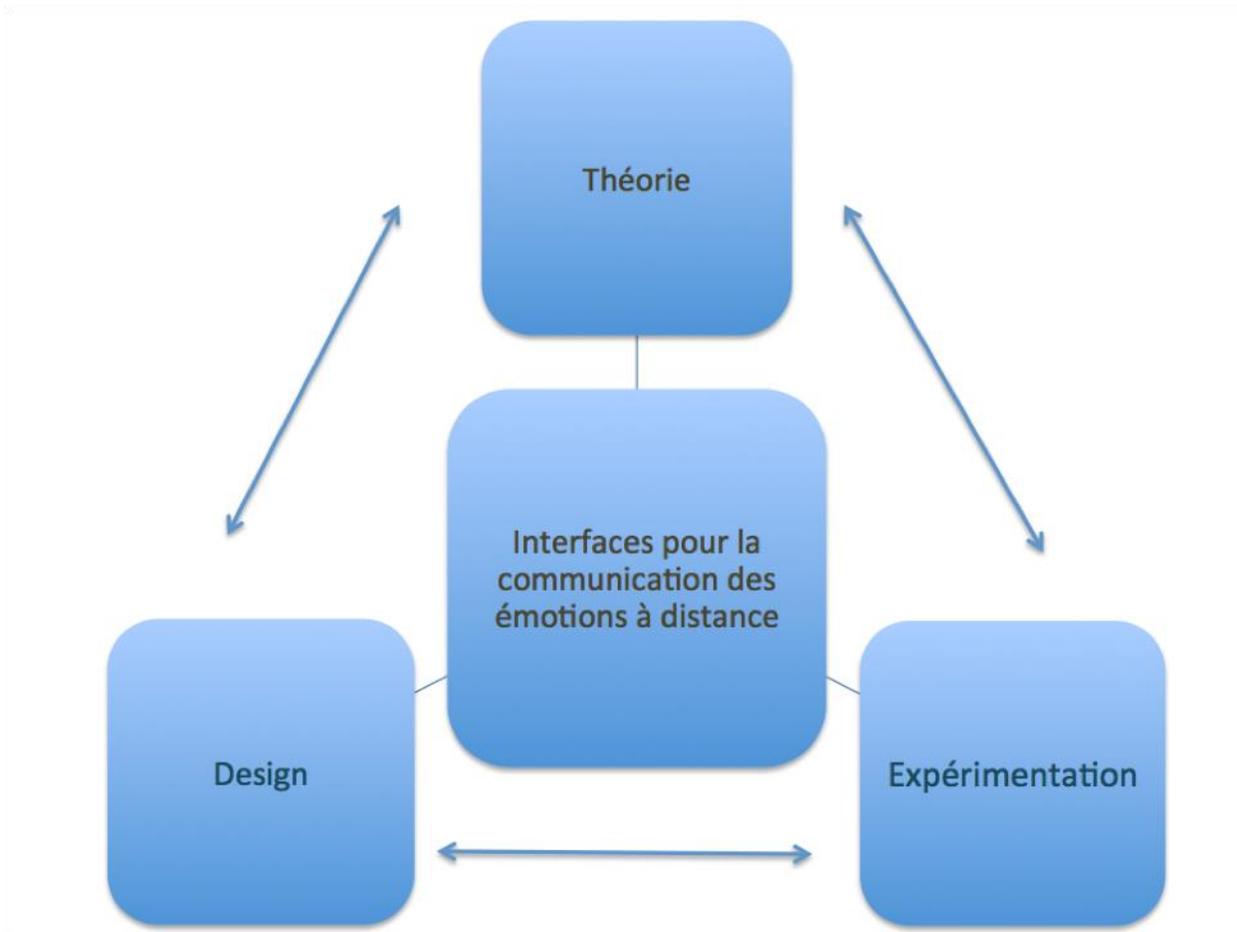


Figure 1. Notre méthode de recherche.

1 Cadre théorique et méthodologique

1.1 Éléments théoriques

1.1.1 De la cognition

1.1.1.1 Énaction et approche sensori-motrice de la perception

1.1.1.1.1 Présentation de l'énaction

Les sciences cognitives ont pour objet l'ensemble des mécanismes de pensée : nos façons de percevoir, de mémoriser, d'apprendre, de raisonner, de calculer, de communiquer, de reconnaître, de guider, mais aussi de définir des valeurs, des appartenances, des désirs, des identités,...

Le courant dominant en sciences cognitives est celui de la computation, qui consiste à considérer l'homme comme une sorte d'ordinateur ; le paradigme computo-représentationnel s'appuie sur la métaphore de l'ordinateur pour décrire les processus mentaux. Il pose d'une part l'hypothèse d'un système cognitif prédéfini et déterminé, et d'autre part l'hypothèse selon laquelle le système cognitif met en place des représentations internes du monde extérieur, lequel est également déterminé. La cognition y est vue comme un *calcul* sur des unités symboliques formellement considérées, à partir de règles de manipulation (des algorithmes). Il n'y a donc pas de cognition sans computation (ce qui n'implique pourtant pas que toute machine computationnelle soit une machine pensante). L'approche cognitiviste s'inscrit dans ce paradigme ; c'est classiquement l'approche dominante en sciences pour la conception. Dans le domaine de l'IHM (Interface Homme-Machine), la référence a longtemps été le modèle MHP (Model Human Processor), proposé par Card, Newell et Moran (1983) et issu de la psychologie cognitive. L'esprit y est vu comme un processus de traitement de l'information, régi par des règles ; la pensée y est comparée à un calcul. Dans ce modèle, le système cognitif humain est décomposé en trois sous-systèmes spécialisés indépendants (perceptif, cognitif, moteur), pourvu chacun d'un processeur et d'une mémoire locale. Il dispose de plus de deux types de mémoire : une mémoire à court terme (ou mémoire de travail), et une mémoire à long terme. Les mémoires sont partitionnées en différents espaces de stockage spécifiques aux différents types de contenu existants : visuel, auditif, tactile, émotionnel,... Le système nerveux est mis en parallèle avec la partie hardware d'un ordinateur ; c'est la partie matérielle. Le sujet est vu comme un individu doté de représentations internes : ce sont des modèles d'un monde extérieur figé et prédéterminé. Ces représentations internes sont supposément stockées dans les divers espaces de mémoire du cerveau ; elles se trouveraient activées par les stimuli issus des capteurs sensoriels dont nous disposons. Les actions motrices effectuées par le sujet correspondraient alors au résultat d'un

traitement cognitif faisant suite à cette activation. Les opérations mentales exécutées sont donc comparables aux logiciels (ou softwares) d'un ordinateur. L'activité humaine est finalement perçue comme étant d'une certaine façon la réalisation d'un programme déterminé, exactement comme l'est un programme informatique.

Lenay (2006a) illustre cette façon d'aborder la cognition par une comparaison avec un sous-marin : les représentations internes sont les représentations du monde disponibles dans l'habitacle du sous-marin (comme par exemple les cartes), les stimuli sont issus des capteurs des appareils (par exemple l'antenne radar), et les tâches (telles que « tourner de 45° à bâbord ») sont découpées en séquences élémentaires (correspondant chacune à une donnée à faire varier grâce à un appareil à bord, telle que la direction avec le gouvernail).

Une alternative à la vision computo-représentationnelle de la cognition est constituée par la théorie de l'énaction, proposée par Varela, Thompson et Rosch (1991). L'énaction met l'accent sur la manière dont les organismes s'organisent eux-mêmes en interaction avec l'environnement. Stewart et al. définissent ainsi l'énaction : « A living organism enacts the world it lives in; its effective, embodied action in the world actually constitutes its perception and thereby grounds its cognition » (Stewart, Gapenne et Di Paolo, 2010, vii)⁵. Les notions d'*autonomie*, c'est-à-dire la capacité à s'adapter à chaque instant à l'environnement (lequel change constamment), d'auto-organisation, d'expérience et d'inscription corporelle sont au cœur de la vision énative de l'esprit. L'énaction place la perception au centre de la boucle sensori-motrice ; notre perception émerge de l'influence de nos actions sur nos sensations (voir figure 2).

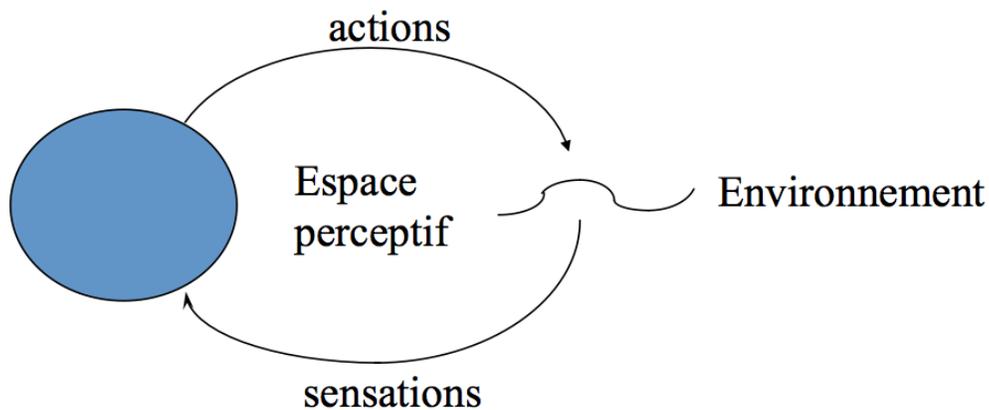


Figure 2. La perception émergeant au centre de la boucle sensori-motrice.

⁵ « Un organisme vivant énonce le monde dans lequel il vit ; son action effective, incarnée dans le monde est en fait ce qui constitue sa perception, et qui par conséquent fonde sa cognition. » (Traduction personnelle.)

Un individu déploie son activité perceptive en fonction des retours sensoriels qui résultent de cette activité (Lenay et al., 2000). On ne peut séparer la perception et l'action. L'individu est en effet engagé dans le monde ; il reçoit des sensations en provenance de ce monde, et ses actions modifient ce dernier (ce qui a pour conséquence de modifier les sensations reçues par l'individu, qui va en conséquence modifier ses actions, etc). La perception du monde ne peut donc se faire que par le biais d'une loi de contingence⁶ sensori-motrice, découverte ou construite par le sujet : il faut qu'il existe une régularité dans le lien entre une action produite et la variation de sensations qui en résulte (O'Regan et Noë, 2001). On ne peut percevoir que parce que l'on peut agir, et sans action, il n'y a pas de perception. Par exemple, si l'on colle une image sur la rétine, et que l'œil ne peut donc plus explorer l'image, alors au bout de quelques secondes l'image s'évanouit : on ne la perçoit plus (O'Regan et Noë, 2001). L'esprit n'est donc pas une entité que l'on peut séparer de l'activité humaine (Racine et Carpendale, 2007).

Gibson (1966) met aussi en avant cette nécessité d'agir pour percevoir. Pour lui, les sens ne sont pas de simples capteurs passifs, produisant des sensations visuelles, auditives, tactiles, odorantes ou gustatives. Ce sont au contraire des mécanismes de recherche *actifs*, permettant de regarder, écouter, toucher, sentir et goûter. La perception est donc active, et non passive ; l'action du sujet, et son activité d'exploration du monde, sont indispensables à la perception. Pour le sens du toucher par exemple, sans une succession d'actions de la main, il est impossible de percevoir les caractéristiques d'un objet, comme par exemple sa dureté (Klatzky et al. 1993). On ne peut reconnaître un objet dans un sac opaque que si l'on peut l'explorer de la main, pour en saisir les contours, la texture, la consistance. De même, on ne peut voir lorsque notre œil est immobilisé (O'Regan et Noë, 2001). Il faut donc agir pour percevoir.

Dans la théorie de l'énaction, la notion de couplage structurel existant entre l'individu percevant et le monde perçu est première. Ainsi, si la computation est une approche comparable à un sous-marin, alors l'énaction se compare au plongeur (Lenay, 2006a) : l'individu est *directement engagé* dans le monde, et il *va* l'explorer. Il n'a pas besoin d'une représentation du monde, étant donné qu'il s'y trouve déjà. Ses mouvements s'adaptent en permanence au monde qui l'entoure, et il ne le perçoit que parce qu'il s'y engage. Dans cette vision, le monde est vécu à chaque instant par le sujet qui le perçoit et qui y agit. En retour, le monde, tel qu'il est constitué par l'individu (en fonction de ses capacités motrices et des capacités de ses sens), guide les actions de l'individu. À chaque instant, le sujet agit dans un environnement qui a du sens pour lui, mais qu'il a lui-même contribué à rendre signifiant. Il n'y a donc pas deux pôles, avec d'un côté sujet et de l'autre objet, entre lesquels une interaction peut éventuellement s'établir ; c'est

⁶ Une contingence se définit par la présence d'une relation temporelle entre deux événements (Moran, Dumas et Symons, 1992).

l'interaction elle-même qui est primordiale. L'énaction pose la cognition comme une activité incarnée et située (Varela, Thompson et Rosch, 1991 ; Stewart, Gapenne et Di Paolo, 2010). Tout sujet est en fait engagé en permanence dans une activité de construction du sens. C'est cette activité qui co-détermine à la fois l'agent et l'environnement, et qui est désignée sous la formulation *sense-making* (De Jaegher et Di Paolo, 2007). La cognition se définit alors par la génération et la transformation de significations que le sujet produit du fait de son couplage avec l'environnement. « They actively participate in the generation of meaning in what matters to them; they enact a world. » (De Jaegher et Di Paolo, 2007, p.488)⁷.

Une vision énaactive du design place les *actions situées* (Suchman, 1987) au cœur de l'interaction entre l'homme et la machine. Ce sont elles qui guident le déroulement de l'interaction, et non un plan ou une procédure prédéterminés que l'utilisateur aurait en tête (ce qui constituerait une vision plus cognitiviste). Les actions situées se constituent dans l'histoire de nos expériences, et fondent ces mêmes expériences. L'utilisateur adapte en permanence son comportement en fonction de la façon dont il interprète les réactions de la machine avec laquelle il interagit et de la perception qu'il a de la situation. Lorsque les caractéristiques de la situation changent, l'utilisateur s'adapte à ces changements et ajuste à la volée ses actions aux contingences environnementales. Il est donc engagé dans une boucle d'action et d'interprétation lorsqu'il interagit avec la machine.

1.1.1.1.2 L'outil et sa saisie dans un cadre énaactif

Les outils modifient nos manières de percevoir et d'agir ; lorsque l'on s'en saisit, on les intègre dans notre boucle sensori-motrice, et ils font entrer en jeu de nouvelles lois de contingences sensori-motrices. Les technologies cognitives ne sont pas que de simples moyens techniques nous permettant d'améliorer nos capacités de calcul, de mémorisation ou nos capacités perceptives ; elles jouent en fait un rôle déterminant dans la constitution de notre expérience (Lenay, Canu et Villon, 1997). L'hypothèse de la cognition étendue (Clark et Chalmers, 1998 ; Clark, 2008) pose le fait que la cognition n'est non seulement pas « dans la tête », mais qu'elle dépend d'éléments environnementaux et d'éléments liés au corps dans son ensemble. La cognition s'étend donc aux autres individus, aux processus de l'environnement et à ceux impliqués dans l'utilisation d'instruments techniques qui supportent notre activité cognitive.

Nous avons vu que l'action joue un rôle déterminant dans la construction de la perception. Nos dispositifs et environnements techniques transforment nos possibilités d'action, et

⁷ Traduction personnelle : « [Les organismes cognitifs naturels] participent activement à la génération de ce qui fait sens pour eux ; ils énaquent leur monde. »

transforment donc notre expérience vécue en offrant de nouvelles capacités de perception, d'imagination, de mémorisation ou encore de raisonnement (Lenay, 2006a).

La technique rend en fait possible les formes les plus générales de l'activité cognitive humaine ; « les outils, les interfaces, les instruments, les organisations matérielles, les technologies et systèmes d'information et de communication que nous concevons, développons et utilisons peuvent affecter nos façons de percevoir, de mémoriser, de raisonner, de définir des valeurs, des appartenances, des désirs, et des identités, mais aussi nos modes de rencontre, nos modalités d'interaction et nos manières d'être et d'agir ensemble. » (Steiner 2010, p.7) La thèse « TAC » pose ainsi la « Technique comme Anthropologiquement Constitutive/Constituante ».

Les outils nous permettent notamment l'adaptation de et à l'environnement. Pour améliorer nos aptitudes, nous disposons en effet selon Kirsh (1996) d'au moins trois stratégies : s'adapter à l'environnement, migrer vers d'autres environnements, ou adapter l'environnement lui-même. Le recours à des outils nous permet de prendre possession efficacement de notre environnement, en améliorant nos capacités et en nous permettant d'adapter l'environnement à nous.

Une interface donnant accès à un contenu dématérialisé constitue elle aussi un dispositif de couplage sensori-moteur, mais il faut dans ce cas distinguer deux espaces perceptifs différents (Auvray et Fuchs, 2007) :

- d'une part un espace perceptif dit *organique*, qui est l'espace où nous percevons sans l'outil ; on se situe dans l'environnement « réel » ;
- un espace perceptif *ouvert par l'outil, distal* ; il émerge du couplage des actions et des sensations à travers l'interface. L'outil sert alors à percevoir. On se situe dans un environnement virtuel, auquel l'interface nous donne accès.

Lorsque l'interface est perçue par le sujet dans l'espace dit organique, c'est alors un objet constitué, présent dans l'environnement. Le sujet peut agir *sur* cette interface, la changer, l'adapter, la manipuler,... Lorsqu'il s'en saisit et s'en sert, il accède alors à un nouvel espace perceptif ; l'interface transforme ses actions et ses sensations, ouvrant un nouveau monde, et le sujet se concentre alors sur ce qu'il peut faire *par le biais* de l'interface (voir figure 3). Elle lui ouvre un monde de possibles auquel il n'avait pas accès auparavant. En cela elle est donc constituante d'une expérience perceptive. Elle devient alors transparente, invisible. L'interface possède finalement deux modes d'existence : tenue/saisie, ou bien déposée/lâchée (Lenay et al. 2007) (voir figure 4). Par exemple, lorsque l'on porte des lunettes adaptées à sa vue, on les oublie et elles deviennent invisibles, s'intégrant à notre corps propre (c'est-à-dire le corps du point de vue du sujet) ; elles sont alors saisies. Par contre, si la correction n'est pas adaptée, ou que les verres sont sales, la conscience de porter les lunettes se fait très présente et dérangeante, les lunettes perdent leur caractère transparent ; elles sont alors déposées, même si on les porte

encore sur le nez. De même, lorsque l'on est à vélo, on oublie la sensation des vibrations des poignées et de la selle, et on a davantage l'impression de percevoir directement sous les roues les irrégularités ou les graviers du revêtement de la route. Le vélo devient une extension de notre corps propre, et c'est au travers de ce dispositif intégré que la perception de l'état de la route émerge. L'état déposé du dispositif n'a pas le caractère transparent de l'objet saisi. L'objet tenu perd d'ailleurs ce caractère transparent en cas de dysfonctionnement ou de rupture de l'expérience perceptive, comme nous l'avons vu dans le cas des lunettes sales, mais on peut aussi par exemple penser à la déconnexion d'une souris lors de l'utilisation d'un ordinateur ; voyant que le curseur ne bouge plus, l'utilisateur aura tendance à produire de grands gestes avec la souris, voir de la soulever et de la frapper sur le tapis, avant de se concentrer sur un éventuel problème de connectique. L'objet souris apparaît alors, tandis que lorsqu'il fonctionnait, il était oublié. Lorsqu'il est tenu, un outil ou une interface est intériorisé(e) ; il fait alors pratiquement partie du corps propre, comme le font les organes biologiques de l'action et de la perception. Mais, contrairement aux organes biologiques, un dispositif peut être extériorisé, détaché de notre corps propre (Leroi-Gourhan, 1943). Le dispositif, outil ou interface, n'est plus alors constituant d'une expérience perceptive, mais un simple objet perçu par le sujet dans son environnement.

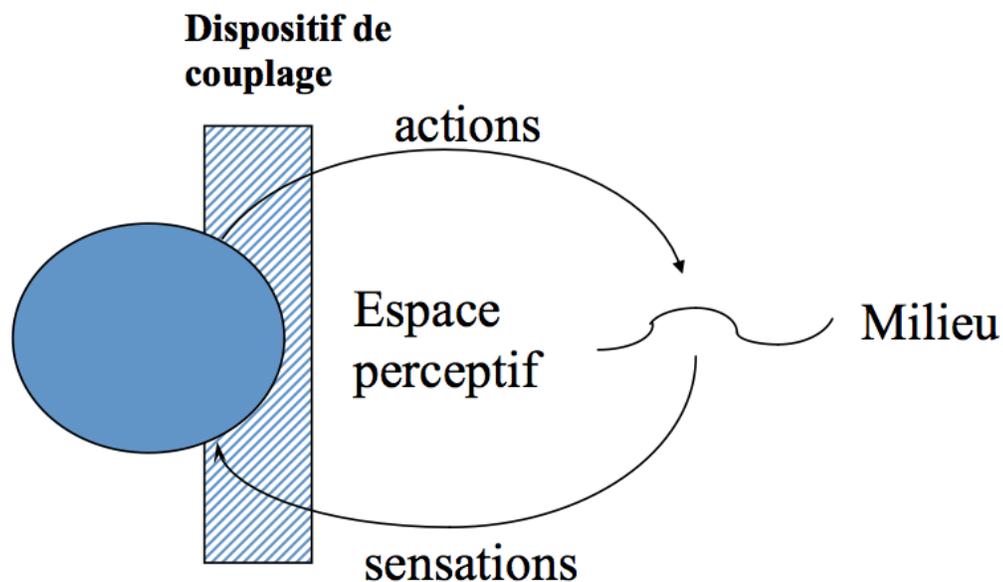


Figure 3. L'intégration de l'outil, qui ouvre de nouvelles possibilités d'agir et de sentir, dans la boucle sensori-motrice.

Saisi, l'objet technique joue un rôle constituant pour nos capacités d'action, de raisonnement ou encore de perception. L'intégration sensori-motrice de l'outil se fait par

l'apprentissage et la maîtrise des relations entre les actions faites avec cet outil et les retours sensoriels dus à cette action (soit la loi de contingence sensori-motrice). L'appropriation sensori-motrice de l'outil va en fait de pair avec une *incorporation* de l'outil dans le corps sentant et agissant : l'usage régulier de l'outil fait disparaître l'impression que le dispositif est une interface tangible entre le sujet et l'environnement. L'outil – lorsqu'il est bien utilisé et bien conçu – devient donc transparent ; sa préhension – s'il y en a une – est *ressentie* comme extension du corps sentant et agissant (Lenay et al. 2007). Ainsi, l'aveugle perçoit le monde avec le bout de sa canne, et non les vibrations de sa canne dans sa main, et il en va de même pour un conducteur qui perçoit directement l'état de la route sur laquelle il se déplace avec les roues de son véhicule (Merleau-Ponty, 1945).

Les outils sont donc pris en permanence dans un cycle de saisie et de lâcher. Le sens que l'on donne à un dispositif est lié à l'expérience perceptive qu'il permet ; ce sens est donc amené à évoluer, à mesure que l'individu découvre les fonctionnalités et les couplages sensori-moteurs permis par le dispositif. L'usage se stabilise lorsque l'individu associe au dispositif une signification stabilisée, et que le monde de possibles que l'individu énonce au travers du dispositif est lui aussi stabilisé.

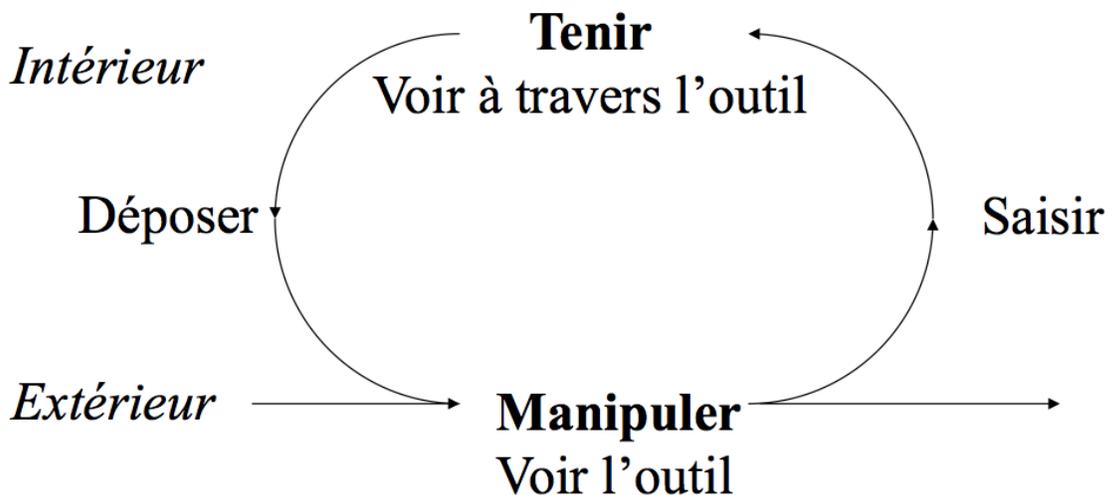


Figure 4. Les deux modes d'existence de l'outil : saisi et déposé (Lenay et al, 2007).

Cela s'applique aussi dans l'espace distal. Par exemple, en jouant à un jeu vidéo, l'utilisateur peut avoir l'impression de réellement se trouver dans l'environnement du jeu, de ressentir les effets de la collision avec des obstacles, ou de la vitesse de déplacement de son avatar. On désigne cet effet par le terme d'*immersion* ; ce terme désigne l'expérience vécue par

l'utilisateur d'un dispositif technique lui ouvrant un nouvel espace perceptif. Auvray et Fuchs (2007) ont mené des travaux sur l'immersion et sur les conditions qui la rendent possibles, en s'appuyant notamment sur l'appropriation de dispositifs de suppléance perceptive (c'est-à-dire fondés sur la conversion des stimuli propres à une modalité sensorielle en des stimuli propres à un autre sens, voir partie 1.1.2). Auvray (2004) pose notamment l'immersion comme étant le résultat d'un processus de saisie de l'outil, divisé en cinq étapes :

- le contact. L'utilisateur manipule l'outil, il agit et ressent des sensations en retour. Il découvre et apprend peu à peu les régularités sensori-motrices nécessaires pour la stabilisation du contact perceptif avec le stimulus. L'utilisateur apprend à faire correspondre une certaine action avec une certaine sensation.

- l'attribution distale. L'existence d'un espace distal est dévoilée à l'utilisateur à mesure qu'il arrive à identifier les relations d'invariance entre ses mouvements et les stimulations sensorielles qui en résultent. Le sujet réalise que l'objet à la source des stimuli qu'il ressent se trouve dans cet espace distal ; ses sensations sont dues à la rencontre avec un objet se trouvant dans l'espace perceptif ouvert par l'outil.

- la maîtrise de l'espace distal. L'utilisateur apprend à faire la relation entre ses actions dans l'espace organique et dans l'espace distal, et à faire la relation entre la réponse du système dans l'espace distal et ses sensations dans l'espace organique. Maîtriser ces relations permet de contrôler les nouvelles régularités sensori-motrices introduites par l'outil. L'anticipation des événements et la spatialisation dans l'espace distal deviennent possibles.

- la localisation distale (voir figure 5). Le sujet maîtrise totalement les lois de contingences sensori-motrices définies par l'outil. Il se saisit de l'interface, qui disparaît et devient transparente pour lui ; l'espace organique est en quelque sorte oublié, et l'utilisateur a le sentiment d'être dans l'espace perceptif ouvert par la médiation technique, dans l'espace distal. L'utilisateur est donc présent là où il agit. (Cette notion s'appuie sur la phénoménologie de la perception de Merleau-Ponty (1945), dont nous parlerons dans la section 2.1.2.4.2.).

- la constitution d'une expérience distale. Cette dernière étape implique notamment le partage de l'expérience perceptive. En effet, c'est la dimension sociale de l'expérience qui permet d'en constituer le sens, l'émotion et les valeurs.

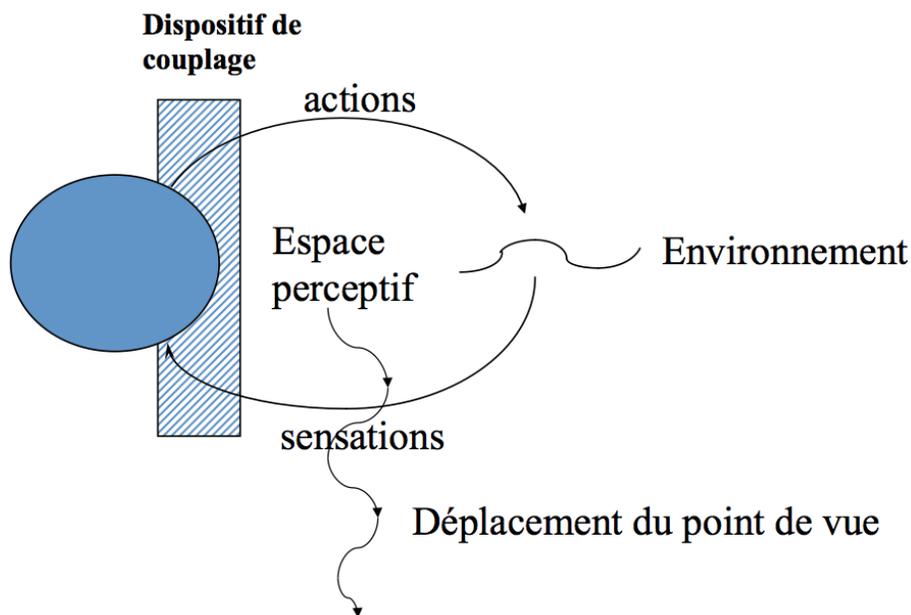


Figure 5. L'intégration d'un dispositif dans la boucle sensori-motrice permet un déplacement du point de vue.

L'immersion ne peut donc se faire que lorsque l'interface est saisie et appropriée. Les affordances jouent un rôle déterminant dans le processus de saisie d'une interface. Cette notion d'affordances a tout d'abord été proposée par Gibson (1986) ; elle désigne les manipulations permises par un objet à un sujet, en fonction des capacités sensori-motrices de ce dernier. Les affordances sont toujours relationnelles, puisqu'elles font entrer en jeu à la fois un objet perçu et un acteur percevant. Deux sujets ayant des capacités perceptives différentes n'auront pas les mêmes affordances vis-à-vis d'un même objet. Par exemple, une surface plate située à hauteur des genoux d'un individu lui permet de s'asseoir dessus ; l'individu perçoit cette surface comme étant « asseyable ». Les affordances de la dite surface dépendent de ses propriétés particulières (hauteur, largeur, solidité,...) en rapport avec celles de l'individu (taille, poids,...). Mais si un objet peut faire un siège pour un individu, il pourra éventuellement constituer une table pour un autre (par exemple un enfant). Pour un éléphant au contraire, il deviendra alors par exemple « saisissable par la trompe », ou encore « écrasable sous la patte ». Les affordances sont donc déterminantes dans la phase de contact, lorsque s'initie le processus de saisie d'un dispositif.

Norman (1988) propose de fonder théoriquement le design sur le concept d'affordances, qu'il abandonnera plus tard (Norman, 2008) au profit du terme *signifiers* (que l'on peut traduire par signifiants) afin d'éviter toute confusion entre l'affordance de Gibson (qui fait référence à des relations, toujours existantes, entre un sujet percevant et son monde) et l'affordance telle qu'il la voit (qui fait référence à des propriétés des objets *telles qu'elles sont perçues par les*

sujets). Autrement dit, les signifiants sont les perceptions de ce que l'on peut ou non faire avec un objet, et ce avant de l'avoir saisi. Par exemple, si l'on est confronté à un trou dans un mur, on saura si l'on peut ou non passer par ce trou sans avoir à véritablement s'y glisser. Pour Norman (1988), il est nécessaire lors de la conception d'un système de pouvoir donner à percevoir les indices nécessaires aux actions que son fonctionnement requiert, mais aussi les conséquences que ces actions ont sur l'état du système. Tout l'enjeu du design d'interaction est de concevoir des éléments « affordants », afin d'optimiser ainsi le processus de saisie en minimisant la quantité d'effort « cognitif » (de planification) nécessaires à l'utilisateur pour se servir du système. Les outils et interfaces doivent être conçus pour répondre à une double exigence : ils doivent d'un côté se faire oublier en mode saisi, pour permettre à l'utilisateur d'agir à travers eux, mais doivent en même temps offrir une possibilité d'être lâché pour pouvoir faire sens. La question du sens découle du fait que l'interface est déposée avant d'être saisie. Il faut donc concevoir des opportunités pour l'action, mais aussi rendre ces opportunités perceptibles.

1.1.1.1.3 Le croisement perceptif

Le croisement perceptif s'effectue entre deux individus ; il consiste en la rencontre de leurs deux activités perceptives, rencontre construisant du sens. Par exemple, lors d'un croisement perceptif par le regard avec quelqu'un d'autre, les deux regards s'accrochent : je vois autrui, qui me voit, et je vois qu'il me voit, il voit que je le vois, et je vois qu'il voit que je le vois, il voit que je vois qu'il me voit, et ainsi de suite... Le croisement perceptif en général est donc une situation dans laquelle deux activités perceptives de même nature interagissent (croisement de regards, caresse). Notons que les croisements perceptifs peuvent prendre des formes très variées suivant les médiations techniques des interactions qui les supportent : téléphone, internet, suppléance perceptive, réalité virtuelle,...

Auvray, Lenay et Stewart (2009) ont mis en place une expérience concernant les croisements perceptifs dans un environnement minimaliste, où deux sujets aveuglés situés dans des pièces différentes interagissent de manière tactile dans un environnement numérique extrêmement simplifié. Il s'agit en effet d'un espace numérique en une dimension, où les sujets déplacent un curseur vers la droite et la gauche à l'aide d'une souris. Un recollement des extrémités de l'espace d'interaction permet d'éviter toute singularité perceptive lorsque le sujet en atteint les limites (l'espace revient finalement à une ligne fermée, c'est un tore géométrique). Le curseur déplacé par chaque sujet est constitué d'une part d'un corps percevant, qui permet aux utilisateurs de percevoir des éléments dans l'espace (lorsque le corps percevant croise un élément,

l'utilisateur reçoit des stimulations tactiles par le biais de cellules brailles), et d'autre part d'un corps-image, qui permet d'être perçu par l'autre sujet.

Le curseur permet donc de « sentir » des objets dans cet espace : lorsque le curseur rencontre un objet, cela provoque pour l'utilisateur un retour tactile en tout ou rien, délivré par deux cellules brailles. Il existe trois types d'objets « touchables » dans le monde, et ils délivrent tous la même stimulation tactile. L'utilisateur peut donc « toucher » soit un objet fixe, soit un objet mobile, soit le curseur de l'autre sujet. L'objet fixe est dit « privé » : il ne peut être perçu que par un seul participant (l'autre participant dispose aussi d'un objet fixe privé, qui est situé à un autre endroit dans l'espace unidimensionnel). Quant à l'objet mobile, c'est en fait un leurre séparé d'une distance constante (de 50 pixels) du curseur de l'autre participant. L'espace mesure 600 pixels, et les corps-images ainsi que les objets mesurent 4 pixels. L'objet mobile se déplace exactement comme le curseur auquel il est attaché (voir figure 6) ; il a donc un comportement aussi complexe que le champ récepteur du partenaire.

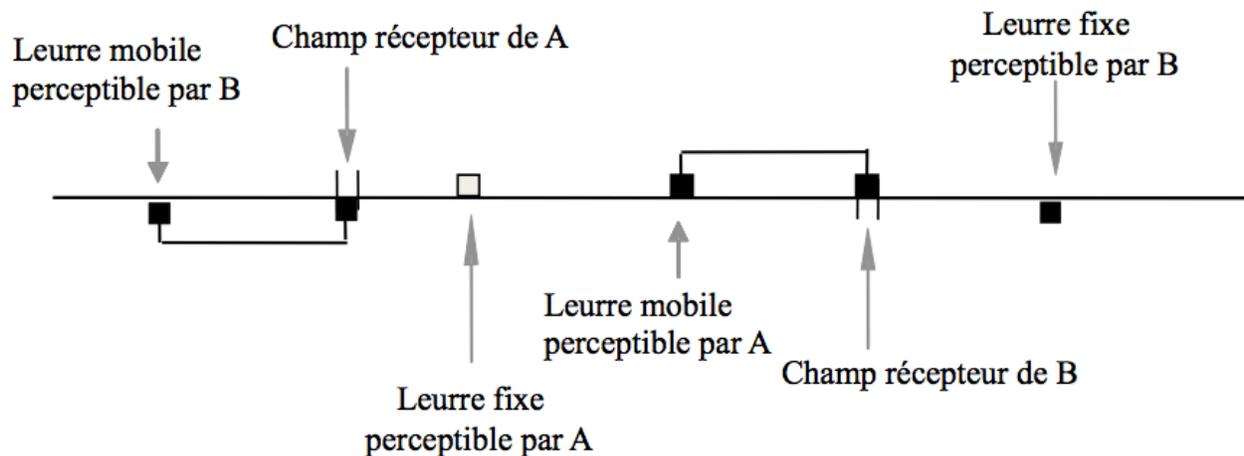


Figure 6. Schéma du dispositif expérimental mis en jeu dans l'expérience de Auvray et al (2009).

Les sujets interagissent par le biais du dispositif pendant trois sessions de cinq minutes chacune, et ils ont pour instruction de cliquer avec leur souris lorsqu'ils pensent que la stimulation tactile qu'ils ressentent est due à la rencontre avec l'autre participant.

Une première analyse des résultats semble montrer que les sujets réussissent la tâche : la fréquence des clics est très élevée lorsque la distance entre les deux participants est faible (voir figure 7). On voit que 62% des clics sont effectués alors que la distance entre les deux sujets est de plus ou moins 30 pixels. La distribution de la fréquence des stimulations reçues par les sujets

montre que 28% des stimulations sont reçues alors que la distance entre les deux sujets est de plus ou moins 30 pixels. On voit clairement apparaître un pic important lorsque la distance entre les deux sujets est nulle, ainsi qu'un pic mineur (indiqué par une flèche sur la figure) lorsque la distance est de 50 pixels, ce qui correspond précisément à la distance entre le curseur de chaque participant et le leurre attaché. On peut donc dans un premier temps supposer que les sujets parviennent à reconnaître le mouvement du curseur de l'autre participant comme étant un mouvement intentionnel, étant donné qu'ils cliquent bien sur le curseur du partenaire. Cependant, un deuxième temps dans l'analyse montre que le pourcentage de stimulations reçues en fonction du type d'objet rencontré dans l'espace est en fait plus important pour le curseur de l'autre participant (52,2%) que pour l'objet fixe (32,7%) et pour l'objet mobile (15,2%) (voir tableau 1). Par conséquent, même si les sujets cliquent effectivement plus souvent sur le curseur de l'autre participant (65,9% des clics) que sur le leurre mobile (23% des clics), il n'y a pas de différence significative de proportion clic/stimulation entre le curseur de l'autre sujet et l'objet mobile (respectivement 1,26 contre 1,51).

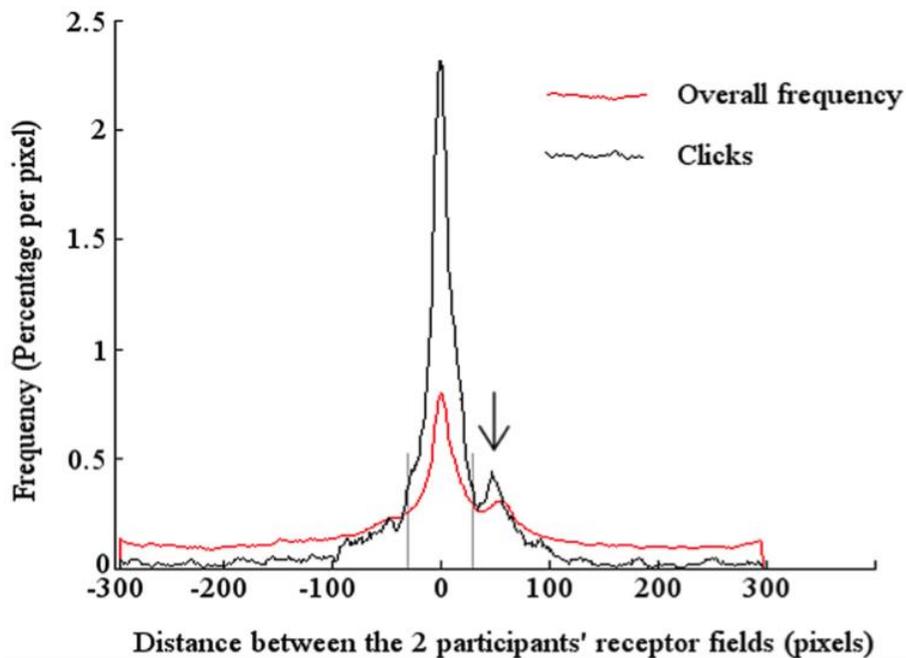


Figure 7. Répartition de la fréquence des clics (courbe noire) et des stimulations (courbe rouge) en fonction de la distance entre les utilisateurs dans l'expérience de Auvray et al (2009).

Tableau 1. Distribution des clics et des stimulations tactiles dans l'expérience de Auvray et al (2009).

	Champ récepteur		Leurre mobile		Leurre fixe	
Pourcentage de clics	65.9 %	± 3.9	23.0 %	± 10.4	11.0 %	± 8.9
Pourcentage de stimuli	52.2 %	± 15.2	15.2 %	± 6.2	32.7 %	± 11.8
Ratio clics / stimuli	1.26		1.51		0.33	

Ces derniers résultats peuvent sembler peu encourageants, mais un troisième temps d'analyse permet de mettre à jour un phénomène très intéressant. L'analyse des trajectoires perceptives des participants montre en effet que la rencontre des deux activités perceptives, lorsqu'elles sont intentionnellement dirigées l'une vers l'autre, fait émerger un attracteur dynamique collectif (Auvray, Lenay et Stewart, 2009 ; Lenay et Stewart, 2012). Il en découle que le pourcentage de stimulations est plus élevé lorsque les curseurs des participants se rencontrent, justement parce que l'activité perceptive de chaque participant attire l'activité de l'autre (Auvray, Lenay et Stewart, 2009 ; Lenay, 2010). Les deux sujets sont à la recherche des sources de stimulation ; lorsqu'une stimulation tactile survient, ils reviennent sur leurs pas pour essayer de rencontrer à nouveau la stimulation. Lorsque la stimulation est liée à la rencontre du curseur de l'autre participant, les sujets sont donc entraînés dans une sorte de « danse » (ibid) : chaque sujet inversant son mouvement, les sujets ne peuvent se déterminer spatialement car ils ne se croisent jamais vraiment au même endroit. Au contraire, lors de la rencontre avec un objet fixe, l'apparition des stimulations est parfaitement contingente avec le mouvement de déplacement de l'utilisateur : l'objet mobile se trouve toujours au même endroit. Le leurre mobile ne peut pas être déterminé, puisqu'il reflète l'activité du curseur auquel il est attaché : la rencontre de l'objet mobile n'est pas contingente avec l'activité de celui qui le perçoit, puisque le mouvement de l'objet mobile ne dépend pas des objets qu'il rencontre. La rencontre des deux curseurs constitue donc un événement particulier, faisant émerger un jeu subtil où chaque sujet essaie de reconnaître l'autre, en tentant d'anticiper la stimulation sans pouvoir la localiser avec précision (ibid). Chacun cherche l'autre et à rencontrer l'autre, sans savoir précisément où il se

trouve. La reconnaissance d'autrui ne se fait pas parce qu'il donne à percevoir un comportement plus ou moins contingent, mais parce qu'il y a engagement des deux sujets dans une activité collective. C'est cette activité qui fait émerger des contingences imparfaites, dont aucun des deux sujets ne contrôle la venue, mais dans lesquelles ils sont tous deux engagés. Les sujets interagissant, même dans un environnement minimaliste, peuvent donc reconnaître quand ils rencontrent leur partenaire bien qu'il n'y ait absolument aucune différence de comportement objectif entre son champ récepteur et le leur mobile.

La différence de fréquences de clics peut donc être attribuée à l'émergence d'un couplage interpersonnel plus ou moins stable à partir de la rencontre des stratégies individuelles ; le processus de reconnaissance d'autrui est donc intrinsèque à l'activité perceptive partagée et à la dynamique mutuelle qui en découle (Auvray, Lenay et Stewart, 2009). L'interaction ne joue donc pas seulement un rôle contextuel ou déclenchant, elle *constitue* la cognition sociale (De Jaegher, Di Paolo et Gallagher, 2010).

L'analyse détaillée des trajectoires perceptives enregistrées a permis de mettre en évidence les stratégies employées par les participants pour réussir la tâche (Lenay 2010). Deux critères doivent être satisfaits pour pouvoir attribuer une stimulation à la présence d'autrui :

- autrui est quelque chose qui se maintient présent ;
- autrui est quelque chose qui résiste à sa détermination spatiale : ce n'est ni un objet fixe, ni un objet obéissant à un mouvement déterminable suivant une règle.

La dynamique d'interaction conduit les sujets à stabiliser la situation de croisement perceptif, chacun oscillant autour de l'autre. Mais, du fait même de cette dynamique, aucun des deux sujets ne peut spatialiser l'autre, puisqu'il est continuellement en mouvement. Autrui ne peut donc être constitué spatialement (ce n'est ni un objet fixe, ni un objet obéissant à un mouvement déterminable suivant une règle), même si sa présence s'impose à ma perception (ibid). Autrui est finalement reconnu comme ce qui résiste à sa constitution précise, et qui pourtant persiste à être présent. Le critère que les sujets semblent adopter pour cliquer, à savoir une stimulation rencontrée fortement présente sans être spatialement déterminée, n'est pas arbitraire mais découle logiquement de la rencontre entre deux intentionnalités qui caractérise le croisement perceptif. Il y a bien quelque chose dans l'interaction qui échappe aux sujets, et seule la dynamique de la rencontre interpersonnelle vient stabiliser l'attracteur de la dynamique collective, alors que l'individu ignore précisément tout de cette dynamique (ibid.). Une nouvelle signification semble donc émerger dans l'interaction, alors qu'elle faisait défaut à l'individu isolé (De Jaegher et Di Paolo, 2007) ; cela correspond à une perspective éactive de l'interaction sociale, à un *participatory sense-making*.

Lenay et Stewart ont mené une nouvelle expérience pour asseoir les résultats précédents (Lenay et Stewart, 2012). Le protocole expérimental est identique, à la différence qu'il n'y a plus

de stimulations tactiles et que c'est un son qui est associé à chacun des trois objets. Il y a donc trois sons différents, chacun respectivement lié à un type d'objet ; les stimulations ne sont pas identiques pour la rencontre des objets, contrairement à l'expérience précédente où la stimulation tactile n'était pas différenciée selon le type d'objet. Après chaque essai de deux minutes, les sujets ont pour tâche d'associer chacun des trois sons avec les objets correspondants. Les résultats font état d'un score élevé de catégorisation correcte, et ce pour les trois types d'objets (ibid). Chaque sujet est capable de distinguer la dynamique du croisement perceptif de la dynamique de la rencontre avec les autres objets, notamment l'objet mobile, étant donné qu'il dispose d'un moyen de différencier les stimulations liées à chaque type d'objets. Le fait que les sons attribués à chaque objet soient distincts permet de lever l'ambiguïté des différentes situations d'interaction, et c'est alors que le croisement perceptif avec le partenaire est bel et bien reconnu comme tel. La dynamique du croisement perceptif n'a pas été modifiée, mais les participants peuvent cette fois l'identifier car l'association des propriétés intrinsèques des objets et des propriétés propres à la dynamique d'interaction peut se faire.

Dans tous les cas, on observe que la situation de croisement perceptif est ressentie comme plus riche et intéressante que la perception d'un objet passif. Même avec une pauvreté sensorielle maximale, une valeur émotionnelle peut être attachée aux perceptions s'il y a la possibilité d'une synchronisation des activités perceptives. Le croisement perceptif est donc une piste de choix pour la communication émotionnelle à distance.

1.1.1.2 Les interactions : supports de la cognition sociale

Une définition générale de l'interaction sociale serait la relation interpersonnelle existant entre deux individus ou plus. Les interactions sociales correspondent aux actions réciproques des individus entre eux, actions qui peuvent être verbales et/ou non verbales. Cette relation présuppose que les individus soient en présence, ainsi qu'une influence réciproque de leurs actions ; ils sont ensemble, et agissent l'un sur l'autre. Cette influence réciproque est classiquement décrite de façon séquentielle. Dans le cas de la communication, il s'agit donc d'une simple transmission (le message est émis par A, transite de A à B, et est reçu par B). Le succès de l'interaction dépend donc dans cette vision des capacités et compétences individuelles des interactants, notamment à traiter et interpréter les informations en provenance de l'environnement social, afin de pouvoir comprendre l'autre participant et interagir avec lui. C'est cette aptitude qui est désignée par *cognition sociale* en psychologie. Elle désigne l'ensemble des processus cognitifs mis en œuvre dans les interactions sociales humaines.

Une sorte d'« individualisme méthodologique » prédomine malheureusement dans l'étude la cognition sociale (Lenay et Stewart, 2012) : les expériences mises en place cherchent à comprendre la compréhension interpersonnelle, mais elles ne permettent paradoxalement pas de réelles interactions entre les sujets. En effet, les sujets sont la plupart du temps passifs, ayant simplement à observer des situations sociales extérieures, puis sont encouragés à porter un jugement sur ces scènes auxquelles ils n'ont pas participé, notamment pour pouvoir observer ce qu'il se passe à un niveau neurologique (voir par exemple Van Overwalle, 2009). On constate donc que ces recherches sont concentrées sur les processus de cognition sociale à la troisième personne (Frith, 2008). L'étude des situations sociales consiste donc à se représenter les actions effectuées par autrui comme étant le produit d'états mentaux intentionnellement dirigés vers un but rationnel (Csibra et Gergely, 1998). Cette approche classique s'explique par le fait que la cognition y est traitée comme une aptitude strictement individuelle dépendant des compétences de chacun, formulée en termes mentalistes. En effet, il y est considéré que des états mentaux internes (que l'on peut mettre en parallèle avec les représentations du monde du computationnalisme) sont à l'origine du comportement, et que ces états mentaux ne sont pas observables ; le comportement d'autrui n'est alors explicable et prédictible que grâce à une « théorie de l'esprit » (Premack et Woodruff, 1978). Deux positions classiques s'opposent quant à la nature de cette théorie. Pour la *théorie de la théorie*, la théorie de l'esprit est un ensemble structuré de connaissances sur l'esprit, qui concerne notamment les intentions, les désirs et les croyances (Astington et Gopnik, 1991). Pour la *théorie de la simulation*, il s'agit en revanche pour le sujet de faire une analogie plutôt qu'une inférence : la théorie de l'esprit est alors une capacité à simuler les états mentaux d'autrui, c'est-à-dire à les reproduire ou à les imiter de façon interne, et ce afin de pouvoir les comprendre par le biais de ses propres états mentaux (Goldman, 1989 ; Gordon, 1986).

Dans les situations d'interaction sociale, l'individu perçoit les manifestations comportementales d'autrui (les expressions corporelles et faciales, comme les sourires, mais aussi les mouvements). Pour les conceptions classiques, l'individu interprète ces manifestations à la lumière d'une théorie de l'esprit préalablement constituée. La *théorie de l'interaction* proposée par Gallagher (2004) et basée sur des études phénoménologiques et développementales constitue une alternative à ces conceptions classiques. Selon cette théorie, nous percevons *directement* les intentions à travers les mouvements corporels, les mimiques faciales et la direction du regard d'autrui. Les informations sociales sont perçues *directement* à travers l'activité perceptive, dans l'immédiateté de l'interaction, et non après-coup via une opération sur des états internes (Neisser, 1988 ; Neisser, 1991 ; Gallagher, 2008).

Pour répondre à la prédominance d'une conception individualiste dans la recherche en cognition sociale, on peut effectuer un renversement épistémologique et adopter une vision

interactionniste : il ne va plus s'agir d'expliquer la cognition sociale par des opérations et représentations internes, mais de remettre au cœur des expériences la participation active des sujets à l'interaction et à la constitution de sens (Varela, Thompson et Rosch, 1991 ; De Jaegher et Di Paolo, 2007). La communication ne dépendra alors plus des capacités de chacun, mais émergera et se construira dans l'interaction. Dans une vision éactive, le monde n'est pas prédonné, et l'individu n'a pas de représentations internes du monde car il y est incarné et agit en son sein (Reddy et Morris, 2004 ; Froese, 2012). Le monde est vu comme le produit d'interactions dynamiques entre le sujet percevant et son environnement, ce qui inclut l'environnement social (Varela, Thompson et Rosch, 1991). C'est donc le système de couplage lui-même (qui recouvre les individus, les interactions dynamiques existant entre eux, ainsi que le contexte dans lequel ces interactions s'inscrivent), et non l'individu, qui doit être l'unité d'analyse de la cognition sociale. Cette conception de la cognition sociale permet de prendre en compte le corps et l'action, la dimension incarnée des individus, ainsi que l'engagement mutuel pour la compréhension interindividuelle (De Jaegher, 2006 ; Krueger, 2011). On ne se focalise donc plus sur l'individu et ses capacités, mais sur les aspects incarnés, engagés, participatifs, situés de la compréhension sociale. Les processus interactifs sont plus qu'un contexte pour la cognition sociale, ils en sont constitutifs : ils viennent compléter, et même remplacer, les mécanismes individuels (De Jaegher, Di Paolo et Gallagher, 2010).

L'action a donc un rôle aussi dans la cognition sociale. Elle en fait une négociation collective, et établit de plus une gestion interactive de l'espace commun (Krueger, 2011). Ce dernier est un espace pratique, ancré dans le corps et dans ses possibilités d'agir. Il est reconfiguré en espace social du fait des interactions interindividuelles, et il s'y joue alors des coordinations intersubjectives et émotionnelles, qui sont structurées de façon dynamique par l'engagement mutuel des interactants. Les actions expressives, comme les regards, les expressions corporelles, les expressions faciales, les vocalisations, les énoncés verbaux ou encore les gestes structurent activement l'espace de l'interaction ; elles sont donc des affordances sociales (ibid). La cognition sociale émerge de l'action spécifique qu'est l'interaction sociale (De Jaegher et Di Paolo, 2007 ; Fuchs et De Jaegher, 2009).

Nous adoptons donc une approche de la cognition sociale qui est à la fois éactive, incarnée et située. Le processus de l'interaction y est placé au cœur de la compréhension interpersonnelle ; il n'est pas posé comme facteur contextuel traité individuellement par des mécanismes internes, comme c'est classiquement le cas, mais comme étant *constitutif* de la cognition sociale.

1.1.1.3 De l'impact des émotions

1.1.1.3.1 La difficile question des émotions

1.1.1.3.1.1 Difficulté de définition

Le mot « émotion » vient du latin « emovere, emotum » (enlever, secouer) et de « movere » (se mouvoir). Le concept d'émotions semble aller de soi tant il est familier et courant dans la vie de tous les jours. Il existe d'ailleurs en ce qui concerne les émotions une distance certaine entre les théories scientifiques d'une part, et d'autre part les théories naïves, instinctives, auxquelles tout un chacun se réfère et avec lesquelles il vit, plus ou moins consciemment. Il nous faut aussi évoquer la complexité des émotions, complexité que l'on retrouve à plusieurs niveaux : l'interaction réciproque avec l'esprit, le langage, le comportement mais aussi la culture, ainsi que le nombre de caractéristiques physiologiques et neurologiques corrélées, sans oublier leur universalité ou leur relativité.

D'un point de vue scientifique, il est avant toute chose difficile de définir ce qu'est une émotion, ce qui pose bien évidemment problème pour en faire l'étude. Il est difficile de trouver une définition formelle et générique de l'émotion. Qu'est-ce qu'une émotion ? Comment et quand survient-elle ? Comment se développent les émotions ?

On peut considérer les émotions selon deux points de vue :

- un point de vue expressif et comportemental, où les émotions modifient le comportement physique et l'enveloppe corporelle de l'individu (gestes, traits du visage, portée de la voix,...) ;
- un point de vue cognitif, où les émotions ont une influence majeure sur certains mécanismes cognitifs (comme par exemple la prise de décision, l'attention ou encore l'interprétation).

Dans le cadre du modèle des processus et composantes des émotions de Scherer (1987), l'émotion est définie comme un épisode de changements synchronisés en interrelation survenant dans tous ou la plupart des cinq sous-systèmes de l'organisme en réponse à l'évaluation d'un stimulus externe ou interne perçu comme un événement pertinent vis à vis des intérêts majeurs de l'organisme (Scherer, 1987 ; Scherer, 2000). Les émotions sont des phénomènes complexes, apparaissant suite à l'évaluation subjective d'un événement, d'un objet ou d'une personne. Des changements de diverse nature se produisent chez l'individu : physiologiques, corporels, faciaux, vocaux,... Par exemple la respiration s'accélère, le cœur s'emballé, la voix devient forte, le visage rougit,... Le processus émotionnel est par ailleurs plus complexe que la seule expression physique des émotions, et il comprend notamment (Scherer, 2000) :

- des changements neurophysiologiques ;
- un sentiment subjectif, souvent conscient ;
- des expressions comportementales ;

- une tendance à certaines actions résultant de l'évaluation de l'événement qui a déclenché l'émotion (par exemple la fuite face à un danger) ;
- une activité cognitive permettant une évaluation des éléments à la source de l'émotion.

Dans le cadre de ce modèle, Scherer s'applique aussi à distinguer les émotions, ainsi définies, d'autres phénomènes affectifs, tels que les sentiments, les humeurs, ou les attitudes (Scherer, 2005)⁸. Il pose que les émotions, bien que difficiles à définir, peuvent être approchées par le biais de leurs caractéristiques et de leurs propriétés spécifiques (Scherer, 2000). C'est ainsi qu'il caractérise les émotions par :

- leur dépendance au temps (aucune tristesse ne dure une seconde, de même que nulle surprise ne dure un an) ;
- leur variabilité, à la fois sur le plan qualitatif, et sur le plan quantitatif (de nombreux niveaux d'intensité existent, par exemple pour aller de la peur à l'effroi, ou de la joie à l'extase) ;
- leurs propriétés physiologiques (par exemple l'accélération du rythme cardiaque, la sueur, le rougissement de la peau,...) ;
- leurs propriétés communicatives et adaptatives (l'émotion est un signal interne et externe) ;
- leurs propriétés rationnelles (les émotions possèdent toujours une cause et/ou un but).

Belzung (2007) procède à un historique des définitions ayant été utilisées pour définir les émotions. D'un point de vue biologique, la définition opératoire de Kleinginna et Kleinginna (1981) pose les émotions comme étant le « résultat de l'interaction de facteurs subjectifs et objectifs, réalisées par des systèmes neuronaux ou endocriniens, qui peuvent :

- induire des expériences telles que des sentiments d'éveil, de plaisir ou de déplaisir ;
- générer des processus cognitifs tels que des réorientations pertinentes sur le plan perceptif, des évaluations, des étiquetages ;
- activer des ajustements physiologiques globaux ;
- induire des comportements qui sont, le plus souvent, expressifs, dirigés vers un but et adaptatifs. » (ibid, p.355)

L'émotion est donc dans cette définition vue comme une réponse adaptative qui présente trois composantes :

- physiologique (rythme cardiaque, pression sanguine, libération d'adrénaline, fréquence respiratoire, pilo-érection, rougissement, contractions de l'estomac, température corporelle, dilatation des pupilles,...),
- comportementale/expressive (comportements de sursaut, d'immobilisation ou d'agression, expressions faciales ou vocales),

⁸ Nous ne prendrons pas de telles précautions, et utilisons « émotion » comme terme générique couvrant émotion, affect, éprouvé, humeur, sentiment, disposition, état d'âme,...

- cognitive/subjective (état interne / sentiment)

En psychologie, il existe plusieurs modèles des émotions, et plusieurs façons de classer les émotions. Le modèle le plus simple consiste à mettre en place une dichotomie selon leur valence : émotions positives (comme la joie) ou bien émotions négatives (comme la tristesse ou le dégoût, ou encore la colère). Le classement et la catégorisation sont des exercices anciens ; Aristote (384-324 av. J.-C.) déjà identifiait la colère, la pitié, la peur et le désir, qu'il associait au plaisir ou bien à la douleur. Pour Descartes (1649), il existe six passions primitives (l'admiration, la haine, l'amour, le désir, la joie et la tristesse) et trente-quatre passions « combinatoires » : on voit apparaître là la notion d'émotions primaires et d'émotions secondaires.

Plusieurs chercheurs ont exploité ce classement en émotions primaires et émotions secondaires. Les émotions secondaires sont vues comme des combinaisons ou des sous-espèces des émotions primaires. Pour Izard (1977), les émotions primaires sont les suivantes : joie, surprise, colère, peur, tristesse, mépris, détresse, intérêt, culpabilité, honte et amour. Plutchik (1980) a réduit cette liste à : acceptation, colère, anticipation, dégoût, joie, peur, tristesse et surprise. Kemper (1981) s'est contenté de la peur, de la colère, de la dépression et de la satisfaction. Ekman (Ekman, Friesen et Ellsworth, 1972 ; Ekman, 1992a ; Ekman, 1992b) a quant à lui identifié colère, peur, tristesse, joie, dégoût et surprise ; il stipule que ces six émotions primaires sont universelles.

Dans le modèle d'Oatley et Johnson-Laird (1987), l'émotion secondaire est formée à partir d'une émotion primaire dont elle garde les propriétés fondamentales, comme la valence positive ou négative (selon qu'elle procure du plaisir ou du déplaisir). S'ajoutent à ce noyau de base des représentations plus ou moins complexes.

Plutchik (1980) a développé un modèle multidimensionnel des émotions. Il y distingue trois aspects :

- l'aspect dimensionnel, les émotions reposant sur trois dimensions : l'intensité (peur-panique), la similitude (la honte est plus proche de la culpabilité que de la joie) et la polarité (la tristesse est opposée à la joie) ;
- la notion de persistance : il y a une différence entre les aspects aigus (souvent les seuls mesurés en situation de laboratoire) et les aspects chroniques (souvent l'objet de la plainte en clinique) ;
- la notion de pureté : les émotions ne sont pas toujours pures, et on a souvent affaire à un mélange.

À partir de là, Plutchik propose son modèle multidimensionnel (voir figure 8) sous forme d'une rosace (version à plat) et d'une demi-sphère (version trois dimensions).

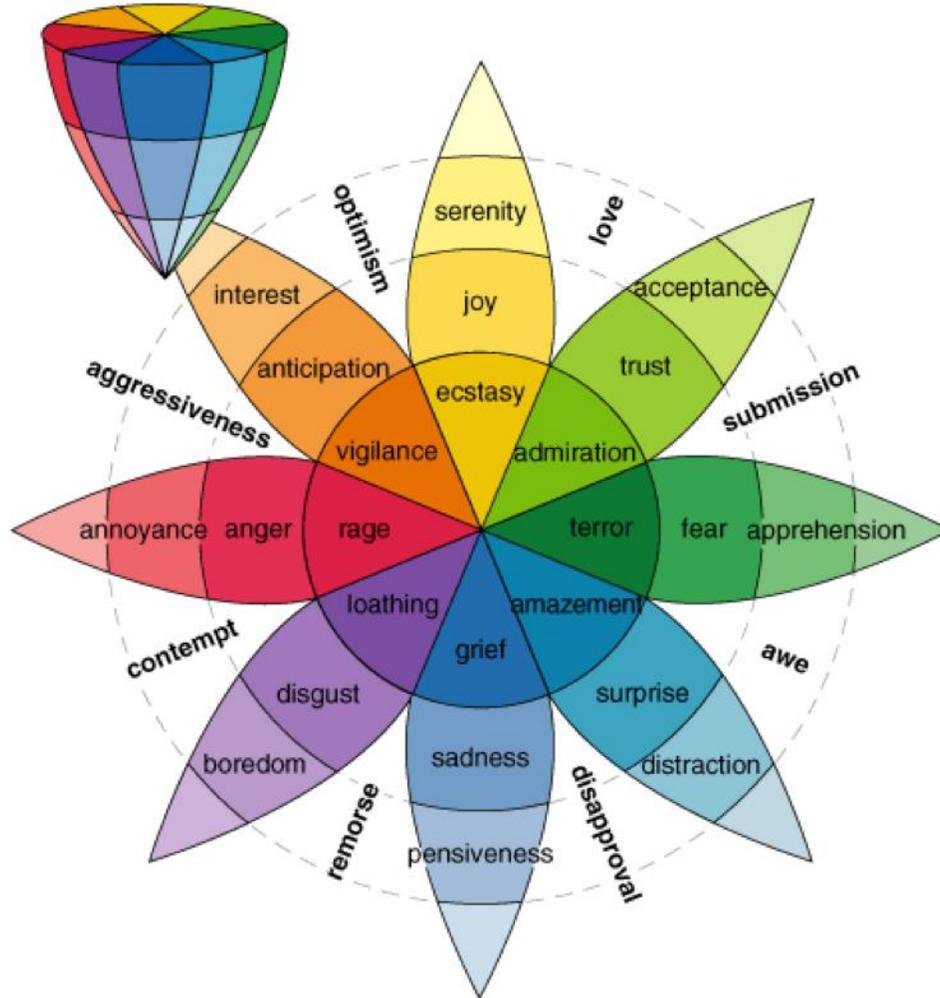


Figure 8. La roue des émotions de Plutchik.

L'aspect dimensionnel découle de l'intensité (anxiété vs angoisse ; peur vs panique), de la similitude (la culpabilité ressemble à la honte, alors qu'elle diffère fortement de la joie) et la polarité (la joie et la tristesse s'opposent). Les huit émotions primaires y apparaissant sont : la peur, la surprise, la tristesse, le dégoût, la colère, l'anticipation, la joie et l'acceptation. Les autres émotions les plus voisines d'une émotion donnée correspondent à celles qui sont situées sur le secteur adjacent (par exemple la peur et la tristesse pour la surprise), et les plus éloignées sont situées sur les secteurs opposés (par exemple la joie et la tristesse). Les sections inférieures et supérieures au niveau des émotions primaires correspondent à des intensités différentes d'une même émotion primaire : par exemple, l'appréhension (intensité faible), la peur (intensité moyenne) et la terreur (intensité forte) ou bien le souci (intensité faible), la tristesse (intensité

moyenne) et le chagrin (intensité forte). Les émotions secondaires résultent elles de la combinaison d'émotions primaires.

En plus des émotions primaires et secondaires, correspondant en général à des événements de durée brève, il existe des émotions dites « émotions d'arrière-plan », qui concernent des états de longue durée comme par exemple le bien-être, la dépression ou encore l'apathie, qui peuvent influencer l'expression des émotions primaires et secondaires.

Nous venons de voir que la question des émotions n'est pas une question simple, et que de multiples définitions et perspectives ont été adoptées pour les étudier. La difficulté suivante, qui en découle en partie, consiste à procéder à la mesure et à l'évaluation des émotions.

1.1.1.3.1.2 Difficulté de mesure

L'affective computing (ou informatique affective en français) est une discipline ayant pour objet l'étude et le développement de dispositifs ayant les capacités de reconnaître, d'exprimer, de synthétiser et de modéliser les émotions humaines. L'un de ses buts premiers est de développer des interfaces informatiques qui détectent automatiquement les émotions des utilisateurs, et y répondent de manière adéquate. Dans cette discipline, la prise en compte les émotions des utilisateurs est donc capitale ; les capacités des systèmes conçus doivent être efficaces dans les domaines de la perception, de l'interaction et de la génération (Pelachaud, 2011).

Ainsi que c'est le cas dans la majorité des sciences humaines, le chercheur dans le domaine des émotions influence sa propre recherche. Kappas (1991) défend cette idée en stipulant qu'il n'y a pas d'observation « neutre » ou « objective » des émotions, dans la mesure où une telle observation ne peut se fonder que sur l'expérience que fait l'observateur de l'état même qu'il veut observer.

La mesure de l'état émotionnel de l'utilisateur est donc critique et délicate ; les trois méthodes les plus utilisées dans ce domaine consistent en l'analyse des expressions faciales, l'analyse de la voix et l'analyse de l'expression corporelle.

Il existe de nombreux travaux concernant les expressions faciales émotionnelles ; elles étaient déjà un sujet d'étude pour Duchenne de Boulogne (1862) et pour Darwin (1872) : ils ont fourni des photographies détaillées de décompositions de mouvements faciaux lors d'expressions faciales émotionnelles. Les études les plus connues sont sans doute celles menées par Ekman (1999) lorsqu'il a mis en évidence ce qu'il pose comme les six émotions de base ; il s'est alors appuyé sur des photos de portraits d'expressions faciales.

L'expression émotionnelle vocale est aussi d'une grande richesse, et fait entrer en jeu de nombreux facteurs mesurables influencés par les émotions : la production vocale (inspiration, vibration des cordes vocales, action des résonateurs), les indices prosodiques (timbre de la voix, rythme, intensité), les indices linguistiques,... (Vaudable, 2012).

En ce qui concerne l'expression corporelle, DeVito (2008) définit les affichages affectifs (affect displays) comme étant des mouvements non seulement du visage, comme par exemple sourire ou froncer les sourcils, mais aussi des mains et du corps en général (tension du corps ou au contraire relâchement, par exemple) qui véhiculent un sens émotionnel. Ils sont souvent inconscients, mais peuvent parfois être effectués consciemment, pour essayer de transmettre par exemple son plaisir ou sa satisfaction.

Riva, Davide et IJsselsteijn (2003), ainsi que IJsselsteijn et al. (2000), examinent les catégories de méthodes utilisées afin de mesurer l'immersion d'un système d'environnement virtuel. Est tout d'abord identifiée la méthode des mesures subjectives : elle s'appuie sur une auto-évaluation par l'utilisateur. Ce dernier répond à des questions concernant le dispositif utilisé, que ce soit à posteriori ou pendant l'utilisation du système. L'un des questionnaires de présence les plus connus est celui développé par Witmer et Singer (1998), mais il en existe bien d'autres, parfois dédiés à la communication (comme le Affective Benefits and Costs in Communication Questionnaire (ABC-Q) développé par Van Baren et al. (2004)). Vient ensuite la méthode des mesures comportementales, qui prend en considération les actions ou attitudes déployées par l'utilisateur qui sont des réponses à des objets ou des événements dans l'environnement virtuel. Enfin, les méthodes physiologiques tentent d'évaluer le caractère immersif d'un système en mesurant les changements des réponses physiologiques du sujet (rythme cardiaque, température de la peau, conductivité de la peau,...)

1.1.1.3.2 Émotions et raison

Les émotions souffrent d'une mauvaise réputation par rapport à la raison, et ce depuis Platon (427-358 av. J.-C.). Il est le premier penseur dont on ait conservé des écrits qui offre une analyse des processus émotionnels, mais sa vision des émotions est en effet négative puisque, selon lui, elles ne font que pervertir la raison et l'empêchent de se développer de manière fructueuse. Ainsi, très tôt rationalité et émotions se sont vues opposées, comme si les secondes empêchaient la première.

Dans L'erreur de Descartes, Antonio Damasio (2006) combat cette idée ; être rationnel, ce n'est en effet pas se couper de ses émotions. Le cerveau qui pense, qui calcule, qui décide n'est pas un autre que celui qui rit, qui pleure, qui aime, qui éprouve du plaisir et du déplaisir.

Damasio s'oppose au dualisme cartésien du corps et de l'esprit, et à tous ceux qui voudraient réduire le fonctionnement de l'esprit humain à de froids calculs dignes d'un superordinateur, et s'appuie pour cela sur les acquis récents de la neurologie, qui démontrent que l'absence d'émotions et de sentiments empêche d'être vraiment rationnel. Il existe bien une relation entre l'émotion et la raison, et ce n'est pas une relation d'opposition. Damasio fait l'hypothèse « des marqueurs somatiques » selon laquelle l'émotion participe à la raison et qu'elle peut assister le processus du raisonnement au lieu de nécessairement le déranger, comme on le supposait couramment. Il se base pour cela sur l'étude neurologique de patients souffrant à la fois de défauts de prise de décision et de troubles de l'émotion.

Bien avant Damasio, Harlow (1868) avait déjà décrit le cas de Phineas Gage. Cet homme avait miraculeusement survécu à un accident survenu en 1848 sur un chantier : suite à une explosion imprévue, une barre de fer lui perfora le crâne de part en part. Mais cela laissa d'importantes séquelles : outre la perte de l'usage de son œil gauche, il voit sa personnalité se transformer du tout au tout. Harlow a suivi Phineas Gage, et a pu assister à son déclin. Avant son accident, Gage jouissait en effet d'une très bonne réputation, mais sa personnalité s'est ensuite transformée de manière drastiquement négative. Cette transformation, et la chute sociale qui l'a accompagnée, sont en partie liées au fait que Gage avait perdu sa capacité à faire usage de ses émotions avant de prendre des décisions. Il ne pouvait par exemple plus faire preuve d'empathie, et n'avait plus aucune honte à être vulgaire. Détecter les émotions d'autrui et y réagir de façon adéquate permet de survivre, et si l'on n'en est pas capable, nos capacités de raisonnement vont s'avérer déficientes.

Cette idée que les émotions jouent un rôle puissant et central dans nos vies, et sur notre cognition, est aujourd'hui de plus en plus répandue (Colombetti, 2008). Les émotions impactent en effet nos croyances, nous permettent de prendre des décisions et guident en large partie la façon dont nous adaptons notre comportement à notre environnement (Gratch et Marsella, 2004 ; Pacherie, 2002). Elles jouent un rôle d'équilibrage de la rationalité dans le comportement et la prise de décision, mais elles ont de plus un rôle fonctionnel dans le comportement en interaction sociale (Bisognin et Pesty, 2004).

1.1.1.3.3 Les émotions en interaction

Les émotions sont classiquement, et ainsi que nous avons pu le voir pour la question de leur définition, traitées de façon internaliste et séquentielle (que ce soit en psychologie ou en neuroscience). Dans le cadre de notre travail, nous nous intéressons aux émotions dans le cadre d'une interaction (et plus précisément d'une interaction à distance). Pour étudier les émotions de

manière pertinente, surtout en situation d'interaction, il convient d'adopter une posture pluridisciplinaire. L'approche linguistique est notamment très importante pour appréhender la question des émotions, ainsi que la considération de la cognition, ou encore l'impact de la culture et de l'environnement. La communication linguistique est en effet un processus interactionnel ; comme le note Bottineau : « Language is not an object, it cannot be acquired. The forming of personal languaging is part and parcel of the forming of the person, just as walking, jumping or flying. In the same way, but at a different scale, the forming of a community's social languaging is part and parcel of the cultural forming of the tribe. » (Bottineau 2010, p.295).

Dans le cadre d'une interaction, la pertinence du contexte dans lequel ont lieu les émotions est une question majeure. Par exemple, Eggs (in Plantin, Doury et Traverso, 2000) met l'accent sur la situation qui provoque un sentiment d'humiliation. La théorie sous-jacente est qu'il y a des situations spécifiques qui causent des émotions déterminées. Eggs relie aussi la « normativité » des émotions au contexte : nous avons des idées normées quant à la manifestation juste et adéquate d'une émotion étant donné un certain contexte (ibid).

Kerbrat-Orecchioni (in Plantin, Doury et Traverso, 2000) étudie le statut des émotions en linguistique au XX^{ème} siècle. Il existe un débat autour de la question des valeurs émotionnelles portées par les déclarations : viennent-elles des caractéristiques internes des énoncés, ou émergent-elles suite à des facteurs plus externes pertinents pour l'énonciation, en l'occurrence la situation, ainsi que la prosodie et les gestes qui lui sont corrélés ? Galati et Sini prennent en compte différents contextes linguistiques et culturels, à savoir des langues Anglo-Saxonnes d'un côté, et les langues européennes néolatines de l'autre. Les groupements de mots émotionnels semblent être cohérents entre le français et l'italien, mais ne correspondent pas à ceux de l'anglais ; il y a donc une façon particulière d'exprimer et de modéliser l'expérience émotionnelle en fonction de la langue (Galati et Sini in Plantin, Doury et Traverso, 2000). Chabrol (in Plantin, Doury et Traverso, 2000), en se concentrant sur les émotions de base, met en avant leur manque de contextualisation lors de la plupart des analyses, et souligne la nécessité de les relier à leur contexte pour en comprendre les mécanismes. Charaudeau met pour sa part en avant l'idée qu'un mot peut changer en fonction de son contexte et de son utilisation (Charaudeau in Plantin, Doury et Traverso, 2000). Bertrand, Matsangos, Périchon et Vion (in Plantin, Doury et Traverso, 2000) soulignent qu'un paramètre contextuel (à savoir le fait que les interlocuteurs connaissaient l'intervieweur) a eu une influence sur le développement conversationnel au cours de l'expérience qu'ils avaient mise en place. Véronique Traverso, qui traite des émotions « dans la confiance » entre deux vieux amis, pose un paramètre contextuel explicite : l'émotion liée à la situation elle-même, et plus précisément au fait que la confiance est une forme de révélation de soi (Traverso in Plantin, Doury et Traverso, 2000). Elle souligne aussi le niveau de l'émotion évoquée et celui de l'émotion interactionnelle (ibid). Bouchard (in

Plantin, Doury et Traverso, 2000) considère de son côté ce que l'on appelle communément les marqueurs du discours.

Le type particulier d'interaction analysé par Marcoccia (in Plantin, Doury et Traverso, 2000), à savoir la communication médiatisée par ordinateur (qui nous intéresse particulièrement étant donné qu'elle est l'un des modes privilégiés pour la communication à distance), est caractérisé par la décontextualisation d'une part de l'identité et d'autre part de la localisation des participants. Internet est un système de communication basé essentiellement sur la transmission de messages écrits, soit pour une interaction synchronisée quasi-instantanée, soit pour une interaction asynchrone (par exemple par l'échange d'emails ; on ne sait quand le destinataire va nous lire, ni quand il va nous répondre). Ses utilisateurs rencontrent les limites du code écrit lorsqu'ils tentent d'accorder autant d'importance à la relation qu'au contenu échangé. Pour pallier les contraintes de l'écrit, c'est souvent la conversation, l'interaction face à face, qui est placée comme idéal de communication. Ce choix est pour le moins paradoxal (il s'agirait de faire du face à face avec de l'écrit), et il a pour conséquence un renouvellement partiel du code écrit. Cette « nouvelle » écriture a un but premier : il s'agit de réussir à trouver un substitut écrit à tout ce qui est paraverbal, aux marques de l'intonation (par exemple n'utiliser que des majuscules pour signifier que l'on crie) et aux expressions mimo-gestuelles produites lors d'une conversation, à savoir principalement celles qui sont liées à l'expression des émotions (les smileys). De manière plus générale, l'utilisation des smileys dans la communication médiatisée par ordinateur met en évidence le fait que l'émotion n'est pas simplement « quelque chose en plus » dans l'interaction. Elle montre aussi que l'expression de certaines émotions de base (joie, colère, tristesse) est indispensable à la construction de la signification d'une intervention, à la définition de la situation, au cadrage, et du même coup au ménagement des faces. Marcoccia identifie quatre fonctions des smileys (ibid) :

1. Fonction illustrative : le smiley exprime l'émotion ressentie par le locuteur ;
2. Fonction interprétative : le smiley permet d'interpréter le message du locuteur ;
3. Le smiley est un marqueur de relation ;
4. Le smiley est un anti-FTA (face threatening act ; acte menaçant la face, selon les terminologies de Goffman (1959) et Brown et Levinson 1978 ; Brown 1987).

Ces quatre fonctions différentes des smileys ne sont pas exclusives les unes des autres. Les smileys rendent conventionnelle l'expression des émotions de manière iconique, et remplacent l'intonation dans de nombreux cas, en une forme ludique de socialisation (Marcoccia in Plantin, Doury et Traverso 2000).

La pluridimensionnalité des émotions est aussi à prendre en compte, en particulier dans une situation d'interaction. Chabrol met l'accent sur une interaction à plusieurs dimensions (Chabrol in Plantin, Doury et Traverso, 2000). Bertrand, Matsangos, Périchon et Vion (in Plantin,

Doury et Traverso, 2000) suggèrent une étude de l'émotion prenant en compte différents modes de communication (« verbal, paraverbal ou mimo-gestuel », *ibid.*, p. 169). Traverso (in Plantin, Doury et Traverso, 2000) pose que l'intonation et les changements de rythme sont des indices significatifs de l'implication et de la compréhension, ce que nous avons déjà évoqué. Maury-Rouan (in Plantin, Doury et Traverso 2000) présente le management des émotions aux niveaux verbal, vocal et mimo-postural-gestuel.

Certains sujets sont typiques de l'approche interactionnelle, comme par exemple la notion d'empathie, le conflit, la relation à la politesse, l'engagement, et les variations interculturelles en général. Un autre sujet qui caractérise une approche interactionnelle, et qui est en total accord avec la théorie de l'énonciation, est la co-production du sens. Bertrand, Matsangos, Périchon, et Vion (in Plantin, Doury et Traverso, 2000, p.178) parlent de la « co-énonciation », qui implique les dimensions de l'action conjointe (nous détaillerons les mécanismes de l'attention conjointe dans la partie 3). Maury-Rouan (in Plantin, Doury et Traverso 2000, p.185) met en avant la « co-construction » des tours de parole. Il y a aussi la « synchronisation » d'Auchlin (in Plantin, Doury et Traverso, 2000, p. 201), la « coopération » de Traverso (in Plantin, Doury et Traverso, 2000, p. 216), et la « co-action » de Bouchard (in Plantin, Doury et Traverso, 2000, p. 228).

Comme le disent Ochs et Schieffelin (1989, p.22, traduction personnelle) : « L'affect imprègne l'intégralité du système linguistique. Presque tous les aspects du système linguistique sont candidats à l'expression de l'affect. »⁹ Les marqueurs linguistiques qui rendent les émotions explicites sont extrêmement nombreux. Plusieurs de ces marqueurs émotionnels sont mis en évidence par Doury dans son chapitre (in Plantin, Doury et Traverso, 2000) sur un débat publié dans le journal scientifique La Recherche. Manno (in Plantin, Doury et Traverso, 2000) pointe quant à lui la syntaxe expressive, c'est-à-dire les phénomènes tels que la répétition et la dislocation, et les dispositifs pragmatiques/rhétoriques comme l'allocutivité et l'emphase. Il écrit d'autre part : « Si l'on ne peut pas affirmer que les manifestations des émotions à l'écrit sont insincères, il est indiscutable que le mode de conception moins spontané et plus prémédité de l'écrit permet d'exercer un contrôle sur l'affectivité. L'expression des sentiments à l'écrit est d'autant plus suspecte que, selon Ekman (1992a), les émotions de base ont en commun un déclenchement rapide, une survenue spontanée, et surtout elles durent quelques secondes » (Manno in Plantin, Doury et Traverso, 2000, p.285).

Les émotions en interaction sont influencées par de multiples facteurs, dont le genre et le statut. Hess et Bourgeois (2010) ont mené deux études pour évaluer l'influence du contexte émotionnel et social, en termes de genre et de statut, sur l'expressivité du locuteur et l'imitation

⁹ « Affect permeates the entire linguistic system. Almost any aspect of the linguistic system [. . .] is a candidate for expressing affect. »

de l'observateur dans le cadre d'une interaction dyadique. 96 dyades composées de personnes du même sexe ont pris part à la première étude, et 72 dyades composées de personnes de sexe opposé ont participé à la deuxième étude. Les résultats montrent que dans toutes les dyades, qu'elles soient de même sexe ou de sexe opposé, les femmes sourient plus que les hommes, et que les membres des dyades de même sexe utilisent plus de sourires de Duchenne que de sourires volontaires pour signaler les intentions sociales. Rappelons que le sourire de Duchenne exprime une joie sincère : il se différencie d'un sourire volontaire (ou forcé) du fait de la contraction d'un muscle que l'on ne contrôle pas situé autour des yeux (*orbicularis oculi*) (voir Frank et Ekman, 1993 ; Frank, Ekman et Friesen, 1993). Dans les dyades de même sexe, l'expressivité et l'imitation faciales étaient déterminées par le contexte à la fois social et émotionnel de la situation. Dans les dyades de sexes opposés, les effets du contexte émotionnel étaient bien maintenus, mais ceux du contexte social disparaissaient.

N'oublions pas le fait que la communication interpersonnelle est largement de nature non-verbale (Argyle, 1988), et que le comportement non-verbal a comme but premier de communiquer les subtilités des états émotionnels entre les individus. En communication en face à face, nous exprimons nos émotions principalement au travers des expressions faciales, de la voix, et du toucher, bien que d'autres gestes comme la posture et des signaux manuels puissent aussi être utilisés (Cassell et Thorisson, 1999 ; Collier et DiCarlo, 1985). Le fait qu'une interface puisse supporter la communication non-verbale en fait un dispositif de choix pour la communication des émotions à distance.

Les émotions émergent au sein des interactions ; elles dépendent de nombreux éléments, qui définissent l'interaction et influent son déroulement. On notera que la notion de *participatory sense-making* (De Jaegher et Di Paolo, 2007) se révèle aussi très utile pour comprendre la question des émotions. En effet, le fait que les interactants participent *ensemble* de la construction du sens permet de comprendre qu'un même élément déclencheur ne provoque pas nécessairement la même émotion chez des individus distincts ; les théories principales sur les émotions ne permettent en effet pas d'expliquer ce phénomène (Arnold, 1960). En introduisant l'idée que le sens associé peut varier, du fait que l'association dépend du contexte et des expériences des individus, il est donc possible d'expliquer ces variations. De même, cela permet d'expliquer pourquoi, selon les cultures, les émotions diffèrent ; elles peuvent en effet varier, exister dans certaines cultures et pas dans d'autres, s'exprimer différemment selon les cultures... Par exemple, les Esquimaux Utka ne connaissent pas le sentiment de la colère ; jamais ils ne l'expriment ni ne s'y réfèrent (Mikolajczak et al. 2009, p. 48).

La question des émotions n'est pas aisée à traiter. Du fait de la difficulté à mettre en évidence une théorie relationnelle des émotions et pour des raisons que nous allons approfondir dans la partie 1.1.3.2, nous allons considérer, comme première hypothèse pour nous guider,

qu'une voie possible pour l'échange d'émotions est de permettre le sentiment de présence d'autrui. Dans un second temps, nous considérerons que le fait d'orienter à deux l'attention perceptive sur un objet commun est un second pas dans l'interaction émotionnelle. Nous allons exposer les raisons théoriques de ces choix, avant de les développer respectivement dans les parties 2 et 3 de ce travail.

1.1.2 La suppléance perceptive comme cadre de recherche

1.1.2.1 Présentation

1.1.2.1.1 La substitution sensorielle

Les dispositifs de substitution sensorielle ont vu le jour dans les années 1960, avec pour objectif de palier à des déficiences sensorielles, notamment visuelles (pour des exemples, voir Auvray (2004) et Ali-Ammar (2005))¹⁰. Parmi eux, le Tactile-Vision Substitution System (TVSS) de Bach-y-Rita est considéré comme le dispositif pionnier et paradigmatique (Bach-y-Rita et al. 1969 ; Bach-y-Rita, 1972). C'est un dispositif de suppléance perceptive visuo-tactile : il transmet des informations de nature visuelle en passant par le sens du toucher. Les dispositifs de suppléance perceptive rendent en effet accessibles des informations provenant d'une modalité sensorielle (le plus souvent la vue) par l'intermédiaire d'une autre modalité sensorielle (le toucher ou l'ouïe, par exemple). Des informations visuelles peuvent donc être converties en informations tactiles. Pour arriver à un résultat de substitution sensorielle, il faut disposer de capteurs, chargés de l'entrée de l'information, de stimulateurs pour la sortie, et d'un système de couplage faisant la liaison technique de ces deux premiers éléments (Lenay et al., 2000). Le TVSS est par exemple un dispositif composé d'une caméra (pour l'entrée de l'information visuelle), d'une matrice 20 x 20 de stimulateurs électro-tactiles (pour la sortie) et d'un convertisseur analogique traduisant les informations visuelles de la caméra en une image tactile affichée par la matrice (Bach-y-Rita et al. 1969) (voir figures 9 et 10).

¹⁰ Pour des raisons que nous détaillerons dans la sous-partie suivante, nous préférons appeler ces systèmes des dispositifs de suppléance perceptive ; voir Lenay et al. (2000).

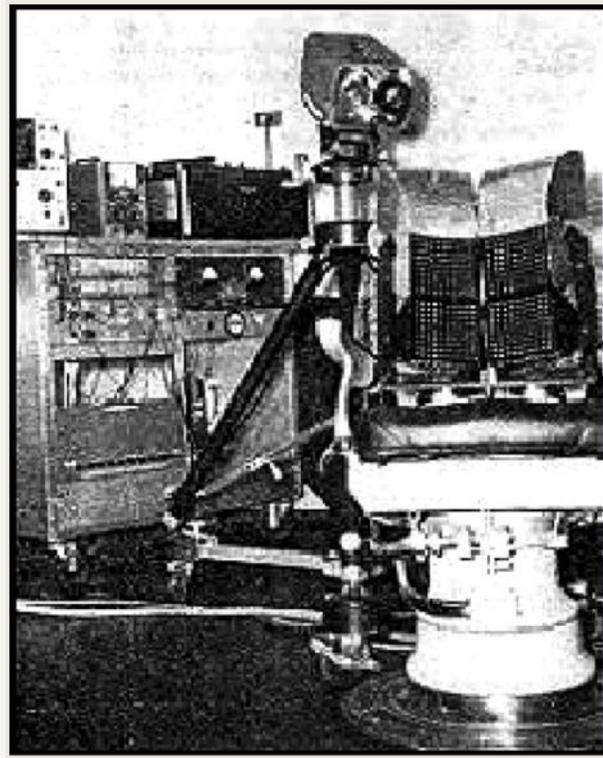


Figure 9. Le TVSS de Bach-y-Rita dans sa première version : caméra fixe, matrice de picots incrustée dans le dossier du fauteuil.

Dans la première version du dispositif, la caméra est immobile, fixée sur un trépied, et la matrice de picots est fixée au dossier d'un fauteuil où l'utilisateur doit prendre place (voir figure 9). Les utilisateurs, non-voyants ou voyants aveuglés, arrivent à détecter des cibles simples et à s'orienter vers ces cibles (Bach-y-Rita, 1972), ou encore à faire la différence entre des lignes orientées de façon différente dans l'espace, ainsi qu'à reconnaître la direction du mouvement de cibles mobiles (ibid). Cependant, ils se heurtent très vite aux limites du dispositif : il est très difficile, et très coûteux en terme d'apprentissage, de reconnaître des formes géométriques simples, ou des objets connus lorsqu'ils sont présentés à la caméra selon différents angles (ibid). De plus, avec cette disposition fixe de la caméra, l'attention des utilisateurs reste fixée sur les stimuli perçus à la surface de la peau ; ces stimuli sont donc vécus comme des sortes de chatouilles, pas nécessairement désagréables, mais toujours perçues (ibid).

Le TVSS a subi plusieurs améliorations. La matrice de picots a été miniaturisée, pouvant ainsi être appliquée sur le thorax, le ventre ou encore le front de l'utilisateur. La caméra a quant à elle été rendue mobile, pouvant être soit tenue dans la main de l'utilisateur, soit fixée à des lunettes portées par l'utilisateur (Collins et Bach-y-Rita, 1973) (voir figure 10).

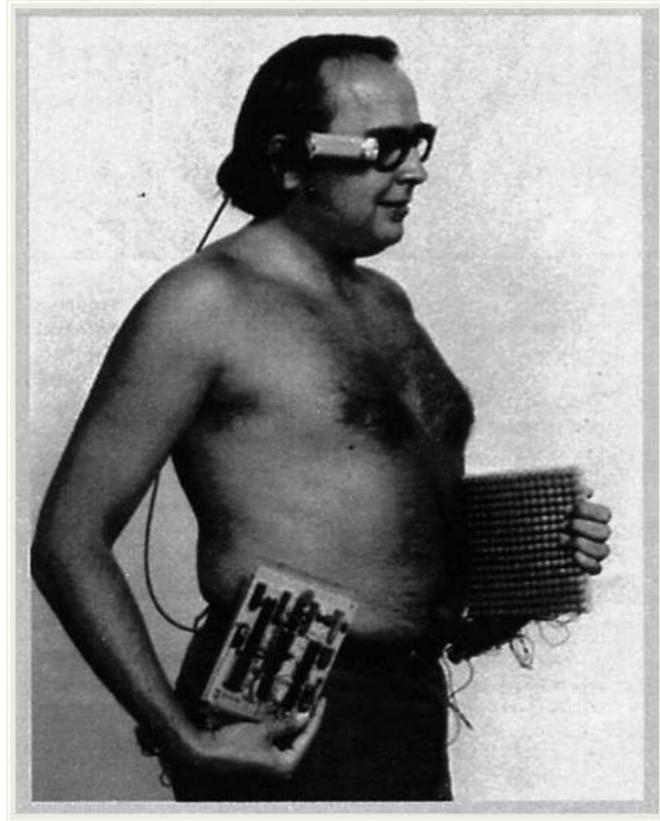


Figure 10. Le TVSS dans une version plus récente : caméra et matrice de picots mobiles.

La manipulation active de la caméra a permis une reconnaissance de formes bien plus efficace que dans la version fixe. Les stimulations tactiles, au lieu d'être perçues à la surface de la peau là où se trouve la matrice de stimulateurs, sont ainsi directement attribuées à l'objet perçu, situé dans l'espace perceptif de l'utilisateur (« là devant moi ») (Pacherie, 1997). Les dispositifs de suppléance perceptive démontrent donc une « mise en extériorité » de ce qui est perçu puisqu'à l'issue de l'apprentissage, l'utilisateur oublie les sensations tactiles sur sa peau, et perçoit des objets localisés à distance, extérieurement à lui (cf figure 3). Le système de suppléance perceptive devient transparent ; l'utilisateur ne le perçoit plus, mais perçoit à travers lui.

Le TVSS a aussi été décliné de façon à ce que la matrice de picots soit remplacée par une matrice d'électrodes à placer sur la langue ; c'est le Tongue Display Unit (Bach-y-Rita et Kaczmarek, 2002). The vOICe est un dispositif similaire au TVSS, excepté que les informations visuelles sont transformées en informations auditives (Meijer 1992). Rappelons que Auvray

(2004) pose que l'immersion dans un dispositif, et donc la capacité d'appropriation active d'un dispositif de suppléance perceptive, se fait en cinq étapes :

1) Le contact. L'utilisateur apprend à détecter les régularités qu'il existe entre ses actions et les retours sensoriels qui en découlent. Ainsi, l'utilisateur peut mettre en place un contact perceptif avec l'objet, le stabiliser et le maintenir dans le temps.

2) L'attribution distale. Par le biais de l'exercice de ce contact perceptif, l'utilisateur comprend que les stimulations qu'il reçoit sont dues à la rencontre d'un objet qui se situe dans l'espace distal. Il y a extériorisation de la source de stimulations.

3) La maîtrise de l'espace distal. L'utilisateur apprend à faire varier ses points de vue en fonction de la variation des stimulations reçues. Le sujet maîtrise peu à peu l'espace distal, même si la localisation des objets reste égocentrée.

4) La localisation distale. L'expérience de l'espace est rendue fiable par une automatisation du couplage entre les actions de l'utilisateur et les stimulations tactiles qui en découlent. La localisation de l'objet n'est plus dépendante du corps lui-même, mais s'inscrit dans un espace perceptif distal.

5) La constitution d'une expérience distale. Cette étape de constitution d'une signification, de valeurs communes et de l'expérience d'émotions se fait par le partage des expériences perceptives ouvertes par l'outil (ibid).

Dans le cas du TVSS, cette immersion a bien lieu : Bach-y-Rita observe un comportement d'évitement caractéristique de la part de l'utilisateur lorsque l'expérimentateur manipule le zoom de la caméra à l'insu de celui-ci (Bach-y-Rita, 1972). L'utilisateur, immergé dans l'espace perceptif, réagit comme s'il percevait l'extension de l'image tactile provoquée par Bach-y-Rita comme un mouvement de l'objet lui-même vers lui, et cherche donc à l'éviter.

Pour pouvoir mettre en extériorité des percepts et par conséquent pouvoir reconnaître des formes, une exploration active de l'espace distal par l'utilisateur est absolument nécessaire. On en revient à l'une des idées fondamentales de l'énaction : il n'y a pas de perception sans action. C'est la possibilité de faire varier les entrées sensorielles en fonction des actions qui permet l'établissement d'un couplage sensoriel minimal (Lenay, Canu et Villon, 1997). Le dispositif de « substitution sensorielle » s'inscrit dans la boucle sensori-motrice et supporte de manière empirique la théorie de la perception active.

1.1.2.1.2 Ou plutôt : la suppléance perceptive

Est-il donc encore adéquat d'utiliser le terme de « substitution sensorielle » s'il faut prendre en considération l'activité perceptive de l'utilisateur ? Modifier la nature de l'entrée

sensorielle ne suffit pas pour permettre l'immersion de l'utilisateur dans une expérience perceptive visuelle. Ce qui est perçu par l'utilisateur « n'est pas issu des invariants de la sensation », mais est issu « des invariants de cercles sensori-moteurs inséparables de l'activité du sujet » (Lenay et al., 2000, p. 296). La perception n'est possible que grâce au couplage des actions et des sensations ; s'il y avait substitution, elle ne serait donc pas sensorielle, mais plutôt sensori-motrice.

Mais en fait, il n'y a pas de substitution : en utilisant un dispositif comme le TVSS, l'expérience perceptive qui nous est proposée ne ressemble pas à une expérience visuelle à proprement parler. Si le dispositif permet bien de saisir des informations de nature visuelle et spatiale (« il y a un poteau juste là, devant moi »), il ne permet pourtant pas de « voir » (au sens le plus commun), ce qui constitue certainement une déception pour les utilisateurs non-voyants, qui pensaient pouvoir accéder au monde des voyants. L'efficacité du dispositif n'est pas nulle, mais ce dernier présente des limites, et ne permet par exemple pas de distinguer les couleurs. Les percepts restent grossiers, les possibilités d'action sont limitées ; la reconnaissance des formes est assez lente et fastidieuse, même suite à un apprentissage. La modalité sensorielle manquante n'est pas remplacée par une autre modalité existante. L'utilisateur se retrouve immergé dans un espace perceptif nouveau, unique, ouvert par le dispositif lui-même, inconnu de toutes les personnes ne l'ayant jamais pris en main (Auvray, 2004). Il s'agit d'une nouvelle forme de couplage sensori-moteur avec le monde, d'une nouvelle façon de percevoir (Lenay et al., 2000). Il n'y a donc pas de substitution, et elle ne serait donc pas sensorielle mais sensori-motrice (Lenay et al. 2000). Lenay et al. (2000) préfèrent donc utiliser le terme de « suppléance perceptive » que le terme « substitution sensorielle » ; les dispositifs ouvrent une modalité perceptive originale et nouvelle, reposant sur des lois de couplage sensori-moteur inédites (ibid.). On notera que cette modalité peut même offrir à sentir des données normalement inaccessibles aux êtres humains ; par exemple, un dispositif de suppléance perceptive peut donner à voir l'infrarouge et l'ultra-violet¹¹.

Les dispositifs de suppléance perceptive peuvent permettre l'accessibilité de l'espace physique, mais aussi de l'espace numérique. Tactos, le dispositif développé par le groupe CRED de l'Université de Technologie de Compiègne, est conçu dans cette deuxième optique : il est destiné aux personnes en situation de handicap visuel, afin de leur donner accès à l'espace numérique, et notamment à la spatialisation de l'information dans cet espace, ce qui n'est pas encore supporté par les techniques actuelles. Pour pouvoir reconnaître une forme numérique en

¹¹ <http://www.blog-lecerveau.org/blog/2013/03/04/un-cyborg-qui-entend-plus-que-ce-que-nous-voyons/>

utilisant le dispositif Tactos, il faut découvrir et exercer des lois de contingences sensori-motrices ; cet apprentissage se retrouve pour tous les dispositifs de suppléance perceptive.

Lenay note au sujet du numérique : « Plutôt qu'une opposition entre monde physique et monde numérique, la suppléance perceptive permet de comprendre, dans la perspective de la cognition située, que la technologie informatique s'insère dans l'ensemble des activités humaines comme transformation du couplage entre les hommes, leur environnement physique et leur mémoire externe numérisée. » (Lenay, 2011). Il pose donc ainsi les outils informatiques comme dispositifs de couplage de plus en plus intégrés et essentiels dans notre vie de tous les jours, nous permettant l'accès à une mémoire externalisée sous forme numérique.

Les résultats issus de la recherche sur la suppléance perceptive sont fondamentaux pour la conception d'outils et de dispositifs dédiés aux interactions perceptives. Ils démontrent en effet d'une part que la richesse de la perception dépend autant des qualités de l'action (mobilité, rapidité,...) que de celles de la sensation (largeur du spectre, nombre de senseurs, ..) (Lenay, Stewart et Gapenne, 2002). D'autre part, ils mettent en évidence le fait que le dispositif d'interaction ne doit pas être considéré comme un « simple » objet *avec* lequel on interagit, mais comme un objet *par* lequel on interagit.

1.1.2.2 Tactos

1.1.2.2.1 Présentation

Revenons plus en détail sur le dispositif Tactos. C'est un dispositif de suppléance perceptive, destiné en premier lieu à des utilisateurs présentant une déficience visuelle afin de leur permettre l'accès aux informations graphiques affichées sur un écran d'ordinateur par le biais de la modalité tactile. Il permet par ailleurs l'étude des phénomènes perceptifs dans une situation minimaliste.

Les technologies actuellement disponibles pour les personnes malvoyantes afin d'avoir accès au numérique sont principalement les logiciels de synthèse vocale, qui lisent à haute voix les informations affichées à l'écran, ainsi que les barrettes de cellules brailles électroniques, qui codent ces mêmes informations en braille. Les informations textuelles affichées à l'écran peuvent donc être transmises à l'utilisateur sous modalité auditive ou sous modalité tactile, lui permettant ainsi d'utiliser un ordinateur adéquatement équipé. Lorsque les métadonnées sont disponibles, l'utilisateur peut même accéder à une description textuelle des images.

On remarquera que l'information délivrée à l'utilisateur malvoyant est de nature strictement linéaire et temporelle : les informations sont délivrées successivement sous forme de mots, ainsi qu'ils sont affichés à l'écran, dans le sens de la lecture. L'organisation spatiale de

l'information est complètement occultée, alors qu'elle est l'une des bases de la navigation pour l'utilisateur voyant (« je vois où est mon curseur, je vois l'icône qui m'intéresse, je fais bouger mon curseur jusqu'à l'icône pour cliquer dessus,... »). La position des fenêtres, des icônes, mais aussi les tableaux, les graphiques, les schémas, les cartes : tous ces objets informatiques sont inaccessibles à l'utilisateur malvoyant. C'est pourtant l'information bidimensionnelle qui fait la richesse des interfaces graphiques, en spatialisant l'information.

Tactos, dispositif constitué d'un logiciel dédié couplé à un dispositif tactile (le Module d'Interaction Tactile, ou MIT), a été développé pour rendre accessible aux malvoyants l'organisation spatiale des informations présentes sur l'écran de l'ordinateur (Hanneton et al. 1999 ; Sribunruangrit et al. 2002). Le système commande des stimulateurs tactiles, à savoir deux cellules Braille piézoélectriques générant l'activation d'un total de seize (4 x 4) picots mécaniques, en fonction des déplacements d'un curseur sur l'écran d'un ordinateur (voir figure 11). Les déplacements du curseur sont commandés par un effecteur que l'utilisateur manipule ; il peut s'agir d'une souris, d'un stylet sur une tablette graphique, du doigt sur une tablette tactile ou un touchpad,... Le curseur est divisé en 16 champs récepteurs élémentaires, chacun déclenchant l'activation en tout ou rien d'un stimulateur tactile (à savoir l'un des picots de l'une des deux cellules brailles du Module d'Interaction Tactile) lorsqu'il croise au moins un pixel noir à l'écran (voir figure 12).

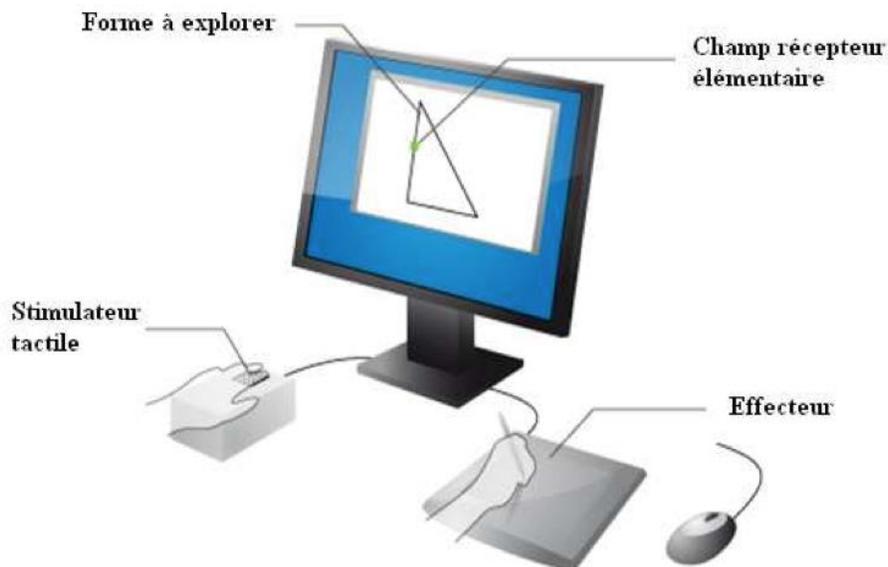


Figure 11. Le dispositif Tactos.

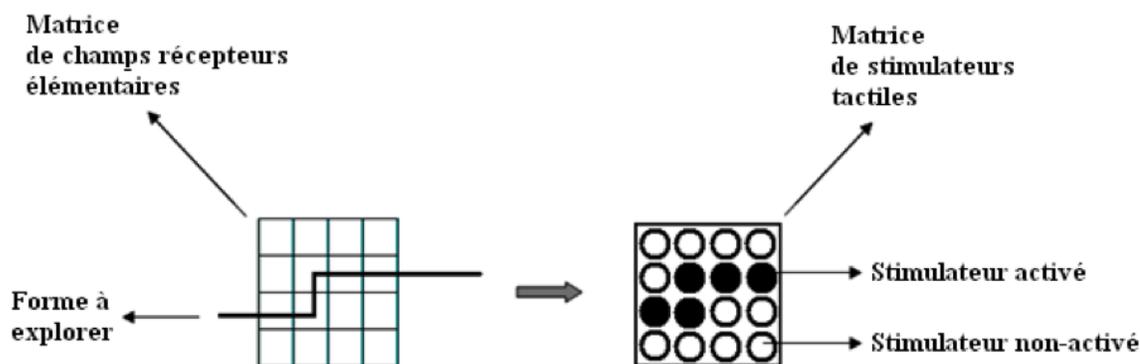


Figure 12. Les 16 champs récepteurs élémentaires forment une matrice 4 x 4. Chaque champ récepteur (à gauche) active le picot correspondant sur la matrice de 16 stimulateurs tactiles (à droite). Il y a donc parallélisme entre les champs récepteurs et les stimulateurs tactiles.

Tactos est destiné aux personnes malvoyantes, mais il peut être utilisé par des personnes non-voyantes, malvoyantes, voyantes avec les yeux bandés ou voyantes sans accès visuel à l'écran (voir figure 13). L'effecteur doit être tenu dans la main dominante, tandis que les cellules brailles doivent être installées sous l'un des doigts de la main non-dominante. C'est par l'effecteur que l'utilisateur peut agir dans le monde numérique, représenté de façon bidimensionnelle sur l'écran. Les sensations sont délivrées par les cellules brailles. Comme il y a coïncidence spatiale entre le point de perception (le champ récepteur) et l'objet perçu (le pixel rencontré), Tactos permet au sujet de faire l'expérience d'une forme de « toucher » des objets affichés à l'écran.

Les effecteurs recommandés sont ceux fonctionnant en absolu, et non ceux qui fonctionnent en relatif (dont la souris fait partie). Il faut en effet que la surface d'action entretienne un rapport bijectif avec l'espace de l'écran : la position de l'effecteur sur la surface d'action doit correspondre à la position du curseur à l'écran. Si je suis en bas à gauche sur ma tablette, alors le curseur est en bas à gauche de l'écran. L'utilisateur peut alors accéder à une connaissance de l'espace en mobilisant des informations proprioceptives directement liées à son geste d'exploration (Auvray et al. 2005). Les effecteurs relatifs ne conviennent donc pas, car ils commandent un déplacement plus qu'une position ; il ne peut alors y avoir de corrélation entre les informations proprioceptives du geste d'exploration et la position exacte du curseur à l'écran (ibid). Cependant, pour certaines tâches, notamment celles basées sur la nature des mouvements (que l'on étudie dans certaines expériences), l'utilisation d'un effecteur relatif peut s'avérer pertinente.

Il existe d'autre part une configuration plus minimale possible du dispositif Tactos, où un unique champ récepteur ne permet l'accès qu'à un seul bit d'information à chaque instant (« mon

curseur touche la forme » ou « mon curseur ne touche pas la forme »). À chaque instant, soit tous les picots sont levés, soit ils sont tous abaissés. Des capacités effectives de reconnaissance de formes sont déjà observées dans cette configuration, mais l'on peut cependant enrichir les possibilités d'agir et/ou de sentir permises par le dispositif. On peut ainsi, comme nous l'avons présenté, augmenter le nombre de champs récepteurs utilisés de façon conjointe : cela permet une forme de « parallélisme », qui améliore les capacités de perception (Sribunruangrit, 2004) (voir figure 14). Le parallélisme des entrées sensorielles engendre une économie de mouvements et de mémoire permettant d'augmenter la vitesse perceptive, et améliorant sensiblement dans le cas de Tactos la qualité de l'exploration et la reconnaissance des formes inscrites à l'écran (Gapenne et al. 2001).

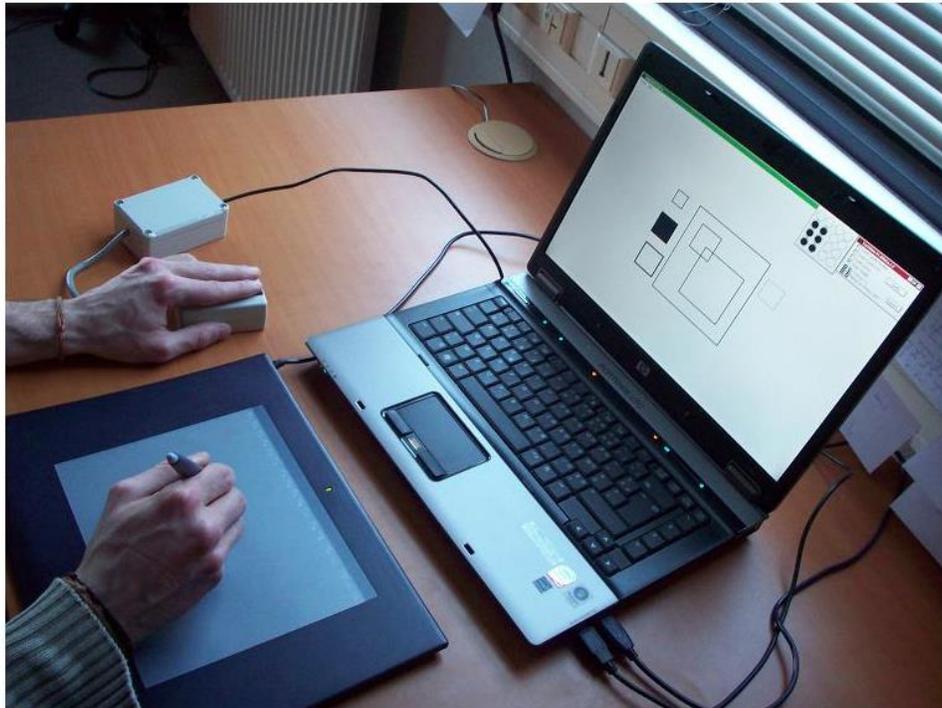


Figure 13. Un utilisateur se servant de Tactos.

Avec Tactos, les sujets sont capables de localiser, d'identifier, de reconnaître et de catégoriser des formes plus ou moins complexes (Gapenne et al. 2003 ; Auvray et al. 2007 ; Loïc Deschamps, 2013). Les formes ne sont pas reçues d'un coup par le système sensoriel, comme ce serait le cas si elles étaient appliquées sur la peau. Il n'y a en effet que 16 champs récepteurs élémentaires, qui ne fournissent chacun qu'une donnée sensorielle très pauvre, correspondant à un seul bit d'information à chaque instant. L'utilisateur ne reçoit donc que 16 informations à

chaque instant puisque les picots n'ont chacun que deux états possibles (levés ou baissés). Les sujets ne peuvent reconnaître les formes se trouvant dans l'espace numérique que parce qu'ils explorent cet espace de façon active, en faisant le lien entre leurs mouvements et les stimulations tactiles qui en résultent. Comme le TVSS avant lui, le dispositif Tactos met en évidence le caractère constitutif de l'action dans la perception.

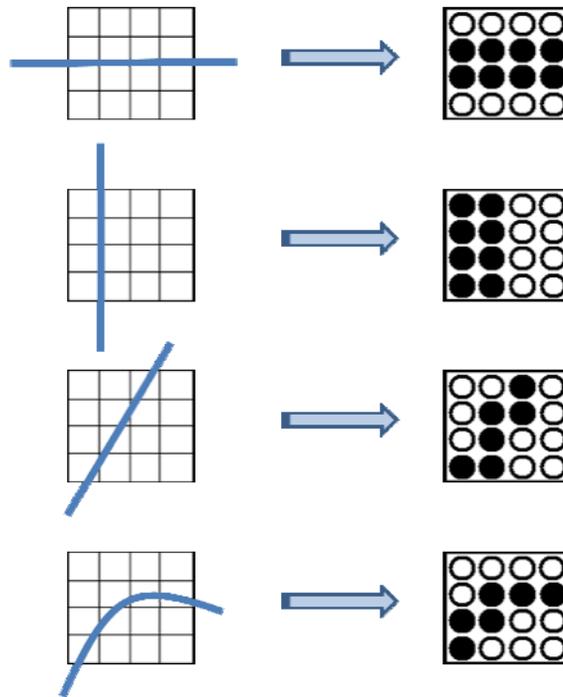


Figure 14. Exemples de stimulations tactiles résultant du croisement du curseur avec différentes courbes.

Tactos permet, grâce à son logiciel, l'enregistrement (à une fréquence de 64 Hz) à chaque instant de la position du champ récepteur et de l'état des picots. On peut donc rejouer ce qu'un utilisateur a fait en utilisant le dispositif, en connaissant aussi les stimulations tactiles qu'il a reçues. À partir des enregistrements, on peut donc procéder à des analyses rigoureuses de l'activité sensori-motrice déployée par les sujets pour percevoir les objets affichés à l'écran.

Ali-Ammar (2005) a mené de nombreux travaux expérimentaux centrés sur la conception de dispositifs de suppléance perceptives destinés à l'enseignement de la géométrie aux jeunes aveugles. Ses travaux montrent qu'après un apprentissage du maniement de Tactos d'une vingtaine d'heures de pratique, et malgré la pauvreté de l'entrée sensorielle, les utilisateurs parviennent à reconnaître des formes affichées à l'écran. La figure 15 montre un enregistrement

de la trajectoire d'un sujet explorant la lettre S. L'utilisateur explore d'abord l'espace au hasard jusqu'à ressentir une stimulation tactile ; le sujet tente alors de la reproduire en effectuant le mouvement inverse. Dès lors, il tente de « s'accrocher » à la forme en mettant en place une stratégie de micro-balayage (un mouvement de va-et-vient autour des lignes qui composent la forme) tout en suivant son contour. Peu à peu, le sujet à parvient à anticiper l'activation du retour sensoriel en fonction du mouvement accompli par le biais de l'effecteur (Lenay, Canu et Villon, 1997). Lorsqu'il perd la forme, l'utilisateur cherche à retourner sur ses pas pour retrouver des stimulations tactiles. Les observations d'individus explorant des formes par le biais de Tactos mettent empiriquement en évidence le fait que les caractéristiques spatiales d'un objet se définissent par des lois de contingences sensori-motrices permettant d'anticiper les conséquences sensorielles de nos propres actions : faire l'expérience d'une forme consiste à faire l'expérience du geste qui permet de la produire (Lenay, Stewart et Gapenne, 2002). Au début, les formes explorées proposées à l'utilisateur sont des figures géométriques simples (ronds, carrés, triangles, lettres,...), mais peu à peu, il développe des stratégies qui lui permettent de reconnaître des formes de plus en plus complexes (des lignes qui comportent des embranchements, des figures imbriquées,...).

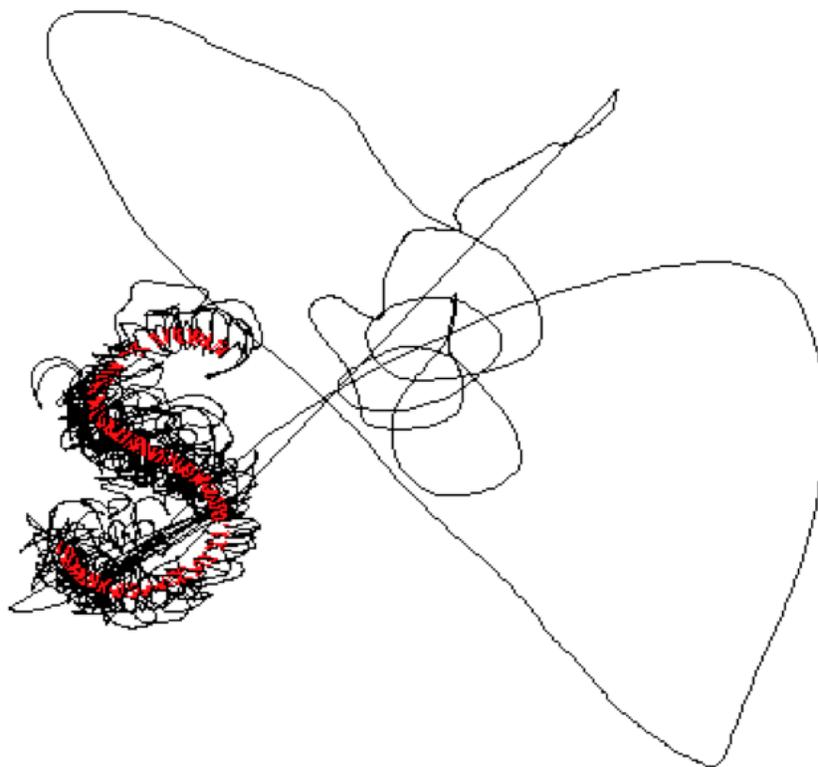


Figure 15. Trajectoire du curseur d'un sujet explorant la lettre S par le biais de Tactos. Les zones en rouge indiquent les moments où l'utilisateur reçoit des stimulations tactiles.

Gapenne et al. (2005) ont utilisé Tactos pour monter une expérience mettant en évidence le fait que les mécanismes d'organisation perceptive chez l'humain ne pouvaient pas être liés à (ni strictement déterminés par) la nature des organes se retrouvant stimulés lors de la situation considérée. En fait, la perception est principalement liée aux lois dynamiques entre les actions et les sensations, ainsi qu'à leurs propriétés. Dans cette expérience, trente sujets avaient pour tâche de percevoir des formes en deux dimensions dépourvues de signification (il y en avait neuf, présentées chacune deux fois au total, dans des ordres différents) à l'aide de Tactos, utilisé avec un effecteur stylet sur tablette graphique. Les auteurs ont aussi fait varier les stimulations transmises aux les sujets, qui recevaient lors de la rencontre avec un objet soit des stimulations tactiles par le biais du MIT, soit des stimulations visuelles (flash lumineux), soit des stimulations auditives (sons), soit toutes ces stimulations à la fois. Les résultats montrent qu'il n'y a pas de différence significative de réponse quelque soit le retour sensoriel mis en jeu.

1.1.2.2.2 État d'avancement et résultats

1.1.2.2.2.1 Analyses d'usage

En parallèle aux expérimentations réalisées dans notre laboratoire, des suivis longitudinaux ont été mis en place dès les premières étapes de la conception de Tactos (voir notamment Ali-Ammar (2005) et Deschamps (2013)). Le but était de pouvoir mettre à l'épreuve le dispositif par les utilisateurs visés, à savoir des personnes malvoyantes, en contexte écologique. La conception d'un tel dispositif nécessite la participation active des utilisateurs ciblés, de façon à pouvoir prendre en compte l'utilisateur final dès les premières étapes de conception d'une part, et à valider l'ensemble des étapes de conception et vérifier leur pertinence dans des contextes d'usage réel d'autre part.

La maîtrise d'un dispositif de suppléance perceptive tel que Tactos nécessite une prise en main et un apprentissage plus ou moins longs. En effet, nous ne sommes pas tous égaux face à l'apprentissage de ce dispositif. On ne peut pas prédire la réussite de l'apprentissage de quelqu'un ; notamment, le fait d'être une personne non-voyante maîtrisant le braille n'est pas nécessairement un avantage. En effet le doigt sera alors habitué aux stimulations délivrées par les cellules brailles, mais la personne sera d'abord concentrée sur les stimulations en elles-mêmes, essayant de « lire » ce qui lui est délivré. De même, une personne voyante aveuglée, qui connaît le fonctionnement visuel d'un curseur et son évolution dans le monde numérique, pourra avoir le plus grand mal à se saisir du dispositif. L'appropriation du dispositif passe par plusieurs étapes, qui nécessitent un accompagnement personnalisé de façon à guider au mieux le processus d'apprentissage, en tenant compte des difficultés spécifiques des utilisateurs, qui n'évoluent pas

tous au même rythme. Ce processus conduit pour tous à l'ouverture d'une modalité perceptive originale, qui permet dans le cas de Tactos d'accéder aux informations spatiales des interfaces graphiques numériques.

1.1.2.2.2 Supports pédagogiques numériques

Tactos permet donc l'exploration de formes dans un espace numérique. Une application simple et aux retombées immédiatement visibles est celle de l'utilisation de Tactos pour fournir des supports pédagogiques numériques aux enfants et adolescents malvoyants. En effet, il est difficile pour ce public d'accéder à des supports pédagogiques efficaces dans certaines matières, notamment celles faisant entrer en jeu des objets aux propriétés spatiales importantes, comme par exemple la géométrie. La géométrie est classiquement enseignée aux aveugles par l'utilisation de papiers thermoformés, qui permettent l'exploration de formes géométriques du bout des doigts, la forme étant imprimée en 3D sur le papier. Ce dispositif est onéreux (coût de la machine à impression, puis coût du papier).

Tactos permet une forme alternative d'accès à la géométrie. Les formes seront générées sur l'écran de l'ordinateur, et les élèves pourront l'explorer grâce au dispositif (voir en figure 16 l'exploration d'un triangle). Cela nécessite bien sûr un investissement, notamment en terme d'heures de formation, mais les avantages sont indéniables : un nombre infini de formes peut notamment être généré sans surcoût et l'élève peut placer de nouveaux traits et points dans la figure (on peut en effet utiliser Tactos en mode dessin), puis explorer l'ajout qu'il vient de faire.

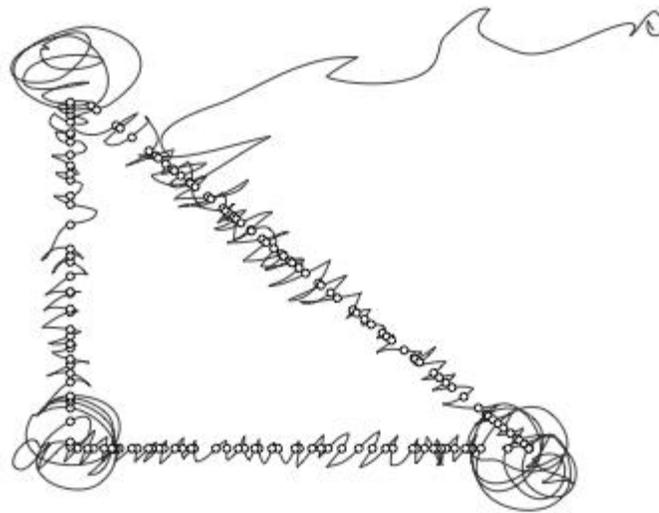


Figure 16. Trajectoire de l'exploration perceptive d'un triangle.

L'utilisation de Tactos comme support de cours de géométrie a été l'utilisation qui a été favorisée lors des suivis longitudinaux, mais il est à noter que Tactos permet aussi l'exploration de cartes (dont la version thermoformée est coûteuse et difficilement accessible), ce qui fournit un support cette fois-ci pour la géographie. Notre dispositif permet alors l'exploration de cartes, mais pas uniquement pour la lecture de points d'information dans un espace (comme une ville sur la carte d'un pays). En effet, il est possible de colorer la carte en zones de différentes couleurs, chacune associée à un élément particulier (par exemple, du bleu pour de l'eau, voir figure 17). Le curseur du dispositif peut être sensibilisé aux couleurs, et ainsi indiquer les différences entre les zones explorées. De nombreuses informations peuvent donc être rendues disponibles : noms de régions, de départements, cours d'eau, mers, océans,... De même, Tactos pourrait fournir des supports pédagogiques pertinents pour les arts.

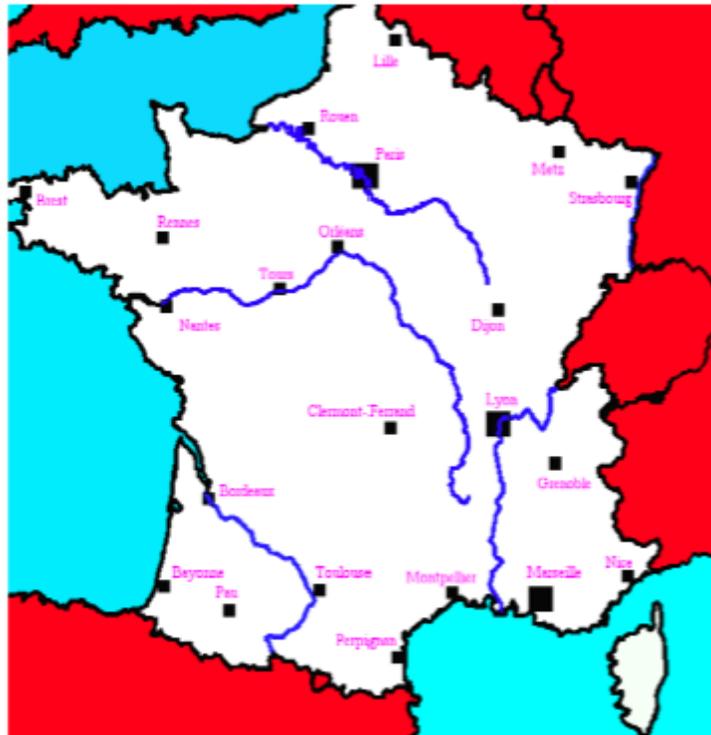


Figure 17. Carte de France enrichie de couleurs et d'informations pour l'exploration par Tactos.

1.1.2.2.3 Émoticônes tactiles

Dans le cadre d'une réflexion sur l'interaction tactile sur le réseau médiatisée par le dispositif Tactos, de nombreuses questions émergent, dont celle de la valeur émotionnelle à attribuer à l'interaction tactile, dans le but de permettre une interaction porteuse de sens entre deux utilisateurs ou plus. Nous consacrerons dans la partie 2 une étude à la valeur émotionnelle de l'interaction tactile, mais dans la perspective de l'utilisation de Tactos, Maillet (2010) a mis en place des émoticônes tactiles, conçues dans le cadre d'un travail exploratoire sur les interactions haptiques (voir figure 18). Il s'interroge sur ce que la modalité tactile peut apporter à la dynamique interactionnelle dans le cadre d'une communication asynchrone comme celle d'un chat. L'objectif de ce travail était de proposer une alternative aux smileys utilisés dans la communication sur Internet. Nous avons déjà évoqué le rôle des smileys dans la partie 1.1.1.3.3 ; ils traduisent notamment l'état émotionnel de l'émetteur. Le but de Maillet est de pouvoir matérialiser les indices non-verbaux et para-verbaux de la communication par le biais de la modalité tactile, afin de pouvoir enrichir les interactions écrites. Ses travaux montrent que les séquences de stimulations proposées pour chaque « taticône » n'ont pas de sens en dehors de ceux qui les ont produites. L'émergence de sens n'est donc possible entre n utilisateurs que dans le cadre d'une convention. Plusieurs questions émergent alors, comme notamment celle des conditions d'émergence d'une signification à partir d'une stimulation tactile abstraite ; nous y reviendrons dans la partie 4.

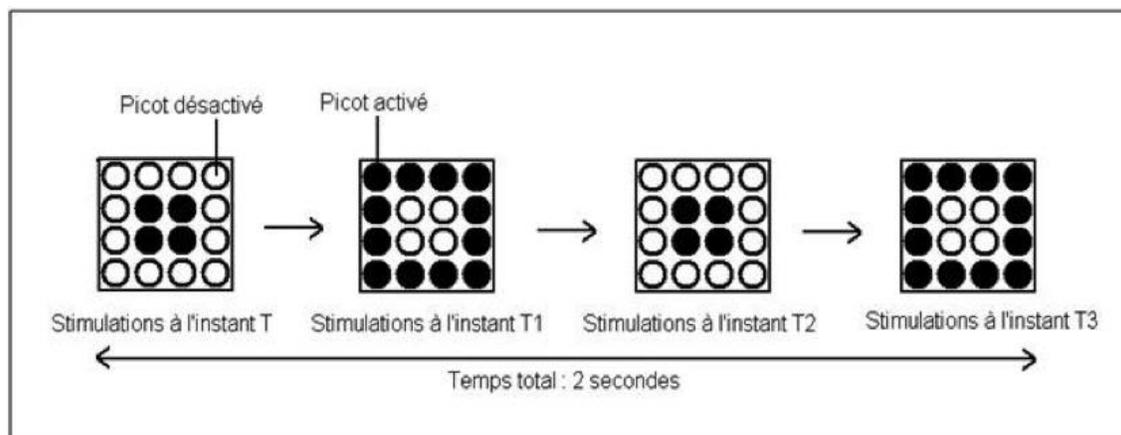


Figure 18. Émoticône tactile proposée par Maillet (2010) représentant la joie.

Il ressort des observations préliminaires de Maillet que les utilisateurs souhaiteraient avoir accès à ces émoticônes tactiles, et éventuellement pouvoir en créer eux-mêmes. Cette

fonction laisserait la possibilité aux utilisateurs de convenir de leur signification dans le cadre d'une interaction.

1.1.2.2.4 *Perspective : Tactos mobile*

Dans le cadre du projet ITOIP, de nouvelles applications pour Tactos sont imaginées, notamment parce qu'une version portative à coupler avec un smartphone est en cours de développement. Ce Tactos portable aura pour but premier de permettre à une personne souffrant de déficience visuelle de pouvoir s'orienter dans la ville grâce aux informations tactiles envoyées par le dispositif. Grâce à la génération de cartes et à la fonction GPS, il sera en effet possible d'indiquer un itinéraire (lequel pourra être enrichi par les utilisateurs eux-mêmes). L'avantage sera de fournir les informations par le biais du canal tactile ; ainsi, l'audition pourra être consacrée à l'attention portée sur l'environnement immédiat.

D'autres applications, notamment dans le domaine ludique, sont en cours de développement. Les retours des premiers utilisateurs sont encourageants ; la bataille navale et le sudoku sont déjà appréciés, et les utilisateurs sont enthousiastes à l'idée de découvrir d'autres jeux.

Une problématique à prendre en considération pour les développements futurs de Tactos est celle de l'acceptation par les utilisateurs. En effet, malgré l'efficacité des dispositifs de suppléance perceptive, ces derniers ont tendance à être délaissés par les utilisateurs. La diffusion de ces dispositifs mène généralement à un échec, tant sur le plan économique que sur le plan social (Lenay, 2006b). Lors de la prise en main du dispositif, les utilisateurs sont tout d'abord enthousiastes, et les résultats sont très encourageants, mais au final les utilisateurs se retrouvent déçus, et ils abandonnent les dispositifs. Ces derniers peuvent en effet d'une part leur sembler inutile ; le monde des aveugles est en lui-même un monde riche et qu'ils vivent comme étant complet. De plus, l'aspect machine peut les déranger, leur donnant le sentiment de présenter aux yeux d'autrui un appareillage monstrueux, suscitant des réactions négatives. Ils ne font de plus pas l'expérience perceptive de la vue par exemple avec le TVSS qu'est celle qu'ont les personnes voyantes. Nous avons en effet déjà vu que tout dispositif de suppléance perceptive ouvre en fait l'accès à une *nouvelle* façon de percevoir. Ce qui fait défaut aux utilisateurs, ce sont aussi les *qualia*, les valeurs émotionnelles des percepts (Bach-y-Rita 2002 ; Lenay, 2006b). Les exemples les plus illustratifs sont le cas de ces adolescents équipés du TVSS à qui l'on fait explorer des images de femmes dénudées ; ils n'ont alors aucune réaction émotionnelle, de même que cet homme qui explore pour la première fois le visage de sa femme (Bach-y-Rita et Kerckel, 2003). Cela s'explique par le fait que les *qualia* ne sont pas des informations présentes

dans le monde et que l'on pourrait capter par le biais d'un canal sensoriel, naturel ou artificiel (Lenay et al., 2000). Les dispositifs de suppléance perceptive proposent de nouveaux percepts, dont le sens et la signification émotionnelle doivent se construire par l'utilisation du nouveau couplage sensori-moteur proposé par les dispositifs, de façon à ce qu'une expérience perceptive riche puisse être mise en place. La mise en place d'un apprentissage précoce et durable pourrait donc être une possibilité pour permettre aux qualia de se constituer.

Mais l'absence de qualia peut aussi s'expliquer par le fait que les dispositifs de suppléance perceptive ne sont jamais utilisés que par des individus, et jamais par des communautés, ce qui conduit à un isolement de l'utilisateur face à la modalité perceptive qui lui est ouverte, inédite pour lui et pour les autres individus (Lenay, 2006b). Les qualia n'émergent que dans le cadre d'une histoire et d'une mémoire collectives, mettant en jeu un réseau complexe d'interactions entre les individus (ibid). Or il n'existe pas de communauté d'utilisateurs de dispositifs de suppléance perceptive, notamment du fait qu'au niveau individuel, les utilisateurs ont du mal à accepter ces sortes de prothèses perceptives, qui nécessitent un appareillage immédiatement identifiable par les autres individus (ibid). Pour que ce défaut puisse être compensé par les qualités du dispositif, il faut pourtant qu'une véritable communauté se constitue, de façon à ce qu'une histoire affective collective et partagée, rendant possible la constitution des qualia, puisse s'écrire (Lenay, Canu et Villon, 1997).

Cette question de l'acceptation par les individus et de la constitution d'une communauté d'utilisateurs est première dans les développements actuels de Tactos ; nous souhaiterions en effet que la version mobile puisse être proposée au plus grand nombre, ce qui inclut les personnes voyantes. Pour cela, les bénéfices à tirer de l'utilisation du dispositif doivent donc compenser la difficulté à maîtriser un outil de ce type, et nous sommes convaincus que la constitution d'une communauté d'utilisateurs est clé dans cette optique. C'est pourquoi nous mettons en place Tactos en réseau : Intertact.

1.1.2.2.2.5 De l'individuel au collectif : Intertact

Au-delà de son usage individuel (navigation dans les interfaces graphiques numériques, exploration de formes inscrites à l'écran,...) ou de son utilisation pour enrichir une conversation de type chat, le dispositif original qu'est Tactos peut également être connecté en réseau et ainsi permettre l'interaction tactile d'utilisateurs distants (voir figure 19). On dénomme la mise de Tactos en réseau « Intertact ».

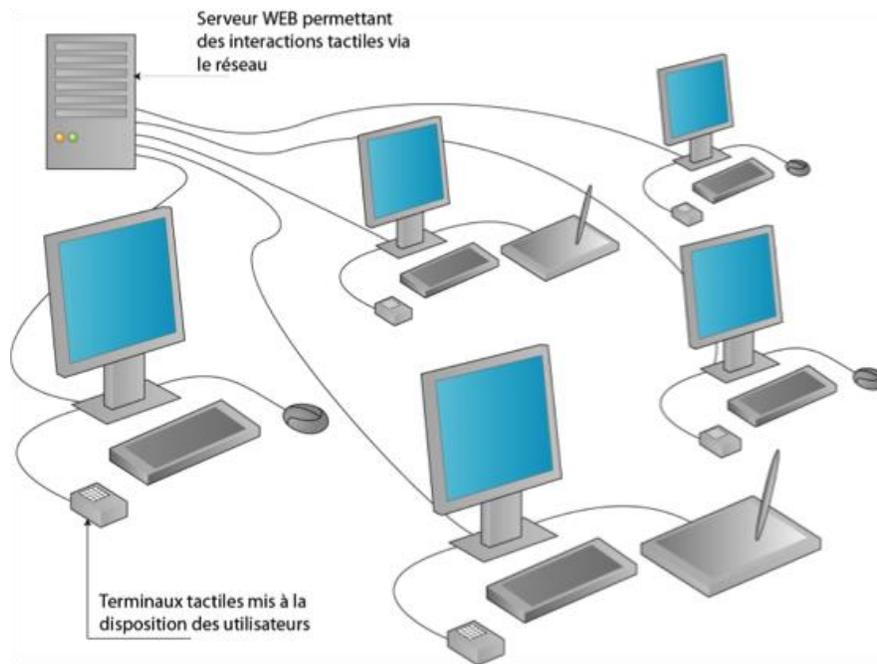


Figure 19. Intertact : un internet tactile en réseau.

Intertact permet la connexion tactile entre plusieurs individus. Entre deux personnes, cela permet notamment une sorte de « caresse » à distance : les utilisateurs peuvent en effet se « toucher » dans un espace numérique partagé (voir figure 20). Intertact permet donc une forme d'interaction tactile à distance. Pour qu'un échange tactile à distance soit possible, le curseur déplacé par chaque utilisateur est augmenté d'un corps-image par rapport à la simple situation d'exploration de formes. En déplaçant son effecteur, l'utilisateur commande son curseur, par lequel il possède un corps percevant (la matrice de champs récepteurs couplée aux stimulateurs tactiles des cellules brailles) et un corps-image, qui est donc un avatar, un corps virtuel, que les autres utilisateurs peuvent percevoir par le biais de leur propre corps percevant. Peuvent être perçus dans l'environnement d'une part les objets présents dans l'espace partagé et d'autre part les avatars (ou corps-images) des autres utilisateurs. Les corps percevants des utilisateurs étant tous deux sensibles aux corps-images des autres, et les corps percevant et corps-image de chacun étant superposés, il en résulte que quand un participant en « touche » un autre, il est en même temps « touché » par cet autre ; il y a donc une forme de croisement perceptif à travers le réseau (Deschamps, Rovira, et Lenay 2008 ; Lenay, 2010).

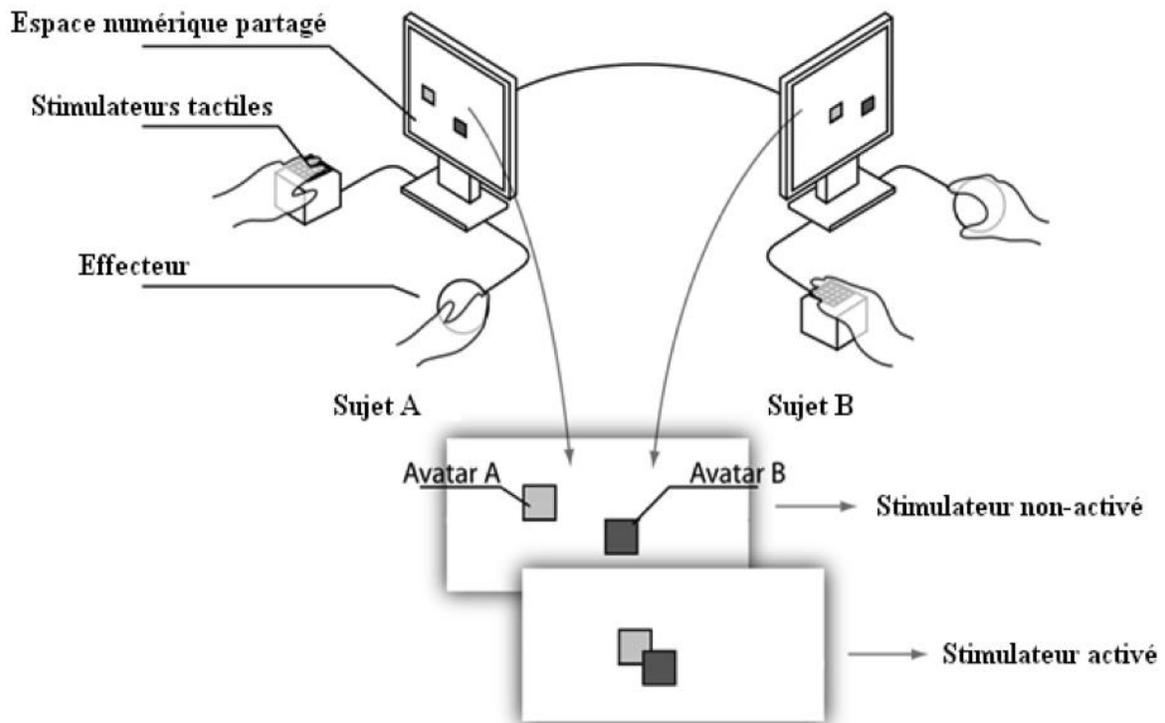


Figure 20. Rencontre tactile entre deux individus dans un espace numérique partagé : lorsque leurs curseurs se croisent, les utilisateurs reçoivent une stimulation tactile.

Cette ouverture d'un usage individuel à un usage collectif rend possible la conception d'espaces numériques pour l'interaction tactile, en proposant des fonctions dédiées notamment à l'apprentissage, à la pratique, à la pédagogie et au jeu (Deschamps, Rovira et Lenay, 2008) (voir figure 21). Cette nouvelle modalité d'interaction tactile ouvre la possibilité de mettre en place une aide pour les personnes mal voyantes, aide pourvue d'une dimension sociale. En parallèle, la conception d'espaces numériques tactiles partagés va de pair avec une réflexion fondamentale sur les interactions perceptives. Nous avons ainsi mis en place une étude expérimentale visant à fournir les étapes-clés nécessaires à la caractérisation des processus d'engagement mutuel dans les interactions¹². Les résultats suggèrent que ces processus constituent des dynamiques mutuelles, ce qui est la base d'une co-construction active de sens (Deschamps et al. 2012 ; pour plus de détails, voir Deschamps, 2013).

¹² Cette étude expérimentale a été exposée dans une communication orale avec actes :
- Deschamps, L., Le Bihan, G., Lenay, C., Rovira, K., Stewart, J. & Aubert, D. (2012). Interpersonal recognition through mediated tactile interaction, Proceedings of IEEE Haptics Symposium, Vancouver, Canada. March 4-7, 2012. pp. 239 - 245

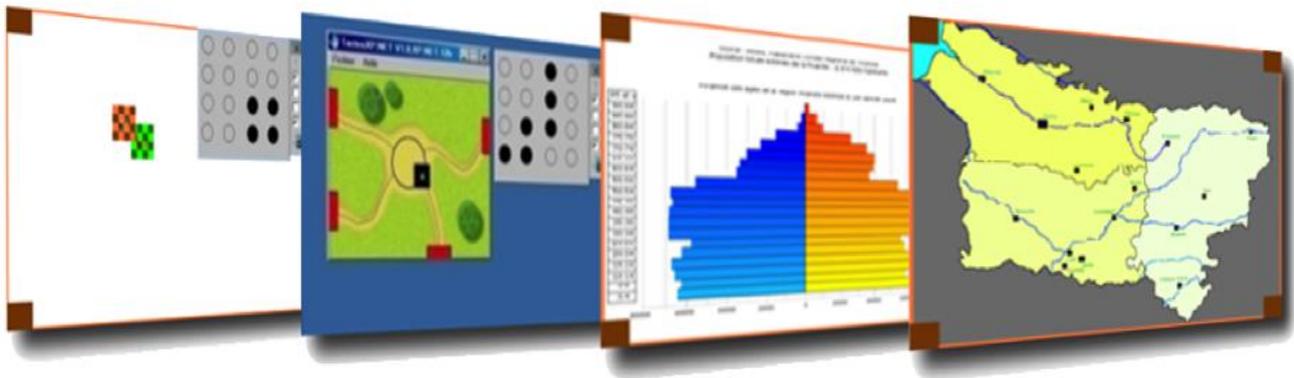


Figure 21. Espaces numériques partagés par Intertact.

En effet, lors d'un usage collectif de Tactos, l'interaction entre des utilisateurs distants est rendue possible. Ils peuvent ainsi accéder à des contenus partagés et mettre en place des communications interpersonnelles au travers de la modalité perceptive nouvelle ouverte par Tactos. Nous souhaitons donc mettre en place des espaces numériques partagés, accessibles de chez soi par le biais d'une connexion internet, et hébergés sur un serveur dédié. Le projet Portintertact, financé par la région Picardie de décembre 2007 à décembre 2009, a permis l'élaboration d'un portail internet d'interaction tactile. Le but de ce projet est de permettre la mise en place d'une communauté d'utilisateurs, où les utilisateurs vont pouvoir écrire ensemble une histoire collective, où significations et valeurs partagées autour du dispositif pourront se développer.

Des analyses d'usage préliminaires ont été menées afin de cerner les questions d'usage et de pouvoir orienter les choix de conception afin que le dispositif Tactos puisse être adapté de façon à permettre des interactions tactiles distales. C'est lors de ces travaux que la version actuelle du MIT (Module d'Interaction Tactile) a été développée (voir figure 22). Il comprend un touchpad absolu (réglé de façon à ce que chaque position sur le touchpad corresponde une position sur l'écran), deux matrices de stimulateurs tactiles (chaque matrice est constituée de deux cellules braille) équipées de boutons, et une nappe de capteurs de pression (FSR – Force Sensing Resistors).

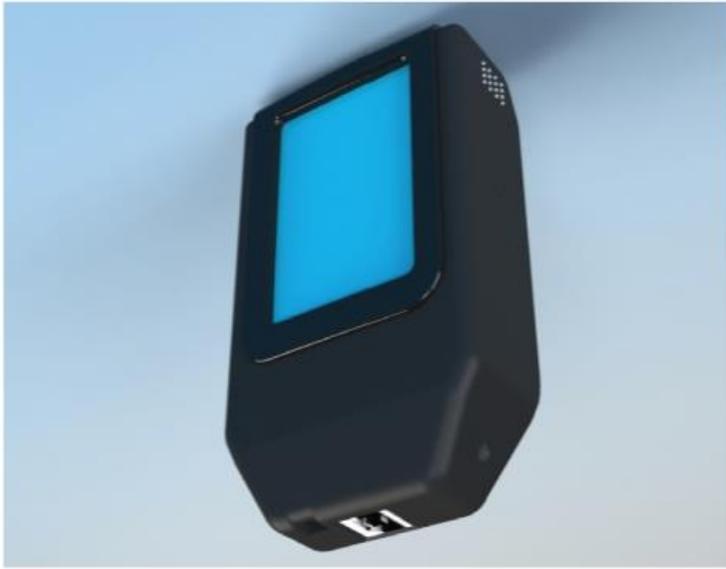


Figure 22. Les deux faces du Module d'Interaction Tactile.

Le projet Intertact rend donc possible la connexion simultanée d'utilisateurs distants, les faisant se rencontrer dans des espaces numériques partagés (voir figure 23). Les versions du prototype sont de plus en plus développées ; un prototype de démonstration est disponible à l'adresse suivante : <http://www.intertact.net/> (section Démonstration). Les versions utilisées au laboratoire avec nos utilisateurs rendent déjà possible l'hébergement de contenus numériques accessibles, l'exploration et le partage de ces informations, et plus généralement, l'interaction tactile entre des utilisateurs distants. Pour en savoir plus quant aux résultats des premières analyses d'usage et aux problématiques de conception qui se posent (comme par exemple celle de la discrimination et la reconnaissance interpersonnelle des utilisateurs), consulter Deschamps (2013).

Les évolutions à venir de Tactos, tant la version mobile que la version en réseau, sont destinées avant tout à des personnes présentant un déficit visuel ; nous les voulons cependant accessibles de la même façon aux personnes voyantes. La question de l'accessibilité est centrale, mais nous ne voulons en aucun cas isoler les utilisateurs malvoyants. Nous aimerions qu'une communauté d'usage se constitue, et que cette communauté dépasse le cadre du handicap. Ce projet est donc pensé avec une perspective de *design for all*. En conception, ce principe consiste à avoir pour objectif de concevoir pour tous (ce qui inclut les générations futures), sans considération pour l'âge, le genre, les capacités ou la culture des utilisateurs.

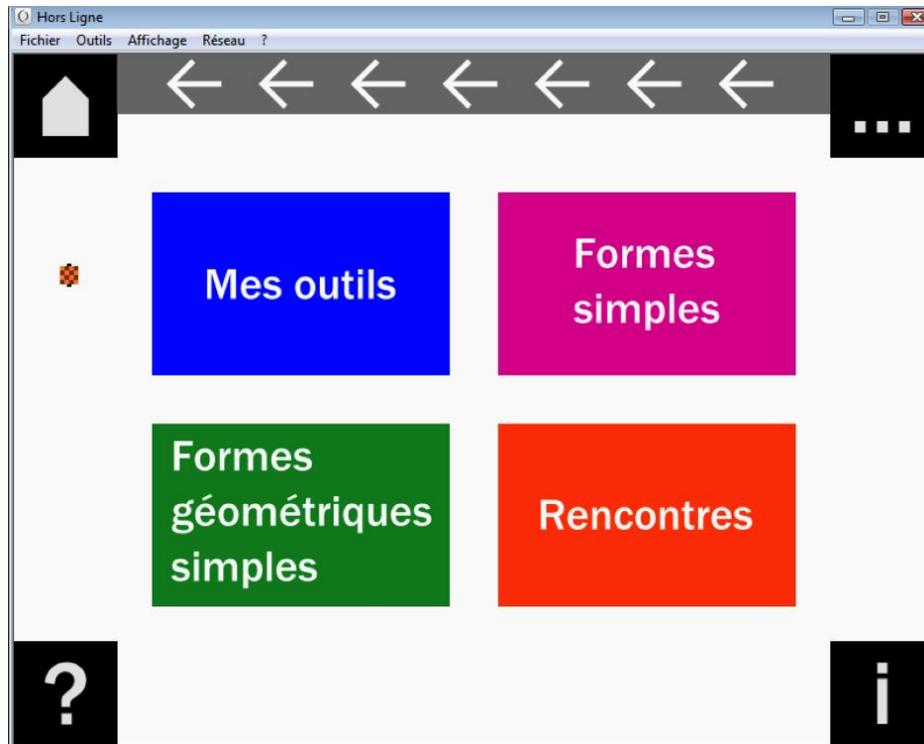


Figure 23. Exemple d'une salle Intertact où se trouve un utilisateur. Chaque cadre représente une porte, où l'utilisateur peut accéder à du contenu. Les coins constituent une structure commune à toutes les pages de navigation et de contenu (salle d'accueil, titre de la salle actuelle, aide, information sur la page actuelle ; le retour à la page précédente se fait en cliquant sur la bande supérieure).

1.1.3 Communication et interaction à distance

1.1.3.1 La médiatisation pour le croisement perceptif

Les dispositifs de communication à distance sont très variés, et les premières versions proposées remontent loin dans le temps (par exemple, l'écriture permet la communication à distance : cette invention majeure remonte à plus de trois millénaires avant notre ère (Herrenschmidt, 1999 ; Leroi-Gourhan, 1964 ; Mugnaioni, 2006)). Il est établi que pour pouvoir communiquer à distance, il faut avoir recours à des outils et techniques permettant de faire transiter un message (quelque soit sa nature) au destinataire, et réciproquement. Il est à noter que l'horizon de nos interactions n'est pas uniquement la transmission d'informations (ce n'est qu'un moyen), mais plutôt le développement de relations. Pour Lévinas (in Ponzio, 1996), le langage remplit avant tout une fonction de mise en relation. Pour lui, ce qui compte dans nos interactions, qu'elles soient orales ou écrites, c'est la *présence* d'autrui. Nous souhaitons à présent nous poser

la question de la médiatisation pour le croisement perceptif, qui consiste en une perception mutuelle de deux interactants. En effet, la situation de croisement perceptif permet une implication particulière dans l'interaction ; elle met autrui au centre de notre perception, tandis que lui ou elle aussi nous met au centre de la sienne. Comment médiatiser cette situation de croisement perceptif, ce qui permettrait une meilleure immersion de l'utilisateur dans l'espace interactif ouvert par le dispositif, en permettant un engagement fort avec l'autre interactant ?

Nous avons vu que par le biais d'Intertact, deux utilisateurs, ou plus, peuvent se connecter en réseau et interagir par le biais de Tactos. Ce dernier constitue un moyen idéal d'étudier la situation de croisement perceptif, puisqu'il permet de savoir ce que les utilisateurs ont senti, ce qu'ils ont fait, et comment ils ont interagi. Tactos permet des rencontres tactiles entre deux individus, ou même plus, dans un espace virtuel partagé. Les expérimentations avec Tactos présentent pour notre travail le très grand intérêt d'avoir lieu à distance. En effet, lors de ces expériences, les sujets sont isolés dans des pièces distantes, et ils ne peuvent interagir que par le biais du dispositif. Tactos constitue donc un nouveau moyen de communication à distance, et nous permet d'étudier la situation de croisement perceptif dans des conditions minimalistes.

Nous souhaitons aider à la conception de dispositifs pour la communication émotionnelle à distance ; l'établissement d'un croisement perceptif constitue un premier pas dans cette direction, afin de permettre aux interactants de s'immerger dans l'interaction en percevant la présence de l'autre. D'autre part, nous nous concentrons plus particulièrement sur l'idée de permettre aux individus une forme de toucher à distance, pour des raisons que nous détaillerons dans la partie 2.

Deux résultats fondamentaux pour la conception d'interfaces haptiques ont pu être apportés par les études expérimentales du croisement perceptif :

- la constitution d'autrui comme d'un alter-ego, l'expérience de sa présence par le biais du croisement perceptif, ne se fait ni dans la perception de son image, ni dans la perception de son comportement objectif. C'est la perception d'une activité perceptive dirigée vers moi qui me permet de reconnaître autrui.
- pour que le croisement perceptif médiatisé au travers d'une interface soit possible, il faut que l'interface permette une rencontre dans un espace d'interaction partagé. La rencontre implique de pouvoir déplacer un avatar composé d'un corps-image (qui peut être perçu par autrui) et un corps percevant (le champ récepteur). Le croisement perceptif est une activité perceptive, et il doit donc se déployer à la fois dans l'espace et dans le temps.

Ces résultats sont majeurs dans l'idée d'établir des recommandations et des spécifications techniques de systèmes d'interactions interpersonnelles à distance. On voit avec Tactos que par un déploiement spatial et temporel de l'activité perceptive, les utilisateurs arrivent à compenser la pauvreté du stimulus, et ce non seulement dans le cas de la perception de formes, mais aussi

dans le cas de la perception du contact avec autrui. Du point de vue de la conception, cela implique qu'il faut se concentrer sur les possibilités offertes par le dispositif à concevoir, sur ce qu'il offre à sentir et à agir, et sur la nouvelle perception permise par la saisie du dispositif.

Maillet (2010) a proposé un stimulateur permettant une version alternative de Tactos. Ce stimulateur permet d'une part de recevoir des signaux tactiles grâce à deux cellules brailles (formant un total de 4 x 4 picots, exactement comme pour le MIT), et de commander les picots grâce à une matrice de 4 x 4 capteurs de pression (FSR – Force Sensing Resistors) (voir figure 24). Pour une utilisation à deux, il existe un mode « caresse » : deux stimulateurs sont reliés, et les cellules brailles de l'un sont connectées avec les capteurs de pression de l'autre participant, et inversement (voir figure 25). Ainsi, un utilisateur peut envoyer des signaux tactiles à son partenaire en appuyant sur la matrice de capteurs ; lorsque ses cellules brailles s'activent, c'est que son partenaire est en train d'appuyer sur sa propre matrice de capteurs. Ainsi, ce stimulateur permet une certaine forme de caresse à distance. La dynamique d'interaction alors mise en place par les sujets, notamment par le biais de jeux de mimétisme, conduit à une situation de croisement perceptif, étant donné que les stimulations reçues par chaque sujet sont associées à l'activité du partenaire, et inversement. L'expérience du contact avec autrui en tant qu'autre n'étant pas celle de son image, ni celle de son comportement, elle se fait via la reconnaissance d'une activité perceptive dirigée vers moi. Avec ce mode caresse, les utilisateurs peuvent constituer une telle forme de contact avec leur partenaire, et ainsi faire l'expérience de sa présence, ce qui participe de l'immersion dans l'interaction. C'est à cette expérience qu'est associée l'émotion, le caractère *touchant* du contact (Lenay, 2010).

Nous avons pour but de permettre l'échange émotionnel à distance ; le croisement perceptif nous semble être une façon de l'atteindre. Nous préconisons donc la conception d'interfaces permettant le croisement perceptif à distance, afin d'enrichir l'immersion des utilisateurs, et de pouvoir leur donner la possibilité de faire l'opportunité de la présence d'autrui, et ce malgré la distance.



Figure 24. Le module d'interaction haptique de Maillet (2010).

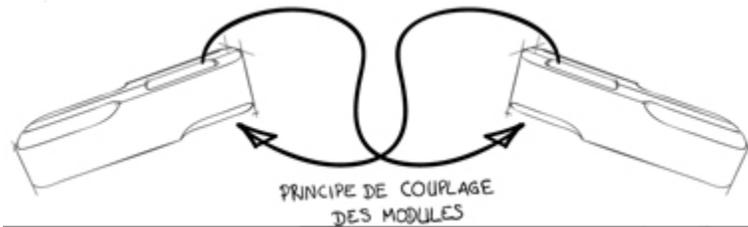


Figure 25. Couplage des modules pour la « caresse ».

1.1.3.2 Sentiment de présence, connectedness, relatedness, social awareness,... : la présence à distance

L'immersion et la communication des émotions fonctionnent de concert. Classiquement, en réalité virtuelle on distingue le sentiment d'immersion (être dans l'espace, agir dans l'espace de ma perception) et le sentiment de présence. Il s'agit avant tout d'une perspective psychologique individuelle, liée à la perception individuelle. Mais la communication des émotions réside en très grande partie dans l'interindividuel. La perception individuelle et l'interaction avec autrui sont toutes deux nécessaires pour une communication émotionnelle à distance.

La notion de présence (*presence*, ou *presence in absence*) est ainsi centrale dans l'étude des interactions à distance. En effet, c'est par elle que l'on exprime le sentiment qu'autrui est

présent, là dans l'interaction, même s'il se trouve géographiquement dans une localisation différente. Sentiment de présence, connectedness, relatedness, social awareness : toutes ces notions se rapportent à l'idée qu'une présence à distance est possible par le biais d'interfaces adaptées. Nous posons que le sentiment de présence est un premier pas possible vers un échange riche sur le plan émotionnel ; il permet d'avoir le sentiment d'être connecté à l'autre interactant.

La notion de présence provient de la recherche sur les outils de travail collaboratif à distance et les espaces virtuels partagés. Elle est définie par Lombard et Ditton (1997) comme une sorte d'illusion perceptive qui consiste à avoir l'impression que les interactions ne sont pas médiatisées. D'un point de vue subjectif, cela consiste à faire l'expérience d'être ensemble dans un même espace ou environnement, même si l'on n'est pas physiquement au même endroit (Witmer et Singer, 1998).

On voit donc que les définitions de la présence peuvent légèrement varier ; la définition de Witmer et Singer fait référence au « sentiment d'être là », tandis que celle de Lombard et Ditton met en avant l'illusion perceptive de la non-médiation.

Short et al définissent quant à eux la présence sociale comme le « degree of salience of the other person in a mediated communication and the consequent salience of their interpersonal interactions » (Short, Williams et Christie, 1976, p.65).¹³

Lombard et Ditton (1997) ont identifié deux catégories principales de concepts autour de la présence : les concepts physiques, et les concepts sociaux. La catégorie physique fait référence au sens d'être physiquement situé quelque part, tandis que la catégorie sociale fait référence au fait d'être ensemble et de communiquer avec la personne avec laquelle on est. La présence est une notion qui peut aussi se discuter en termes de contact (le niveau social), de contenu (les informations) et de contexte (les circonstances de l'expérience) (Howard et al. 2006). Toutes ces facettes sont à prendre en compte pour concevoir des dispositifs permettant la présence.

La notion de présence met bien en avant l'écart qu'il existe entre l'utilisation de dispositifs d'interaction à distance et la richesse vécue au cours d'interactions en face à face (Olson et Olson, 2000). Plusieurs travaux (Short, Williams et Christie, 1976 ; Minsky, 1980 ; Zahorik et Jenison, 1998 ; Witmer et Singer, 1998) mettent en avant le fait que les utilisateurs font plus facilement l'expérience de la présence à distance si l'interface leur offre la possibilité d'avoir des contacts informels et impromptus ; c'est sans doute ce qui explique le succès de la fonction poke de Facebook, que l'on pourrait rapprocher d'une tape amicale virtuelle. La dimension relationnelle des interactions contribue aussi à faire disparaître la frontière entre interactions médiatisées et interactions en face à face.

¹³ Traduction personnelle : « Le degré de saillance de l'autre individu lors d'une communication médiatisée, et par conséquent la saillance des interactions interpersonnelles. »

La notion de *connectedness*, ou plus exactement de *connectedness oriented communication* (communication orientée vers le sentiment d'être connecté) a été proposée par Kuwabara et al. (2002) ; elle vise à mettre en avant les aspects de la communication permettant et enrichissant les relations sociales entre les individus en encourageant le sentiment qu'il existe une véritable connexion entre eux. L'idée est donc non pas de se focaliser sur la transmission du contenu de messages (ainsi que le font les réseaux de communication traditionnels), mais sur le sentiment d'être connecté, sentiment qui est encouragé par l'échange d'information de présence et de statut, ainsi que par des interfaces qui se font oublier (Kuwabara et al. 2002 ; Rettie, 2003 ; Octavia, van den Hoven et De Mondt, 2007).

Pettigrew (2009) s'est penché sur la question des SMS (Short Message Service) et a comparé ce mode de communication à l'échange verbal téléphonique. Il montre que les SMS sont avant tout utilisés pour engager, enrichir, maintenir, en bref impacter les relations interpersonnelles, et il s'intéresse à l'usage des SMS entre des dyades composées de personnes ayant une relation interpersonnelle très proche. Les résultats de cette étude montrent que :

- les utilisateurs trouvent que l'échange de SMS est plus régulier et privé que les interactions vocales au téléphone ;
- les SMS peuvent être utilisés à la fois pour affirmer son autonomie et pour maintenir le sentiment d'être connecté ;
- les dyades dont les membres sont des couples et celles dont les membres ne le sont pas perçoivent les bénéfices de l'échange de SMS de manière différente (ibid).

On notera que le besoin de se sentir connecté aux autres individus (pas spécifiquement dans une optique d'interaction à distance) est un besoin identifié et reconnu en psychologie. Adler et Brett (1999) proposent le concept d'intérêt social, ou *social connectedness*, qui se réfère aux attitudes d'un individu et à ses relations en société ; ce sont en effet ces éléments qui vont déterminer à la fois son succès et sa santé mentale. En psychologie sociale, Smith et Mackie (2007) posent la recherche de *connectedness* comme étant l'un des trois principes basiques de motivation qui sous-tendent le comportement social. On comprend donc qu'une possibilité à garder en tête pour la conception d'interfaces pour la communication émotionnelle à distance est de proposer des interfaces qui répondent à ce besoin de *connectedness*.

Le fait de se sentir proches, liés, ou *relatedness*, est aussi important pour le bien-être psychologique des individus (Hassenzahl et al. 2012). Sheldon et al. (2001, p. 339) définissent que la *relatedness* est accomplie lorsque que l'on a le « sentiment que l'on a des contacts intimes

réguliers avec des gens qui se soucient de nous, par opposition à l'impression de se sentir seul et délaissé » (traduction personnelle)¹⁴.

L'*awareness* est définie par Dourish et Bly (1992) comme la compréhension de l'activité des autres, ce qui fournit un contexte pour sa propre activité. C'est donc « the state of knowing about the environment in which you exist; about your surroundings, and the presence and activities of others » (Wisneski et al. 1998, p. 24)¹⁵. Ce concept est particulièrement mis en avant dans la recherche en CSCW, pour Computer Supported Cooperative Work (Heath et al. 2002 ; Luff, Heath et Sanchez Svensson, 2008).

En général, les dispositifs ambiants (voir Dunne et Raby, 1994 ; Ishii et Ullmer, 1997 ; Wisneski et al. 1998) sont liés de près à l'*awareness*, étant donné qu'ils visent une expérience périphérique, implicite. En effet, et comme l'implique l'idée de périphérique, ces dispositifs sont conçus de façon à ne pas requérir d'attention primaire et sont donc discrets. Ils s'insèrent par conséquent dans les routines et activités quotidiennes sans causer de rupture, ni pour l'expéditeur ni pour le receveur. Ishii (2008) oppose ces dispositifs aux technologies explicites basées sur la conversation, comme par exemple le téléphone. Les systèmes orientés pour l'*awareness* portent plus d'attention au fait de donner à vivre un sentiment de connexion, ou un sentiment d'être en contact, qu'au contenu même de la communication.

Durlach et Slater (2000) posent le concept de *togetherness* (le fait d'être ensemble) dans un environnement virtuel. Ils le définissent comme « the sense of people being together in a shared space, which is the counterpart for shared VEs to the presence of an individual in a VE » (Durlach et Slater, 2000, p.214)¹⁶.

L'idée d'intimité n'est pas définie de façon unanime, mais Usoh et al. (2000), suite à une analyse des recherches autour de ce concept, listent huit aspects remarquables de l'intimité, qui sont : l'intimité physique, la communication non verbale, la révélation de soi, la présence, l'intimité cognitive, l'intimité affective, l'engagement et la réciprocité.

Toutes ces notions de présence à distance et de connexion entre les individus apportent une aide certaine pour la conception d'interfaces pour la communication émotionnelle à distance ; elles offrent en effet de nombreuses pistes d'éléments à prendre en considération pour une telle conception. Notons que des interfaces à retour tactile ont été développées en vue

¹⁴ « Feeling that you have regular intimate contact with people who care about you rather than feeling lonely and uncared for » (Sheldon et al, 2001, p.339).

¹⁵ Traduction personnelle : « L'état de connaissance de l'environnement dans lequel on existe ; ce qui nous entoure, ainsi que la présence et l'activité des autres ».

¹⁶ Traduction personnelle : « Le sentiment pour les individus d'être ensemble dans un espace partagé, ce qui en fait la contrepartie de l'immersion pour un individu dans un environnement virtuel, mais étendue à un environnement virtuel partagé ».

d'augmenter le sentiment de présence (Oakley, ABrewster et Gray, 2001) et pour exprimer et interpréter les émotions (Bailenson et al. 2007) ; nous en ferons un état de l'art dans la partie 2.2.1.3.1.1.

1.1.3.3 La communication des émotions à distance

Nous avons déjà traité de la communication à distance au travers de l'écrit (voir partie 1.1.1.3.3). C'est actuellement une des formes de communication des émotions à distance les plus répandues, avec le téléphone et la visiophonie, mais nous avons vu que malheureusement ces types de communication n'ont pas été conçus avec l'idée que transmettre les émotions à distance était un but ; ainsi la transmission des émotions est possible dans une certaine mesure, mais elle n'est pas très efficace et doit être améliorée.

L'exemple de EmoHearts (Neviarouskaya, Prendinger et Ishizuka, 2010), développé pour Second Life, est une proposition originale d'explicitation des émotions véhiculées au travers de l'écrit. Second Life est un jeu vidéo mettant en scène un univers virtuel en 3D, sorti en France en 2003 ; les joueurs y bâtissent une « seconde vie » virtuelle. Ils sont encouragés à établir et renforcer des relations interpersonnelles, à partager des idées, à faire de nouvelles expériences et bien sûr à ressentir d'authentiques émotions à mesure qu'ils vivent les aventures de la réalité virtuelle proposée par le jeu. Étant donné que les émotions jouent un rôle puissant dans la communication, les auteurs se posent la question de savoir comment déclencher un affichage visuel de l'état émotionnel des utilisateurs dans un monde virtuel. Il faudrait à priori que le joueur assigne manuellement une expression faciale ou un geste à son propre avatar ; cela serait comme choisir son statut dans une liste, ce que l'on fait avec les logiciels de messagerie instantanée. La détection de l'affect à partir du texte, qui permettrait l'expression automatique des émotions dans un monde virtuel, est donc une méthode qui évite le contrôle manuel par l'utilisateur et qui enrichit sans effort les communications à distance. Les auteurs ont mis en place une approche basée sur les règles lexicales pour pouvoir reconnaître les émotions à partir d'un texte et ont conçu une application Second Life du Modèle d'Analyse Affective développé. Cette application, appelée EmoHearts, se base sur les résultats du modèle en fonction de ce qu'écrit l'utilisateur, pour déclencher les animations des expressions faciales de l'avatar, et figure les émotions par des textures en forme de cœur (voir figure 26).

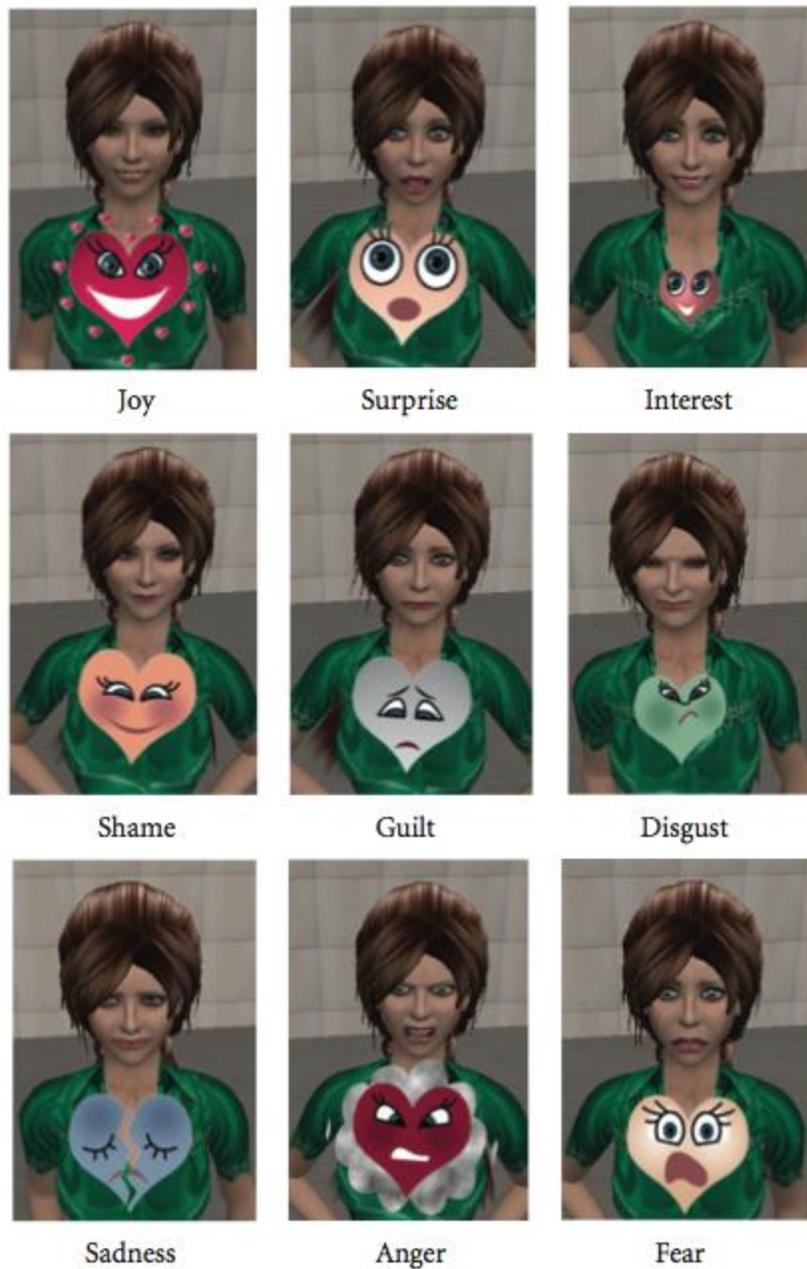


Figure 26. L'expression des émotions par EmoHearts dans Second Life à partir du texte tapé par l'utilisateur.

Dans le domaine de l'affective computing, D'Mello et Calvo analysent la place des émotions de base, ou primaires, comme étant trop centrale ; d'après leurs résultats, des émotions secondaires telles que l'ennui, la confusion et la frustration se produisent cinq fois plus souvent que les émotions de base dans une situation d'utilisation d'un ordinateur (D'Mello et Calvo,

2013). Mettre l'accent sur les émotions secondaires permettrait une utilisation des nouvelles technologies plus émotionnelle, tant lors d'une utilisation individuelle qu'en interaction à distance avec d'autres.

La question de la communication des émotions à distance présente donc des pistes d'exploration, que ce soit en ce qui concerne la gestion du texte ou le type d'émotions à prendre en considération. Nous choisissons la piste de l'intégration du croisement perceptif, et la déclinerons en essayant notamment de proposer une forme de toucher à distance.

1.2 Méthodologie

1.2.1 Méthodologie globale : au croisement des disciplines

Ainsi que nous l'avons présenté dans la partie 0.4, notre recherche s'organise autour de trois pôles : théorie, expérimentation, et design. Il s'agit d'articuler une recherche appliquée, centrée sur la conception de dispositifs et sur leurs usages, à une recherche fondamentale, centrée sur les questions théoriques soulevées par l'usage, et mise à l'épreuve par le biais de l'expérimentation. Notre démarche est donc résolument au croisement de plusieurs disciplines et pratiques. Il s'agit pour nous d'articuler les divers points de vue et de les enrichir les uns les autres.

Le bénéfice premier est d'avoir une approche globale de la question de la transmission des émotions à distance. La diversité des points de vue permet la prise en compte de différents aspects et de répondre à des questions de natures différentes. En design, la pluridisciplinarité permet la génération d'idées stimulantes et innovantes (Bødker, Nielsen et Petersen, 2000). Tout nouvel objet ouvrant un espace d'usages possibles, la conception de l'objet et du monde qu'il ouvre nécessite une analyse de divers points de vue. L'étude de la communication est de plus un espace scientifique fertile et central, qui nécessite d'adopter de multiples points de vue.

Nous avons présenté jusqu'ici dans cette partie 1 les éléments de sciences cognitives que nous mettons en jeu dans notre recherche. Nous allons à présent exposer les aspects design et expérimentation, avant de présenter le cadre minimaliste dans lequel nos travaux s'intègrent.

1.2.2 Méthodologies design

Jocelyn de Noblet explique qu'à l'origine, le mot *design* provient de l'anglais et signifie une conception, non plus artisanale, mais industrielle (Noblet, 1988), mais que la notion de design comme on l'entend en français n'existe, historiquement parlant, qu'à partir du moment où

l'on commence à parler d'esthétique industrielle (ibid). Il complète qu'il faut ici entendre le mot *esthétique* dans son sens originaire : « ce que je perçois » (ibid). Le design pose une problématique liant à la fois esthétique et industrie, dans un contexte de production de masse, de qualité et de modernité. Le design concerne le processus global de la conception. Il ne s'intéresse pas uniquement à l'aspect esthétique de l'objet à concevoir, mais prend aussi en compte bien d'autres considérations et dimensions (fonctionnelles, techniques, ergonomiques, environnementales, biologiques, économiques, sociales,...). Tout objet s'inscrit dans la culture et la société, et se retrouve donc dans un contexte d'une grande complexité.

La conception est définie comme une activité de création qui part des besoins exprimés et des connaissances existantes, et aboutit à la définition d'un produit satisfaisant ces besoins d'une part, et étant industriellement réalisable d'autre part.

L'objet du design n'est donc pas que la mise en jeu de la créativité. La participation et l'expérience font aussi partie du processus (Binder et al. 2012). Ainsi, il existe des méthodes en design ; le travail du designer n'est pas un travail de créativité pure.

Nous nous concentrons avant tout sur la conception d'interfaces permettant la communication à distance, notamment sur leur état tenu (mais aussi sur leur état déposé). Ces interfaces que nous considérons sont des interfaces pour l'interaction, notamment pour l'interaction tactile. Nous considérons le rôle de l'utilisateur comme étant central dans la conception, notamment du fait que son intégration engendre une itération plus efficace, permettant le développement d'un meilleur dispositif.

1.2.2.1 Conception et appropriation des interfaces

Les outils ne sont pas donnés d'emblée ; les utilisateurs les investissent peu à peu, s'en saisissent et se les approprient. Les usages ne sont pas uniquement déterminés par les aspects fonctionnels et techniques de l'outil spécifiés par le concepteur ; les utilisateurs inventent de nouvelles fonctionnalités et donnent des sens nouveaux aux dispositifs (Rabardel, 1995). La prise en compte des usages dans la conception nous paraît donc d'une grande importance, ainsi que la compréhension du processus d'appropriation par les usagers.

Le processus de conception peut se dérouler de plusieurs manières. Une des approches possibles consiste à passer par plusieurs évolutions et transformations de l'objet. Cette approche a été modélisée par Pahl et Beitz (1996), qui décomposent le processus en quatre étapes :

- Analyse du besoin : on recense les exigences que le produit doit satisfaire.

- Recherche de concepts : c'est la phase de créativité. On propose des solutions possibles qui satisfont tout ou partie des exigences, on les combine, et on en fait une sélection, afin de ne garder qu'une solution de principe à l'issue de l'étape.
- Conception architecturale : on détermine l'architecture du produit, en choisissant ses composants, leur organisation et leurs relations. C'est là que l'on choisit les formes, les matériaux, les procédés de fabrication et d'assemblage. (Il faut veiller à ce que les choix permettent de respecter l'ensemble des contraintes liées au coût, aux délais, à la qualité,...)
- Conception détaillée : on réalise les plans du produit et on met en place les processus de fabrication et de contrôle de la qualité. On évalue alors la solution pour s'assurer qu'elle réponde bien à l'analyse du besoin qui avait été menée en première étape.

Ce modèle fait office de référence, et présente l'avantage d'être bien structuré, mais il est aussi grandement critiqué. Tout d'abord, il met en avant un processus linéaire, assez éloigné de l'activité réellement déployée par les concepteurs (Micaëlli et Forest, 2003). Mais surtout, on y voit que les choix de conception mis en jeu au départ ne sont remis en cause et évalués qu'à la toute fin du processus de conception, ce qui retarde d'autant l'acceptation ou le rejet de la solution.

Notons qu'il existe ces dernières années une tendance à voir s'effectuer un transfert de disciplines associées aux sciences humaines vers les phases les plus amont de la conception (Duchamp, 1988). C'est par exemple le cas de l'ergonomie et des sciences cognitives, mais aussi de la psychologie ou encore de l'ethnométhodologie. De même, Deckers et al. (2011) soulignent l'aide apportée par la phénoménologie et l'importance du croisement perceptif en design pour parvenir à l'implication de l'utilisateur. Le but de l'articulation de ces disciplines autour du processus de conception est de parvenir à innover d'une part, et à anticiper d'autre part, en vue de parvenir à produire un dispositif le plus efficace possible de la manière la plus efficace possible.

Nous souhaitons concevoir un dispositif de communication à distance riche sur le plan émotionnel. Le système doit donc permettre de mettre en relation au moins deux individus qui ne se trouvent pas dans le même espace physique, en proposant un couplage entre les actions faites par les utilisateurs et les sensations qu'ils reçoivent. La question de l'appropriation de l'interface se pose alors. Nous avons déjà évoqué dans la partie 1.1.1.1 le caractère actif de la perception. Nous avons aussi mis en avant la dualité entre l'interface tenue, constituante pour le sujet d'une expérience perceptive, et l'interface lâchée, constituée dans le milieu perceptif du sujet (voir partie 1.1.1.2). Nous envisageons donc le processus d'appropriation sous la forme de cycles entre ces deux modes d'existence de l'interface. L'intégration sensori-motrice de l'outil se met en place par l'apprentissage et la maîtrise des relations entre les actions faites avec cet outil et les

retours sensoriels dus à ces actions. L'appropriation de l'interface s'articule donc en un cycle alternant les moments où l'interface est tenue, et ceux où elle est déposée.

1.2.2.2 Concevoir pour l'interaction

Rabardel (1995) prône le fait que la conception doit continuer dans l'usage, par le biais d'observations en situation (et non uniquement pour test dans des situations expérimentales). En effet, l'usage est toujours relatif à un certain contexte, il est toujours situé ; les situations expérimentales constituent un autre contexte, différent, ne permettant pas l'accès à toutes les actions des utilisateurs, qui font toujours preuve de créativité et d'inventivité, ne se saisissant pas nécessairement de l'objet conçu ainsi que cela avait été pensé. Introduire un objet en situation permet d'observer la façon dont il modifie l'activité des individus. Les utilisateurs sont directement placés dans un contexte de construction d'un nouvel usage, et ils sont intégrés dans le cycle de conception. Une approche centrée sur l'action offre la possibilité d'identifier le type d'expérience perceptive rendue possible par l'interface, et de spécifier un espace sensori-moteur qui fasse sens pour l'utilisateur.

Bill Verplank et Bill Moggridge, les fondateurs du terme *interaction design* (un design focalisé sur l'interaction entre homme et ordinateur, et par extension entre homme et contenu dématérialisé, et entre hommes par le biais d'ordinateur), posent l'utilisateur comme étant engagé dans le monde. Le travail du designer est de s'interroger sur les capacités d'agir et de sentir de l'utilisateur, et de concevoir des interfaces s'intégrant dans la boucle sensori-motrice et donnant accès à une nouvelle expérience (voir figure 27) (Moggridge, 2006). L'interaction design nécessite donc pour le designer de répondre à trois questions :

- Comment fait-on ? De quelle façon affecte-t-on le monde : en le touchant du doigt, en le manipulant, en s'asseyant dessus ?...
- Comment sent-on ? Que sent-on du monde, et quelles sont les qualités sensorielles qui donnent forme au numérique ?
- Que sait-on ? De quelles façons peut-on apprendre et faire des projets, ou encore, que veut-on donner à penser ?

Cette approche trouve bien sûr écho avec la boucle sensori-motrice de l'énaction, excepté pour la partie « savoir », que l'énaction remplace par la maîtrise des lois de contingences sensori-motrices. De plus, l'idée qu'un dispositif puisse donner à penser reste parfaitement en adéquation avec l'approche énactive.

INTERACTION DESIGN



Figure 27. L'interaction design selon Bill Verplank (Moggridge, 2006).

Étant donné que notre recherche est centrée sur l'échange émotionnel à distance, nous nous intéressons aux conditions de l'expérience perceptive du contact avec autrui. Pour cela, nous souhaitons mettre en avant l'importance de la tangibilité et du fait d'être incarné dans le design d'interaction. Kirsh (2013) expose quatre idées apportés par les théories de la cognition incarnée (dont l'énaction fait partie) pouvant fournir de nouvelles idées sur l'interaction (voir aussi partie 1.1.1.1.2) :

- l'interaction avec les outils modifie nos façons de penser et de percevoir – lorsqu'on les manipule, les outils sont rapidement incorporés dans le corps propre, et cette assimilation entraîne des changements fondamentaux dans notre façon de percevoir et de concevoir notre environnement ;
- nous pensons avec notre corps, pas simplement avec notre cerveau ;
- on apprend plus en faisant qu'en regardant – il est des fois où accomplir physiquement une activité est mieux que de la voir exécutée par quelqu'un d'autre, même si notre système de résonance moteur (motor resonance system) s'enclenche avec force lors de l'observation d'une autre personne ;
- il est des fois où nous pensons littéralement *avec* des choses.

Ces quatre idées ont des implications majeures pour le design d'interaction, en particulier pour le design de systèmes tangibles, physiques, prenant en compte le contexte, et de systèmes

de téléprésence. La piste des interfaces matérielles en vue de permettre l'interaction avec du contenu numérique est en effet fertile. Elles sont basées sur l'augmentation d'objets physiques existant par l'ajout d'information numérique ; ce sont des types d'interfaces basées sur la réalité, qui tirent profit des indices matériels du monde réel et physique amenés par les objets existants (Chang et Ishii, 2006). Le déploiement des représentations des fonctions chez les interfaces tangibles est apprécié des utilisateurs (Maillet et al. 2005).

La communication des émotions ne se fait pas par la sortie du dispositif, mais dans la dynamique d'utilisation (Gaver, 2002). Il y a en effet des aspects spécifiques de la médiation de l'intime, comme le soulignent Howard et al. (2004) ; ils proposent par conséquent plusieurs façons de supporter les actes intimes à distance :

- permettre la révélation de soi ;
- communiquer les émotions, pas forcément par le biais de mots ou de texte, mais au contraire plutôt sans avoir recours à la prononciation ;
- transmettre un sentiment de présence en absence, au travers de mécanismes d'awareness périphériques ;
- jouer avec l'ambiguïté et l'inachèvement afin de permettre des échanges nuancés ;
- permettre l'alignement des actes intimes avec les échanges instrumentaux, afin de pouvoir saisir l'opportunité d'exprimer l'amour dans les actes du quotidien ;
- faire usage des significations reçues et des langages privés qui évoluent de façon continue au sein des relations riches ;
- permettre la mutualité, sans exiger pour autant la réciprocité de façon symétrique.

Plusieurs chercheurs (Motamedi, 2007 ; Sonneveld, 2002) notent de plus l'apport du sens du toucher dans le design d'interaction, et prônent une esthétique de ce sens, ce que nous détaillerons dans la partie 2.

1.2.2.3 Conception centrée utilisateurs

Inclure l'utilisateur dans le processus de conception présente d'indéniables avantages. Un cycle de proposition-évaluation-choix aboutit à la matérialisation de dispositifs permettant des actions organiques faisant sens pour l'utilisateur. Une démarche de conception participative intègre en effet le facteur humain dans le processus même de l'élaboration du système technique. Prendre en compte l'utilisateur dès le départ et de façon cyclique s'oppose aux processus de conception linéaires (qui ont pour point de départ l'analyse d'un besoin et se terminent par un processus d'évaluation, et où la conception est guidée par des considérations techniques). Cette méthode où ce sont les besoins des utilisateurs que l'on prend en compte est désignée sous le

terme de « conception centrée utilisateur » (norme ISO 13407, 1999). Elle permet de penser en même temps l'objet technique et son usage (Jouët, 1993). Les designers ont besoin d'anticiper le plus tôt possible l'expérience vécue par l'utilisateur lors de l'usage du dispositif conçu (Lenay et al. 2007). L'intégration des utilisateurs le plus tôt possible dans le cycle de conception permet aux designers d'avoir accès à cette expérience. On co-construit ainsi l'objet technique et son usage.

L'objectif est de comprendre le fonctionnement d'un dispositif technique et son intégration par l'utilisateur. L'analyse d'usage se fait par l'observation et l'interprétation des actions et remarques des utilisateurs, dans des contextes d'emploi réels. L'investigation porte donc sur l'usage lui-même, puisque les observations se font en situation d'usage libre. Cette façon de procéder permet d'enrichir le processus de conception, en approfondissant et diversifiant les hypothèses de conception et les pistes de solutions pertinentes, et en se donnant les moyens de préciser parallèlement l'utilité et l'utilisabilité par le biais de la mise à l'épreuve du concret (Maillet, 2010). On peut ainsi étudier l'appropriation du dispositif, de la prise en main à l'intégration, en passant par l'apprentissage. Cette démarche permet la spécification de la pertinence écologique du dispositif d'une part, mais aussi de préciser les situations d'usage, leur évolution, ainsi que parfois l'apparition d'usages non anticipés et non prévus à mesure que les utilisateurs s'approprient le dispositif (Von Hippel, 2009).

La conception centrée utilisateur met donc l'utilisateur au centre de sa démarche, en le faisant participer au processus de conception, mais elle est aussi basée sur une méthode itérative et une approche pluridisciplinaire. L'itération permet de suivre l'évolution de la problématique de conception à mesure que la conception elle-même avance, et de s'adapter à cette évolution.

En permettant aux utilisateurs d'évaluer le dispositif, on peut orienter les choix de conception vers la solution qui s'adaptera le mieux à eux. Le recueil de données qualitatives et quantitatives de l'usage d'un dispositif, et leur analyse, permettent alors de proposer des spécifications matérielles et fonctionnelles. Il faut d'ailleurs prendre en compte l'aspect du *plaisir de l'usage* lors de la conception (Overbeeke et al., 2002). La dimension émotionnelle de l'interaction est en effet un aspect essentiel des interactions entre les hommes et les objets.

En ce qui nous concerne, nous ne partons pas d'un besoin spécifique, mais cherchons à explorer la thématique de la communication émotionnelle à distance, et allons intégrer l'utilisateur à notre processus. Nous nous concentrons ainsi sur l'utilisation des dispositifs par les utilisateurs, et allons leur proposer nos propres interfaces. Ils seront bien entendu intégrés au processus de conception ; nous optons pour une conception itérative augmentée des retours utilisateurs. Cependant notre démarche inclut un processus de design enrichi de la démarche expérimentale. Ainsi, nous mettons en œuvre un processus d'innovation radicale, où l'on ne s'appuie pas sur un simple besoin identifié. Ce sont les outils, par le biais des dispositifs existants

ainsi que ceux que nous allons proposer, qui vont créer les conditions d'un besoin qu'il s'agit pour nous d'explorer. En effet, une innovation radicale ne peut découler de besoins identifiés chez les utilisateurs, étant donné qu'elle va en fait être la source de nouveaux besoins, qui n'existent donc pas encore. Notre piste pour ce type d'innovation est de prendre en compte la recherche fondamentale ; grâce à une approche minimaliste, nous allons mener cette recherche autour du thème de la communication émotionnelle à distance, pour ensuite proposer des solutions innovantes. Nous allons à présent exposer la méthodologie expérimentale que nous allons mettre en œuvre dans notre processus d'innovation ; notons que les expériences psychologiques sont déjà en elles-mêmes porteuses d'innovation, en ce qu'elles sont heuristiques.

1.2.3 Méthodologie expérimentale

La méthodologie expérimentale est la méthode dominante pour la recherche en psychologie ; elle donne une importance particulière aux données quantitatives (Searle, 1999).

La méthodologie expérimentale consiste pour l'expérimentateur à altérer un élément, ou variable, afin d'observer si cela entraîne un effet sur une autre variable, dans une situation standard et répétable. La répartition des sujets doit se faire de manière aléatoire selon les différentes conditions. Il existe deux types d'expérimentations : les expériences en laboratoire, et les expériences sur le terrain. L'objectif est pour nous d'étudier scientifiquement l'influence des médiations techniques sur les phénomènes de perception et d'interaction. La méthodologie expérimentale permet de valider des hypothèses et de confirmer des prédictions en étudiant le fonctionnement d'un processus hypothétique, et se trouve finalement être une source d'inspiration pour l'innovation. Il s'agit d'analyser les effets de facteurs ou de leur interactions.

Nous avons mis en œuvre des expériences en laboratoire, c'est-à-dire ayant lieu dans un environnement où nous pouvions contrôler le plus grand nombre possible de variables. Tous les sujets ont passé les expériences dans les mêmes conditions. La psychologie expérimentale est en effet une méthodologie systématique d'observation et d'acquisition de connaissances sur les attitudes et les comportements, étant basée sur l'expérimentation. Elle se fixe pour objectif de contrôler les différents paramètres d'une situation pour en évaluer, de façon systématique, l'impact sur les attitudes et comportements. La première étape du processus consiste à formuler une hypothèse que l'on souhaite mettre à l'épreuve. Il faut ensuite mettre en place une situation expérimentale, dont il faudra contrôler toutes les variables possibles. Expérimenter, c'est créer (ou isoler) des sources de variation et vérifier que les variations induites sont importantes statistiquement par rapport au simple hasard. Utiliser la méthodologie expérimentale va consister

à créer une situation particulière qui va permettre de tester une hypothèse causale concernant la mesure d'un phénomène précis en fonction de la manipulation d'un ou plusieurs facteurs.

Notons que si notre travail comporte des expérimentations psychologiques, nous avons aussi mis en œuvre des observations, des questionnaires, des entretiens, des analyses du discours, des analyses d'usage ainsi que des mises en situation non contrôlée, qui permettent une analyse plus qualitative. En effet, les techniques prenant en compte les informations ethnographiques et contextuelles, ainsi que les cultural probes, peuvent être mises en place pour étudier la façon dont les techniques interactives sont utilisées pour la médiation des relations intimes (Vetere et al., 2005).

1.2.4 Le minimalisme comme cadre de travail

Le minimalisme est un principe qui consiste à employer des systèmes d'interaction qui appauvrissent considérablement le couplage entre un individu et son environnement, dans le but d'expérimenter et de mieux comprendre les phénomènes perceptifs et cognitifs, individuels ou collectifs. L'étude minimaliste est rendue possible par la mobilisation de médiations techniques telles que le dispositif de suppléance perceptive Tactos.

Le minimalisme constitue un cadre d'analyse particulier, qui nous semble pertinent dans le cadre de notre recherche. Il donne les moyens d'étudier concrètement la perception (Auvray et al. 2007) et le processus de l'interaction (Lenay, 2006c). Il permet notamment d'observer une néo-genèse de l'acte perceptif, de la cognition sociale et des mécanismes d'interaction. Le fait de passer par des médiations techniques, comme Tactos, offre la possibilité d'analyser la constitution même du processus d'interaction et de la co-construction de sens dans cette interaction, ainsi que de comprendre comment les interfaces perceptives rendent possible un mode de perception. Les sujets découvrent un nouvel espace de couplage, où les savoir-faire ne sont pas déjà donnés (ibid). Pour le chercheur, la médiation technique permet l'observation et l'enregistrement de tout ce qu'ont fait les sujets pendant la tâche. La réduction à l'extrême du répertoire des actions et des sensations permet l'analyse précise par l'étude du déploiement des engagements individuels et collectifs (ibid ; Ali-Ammar, 2005).

Lenay et Stewart (2012) remarquent en effet que les travaux concernant la cognition sociale dans une perspective interactionniste font face à des difficultés considérables, tant au niveau théorique qu'au niveau méthodologique, telles que l'identification des variables comportementales significatives, leur enregistrement sans perturbation de l'interaction ainsi que la distinction entre :

- 1) les contributions nécessaires et suffisantes de chacun des partenaires à un niveau individuel pour qu'une dynamique collective émerge ;
- 2) les particularités qui découlent de cette dynamique collective et qui échappent au contrôle des partenaires à un niveau individuel ;
- 3) les phénomènes émergents de cette dynamique collective qui sont par la suite appropriés et utilisés par les partenaires.

Lenay et Stewart (ibid) proposent par conséquent un paradigme expérimental minimaliste : en réduisant les entrées sensorielles à un strict minimum, ils forcent un déploiement spatial et temporel des activités perceptives, lequel rend possibles l'obtention d'un enregistrement complet et le contrôle des dynamiques d'interaction.

Le minimalisme expérimental est donc une forme de réduction expérimentale, qui reprend les principes en poussant plus avant le raisonnement. La réduction expérimentale consiste à réduire la complexité d'une situation ou d'un phénomène en neutralisant certaines variables, et en contrôlant les variables restantes. On cherche donc à simplifier pour mieux comprendre ce que l'on étudie (Ali-Ammar, 2005). Avec l'utilisation d'un dispositif comme Tactos qui contraint les entrées sensorielles et les moyens d'action, on a la possibilité d'observer le déploiement de l'activité perceptive et cognitive des sujets. On peut donc externaliser ces activités et les rendre ainsi accessibles, pour les analyser. On identifiera alors les conditions minimales permettant l'émergence d'un phénomène en particulier. C'est après avoir identifié ces conditions que l'on pourra peu à peu complexifier la situation, les entrées sensorielles ou les moyens d'action.

Les conditions très minimales mises en œuvre dans ces situations expérimentales permettent d'obtenir des résultats ayant une plus grande valeur générale. Ces résultats concernent le sujet de la perception des autres individus comme êtres intentionnels, ainsi que celui de l'expression et de la compréhension mutuelles des intentions et des émotions (Maillet, Guénand, et Lenay, 2006 ; Maillet, Lenay, et Guénand, 2008).

Le minimalisme permet d'étudier les conditions techniques de l'émergence de percepts. Il offre en effet des possibilités expérimentales particulières par le biais de dispositifs comme Tactos. Les modalités d'action (déplacement dans un monde à une ou deux dimensions) et de sensations rendues possibles (retour haptique en tout ou rien ou en parallélisme sur 16 picots) étant élémentaires, les actions effectuées, les gestes et les stratégies d'exploration peuvent être étudiées, tout comme la mise en place d'un couplage sensori-moteur. On force un déploiement spatial et temporel de l'activité perceptive.

En tant que principe épistémologique, le minimalisme peut servir de principe de conception d'interfaces dédiées à un toucher à distance. L'avantage de l'approche minimaliste est qu'elle offre aux designers la possibilité de se concentrer sur la nature de l'expérience

perceptive que l'on cherche à médiatiser par le biais de l'interface. Elle permet d'acquérir les conditions techniques qui rendent possible une expérience perceptive visée ; il reste alors à décider de l'enrichissement de l'interface en fonction de la situation d'usage réelle envisagée, au fur et à mesure qu'elle se précise. Chang et al. (2007) prônent d'ailleurs un design d'interaction où le but est d'atteindre la simplicité (et passent par la tangibilité pour essayer de l'atteindre). Leur approche repose sur l'exercice d'un design minimaliste, en vue d'explorer la capacité à communiquer des éléments composants l'interaction. Il en résulte des solutions de design expressives, des perspectives utiles sur le design d'interaction et de nouvelles techniques d'interaction (ibid).

1.2.5 L'approche phénoménologique

La phénoménologie est une étude rigoureuse et systématique de la conscience, introduite par Husserl (1900). La question que la phénoménologie se pose est : qu'est-ce qui nous apparaît effectivement, qu'est-ce qui fonde notre connaissance, avant la science ? La phénoménologie considère en effet la façon dont les phénomènes *nous apparaissent*, la façon dont nous les vivons en première personne. La phénoménologie examine les façons dont nos environnements physiques et sociaux, incluant les objets et les instruments, impactent nos expériences, notre cognition, notre façon de résoudre les problèmes, et façonnent nos interactions intersubjectives et sociales. Le corps est posé comme ce qui appartient à la présence la plus immédiate au monde. Le vécu corporel appartient au monde de la sensation, tandis que l'image du corps appartient au monde de la perception.

La phénoménologie a subi plusieurs évolutions. Husserl la pose comme un projet ayant pour but de décrire les structures de base de la conscience, c'est-à-dire les caractéristiques de la conscience qui influent la façon dont divers objets du monde nous apparaissent. Heidegger propose sa propre caractérisation de la phénoménologie, légèrement différente : pour lui, le but de la phénoménologie est de fournir une analyse ontologique fondamentale de l'existence humaine en tant qu'« être au monde » (Heidegger, 1927). Il montre que notre façon première d'être au monde implique une façon pragmatique, orientée sur l'action, d'être relié à notre environnement.

Merleau-Ponty a proposé un troisième développement de l'analyse phénoménologique, en mettant l'accent sur l'incarnation (Merleau-Ponty, 1945). Il a développé plusieurs idées esquissées par Husserl et a approfondi l'analyse concernant la perception incarnée, en y intégrant la psychologie et la neuroscience, posant ainsi la base d'approches nouvelles de la cognition incarnée. Notre travail s'inscrit dans la vision de Merleau-Ponty sur la phénoménologie.

Merleau-Ponty base sa phénoménologie sur la perception : « Il y a bien un acte humain qui d'un seul coup traverse tous les doutes possibles pour s'installer en pleine vérité : cet acte est la perception, au sens large de connaissance des existences » (ibid, p. 66). C'est l'essence même de la Phénoménologie de la perception : un retour aux phénomènes, et plus particulièrement au phénomène de la perception, pour comprendre d'où viennent notre connaissance du monde, notre savoir, comment nous avons accès au monde. Merleau-Ponty met en effet l'accent sur le corps, et sur son rôle pour notre perception et notre cognition, ce qui s'oppose à la vision traditionnelle de la philosophie dans les pays occidentaux, où l'esprit (ou l'âme) domine. Pour Merleau-Ponty, c'est au contraire par le corps que l'on perçoit, que l'on sait et que l'on agit. L'intentionnalité est posée comme étant avant tout une intentionnalité motrice. Le monde fait sens parce que nous pouvons y agir, et ce sens découle de nos actions. Notre relation première au monde est pragmatique et à portée de main ; c'est avant tout une relation incarnée, où nos mains et notre système moteur jouent nécessairement un rôle essentiel. Merleau-Ponty explique cette notion d'être au monde par le corps lui-même : « Le corps est le véhicule de l'être au monde, et avoir un corps c'est pour un vivant se joindre à un milieu défini, se confondre avec certains projets et s'y engager continuellement » (ibid, p. 111).

Il s'empare de la distinction phénoménologique entre le corps vécu, ou corps propre (Leib) et le corps biologique, objectif (Körper), distinction posée par Husserl. Le Leib est le corps vivant, lieu des sensations et des émotions, tandis que le Körper est le corps en tant qu'organisme, de forme pour autrui. On distingue donc le corps comme sujet (ou comme agent), et le corps comme objet. Nous sommes engagés dans le monde, et l'on y rencontre les autres sujets individus intersubjectivement, en tant qu'acteurs incarnés plutôt qu'en tant qu'observateurs passifs ou qu'esprits détachés. Notre corps n'est pas un objet – ni pour nous, ni pour les autres individus –, mais un corps agissant qui exprime sa subjectivité par le biais de postures, de gestes, d'actions, de mouvements et d'expressions. Mon corps, en tant qu'agent percevant, est engagé dans le monde, et participe avec les autres corps à ce que le monde fasse sens. Ma capacité à comprendre autrui implique un ensemble de processus incarnés, sensori-moteurs et contextuels, plutôt que fondamentalement inférentiels ou découlant d'une simulation. Je ne suis pas fondamentalement un observateur en troisième personne des autres, au contraire je m'engage et interagis avec eux dans des contextes socialement définis.

La phénoménologie peut donc être une source d'inspiration pour le designer (Hummels, 2013) ; elle est pertinente pour le design étant donné que les objets, outils et technologies fabriqués par l'Homme affectent la façon dont nous faisons l'expérience de notre environnement (ce qui est précisément ce que la phénoménologie étudie).

1.3 Conclusion : l'articulation Théorie-Design-Expérimentation comme enjeu de recherche

Cette partie nous a permis de poser les cadres théoriques et méthodologiques dans lesquels ce travail s'inscrit, et ce pour les multiples points de vue disciplinaires adoptés ici.

Notre recherche se concentre sur l'interaction émotionnelle médiatisée, et sur les conditions qui la rendent possible. Il s'agit pour nous d'articuler théorie, design et expérimentation. Si la théorie et la pratique expérimentale sont déjà établies et accordées comme posture de recherche, il n'en va pas de même pour le design. Celui-ci peut pourtant offrir une grande richesse de perspective pour la recherche, par la conception non seulement d'outils nouveaux, mais aussi de méthodes. De même, la théorie peut venir enrichir la posture du design, comme le permet par exemple l'approche phénoménologique. Ces deux pôles peuvent ainsi se compléter, mais cette relation n'est pas encore parfaitement établie, l'aspect appliqué du design et fondamental de la théorie restant encore peu articulés. De même, le design et l'expérimentation ne forment pas une paire classique, alors que la conception peut fournir à l'expérimentation de nouveaux dispositifs et méthodes, tandis que la méthodologie expérimentale peut constituer un outil pour le designer, et lui fournir des indications et de nouvelles pistes.

Dans la partie suivante, nous allons nous baser sur le socle théorique présenté ici afin de procéder à un premier cycle : théorie -> design -> expérimentation (voir figure 28). En nous basant sur les recherches concernant le sens du toucher, nous allons proposer une première interface pour la communication émotionnelle à distance. Cette interface nous servira d'outil pour une expérimentation, visant à différencier le partenaire d'un robot. Nous posons l'hypothèse que l'activité perceptive du partenaire permettra une nouvelle fois de le différencier d'un objet mobile et qu'une interaction touchante sera proposée aux utilisateurs. Ces résultats permettront de revenir sur la question de l'interaction émotionnelle à distance et des dispositifs de suppléance perceptive.

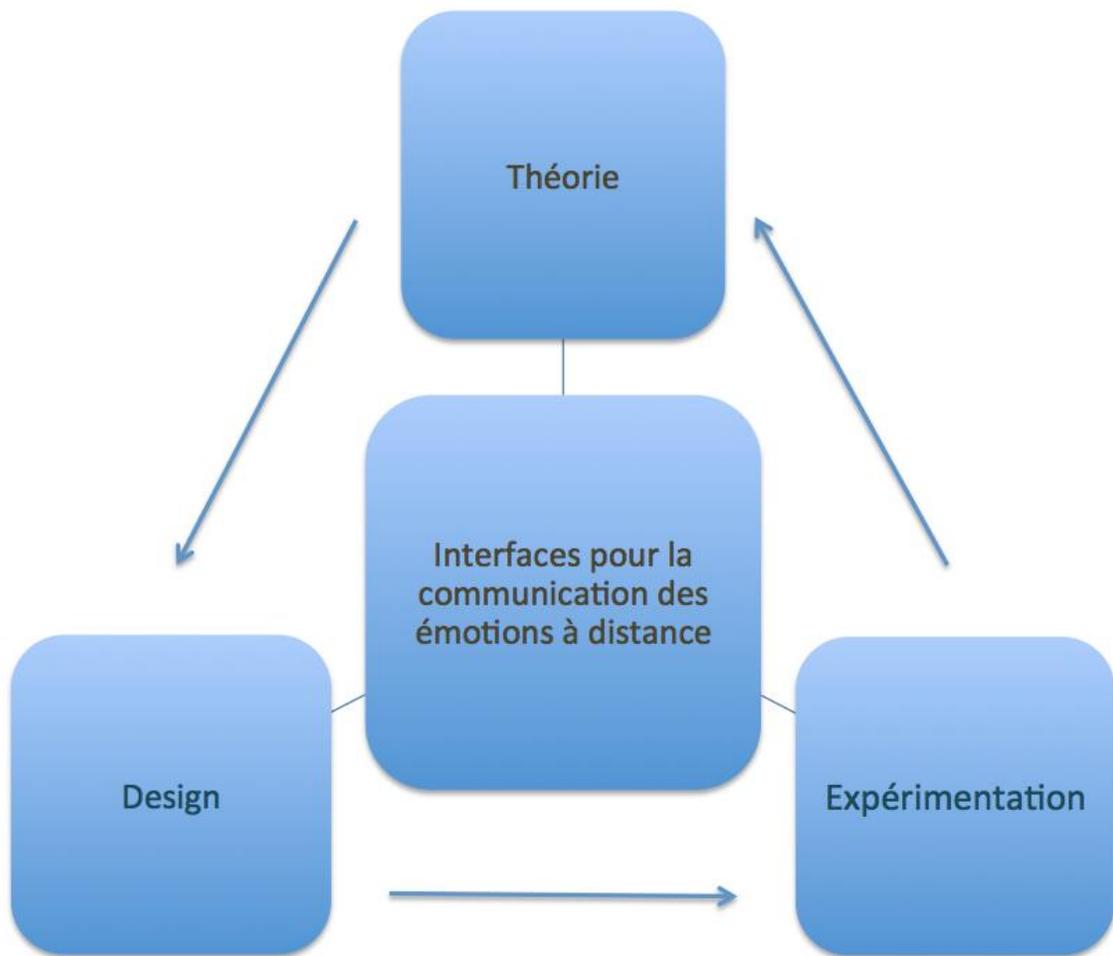


Figure 28. Un premier cycle théorie-design-expérimentation.

2 La présence à distance : du théorique vers l'expérimentation en passant par le design

2.1 Introduction

Dans cette partie, nous allons nous intéresser au sens du toucher pour l'interaction émotionnelle à distance. Nous allons dans un premier temps exposer les particularités de ce sens, et exposer en quoi il constitue un candidat idéal pour la communication émotionnelle. Dans un second temps, nous allons présenter Touch Through, notre première proposition d'interface pour la communication émotionnelle à distance, basée sur une suppléance perceptive du toucher. Enfin, nous allons présenter une expérimentation menée avec Touch Through, permettant d'étudier le sentiment de présence d'autrui permis par un dispositif de ce type.

2.2 Contexte théorique

2.2.1 Interactions interpersonnelles mobilisant le toucher

2.2.1.1 Sens du toucher

2.2.1.1.1 Propriétés

2.2.1.1.1.1 *Le toucher¹⁷ au cours du développement*

Nous avons tendance à considérer le sens du toucher comme étant un sens secondaire, aisé à oublier. En effet, notre vue et notre ouïe nous semblent être les principales informations sensorielles nous parvenant du monde extérieur. Et s'il est vrai qu'être enfermé dans une pièce totalement obscure pendant une semaine nous permet d'aiguiser non seulement notre ouïe, mais aussi notre sensibilité au toucher (ainsi que notre sensibilité à la douleur), il n'empêche que non seulement nous recevons constamment des informations tactiles, mais qu'en fait, le sens du toucher est même nécessaire à notre survie. C'est là une caractéristique unique : aucun de nos autres sens n'est aussi vital que le toucher, qui nous permet d'ailleurs en partie d'accéder à l'expérience que nous faisons de la vision (Pacherie, 1997).

Le toucher met en jeu notre plus grand organe sensoriel ; la peau est en effet le plus grands des organes du corps (Montagu l'estime à près de 2 mètres carrés : « 18 square feet of skin » (Montagu, 1986, p.270)). La peau est non seulement le plus grand de nos organes

¹⁷ On notera les différences de langue : en français, sentir peut se référer soit au sens du toucher, soit au sens de l'odorat, soit au ressenti, quand en anglais on dispose à la fois de « sense », de « smell » et de « feel ».

sensoriels, mais c'est aussi le plus ancien : c'est le premier à se développer chez le fœtus (Field, 2003 ; Montagu, 1986). Les premiers inputs sensoriels que l'on reçoit nous viennent du sens du toucher, alors que nous sommes encore dans l'utérus de notre mère, et c'est ce sens qui continue à être le moyen premier de faire l'expérience du monde durant la prime enfance et une grande partie de l'enfance, et même à nouveau lors du vieillissement (Field, 2003 ; DeVito, 2008). Il contribue au développement cognitif et socio-émotionnel (Hertenstein, 2002 ; Stack, 2008). C'est aussi un sens constamment à l'affût, que l'on ne peut ignorer : la peau est constamment en état de recevoir des messages. On ne peut « fermer sa peau » comme on fermerait les yeux ou comme on se boucherait les oreilles (mais on peut porter des vêtements).

Le toucher est un sens extrêmement important dont nous avons besoin pour survivre. Le toucher est critique pour la croissance, le développement, la communication et l'apprentissage, mais il sert aussi à réconforter, à donner de l'assurance et de la confiance en soi (Field, 2003). L'amour et l'affection sont appris par le biais du toucher ; les premiers liens émotionnels sont construits par le contact physique, et forment la base d'un développement futur sur les plans émotionnels et intellectuels (Field, 2003).

L'apprentissage et l'exploration du monde dépendent du toucher. Lorsqu'il vient au monde, le nourrisson appréhende en premier lieu son environnement matériel et social par ses sensations tactiles. Sans le sens du toucher, il serait impossible de se déplacer dans le monde. On pense souvent que ce sont les mains qui nous fournissent le plus d'informations tactiles étant donné que nous les utilisons pour manipuler les objets, mais tout ce que nous faisons (ce qui inclut s'asseoir, marcher, s'embrasser, ressentir de la douleur,...) dépend de notre sens du toucher. Nous nous en apercevons bien plus lorsqu'il nous faut négocier une rue rendue glissante par le givre, ou arpenter un chemin fait de cailloux. C'est par le toucher qu'on apprend ce qui est rugueux et ce qui est lisse, ce qui est froid et ce qui est chaud (savoir qui est critique en vue d'éviter les coupures, échardes et autres brûlures). Et, sans le sens du toucher, le plaisir du contact peau à peau, ou de caresser un animal, nous serait inconnu.

Le toucher est important dès les premières interactions ; le toucher d'une mère sur son enfant nouveau-né peut avoir de nombreux impacts positifs (meilleure prise de poids, moins de pleurs, éveil plus précoce, moins de stress,...) (Field, 2003). Tronick et ses collègues ont montré qu'aux Etats-Unis, les mères passaient 60% du temps d'interaction avec leur bébé à le toucher, et que la plus grande proportion de ce temps était passé à caresser l'enfant en rythme et à le porter. Les chatouilles et les baisers ont lieu moins souvent et de façon plus discrète, mais à l'âge de deux ans, les bébés savent déjà prendre dans les bras et embrasser. Les mêmes études réalisées auprès de mères dépressives ont montré qu'elles ont tendance à moins toucher leur enfant, et qu'elles tendent plus à taquiner l'enfant du doigt, ce qui fait se détourner ce dernier et le rend

irritable (Tronick, 1995). On constate des retards dans le développement social et culturel des enfants qui ont souffert d'un manque de contacts tactiles avec leurs parents.

Considérons à présent le toucher durant l'enfance et l'adolescence. Dans la culture américaine, les stimulations apaisantes des parents et les rythmes d'autostimulation de l'enfant sont remplacés par le jeu avec les pairs (des jeux un peu brusques, comme se faire tomber à la renverse, à la crèche et à la maternelle, suivis par les jeux de contact tout au long de l'école, jusqu'au lycée). Plus les enfants grandissent, plus le contact physique devient tabou¹⁸, du moins lorsqu'il vient d'un parent ou d'un adulte (Field, 2003). Au moment où les enfants entrent au collège, ils ne reçoivent plus que moitié moins de contacts que lorsqu'ils sont entrés à l'école primaire, et le toucher est d'une nature différente – plus de contacts d'épaule à épaule, coude à coude, au lieu des contacts directs de la main (Field, 2003). À l'adolescence, le taux de toucher peut remonter si des relations intimes entre pairs se mettent en place. L'adolescent se met aussi de manière assez typique à traiter les membres de sa famille comme s'ils étaient à éviter. Étant donné les tabous concernant le toucher, il n'est pas surprenant que des adolescents se retrouvent privés de contacts, ce qui impacte leur bien-être émotionnel (ibid).

À l'âge adulte, l'aspect sexuel prend toute son ampleur, véhiculé essentiellement par le toucher. Il est l'un des acteurs majeur de l'intimité (Gibbs et al., 2006). Montagu note : « La conception française de la relation sexuelle comme étant 'l'harmonie de deux âmes et le contact de deux épidermes', souligne avec élégance une vérité profonde – l'implication massive de la peau dans les rapports sexuels. Cette vérité est qu'il n'est pas d'autre relation où la peau soit aussi totalement impliquée. De fait, le sexe désigne la plus haute forme de toucher. Dans le sens le plus profond du terme, le toucher est le véritable langage du sexe. Les lèvres et les parties génitales sont particulièrement bien pourvues en terminaux nerveux sensoriels ramifiés en disques et concaves. » (Montagu, 1986, p. 204, traduction personnelle)¹⁹.

Notons que la fonction même du baiser reste difficile à définir. Desmond Morris suggère : « Dans les premières sociétés humaines, avant que l'on n'invente la nourriture pour bébé, les mères sevrèrent leurs bébés en pré-mâchant leur nourriture puis en la faisant entrer dans

¹⁸ La diffusion des normes sociales et des règles de bienséance des classes les plus élevées vers les plus populaires a participé de façon très importante à la dévalorisation du toucher : elles interdisent de toucher, de nous gratter en public, de manger avec les doigts, etc. (Eliás, 1939).

¹⁹ « The French wit who defined sexual intercourse as 'the harmony of two souls and the contact of two epidermis,' elegantly emphasized a basic truth – the massive involvement of the skin and sexual congress. The truth is that in no other relationship is the skin so totally involved as in sexual intercourse. Sex, indeed, has been called the highest form of touch. In the profoundest sense, touch is the true language of sex. The lips and the external genitalia are especially well-supplied with concave, disk-like branched sensory nerve endings. »

la bouche de l'enfant lors d'un contact lèvres à lèvres – geste qui impliquait naturellement une grande activité linguale et la nécessité de faire pression dans la bouche. Cette forme de système, qui ressemble à celui qu'utilisent les oiseaux pour nourrir leurs petits, nous semble étrange et étranger aujourd'hui, mais notre espèce l'a sans doute pratiqué pendant un million d'années, voir plus, et les baisers érotiques adultes d'aujourd'hui sont presque certainement un geste relique issu de ces origines » (Morris, 1977., p. 51-52)²⁰.

Une autre possibilité est que l'on s'embrasse parce que les lèvres sont particulièrement bien fournies en glandes sébacées, et que le sébum (qui possède aussi un rôle de phéromone) faciliterait l'attachement entre mère et enfant, ainsi qu'entre amants. Pour certains, la partie la plus significative de l'intimité sexuelle est tout simplement le fait d'être proche et de se toucher. Dans un sondage réalisé auprès de 100.000 personnes et mené par Ann Landers, on a posé la question « Seriez-vous satisfait si l'on vous tenait fort contre soi et vous traitait tendrement, et qu'il n'y ait plus d'acte sexuel ? » 72% des répondants ont répondu oui – et 40% de ces répondants avaient moins de quarante ans (Landers, 1985, p.131-132).

Le contact corporel est un acte très intime, en particulier pour les femmes. À tous âges, et ce depuis la naissance, les femmes ont des seuils de douleur et de contact plus bas que les hommes, ce qui peut expliquer pourquoi les femmes sont plus réceptives au toucher que les hommes (Field, 2003). On manipule moins les petits garçons, on les caresse moins souvent, et on les porte moins longtemps que les filles, ce qui pourrait expliquer pourquoi ils sont moins réactifs que les filles au toucher. À un âge plus avancé, cependant, selon le Hite report on male sexuality, les hommes rattrapent les femmes et désirent aussi plus de toucher non-lié au sexe de la part des femmes (Hite, 1982). Dans le Sensate Focus System de sex therapy conçu par Masters et Johnson, aucune relation sexuelle n'est autorisée durant les premières semaines de sex therapy. À la place, les membres du couple doivent simplement toucher le corps de leur conjoint. Cela permet de faire diminuer l'anxiété liée à l'obligation de performance et de rehausser le désir (Masters et Johnson, 1970).

La privation du toucher peut entraîner de nombreuses conséquences néfastes, ou accentuer certaines maladies : violence, troubles du sommeil, stress, manque d'énergie, réponse immunitaire affaiblie, retard dans la croissance, manque de certaines fonctions cognitives, problèmes moteurs, sensibilité au toucher moindre, allergies (dermatites, asthme), maladies

²⁰ « In early human societies, before commercial baby food was invented, mothers weaned their children by chewing up their food and then passing it into the infantile mouth by a lip-to-lip contact – which naturally involved a lot of tonguing and intramouth pressure. A form of bird-like system of parental care seems strange and alien to us today, but our species probably practiced it for a million year or more, and adult erotic kissing today is almost certainly a relic gesture stemming from these origins. »

cardiovasculaires,... (Field, 2003). Dans son article The no-touching epidemic – an English disease, Heylings (1973) décrit les symptômes découlant d'une privation du toucher des autres. Ils incluent un sentiment de solitude et d'isolement, des doutes sur la loyauté d'autrui, un sentiment d'insécurité, des inhibitions émotionnelles, des réactions inhabituelles tant au fait d'être touché que de toucher par inadvertance, l'incapacité de communiquer avec des individus se tenant à proximité, et un refus des massages comme forme de thérapie. Le besoin d'un toucher aimant perdure tout au long de la vie. Ce besoin est tellement primordial que l'on en parle comme d'une « faim » d'être touché ; on peut avoir faim de toucher ou d'être touché (Field, 2003), et, de même que pour le besoin de nourriture, on peut dépérir lorsque cette faim n'est pas apaisée. Le toucher est primordial pour le développement physique et mental. Plusieurs expérimentations réalisées sur des rats ont montré que la privation tactile menait à des retards dans le développement et au dépérissement (Montagu, 1986). Cela s'est aussi confirmé pour les êtres humains par l'observation d'enfants ayant grandi dans des orphelinats roumains, où le toucher était rare du fait du manque d'effectifs au niveau du personnel (Field, 2003). Ce besoin du toucher semble être primordial au point que des bébés chimpanzés privés de leurs mères préfèrent ne pas quitter un substitut de mère en tissu éponge ne donnant pas de lait plutôt qu'un substitut fait de grillage mais fournissant du lait (ibid).

C'est au travers du toucher qu'un individu a besoin de faire l'expérience d'être en sécurité et d'être choyé. Durant les premières phases de la vie, le toucher doit être de nature aimante et protectrice enfin que nous puissions devenir des êtres humains sains et empathiques. Des touchers négligents ou brutaux prodigués de manière fréquente coupent les nourrissons de leur capacité à développer leur vie affective. L'expérimentation des singes montre aussi que les bébés s'étant retrouvés sans aucun substitut de mère ont développé un comportement d'épouillage anormal (ibid).

Montagu relate au sujet de la privation du toucher l'expérience d'un de ses patients, un homme adulte ayant dû être placé dans une bulle stérile l'isolant du monde extérieur : « Il y a à peu près une semaine, ça a commencé à me taper sur les nerfs... de ne pas pouvoir toucher d'autres gens et d'espérer pouvoir sortir bientôt. J'avais l'impression oppressante que l'espace se resserrait autour de moi et je ne pouvais plus le supporter. Il fallait juste que je puisse *sentir* d'autres gens, je voulais sentir quelqu'un, toucher une autre personne. Si j'avais pu le faire, j'aurais pu tenir plus longtemps... mais je ne pouvais pas, il n'y avait aucun moyen pour que je puisse toucher qui que ce soit ou exprimer mes émotions à quelqu'un juste en lui touchant la main, ou en la serrant. C'est très difficile à expliquer – c'est une situation pour laquelle les mots me manquent. On a le sentiment d'être complètement seul au monde, et que tout est froid. Il n'y

a aucune chaleur. La chaleur disparaît complètement, et on a juste l'impression qu'il n'y a plus rien d'autre que nous. » (Montagu, 1986, p. 266, traduction personnelle)²¹.

On voit apparaître au travers de ce témoignage l'importance du toucher pour la communication des émotions, et le fait que le toucher soit absolument nécessaire pour le bien-être. Il est établi que les êtres humains ont besoin d'être touchés, pour différentes raisons, et que ce besoin ne peut pas être compensé par le biais des autres sens.

2.2.1.1.1.2 *Le toucher : rencontres physiques et conscience de soi*

L'interaction physique avec le monde et les autres individus ne se limite bien sûr pas seulement aux mains. Ces dernières jouent un rôle majeur, mais c'est en fait le corps tout entier qui se retrouve impliqué dans le toucher. L'engagement physique dans le monde, ainsi que la conscience de toucher et d'être touché, permettent aux individus d'être conscients d'être eux-mêmes un corps physique, qui est engagé dans un monde physique, et qui partage ce monde avec d'autres objets, eux aussi physiques. C'est au cœur de cette rencontre incarnée que le « je » peut faire l'expérience simultanée de lui-même et du monde qui l'entoure, faisant de cette rencontre la base de la conscience de soi (Bermúdez, Marcel et Eilan, 1998 ; Serino et Haggard, 2010). Selon Merleau-Ponty, cette conscience de soi est pré-réflexive, et en tant que telle est la base d'une conscience réflexive de soi et du monde extérieur (Merleau-Ponty, 1945). Le toucher joue ainsi un rôle central dans notre rapport au monde.

Le sens du toucher permet en effet de directement sentir et faire l'expérience des frontières existant entre le soi et le monde extérieur, ainsi que de l'interaction qui se joue entre les deux, autour de ces frontières. Les individus peuvent partiellement voir leur propre corps, mais ils ont surtout besoin de le sentir afin de réussir à prendre conscience d'eux-mêmes. Le neurologue Oliver Sacks pose cet aspect de l'expérience physique du corps comme étant au fondement de la conscience de soi lorsqu'il décrit certains de ses patients dont les perceptions de soi sont perturbées. Par exemple, une patiente ayant perdu le sens de la proprioception se sentait « désincarnée » ; ou encore, un patient qui ne ressentait plus sa jambe comme faisant partie de lui

²¹ « About a week ago, it started to get on my nerves... not being able to feel other people and hoping I could soon come out. I felt like everything was closing in on me and I couldn't stand it anymore. I just had to *feel* other people, I wanted to feel somebody, touch another human being. If I could have done this, I could have stuck it out longer... but I couldn't, there was no way I could touch anyone or in any way express my feelings toward somebody just by touching their hand or squeezing it. This is very difficult to explain – it leaves you at a loss for words. You just feel you are all alone in the world and everything is cold. There is no warmth. The warmth is all gone, and you just feel like there isn't anything. »

tentait de jeter cette jambe, qui lui était étrangère, hors de son lit (Sacks, 1986 ; Sacks, 1992). De telles situations, où les sensations tactiles sont perturbées, peuvent aussi se produire dans des cas non-pathologiques, comme par exemple lorsque l'on se fait anesthésier une dent chez le dentiste ou encore lorsque l'on se retrouve avec un membre engourdi, qui nous semble ne plus vraiment nous appartenir. Lorsque l'on touche alors ces parties du corps, elles nous paraissent en effet étrangères, comme si elles ne faisaient pas partie de nous ; l'expérience que l'on en a est alors celle d'une matière qui est comme « morte », par opposition au corps vivant que nous habitons et contrôlons.

Solliciter le sens du toucher nous donne la sensation d'être en contact avec le monde, et mène à un sentiment de conscience de soi. Le toucher permet aux individus de prendre conscience d'eux-mêmes ainsi que du monde extérieur, en leur donnant la possibilité de faire l'expérience des limites existantes entre les deux. Dans le même temps, il questionne ces limites et par extension les limites de la physicalité propre à chacun, étant donné que l'on ne fait pas l'expérience de ces limites comme étant des limites fixes. En effet, le sens du toucher peut paradoxalement aussi rendre floue la limite entre le soi dont on fait l'expérience et le monde. Un phénomène illustrant l'expérience de la limite entre le soi et le monde extérieur est la conséquence de la capacité des individus à sentir au travers d'objets, en intégrant ces objets dans leur corps propre (Polanyi, 1967). L'objet devient alors transparent pour nous et nous semble faire partie de nous (nous avons déjà traité cette question dans la partie 1.1.1.1.2). L'illusion de la main de caoutchouc fournit un autre exemple de cette idée de limite floue entre le soi dont on fait l'expérience et le monde : lorsqu'une personne regarde une main de caoutchouc se faire caresser, alors que sa propre main (cachée) est elle-même caressée en rythme avec la main de caoutchouc, la personne finit par attribuer la main de caoutchouc à son corps propre et à avoir l'impression que la main de caoutchouc est en fait sa propre main (Tsakiris et Haggard, 2005). Le sens du toucher nous confronte d'une part à la matérialité du monde, et d'autre part à notre propre matérialité. Ainsi il nous permet de faire l'expérience du monde comme étant réel et de dire « j'existe » (Damasio, 1999).

2.2.1.1.1.3 Fonctionnement

Le toucher est défini par la stimulation de la peau par des stimuli thermiques, mécaniques, chimiques ou électriques. Tous ces stimuli provoquent des changements au niveau de la peau, qui nous permettent d'avoir les sensations de la pression, de la chaleur et de la vibration ; cependant, après une journée à pratiquer la même activité, comme écrire avec un stylo, ou même

simplement porter nos vêtements, nous avons tendance à devenir moins sensibles à ces sensations.

Loomis et Lederman (1986) présentent trois aspects de l'interprétation de l'information par le biais du toucher, qui forment la perception tactile. Ils établissent en <position relative du corps (tête, tronc, membres,...). La kinesthésie concerne donc aussi la proprioception, la perception consciente de la position et des mouvements des différentes parties du corps. Lorsqu'une perception implique le sens cutané et/ou la proprioception, on peut donc parler de perception tactile ; par conséquent, Loomis et Lederman (ibid) identifient trois formes de perception tactile :

- la perception du tact, qui ne dépend que de variations des stimulations cutanées causées par des actions telles que le traçage d'un motif avec le doigt sur la peau d'un individu. Pour que seule cette perception soit en jeu, cela nécessite que l'individu en question soit immobile, sous peine que le sens kinesthésique ne soit impliqué.

- la perception proprioceptive concerne des variations au niveau de la stimulation kinesthésique. Cependant, pour qu'elle advienne sans que le sens cutané ne soit sollicité, il faut mettre en place des situations contraintes, en utilisant par exemple un anesthésique.

- la perception haptique est désignée par Loomis et Lederman comme étant la forme de perception tactile impliquant à la fois la stimulation cutanée et la stimulation kinesthésique. C'est la forme de toucher avec laquelle nous sommes le plus familiarisés, et nous l'utilisons chaque jour afin d'explorer et de comprendre notre environnement par le toucher.

Klatzky et Lederman (2002) désignent quant à eux le toucher comme étant un phénomène physiologique pouvant être classifié de deux façons. La première classification se réfère à la perception sensorielle du toucher, qu'elle soit cutanée ou kinesthésique. (La perception cutanée découle des capteurs de la peau, qui sont sensibles à la texture, la température, la douleur, et la perception kinesthésique est ressentie au travers de récepteurs se trouvant dans les muscles et les articulations, permettant ainsi de sentir la position relative des membres, et fournissant aussi la capacité à percevoir des forces.) La seconde classification est liée au fait que le toucher soit actif ou passif (Loomis et Lederman, 1984). Le toucher actif est perçu par la personne qui initie le toucher, tandis que le toucher passif est ressenti par la personne qui est touchée. Dans ses travaux sur la neurobiologie du toucher, Cullen établit que la raison de faire la distinction entre ces deux modes de toucher découle du fait qu'ils sont des processus traités de manière différente (Cullen, 2004). On notera que Merleau-Ponty posait déjà la distinction entre toucher et être touché et la réversibilité touchant-touché (Merleau-Ponty, 1964). En effet, du fait du contact physique nécessaire pour toucher, il y a une forme de réciprocité entre le corps perçu et le corps percevant : toucher, c'est aussi être touché. Cette inévitable réciprocité est caractéristique du sens du toucher. Voir n'implique pas nécessairement d'être vu, de même qu'entendre n'implique pas

nécessairement d'être entendu. Mais contrairement aux autres sens, toucher implique simultanément d'être touché. Toucher et être touché s'intègrent ensemble dans un même phénomène : c'est l'expérience du toucher. Il est d'ailleurs possible, en se concentrant sur la stimulation tactile, de se sentir en touchant : par exemple, en promenant le doigt sur une arête suffisamment nette, il est possible de sentir les aspérités et reliefs de la peau de la pulpe de notre doigt.

McGlone et Reilly (2010) suggèrent que le sens cutané, auquel quatre modalités sont traditionnellement associées (tactile, thermique, douleur et démangeaison), possède aussi une cinquième modalité chargée des propriétés plaisantes, de l'affect positif, du toucher. Les canaux sensoriels du toucher peuvent alors être considérés comme ayant des fonctions soit discriminatoires, soit émotionnelles. En effet on peut d'une part être informés de la localisation spatiale et temporelle d'événements se produisant à la surface du corps (comme la présence d'un insecte ou la température du vent), et d'autre part recevoir des expériences émotionnelles (négatives dans le cas de la douleur, mais aussi positives par exemple lors d'une caresse).

2.2.1.1.2 Caractéristiques du toucher

Le sens du toucher présente des caractéristiques qui lui sont propres et font de lui un sens à utiliser dans le cadre d'interactions motivées par la dimension émotionnelle et relationnelle des contacts.

D'une part, le toucher est un sens très riche, qui, par la perception active, nous permet non seulement de reconnaître un objet, mais aussi de le caractériser (texture, souplesse, dureté, température,...) (Klatzky, Lederman et Metzger, 1985). Nous pouvons identifier, caractériser, manipuler un grand nombre d'éléments de notre environnement par le biais du toucher.

D'autre part, le toucher est un sens privé, dont l'utilisation permet la discrétion. En effet, le toucher s'oppose de ce point de vue à la vue et à l'ouïe : ce que l'on voit ou entend est nécessairement public, et peut être perçu par d'autres se trouvant proches de soi. (Notons que l'on peut rendre privé un objet visuel ou auditif, mais cela nécessite le recours à un dispositif, comme par exemple un filtre de confidentialité ou des écouteurs, ce qui implique une situation d'usage spécifique pas toujours disponible). La perception par le toucher est au contraire privée : je suis précisément le seul à toucher ce que je touche. Il n'y a en effet qu'une seule surface de contact possible entre l'objet touché et le corps touchant (Pacherie, 1997). De ce caractère privé découle certainement la valeur émotionnelle associée à ce sens.

C'est d'ailleurs cette valeur émotionnelle qui nous intéresse plus particulièrement. Notons qu'à partir d'un sens du contact purement physique, le champ lexical du toucher a été étendu à

des répercussions psychiques et émotives (voir Lenay, 2010). Ainsi, le vocabulaire de l'affect et des émotions fait fréquemment référence au toucher. Nous détaillerons cet aspect dans la partie 2.2.1.2.2. Le toucher permet de plus les marques d'affect ; nous nous embrassons, nous serrons la main, nous prenons dans les bras,... Tous ces gestes sont chargés d'une valeur émotionnelle. Le toucher participe ainsi très largement à notre communication non-verbale ; nous allons détailler dans la sous-partie suivante tous les aspects communicationnels du toucher, ce qui inclut la valeur émotionnelle.

Notons que, même si les caractéristiques du toucher (qui est par définition un sens du contact direct) en font une piste de choix pour la communication émotionnelle, le fait de devoir le médiatiser implique la possible remise en question de ses caractéristiques.

2.2.1.2 Communication par le toucher

2.2.1.2.1 Le toucher : un canal de communication spécifique

Le toucher est notre sens le plus social. Il implique en effet un contact (soit l'interaction la plus physiquement proche) avec quelqu'un d'autre, et il est de ce fait extrêmement important pour les interactions sociales. C'est un aspect fondamental de la communication interpersonnelle, qui permet notamment un sentiment de connexion, l'indication de l'intention et l'expression des émotions (Brave et Dahley, 1997).

Il convient de noter qu'il existe de grandes différences culturelles concernant le toucher. Selon la culture, les formes de toucher considérées comme appropriées varient (Field, 2003). On peut noter comme formes de toucher répandues les poignées de main, les bises, les embrassades, les accolades, les étreintes, les baisers, ou encore le fait de se frotter le nez ou même de s'asseoir sur les genoux. Tous ces gestes peuvent se décliner. Par exemple, il existe divers types de poignées de main : la poignée de main où l'on nous encercle la main des deux mains (connue sous le nom de « poignée de main du politicien »), celle où l'on nous écrase les phalanges (intimidation), celle où l'autre interactant garde le bras tendu (pour nous tenir à distance), celle où l'on s'attrape par les avant-bras, celle encore où la personne s'incline en nous serrant la main... Une poignée de main faible et une poignée de main très vigoureuse seront interprétées différemment, et véhiculeront un message différent.

Les différences n'existent pas qu'au niveau de la culture ; on les trouve en effet aussi au niveau du genre. Dans une étude réalisée dans un hôpital par Fischer et Gallant (1990), 85% des patients ayant été touchés ont donné un avis positif sur l'hôpital et son personnel, contre seulement 53% des patients n'ayant pas été touchés ; ceux qui avaient été touchés récupéraient de leur traitement aussi plus facilement. Cependant, Fisher et Gallant ont procédé à une étude

plus contrôlée, où ils ont aussi pris en compte le genre de la personne qui était touchée. Les femmes touchées faisaient preuve de moins d'anxiété vis-à-vis de la chirurgie que celles n'ayant pas été touchées, alors que les hommes touchés étaient plus anxieux. Les femmes touchées avaient aussi plus tendance à tendre la main et à se saisir de celle des infirmières que les hommes, et avaient en moyenne une tension plus basse suite à l'intervention que les hommes touchés. Les auteurs suggèrent qu'être touché a pour effet sur les hommes de se sentir plus vulnérable et plus dépendant.

Les réactions différentes présentées par les hommes et les femmes au toucher s'expliquent en partie par la socialisation. Ashley Montagu suggère que les femmes étant souvent considérées et traitées comme étant inférieures aux hommes, elles sont plus souvent réceptrices qu'initiatrices du toucher (Montagu, 1986, p. 341). On retrouve cette notion dans une étude de Nancy Henley : en public, les hommes initient un toucher envers les femmes plus souvent que l'inverse, le plus souvent parce qu'ils sont considérés comme ayant un statut supérieur (Henley, 1973). De même, elle retrouve ces observations : en public, les personnes âgées initient plus souvent le toucher envers les jeunes que l'inverse, et les personnes ayant un revenu plus élevé initient plus souvent un toucher envers celles ayant un revenu moins élevé que l'inverse. Une étude de Jourard montre que les bébés filles sont plus fréquemment touchés que les garçons, et ce par les deux parents, et que les filles touchent leurs deux parents plus souvent que les garçons (Jourard, 1966). Dans une autre étude, il montre que les mères comme les pères touchent plus de zones différentes du corps des filles que des garçons, et qu'en retour les filles touchent leurs parents sur plus de zones que les fils (ibid). On peut facilement imaginer que les différences genrées observées à l'âge adulte résultent de ces différences précoces dans la façon de toucher. Le comportement masculin plus agressif pourrait d'ailleurs être lié à cela (ibid).

Le toucher permet d'influencer les comportements, attitudes ou sentiments des autres. On peut par exemple toucher quelqu'un pour lui enjoindre de se dépêcher, de se déplacer, de rester où il est, ou encore de faire quelque chose. On peut se servir du toucher pour attirer l'attention, ou pour attirer le regard dans une certaine direction. Certains touchers consistent en des rituels, par exemple les touchers liés à l'arrivée ou au départ, comme les poignées de main ou les bises.

Il existe différentes façons de toucher. Thayer (1982) classifie les touchers en différents types, établissant ainsi une taxinomie du toucher : le type fonctionnel-professionnel (le docteur qui touche son patient pour l'examiner), le type social-poli (la poignée de main), le type amical-chaleureux, le type amoureux-intime et le type sexuel. Cette taxinomie est basée sur et illustrée par les relations des individus. Thayer observe que cette façon d'ordonner les catégories peut être caractérisée par un niveau d'intimité croissant : plus de parties du corps deviennent accessibles, les touchers deviennent plus longs et plus fréquents, ainsi qu'une variété croissante des types de touchers impliqués. Fagan (1998) postule l'existence de différents types de touchers, qui vont du

toucher public et formalisé au toucher intensément personnel ; toucher rituel, toucher sportif, toucher de punition, toucher éducatif, toucher évocateur d'intimité, ou encore toucher sexuel. Ces catégories ne sont pas mutuellement exclusives, et il existe vis-à-vis du toucher de nombreuses significations et de nombreux besoins sous-jacents.

Le toucher comme coïncidence du corps qui perçoit et du corps perçu est le sens du contact interpersonnel. À l'intérieur de ce contact, ce sens constitue un fondement fort pour le développement de l'affection et de l'intimité (Field, 2003 ; Montagu, 1986). Le toucher est synonyme de contact, et par conséquent d'une implication physique, tandis que voir et entendre sont des sens plus à même de créer une distance, une séparation entre le sujet percevant et l'objet perçu. Dans les études sociales, le sens du toucher est par conséquent souvent considéré comme notre sens le plus social (Field, 2003). Lors du toucher, la communication est effectivement incarnée, et nos interactions tactiles sont porteuses d'un message tangible d'interconnexion : celui de ne pas être seuls (ibid). Le toucher interpersonnel nous indique si nous sommes en sécurité, si l'on est choyés, et si nous avons de la valeur (Field, 2003 ; Finnegan, 2002).

Le toucher nous permet notamment d'enrichir nos relations sociales, essentielles à notre survie (Dunbar, 2010). Le toucher constitue en fait le canal de communication interpersonnelle le plus intime dont nous disposons (Thayer, 1986). Fagan (1998) suggère même que le toucher est le premier langage par lequel nous apprenons à communiquer l'affection interpersonnelle. Afin de pouvoir fonctionner correctement dans le cadre des contacts sociaux, il faut être capable de comprendre (et d'exprimer) le langage du toucher correctement.

Notons que la signification du toucher est considérablement liée au contexte. Jones et Yarbrough (1985) posent qu'il n'existe pas qu'une seule relation de correspondance entre significations et attitudes tactiles. D'une part, le même comportement au niveau tactile peut en effet avoir plusieurs significations, et d'autre part différentes attitudes tactiles peuvent avoir la même signification. Ils insistent sur le type de toucher, le type de parties du corps touchées et le rôle du contexte (qui inclut l'âge, le genre, la relation, le statut, l'environnement, la culture, le timing du toucher, les expressions faciales accompagnant le toucher, etc). Le contexte possède en effet un rôle certain dans la construction de la signification du toucher. On retrouve ici l'idée du *participatory sense-making* de De Jaegher et Di Paolo (2007).

Les aspects sociaux du toucher interpersonnel peuvent aussi être importants pour la signification du toucher dans l'interaction homme-produit. Lorsque nous sommes en relation avec des objets, le toucher peut aussi être considéré comme un canal de communication, qui fait entrer en jeu différents styles affectifs, qui exprime différents types de relations affectives, allant des plus fonctionnelles-professionnelles aux plus intimes.

2.2.1.2.2 Le toucher comme sens du contact et du touchant

Dans la vie de tous les jours, le toucher est en fait un canal parmi les plus importants pour la communication des émotions. D'après Saul Schanberg, le toucher est notre contact le plus fort : « Le toucher est dix fois plus puissant que le contact verbal ou émotionnel, et il affecte pratiquement tout ce que nous faisons. Aucun autre sens ne peut être aussi excitant que le toucher. C'est une chose connue depuis longtemps, mais dont nous n'avions pas réalisé qu'elle avait une base biologique. Si le toucher ne procurait pas de sensations agréables, il n'y aurait ni espèces, ni paternité ou maternité, ni survie. La mère ne toucherait pas son enfant de la bonne façon si elle ne recevait pas de plaisir à le faire. Si nous n'aimions pas le fait de nous toucher ou nous caresser les uns les autres, nous n'aurions pas de relations sexuelles. Les animaux qui ont d'instinct plus touché leurs petits ont produit une progéniture qui survivait mieux et avait plus d'énergie, et par conséquent ont transmis leur tendance au toucher, qui s'en est vue renforcée. Nous oublions que le toucher est non seulement élémentaire pour notre espèce, mais qu'il en est la clé même. » (Schanberg, 1995, in Field, 2003, p. 57, traduction personnelle)²².

En ce qui concerne la communication des émotions, le toucher permet de communiquer le ton hédoniste de l'émotion, à savoir soit des éléments positivement connotés (chaleur, intimité), soit des éléments négativement connotés (douleur ou inconfort) (Hertenstein, 2005 ; Hertenstein et Campos, 2001 ; Jones et Yarbrough 1985 ; Knapp et Hall, 1997). Le toucher peut aussi servir à intensifier la communication relative aux émotions (Knapp et Hall, 1997). Wang et Quek (2010) posent que le toucher est porteur d'une information affective d'un genre unique. Ils introduisent l'idée d'immédiateté du toucher ; le toucher affectif ne passe pour eux par aucune transposition par une expression symbolique. Ainsi, l'initiateur du toucher ne pense pas à un message qu'il voudrait exprimer par le biais du toucher. Il n'essaie pas de dire quoi que ce soit, et le récepteur du toucher n'interprète pas un message – la réception de l'affect est immédiate. Cette idée, proche du sens commun, se retrouve dans l'exemple du baiser entre deux amants – ils s'embrassent simplement, et ne sont pas en train d'essayer d'affirmer quelque chose chacun de leur côté.

²² « Touch is ten times stronger than verbal or emotional contact, and it affects damned near everything we do. No other sense can arouse you like touch. We always knew that, but we never realized it had a biological basis. If touch did not feel good, there would be no species, parenthood, or survival. The mother would not touch her baby in the right way unless the mother felt pleasure in doing it. If we did not like the feel of touching and patting one another, we would not have had sex. Those animals that did more touching instinctively produced offspring which survived and had more energy, and so passed around their tendency to touch which became even stronger. We forget that *touch is not only basic to our species, but the key to it.* »

Hertenstein et al. (2006 ; 2009) ont mené des expériences et ont mis en avant le fait que les individus peuvent reconnaître des émotions (notamment la colère, la peur, le dégoût, l'amour, la gratitude, la sympathie, le bonheur, la tristesse) simplement en faisant l'expérience d'être touché sur le bras ou sur tout le corps par une personne qu'ils ne peuvent pas voir et ne connaissent pas. Les taux de reconnaissance varient de 48% à 83%, ce qui est comparable à la précision du décodage d'émotions transmises par la voix et les expressions faciales (Elfenbein et Ambady, 2002). Ces résultats sont étendus par Bailenson et al. (2007) dans le cadre d'une médiation du toucher ; au cours d'une étude expérimentale, ces chercheurs ont évalué les capacités pour les sujets d'exprimer sept émotions (dégoût, colère, tristesse, joie, peur, intérêt et surprise) par la manipulation d'un système à retour d'effort et de reconnaître ces émotions lorsqu'elles étaient rejouées par le système. Ils ont tiré de leurs observations des descripteurs des mouvements utilisés pour chaque émotion (vitesse, régularité, amplitude,...), et ont mis en avant le fait que les sujets reconnaissaient les émotions jouées dans une proportion supérieure au hasard. Ces résultats sont aussi trouvés par Smith et MacLean (2007), qui ont fait interagir des dyades par le biais d'un dispositif de médiation du toucher. Ils observent que les émotions peuvent être transmises par le biais du toucher (ibid). Kaitz (1992) a d'autre part montré que des amants dont on aurait bandé les yeux peuvent se reconnaître simplement en se touchant la main.

Gallace et Spence proposent un récapitulatif des différentes études concernant le toucher entre les personnes (Gallace et Spence, 2010). Ils posent la surface du corps humain comme étant non seulement un organe de perception mais aussi comme un système de communication interpersonnel très puissant. Le toucher social joue en effet un rôle des plus importants (si ce n'est le plus important) dans le développement psychologique des hommes. Malgré cela, les aspects les plus perceptifs du toucher ont bien plus souvent été étudiés que ses aspects plus interpersonnels et sociaux. L'article fait état des différentes recherches concernant ce sujet et montre que :

- le toucher peut amener les individus à plus facilement accéder à des requêtes (un phénomène connu sous le nom d'effet « Midas touch ») et à modifier leurs jugements de certains services. Il est étonnant de noter que cet effet a lieu que les personnes puissent se souvenir du contact tactile ou non ;
- le toucher peut aider à créer des liens entre les personnes (dans un couple, dans des groupes, etc). Cet effet est au moins en partie dû à l'émission hormonale d'ocytocine qui suit les contacts interpersonnels incluant une stimulation tactile ;
- le toucher est aussi efficace que la vision pour véhiculer des messages émotionnels. En particulier, la colère, la peur, le dégoût, l'amour, la gratitude et l'empathie peuvent être facilement communiqués uniquement par le biais du toucher.

Gallace et Spence soulignent l'importance du toucher dans les interactions sociales de tous les jours (ibid). Ils posent le fait que la stimulation tactile interpersonnelle constitue un moyen effectif d'influencer le comportement social des individus. Le toucher interpersonnel joue un rôle important dans le contrôle de notre bien-être émotionnel ; ce point est critique dans la conception d'interfaces pour communiquer ses émotions à distance. Si l'on réussit à proposer une alternative qui permette au toucher d'exercer les mêmes fonctions que dans l'interaction directe, le fait de pouvoir agir sur le bien-être émotionnel de notre interlocuteur est tout à fait fondamental.

La peau contient des récepteurs pouvant provoquer en eux-mêmes des réponses émotionnelles. Le toucher interpersonnel correspond à ce qu'il y a de plus émotionnel dans nos expériences tactiles.

Compenser le manque du contact « chair à chair » relève de la gageure dans les communications et les relations virtuelles ; d'ailleurs on reproche aux environnements virtuels collaboratifs de manquer de chaleur émotionnelle et d'intimité non-verbale. Suite à une étude phénoménologique menée par Register et Henley (1992), il a d'ailleurs été mis en lumière que la communication non-verbale est critique pour faire l'expérience de l'intimité, et que cette expérience est difficile à expliquer ou à décrire avec des mots. « It seems that there is an element of intimacy that is more accurately expressed via other sensory modalities, such as sight and touch, and in the absence of language. » (ibid, p. 473)²³.

Lenay (2010) relève que le sens du toucher est utilisé dans le vocabulaire émotionnel (par exemple : « touchant », « poignant », « saisissant », « être touché au cœur »,...). Le champ sémantique de la fonction perceptive tactile, du contact avec le tangible, permet une métaphore au niveau émotionnel. Lenay définit un contact touchant comme étant à la source d'une émotion ressentie suite à la perception de l'activité du partenaire avec lequel on interagit. C'est un contact interpersonnel et émouvant. Analysant l'émotion comme une force qui met en mouvement (émotion), il pose que les contacts touchants sont basés sur la dualité corps percevant (Leib) et corps perçu (Körper). Lors d'un contact touchant, une rupture de symétrie perceptive se produit : l'ignorance du sujet de son corps perçu par autrui est alors révélée (ibid). Si l'on peut émouvoir quelqu'un, c'est parce qu'il y a une rupture entre ce que l'on perçoit et ce que l'on donne à percevoir (autrui peut percevoir des choses qui ne me sont pas accessibles, comme par exemple mon propre visage). Réciproquement, ce qui m'émeut, c'est que la dynamique de l'activité

²³ Traduction personnelle : « Il semblerait qu'il existe une composante de l'intimité qui est plus précisément exprimée par le biais d'autres modalités sensorielles, comme la vue et le toucher, et en l'absence de langage. »

perceptive d'autrui lui échappe lors de notre rencontre. Autrui et moi ne sommes en effet pas à la même place : je peux percevoir autrui d'une manière dont il ne peut lui-même se percevoir, et réciproquement. Si je ne perçois pas ce qui me permet de percevoir (mon Leib), je ne perçois pas non plus ce que je donne à percevoir (mon Körper ; je ne vois pas mes yeux tant qu'ils me donnent à voir), mais autrui peut par contre percevoir mon Körper (mes yeux). Ainsi, lors d'un croisement perceptif, ma propre activité perceptive m'échappe. C'est même toute la dynamique collective de nos interactions qui nous échappe, étant donné qu'elle dépend de nos images (images que nous ne percevons pas nous-mêmes). Il y a une forme d'asymétrie d'engagement dans la relation : chacun agit dans l'ignorance de ce qu'il fait pour autrui. Cette asymétrie est réciproque : j'ignore ce qu'autrui rencontre de moi, tandis qu'autrui ignore ce que je rencontre de lui. C'est là que se situe la possibilité d'un contact touchant. « On comprend aussi que [cette asymétrie réciproque des engagements dans l'interaction] donne une valeur émotionnelle au contact. En effet, cela signifie que je suis nu, livré au regard d'autrui, et qu'il est nu, offert à mon regard. Notre visage est nu, non pas au sens où il ne serait pas protégé par un masque ou un maquillage, mais parce que nous le présentons aux regards des autres sans pouvoir le contrôler comme on le ferait d'une chose devant soi. Au contraire, il est pris dans une dynamique qui lui échappe. » (ibid, p. 385) Lenay met ainsi en évidence un paradoxe du touchant : ce qui est touchant chez celui qui nous touche, c'est en fait ce qui lui échappe, et qui le révèle donc véritablement.

2.2.1.3 Médiation des interactions haptiques

Les qualités du toucher et son impact dans la communication émotionnelle en font donc une piste de choix pour la communication émotionnelle à distance. On notera cependant que les interfaces haptiques ne sont pas très répandues dans les dispositifs accessibles au grand public. Dans les téléphones mobiles, on trouve des vibreurs, qui se déclenchent pour annoncer un message ou un appel ; non couplés à la sonnerie, ils sont utilisés pour le mode « discret » du téléphone. Les vibreurs sont aussi intégrés dans les jeux vidéos, notamment pour participer à l'effet d'immersion dans le jeu ; ils permettent en effet de simuler par exemple les impacts avec des obstacles ou encore les battements de cœur de l'avatar.

La notion de technologies phatiques (phatic technologies), développée par Vetere, Howard et Gibbs (2005) et Gibbs et al. (2005), désigne les systèmes permettant d'établir et de maintenir les interactions sociales. Ces systèmes ne cherchent donc pas à capter et communiquer l'information en tant que telle, mais à permettre la construction et l'entretien de relations.

Hassenzahl et al. (2012) notent qu'un nombre croissant de couples sont forcés de vivre éloignés, et que par conséquent, de plus en plus de chercheurs et designers proposent des dispositifs pour la communication à distance, permettant de soutenir la présence, l'awareness, la connectedness, la relatedness, l'intimité,... (Nous avons détaillé ces concepts en partie 1.1.3.2.) Ils notent que l'usage explicite de connaissances théoriques et empiriques (notamment psychologiques) ne se fait que dans un cas sur deux lors de la conception. Les sujets les plus fréquemment évoqués sont : l'intimité (Chung, Lee et Selker, 2006 ; Vetere et al., 2005), l'amour (Pujol et Umemuro, 2009 ; Saslis-Lagoudakis et al., 2006), la communication (Lindley, Harper et Sellen, 2009 ; Tsujita, Tsukada et Siio, 2009), l'émotion (Li et Jianting, 2009 ; Tollmar et Persson, 2002), le toucher (Chang et al., 2002 ; Motamedi, 2007) et le jeu (Feltham, Vetere et Wensveen, 2007). En faisant la revue de 143 artefacts permettant la médiation de relations proches et ayant mené à une publication (notons que tous ces dispositifs n'ont pas pour vocation de médiatiser le toucher), ils identifient six stratégies utilisées lors de la conception en vue de permettre une expérience de relatedness à distance :

- l'awareness (le fait d'avoir connaissance, d'être conscient). Ces dispositifs créent un sentiment d'awareness cognitive et de continuité en permettant le partage de différents types d'information (notamment ambiante) concernant les activités en cours ou les humeurs du moment entre les partenaires (sans qu'ils ne conversent ou ne partagent une quelconque activité). Les sous-stratégies consistent à exposer la présence, l'activité ou l'humeur. Les principes psychologiques et les prérequis nécessaires d'un point de vue design sont la révélation de soi réciproque et l'ambiguïté.

- l'expressivité. Ces dispositifs mettent l'accent sur l'aspect affectif et émotionnel de l'intimité. Ils permettent aux partenaires d'exprimer leurs sentiments et leurs émotions par des biais très variés, comme par exemple en développant leur propre langage, ou en utilisant le langage habituel de manière ambiguë. Les sous-stratégies incluent des symboles. Les principes psychologiques et les prérequis nécessaires d'un point de vue design sont une expression enrichie des émotions, la réciprocité, l'intégration dans les habitudes quotidiennes, l'ouverture de l'interprétation et la communication phatique.

- la physicalité. Ces dispositifs permettent la médiation d'un sentiment d'intimité physique. Ils simulent soit des effets secondaires de la proximité physique (par exemple la chaleur corporelle, ou les battements du cœur) soit des gestes signifiants (comme les câlins, ou les caresses). Les sous-catégories sont les paramètres physiologiques et les gestes. Les principes psychologiques et les prérequis nécessaires d'un point de vue design sont la réciprocité, la simultanéité et les contraintes contextuelles.

- le don de présents. Ces dispositifs permettent de faire la démonstration de son affection et de la valeur que l'on porte à l'autre participant par le biais d'un présent. Les principes psychologiques

et les prérequis nécessaires d'un point de vue design sont la réflexion, l'effort, l'appréciation, la prévenance et la similarité, la communication symbolique.

- l'action conjointe. Ces dispositifs permettent de réaliser une action ensemble, de façon conjointe, ce qui nécessite normalement de se situer dans le même espace. Les sous-stratégies incluent des habitudes, établies et nouvelles. Les principes psychologiques et les prérequis nécessaires d'un point de vue design sont la communication activante, l'interdépendance du comportement, la sélection des activités et la sérendipité.

- les souvenirs. Ces dispositifs gardent la trace d'activités passées et de moments particuliers de la relation. Les principes psychologiques et les prérequis nécessaires d'un point de vue design sont les souvenirs (memorabilia), l'engagement et la tangibilité.

La médiatisation haptique d'interactions interpersonnelles est un enjeu de plus en plus important. Postulant que le sens du toucher est un vecteur majeur pour la communication des émotions, nous allons dans la sous-partie suivante procéder à un état de l'art des dispositifs de médiatisation du sens du toucher. Nous verrons les diverses approches adoptées. En effet, selon les aspects du toucher que les concepteurs cherchent à transmettre, les propositions sont extrêmement variées : retour de force, texture, vibration, chaleur,...

2.2.1.3.1 État de l'art des dispositifs pour la communication émotionnelle à distance²⁴

Les dispositifs haptiques permettent de se toucher à distance, ou permettent en tous cas une forme de toucher à distance. Ce sont eux qui nous intéressent tout d'abord ici ; nous allons en effet ici présenter un état de l'art des dispositifs pour la communication émotionnelle à distance. L'effervescence actuelle attachée au développement de ce type d'interfaces ne nous permet pas d'en faire un état de l'art exhaustif ; nous présentons ici les dispositifs qui nous semblent les plus fondateurs et les plus proches de nos propres travaux. On constatera que beaucoup de ces dispositifs n'existent qu'à l'état de prototypes, et qu'aucun n'est répandu auprès du grand public.

Nous allons présenter en premier lieu des systèmes de médiation haptique, avant de faire, dans un second temps, un état des lieux des autres systèmes permettant la communication émotionnelle à distance.

²⁴ Pour une revue des dispositifs haptiques à distance, voir Benali-Khoudja et al. (2004), Fuchs (2007) et Pasquero (2008), ainsi que Haans et IJsselsteijn (2006). Pour une revue de dispositifs mobiles intégrant de l'haptique, voir Kwon et Kim (2008). Pour une revue des dispositifs soutenant la médiation de relations intimes à distance, voir Hassenzahl et al. (2012) et Howard et al. (2006).

2.2.1.3.1.1 Systèmes de médiation haptique

Nous allons ici exposer les propositions pour un toucher à distance.

Le Telephonic Arm Wrestling (voir figure 29), conçu par Norman White et Doug Back en 1986 permettait en théorie de s'adonner à un bras de fer à distance (le système souffrait en pratique de problèmes liés aux délais de réseau).

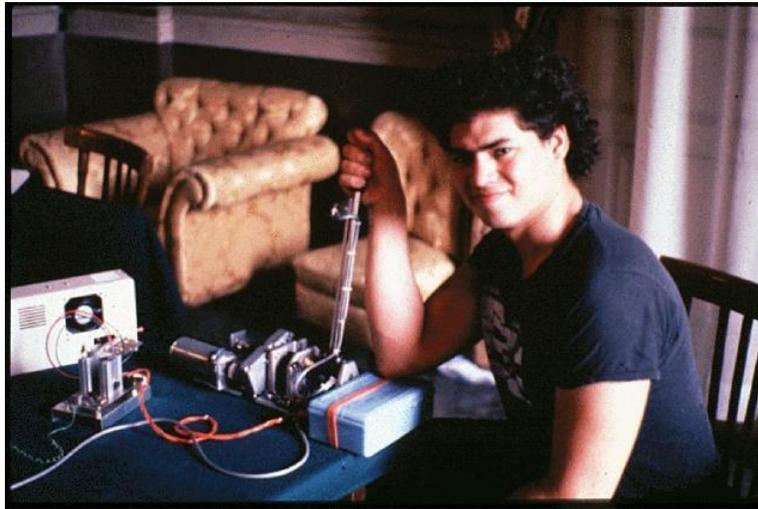


Figure 29. Le Telephonic Arm Wrestling (White et Back, 1986).

inTouch est un système de médiation haptique développé par Brave et Dahley (1997). L'idée derrière ce prototype est de créer l'illusion que deux personnes à distance interagissent avec un objet physique partagé. En réalité, chaque utilisateur interagit avec son propre objet, mais lorsque l'un des dispositifs est manipulé, les deux objets sont affectés. Les deux objets connectés sont trois rouleaux montés sur une base (voir figure 30). Quand l'un des rouleaux est tourné, le rouleau correspondant tourne de la même façon sur l'autre dispositif.



Figure 30. Les rouleaux du dispositif inTouch (Brave and Dahley 1997).

Brave, Ishii et Dahley (1998) ont développé les travaux engagés avec inTouch. Ils préconisent l'utilisation d'interfaces tangibles pour enrichir la communication et la collaboration à distance (voir figure 31). Cette approche, basée sur le toucher et la physicalité, s'appuie sur le concept d'Objets Physiques Distribués et Synchronisés (Synchronized Distributed Physical Objects), qui utilisent la télémanipulation afin de créer l'illusion que des utilisateurs à distance interagissent avec des objets physiques partagés. Deux applications basées sur cette approche sont présentées : PSyBench, un espace de travail partagé (voir figure 32), et inTouch, que nous avons présenté juste au-dessus.

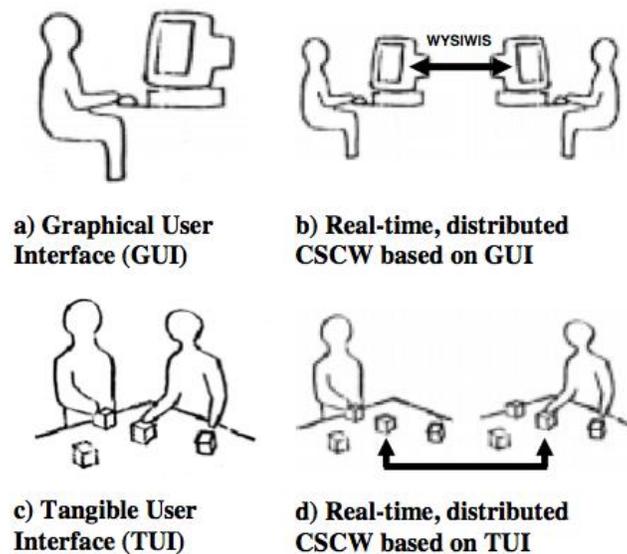


Figure 31. L'opposition entre interfaces graphiques et interfaces tangibles pour la communication à distance (Brave, Ishii et Dahley, 1998).

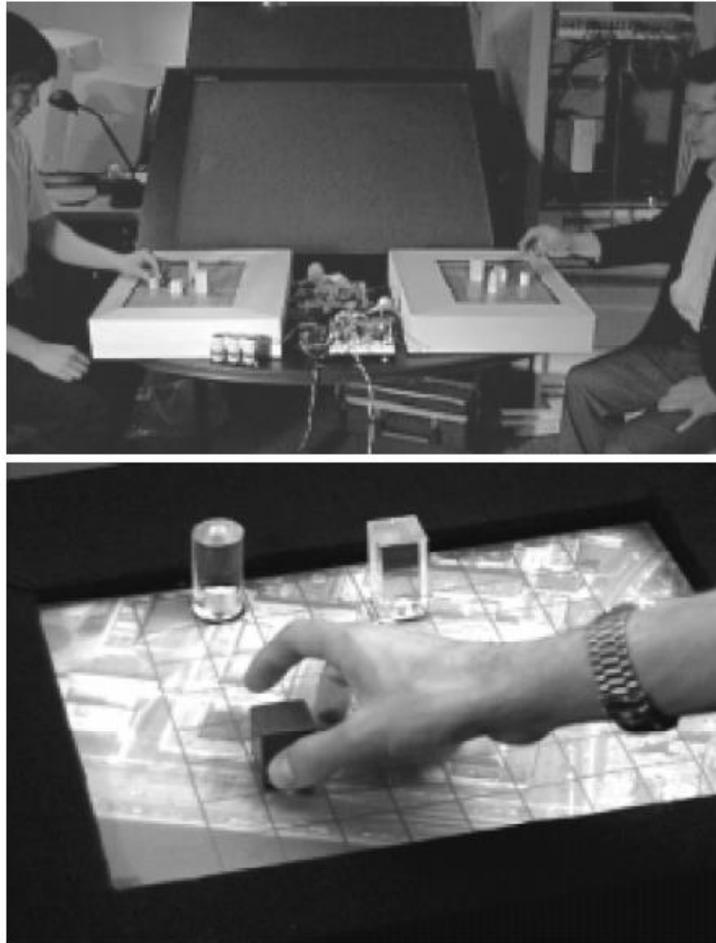


Figure 32. Un prototype de PSyBench (Brave, Ishii et Dahley, 1998).

Dobson et al. (2001) ont quant à eux développé un dispositif de communication interpersonnelle par du vibrotactile et de la chaleur (VibroBod, équipé de vibrateurs, de capteurs de pression et de micros, voir figure 33) ainsi qu'un gant vibrant et chauffant pour la navigation dans des newsgroups (What's shaking) en vue d'enrichir l'interaction interpersonnelle dans un



Figure 33. VibroBod de (Dobson et al., 2001).

espace digital. Ils ont posé le fait qu'ajouter de la température et de la vibration dans les communications interpersonnelles à distance facilitait l'échange de contenu émotionnel et social.

Oakley et O'Modhrain (2002) ont développé Contact IM. C'est un système qui permet d'accompagner l'échange de messages (envoyés par une messagerie instantanée) d'une activité haptique. Il s'agit de la réplique d'une activité physique coopérative simple, à savoir l'échange d'une balle. Lorsqu'un des deux utilisateurs souhaite envoyer un message, il le rédige puis la moitié d'un court de tennis lui est présentée. Trois objets signifiants s'y trouvent : une icône de main (représentant l'utilisateur), une balle (représentant le message) et un filet. En utilisant un dispositif à retour de force, l'utilisateur peut faire bouger la main jusqu'à la balle, et la « ramasser » en appuyant sur un bouton. Une fois ramassée, la balle est attachée au curseur en forme de main de l'utilisateur par un élastique (voir figure 34). L'utilisateur peut alors faire aller et venir la balle, afin de décider l'élan qu'il va lui appliquer. La balle peut être lancée en relâchant le bouton ; elle va rebondir jusqu'à passer par dessus le filet, disparaissant alors sur cet écran et apparaissant sur l'écran du destinataire où se trouve l'autre moitié du court. Le partenaire est équipé du même dispositif, grâce auquel il peut attraper la balle et ainsi accéder au message, et la renvoyer à l'autre utilisateur avec sa réponse.



Figure 34. Contac IM : la balle attrapée est sur le point d'être lancée (Oakley et O'Modhrain, 2002).

Shin et al. (2007) se sont aussi penchés sur l'enrichissement tactile des messageries instantanées. Ils proposent une interface tactile basée sur des émoticônes visuelles et tactiles (TCONS, pour TouchCON, Touch Emoticon) ; ces émoticônes peuvent être définies par les utilisateurs eux-mêmes. Les émoticônes sont émises par le biais du TCON Display, composé de

deux dispositifs ayant la forme grossière de mains, et d'un dispositif ayant la forme de lèvres. Ces dispositifs sont équipés de vibrateurs, de picots, de retours thermiques, de capteurs de pression, de boutons et de LED (voir figure 35).

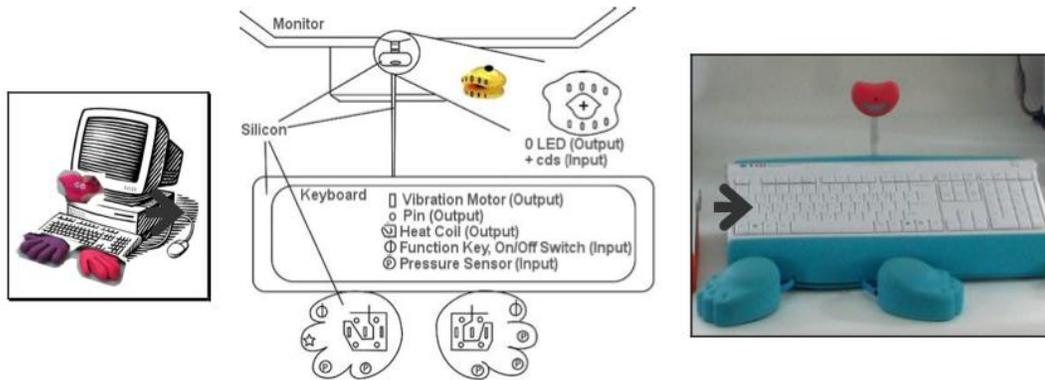


Figure 35. Le TCON display de Shin et al (2007).

eMoto, développé par Sundström, Ståhl et Höök (2005), est un système basé sur l'échange de messages affectifs. Un utilisateur rédige un message qu'il accompagne d'une expression graphique ; cette expression est choisie suite à l'écriture du message, par le biais des gestes (pression et secouement) de l'utilisateur sur le stylet (équipé d'un accéléromètre et d'un capteur de pression) accompagnant son téléphone (voir figure 36).

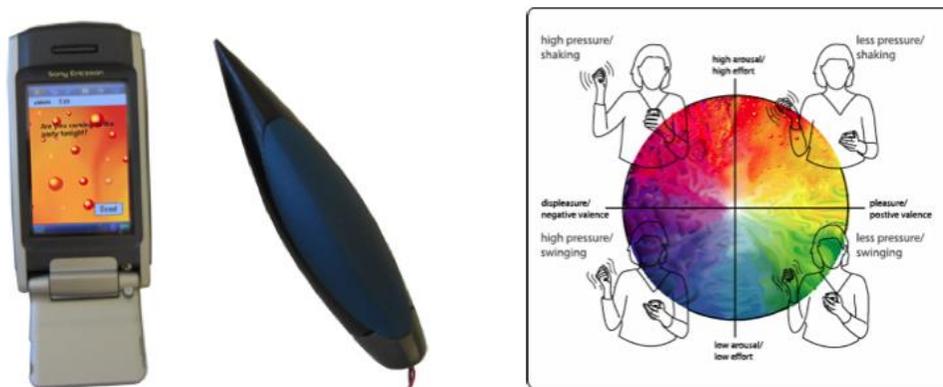


Figure 36. eMoto : le stylet et les possibles expressions à envoyer selon les gestes (Sundström, Ståh et Höök, 2005).

ComTouch est un prototype développé par Chang et al. (2002). C'est un dispositif destiné à l'augmentation des communications téléphoniques par l'ajout d'un retour tactile. La pression de la main des utilisateurs est convertie en vibrations d'intensité différente qui sont transmises par le biais d'un gant (voir figure 37).



Figure 37. Le prototype de ComTouch (Chang et al, 2002).

POKE (Park, Hwang et Nam, 2011 ; Park, Baek et Nam, 2013 ; Park et Nam, 2013), tout d'abord nommé CheekTouch (Park, Lim et Nam, 2010 ; Park, Bae et Nam, 2012), est un dispositif permettant un retour tactile sur la joue lors d'échanges téléphoniques par téléphone portable (voir figures 38 et 39). Le retour tactile s'effectue par le biais d'une surface gonflable ; elle se gonfle contre la joue selon les gestes et la pression appliquée par l'autre utilisateur.



Figure 39. POKE en utilisation ; la surface se gonfle plus ou moins selon la force appliquée (Park, Lim et Nam, 2010).

Hemmert et al. (2011) proposent l'intégration de nouvelles façons de communiquer par le biais de téléphones mobiles en utilisant des indices de nature physique, comme le fait de saisir, d'embrasser et de chuchoter. La prise est simulée par un système sensible à la pression se trouvant autour du téléphone ; lorsque le téléphone est serré, un système en forme de boucle à température de la peau se resserre proportionnellement autour de la main de l'interlocuteur (voir figure 40).



Figure 40. La simulation de la saisie (de la main) par le prototype de Hemmert et al. (2011).

Le baiser est simulé grâce à un capteur d'humidité sur le téléphone de l'envoyeur, et d'une membrane semi-perméable recouvrant une éponge humide sur le téléphone du receveur (voir figure 41). Plus le capteur de l'envoyeur est mouillé, plus le téléphone du receveur va laisser échapper d'eau. Le système permet donc la télécommunication de baisers, mais aussi de larmes ou de transpiration, selon le contexte.



Figure 41. La simulation du baiser par le prototype de Hemmert et al. (2011).

La troisième proposition, celle simulant le chuchotement, consiste en des capteurs de courant d'air sur le téléphone de l'envoyeur, et en un système de jet d'air à faible pression sur le téléphone du receveur (voir figure 42). Le capteur est sensible à l'air expulsé par les narines et par la vocalisation, tandis que le jet d'air est envoyé avec une force proportionnelle à ce qui est capté.



Figure 42. La simulation du souffle par le prototype de Hemmert et al. (2011).

Hoggan et al. (2012) proposent ForcePhone (voir figure 43). Ce système de communication permet un échange haptique pendant une conversation téléphonique. Lorsqu'un utilisateur serre son téléphone, le niveau de pression est converti en vibrations du téléphone de l'autre utilisateur. Les auteurs désignent les messages liés à la pression et au vibrotactile du ForcePhone par le terme pressages.



Figure 43. Le ForcePhone de Hoggan et al. (2012).

Un autre dispositif permettant d'enrichir les interactions par le biais de smartphone est le brassard développé par Wang et al. (2010 ; 2012, voir figure 44). Il est équipé d'un mécanisme

moteur et est couplé à une coque de téléphone comportant un capteur de pression. Lorsqu'un utilisateur serre son téléphone, son interlocuteur sent son brassard se serrer autour de son bras.

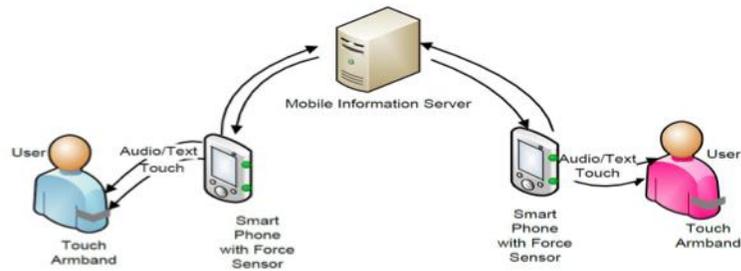


Figure 44. Le système de Wang et al. (2010 ; 2012)

KUSUGURI est une application smartphone développée par Furukawa, Kajimoto et Tachi (2012). L'interface tactile proposée permet de partager une partie de son corps, la main, avec un autre utilisateur à distance, et de ressentir le contact de l'autre individu ; on a l'illusion qu'il nous chatouille (voir figure 45). L'application permet le contact direct, le transfert de la sensation de chatouille et la bidirectionnalité. Le contact direct se fait dans la mesure où l'un des utilisateurs peut voir un doigt de l'autre participant comme s'il était directement sur sa main et s'y déplaçait. Les chatouilles découlent de la vibration du téléphone dans la main de l'utilisateur, ce qui, couplé aux déplacements du doigt de l'autre interactant, entraîne l'illusion de la perception de chatouilles. Enfin, les chatouilles peuvent être bidirectionnelles étant donné que l'utilisateur chatouillé peut aussi chatouiller l'autre participant.



Figure 45. Un utilisateur de KUSUGURI ayant la sensation d’être chatouillé par un utilisateur distant. Le doigt de l’autre participant se déplace sur l’image de sa propre paume et il reçoit une faible vibration synchronisée avec les mouvements du doigt, ce qui donne l’illusion d’être chatouillé (Furukawa, Kajimoto et Tachi, 2012).

CoupleVIBE est une application smartphone développée par Bales, Li et Griwsold (2011), visant à améliorer l’awareness entre les couples à longue distance. L’application envoie automatiquement les informations de localisation d’un utilisateur à son ou sa partenaire par le biais de signaux vibratoires différenciés (voir figure 46).

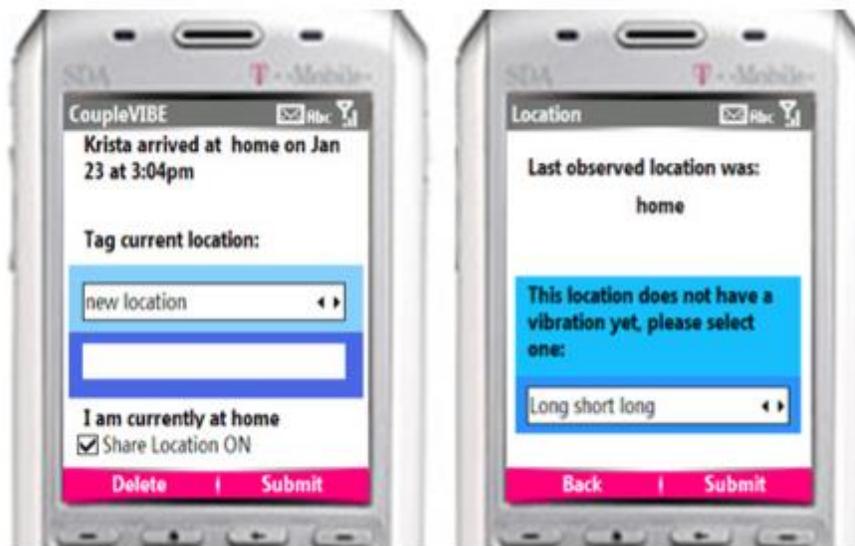


Figure 46. L’application CoupleVIBE (Bales, Li et Griwsold (2011)).

DiSalvo et al. (2003) ont développé The Hug (L'Étreinte, ou Le Câlin), un produit permettant la communication intime à distance par le biais d'une peluche (voir figure 47). Lorsqu'un utilisateur enlace son Hug, la peluche possédée par l'autre interactant émet de la musique et une lumière. Les interactants peuvent communiquer oralement par le biais des peluches et en les caressant, ils font vibrer celle de leur partenaire. Plus l'interaction dure, plus la peluche se réchauffe. À la fin de l'interaction, la peluche se met à nouveau à émettre de la musique et de la lumière, ce qui indique cette fois la fin de l'interaction. Si lors de l'activation d'une peluche l'autre interactant n'est pas présent, il est alors possible de laisser un message par le biais du dispositif ; sont alors enregistrées la voix et les caresses délivrées sur le Hug. L'autre interactant pourra jouer le message lorsqu'il sera de retour et disponible.



Figure 47. Une utilisatrice tenant son Hug (DiSalvo et al., 2003).

Mueller et al. (2005) proposent Hug over a distance (une étreinte à distance). Ce système se compose d'un koala en peluche et d'une veste dans laquelle se trouve des petits compartiments d'air (voir figure 48). Lorsque l'interactant qui possède le koala touche l'écran tactile du PDA incrusté dans le ventre de la peluche, cela envoie un signal à la veste de l'autre interactant, dont les compartiments d'air se gonflent. Ainsi, celui qui porte la veste a la sensation de recevoir une étreinte.



Figure 48. Hug over a distance : le koala en peluche et la veste gonflable (Mueller et al., 2005).

Takahashi et al. (2011) proposent aussi un dispositif, Sense-Roid, permettant de faire l'expérience d'une étreinte, à la différence qu'ils préconisent de l'utiliser pour faire l'expérience de sa propre étreinte (voir figures 49 et 50). Le système est composé d'un mannequin équipé de capteurs tactiles afin de pouvoir enregistrer son étreinte (caresses dans le dos incluses), et d'une veste équipée de vibrateurs et de muscles artificiels qui rejoue en temps réel l'étreinte enregistrée par le mannequin. Ce système est donc pensé pour être utilisé seul, mais l'exemple précédent nous montre qu'il pourrait être utilisé pour mettre en place une étreinte à distance.



Figure 50. Le système Sense-Roid : un mannequin équipé de capteurs, et une veste tactile (Takahashi et al., 2011).



Figure 49. Une utilisatrice faisant l'expérience d'une « étreinte de soi à soi » par le biais de Sense-Roid (Takahashi et al., 2011).

Kissenger (Kiss Messenger) est un dispositif développé par Samani et al. (2012) ; ce dispositif interactif fournit une interface physique pour la transmission d'un baiser entre deux personnes se trouvant à distance (voir figure 51). Deux dispositifs sont liés entre eux ; ils sont sensibles à la force des lèvres appliquée par l'autre utilisateur. Cette force est transmise par le biais de vibrateurs. Ce dispositif est conçu pour enrichir d'autres modes de communication à distance, notamment la vidéoconférence (voir figure 52).

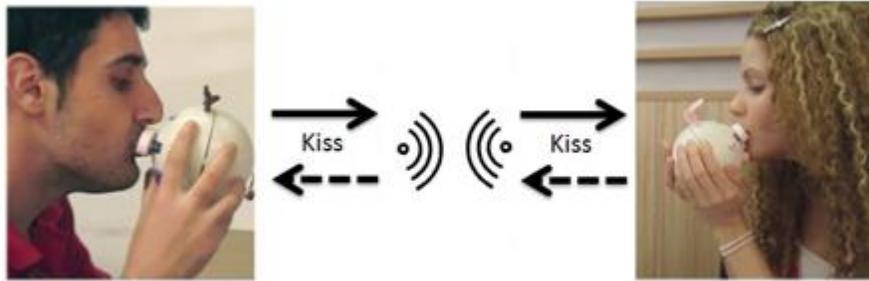


Figure 51. Le concept du système Kissenger (Samani et al., 2012).



Figure 52. Le système Kissenger en utilisation lors d'un vidéochat (Samani et al., 2012).

Afin de donner l'impression de se saisir la main à distance, Gooch et Watts (2012) ont développé plusieurs prototypes. YourGlove (TonGant) est un système composé d'une main de robot habillée d'un gant qui peut se contracter (voir figure 53).



Figure 53. Le prototype YourHand (Gooch et Watts, 2012).

Le prototype HotHands (MainsChaudes) se base non plus sur les mouvements mais sur la chaleur. Chaque utilisateur possède une main de mannequin équipée d'une pompe à chaleur et d'un capteur de chaleur (voir figure 54). Ainsi, lorsqu'un utilisateur place ses mains sur la main qu'il possède, la main possédée par l'autre interactant se réchauffe.



Figure 54. Le prototype HotHands (Gooch et Watts, 2012).

Enfin, le prototype HotMitts (MoufflesChaudes) est composé non pas de mains de mannequin, mais de l’empreinte moulée de la main de l’autre interactant (voir figure 55). Le système fonctionne par ailleurs de la même façon que HotHands.



Figure 55. Le prototype HotMitts (Gooch et Watts, 2012).

Prattichizzo et al. (2010) proposent RemoTouch, un système de toucher à distance composé d’une personne à distance équipée d’un gant spécial, d’une caméra et d’un micro, et d’un utilisateur équipé de stimulateurs tactiles permettant de sentir l’interaction tactile à distance, ainsi que d’enceintes et de lunettes affichant une image (voir figures 56 et 57). C’est donc la personne à distance qui va être en charge de collecter les informations tactiles à envoyer à l’utilisateur grâce aux capteurs de force dont elle est équipée. L’utilisateur est équipé d’un gant où se trouvent les stimulateurs tactiles, mais aussi des lunettes affichant ce qu’une caméra portée sur la tête de l’autre personne enregistre. Sont aussi retransmis les sons de la scène de la personne à distance. Ainsi, l’utilisateur voit, entend et sent ce que la personne à distance voit, entend et sent (voir figures 56 et 57).



Figure 56. Un homme portant RemoTouch fait l'expérience à distance de tenir son bébé par l'intermédiaire de sa femme. Il porte des lunettes affichant l'image captée par la caméra fixée sur la tête de sa femme, entend ce qui est enregistré dans son environnement et reçoit des stimulations tactiles correspondant à ce que les capteurs de force qu'elle porte enregistrent (Prattichizzo et al., 2010).



Figure 57. Un homme portant RemoTouch fait l'expérience de jouer du piano à distance par l'intermédiaire d'un joueur (Prattichizzo et al., 2010).

Le dispositif Keep in Touch, développé par (Motamedi, 2007), est un écran tactile en tissu mis en réseau conçu pour supporter et maintenir l'intimité entre des couples à distance (voir figure 58). Chacun des partenaires possède l'un des tissus. Sur chaque tissu est projetée l'image floue de l'endroit où se trouve l'autre tissu. Lorsque les interactants touchent le tissu, l'image devient alors nette ; si les deux partenaires interagissent en même temps, ils se voient alors presque comme s'ils étaient face à face.



Figure 58. Le système Keep in Touch (Motamedi, 2007).

Laschke, Hassenzahl et Mehnert (2010) se sont penchés sur la question de la communication entre garçons adolescents, et ont identifié comme moyens pour eux d'exprimer leur besoin de relatedness et d'échanges physiques le fait de se bagarrer, de se chamailler (to squabble). Afin de supporter ce type d'interaction à distance, les auteurs ont développé linked., un dispositif ressemblant à un oreiller permettant de se bagarrer à distance. Ils proposent deux versions de leur prototype (voir figures 59 et 60), toutes deux basées sur le même concept : chacun des interactants a en sa possession un oreiller, qui se gonfle et se dégonfle selon la façon dont l'autre utilisateur frappe son propre oreiller.



Figure 59. Version 1 de *linked*. (Laschke, Hassenzahl et Mehnert, 2010).



Figure 60. Version 2 de *linked*. (Laschke, Hassenzahl et Mehnert, 2010).

Emotional Touch est un dispositif développé par Hashimoto et Kajimoto (2008), qui consiste en deux enceintes entourées d'une bande de caoutchouc, qui permet d'envoyer un retour tactile à l'utilisateur lorsque ce dernier applique ses mains sur chaque haut-parleur (voir figure 61). Les vibrations reçues dépendent de la fréquence d'émission et peuvent donner un retour donnant différentes impression (tenir quelque chose de vivant, douceur,...).



Figure 61. Le dispositif vibrant de Hashimoto et Kajimoto (2008).

Dans la même lignée, les laboratoires Disney viennent de mettre au point un retour haptique destiné à l'enrichissement de l'expérience de jeu, Aereal, basé sur de l'air pulsé (voir figure 62). Aereal peut simuler la présence d'un objet, des textures et des mouvements d'air.

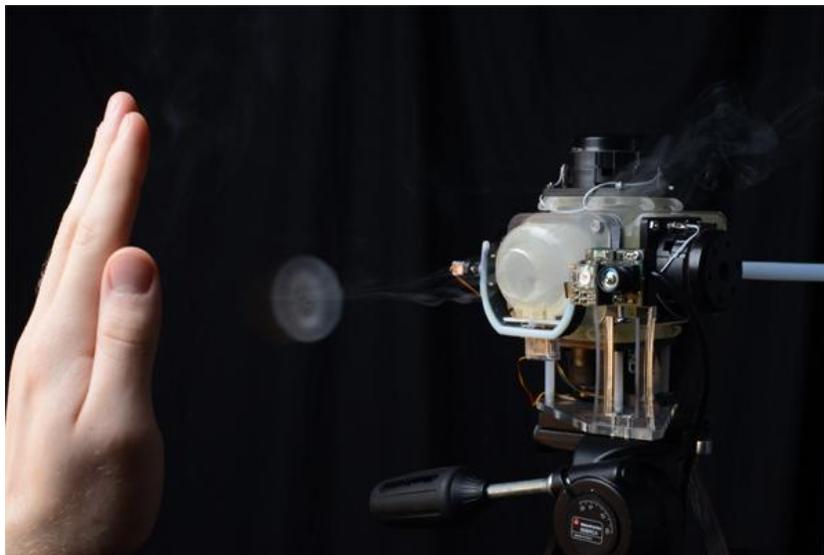


Figure 62. Aereal, développé par Disney Research (Sodhi et al. 2013).

On notera que ces deux dispositifs n'ont pas été conçus pour l'interaction, mais plus pour une expérience individuelle ; on pourrait ouvrir leur utilisation à une pratique interindividuelle.

Enfin, le dispositif Tactos, que nous avons présenté dans la partie 1.1.2.2, permet lui aussi une forme de caresse distale.

Après avoir pu échanger des signaux à distance, au travers par exemple des télégraphes optiques puis électriques, nous avons ensuite pu nous parler à distance, avec le développement du téléphone, et nous voir à distance, grâce aux progrès de l'informatique, des réseaux et de la visiophonie. Les interactions haptiques quant à elles nous permettent donc de nous « toucher » à distance. L'introduction de l'haptique dans la communication à distance semble permettre la transmission des émotions d'une manière plus efficace comparée à l'échange de messages écrits ou à l'interaction par vidéoconférence. En effet, le sentiment de présence est plus puissant avec le toucher.

2.2.1.3.1.2 Autres systèmes pour la communication émotionnelle

Les dispositifs que nous allons à présent mettre en avant ne font pas entrer en jeu le sens du toucher, mais ils supportent le sentiment de présence, en proposant des interfaces pour la communication émotionnelle à distance.

Le framework KAN-G (voir figure 63), développé par Liechti et Ichikawa (2000), vise un type d'awareness affective en permettant aux utilisateurs de s'échanger des photos numériques ainsi que leurs réactions à ces photos ; sont possibles la capture, la distribution, l'observation et l'annotation des photographies. Les images sont échangées, ainsi que les réactions aux images (une photographie du visage du receveur est prise quand il visionne la photo), ce qui permet aux utilisateurs de se sentir connectés. Les photographies sont en effet un excellent support pour la communication émotionnelle, souvent plus efficaces que les vidéos, dans la mesure où elles font plus appel à la mémoire et à l'imagination que ces dernières.

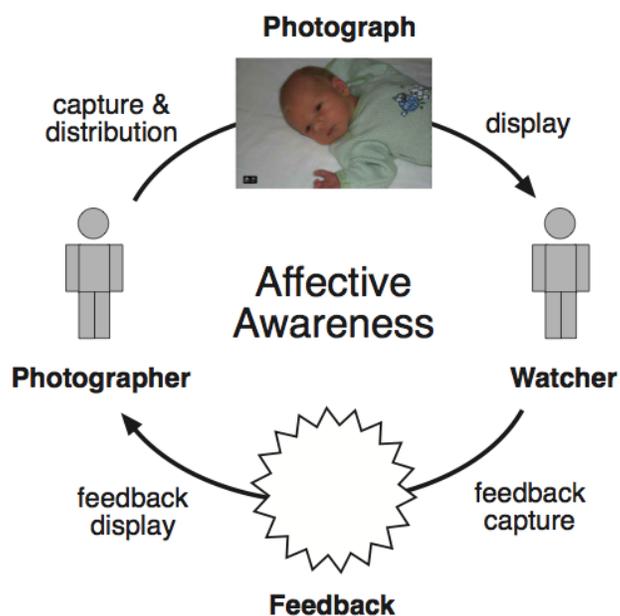


Figure 63. Le framework KAN-G de Liechti et Ichikawa (2000).

Le système ASTRA (voir figure 64) développé par Markopoulos et al. (2004) est basé sur un concept similaire. Ce système d'awareness est destiné aux individus d'une même famille ; certains sont en déplacement tandis que les autres restent à la maison. La communication passe par des photos, des messages courts et l'information de la disponibilité de chacun. À l'aide un dispositif mobile, la personne en déplacement peut prendre des photos, mais aussi dessiner ou écrire à la main. Ce qui est enregistré est partagé avec les personnes restées à la maison, par exemple par le biais d'une tablette tactile.



Figure 64. Le système ASTRA (Markopoulos et al., 2004).

LumiTouch (Chang et al., 2001) est un autre dispositif de communication à distance passant par les photographies. C'est un système composé de deux cadres photo interactifs, reliés par le biais d'une connexion internet. Quand l'un des utilisateurs saisit son cadre, l'autre cadre s'illumine (voir figure 65). Il existe une forme de communication passive par le biais du dispositif : lorsque l'un des utilisateurs se trouve à proximité de son cadre sans le saisir, le cadre de l'autre personne émet une faible lumière ambiante, indiquant la présence de la personne à distance. La communication active a lieu lorsque l'un des utilisateurs saisit son cadre ; l'autre cadre s'illumine alors, selon l'endroit où le cadre a été saisi, avec quelle force, selon quelle durée,...

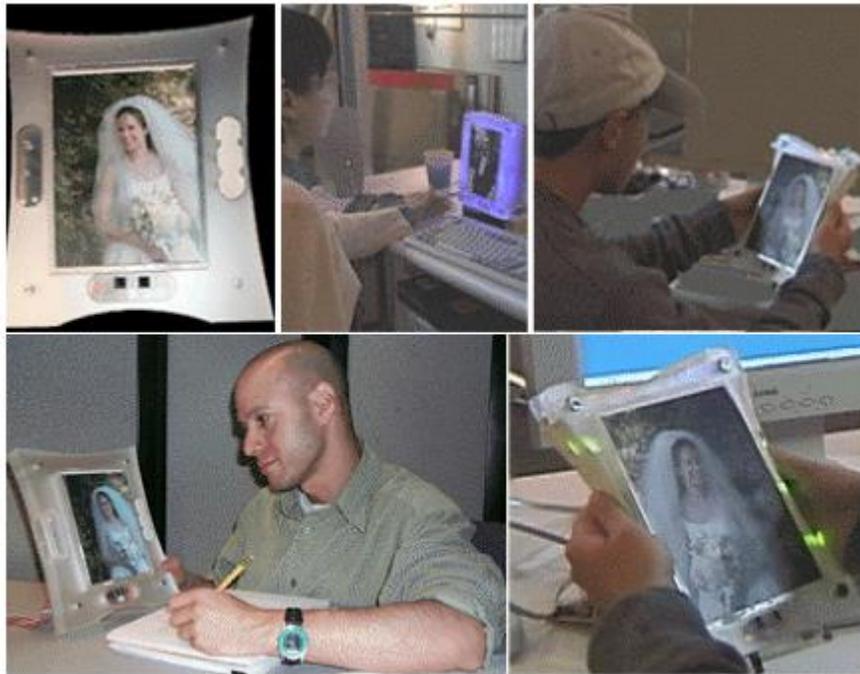


Figure 65. LumiTouch (Chang et al., 2001).

SnowGlobe est un dispositif proposé par Visser, Vastenburt et Keyson (2011, voir figure 66), visant à supporter l'awareness. SnowGlobe est une lampe qui traduit les mouvements d'un utilisateur à distance en se mettant à briller plus fort. Les utilisateurs peuvent se saisir de leur SnowGlobe et le secouer ; le SnowGlobe de l'autre interagissant se mettra alors à clignoter.



Figure 66. SnowGlobe ; détail et mise en situation (Visser, Vastenburg et Keyson, 2011).

Yarosh et al. (2009) proposent ShareTable (voir figure 67), une table qui offre un espace pour les interactions synchrones entre parents et enfants se trouvant à distance. Accompagnant une visioconférence, elle est équipée d'un système caméra-projecteur qui permet le partage visuel d'objets physiques.

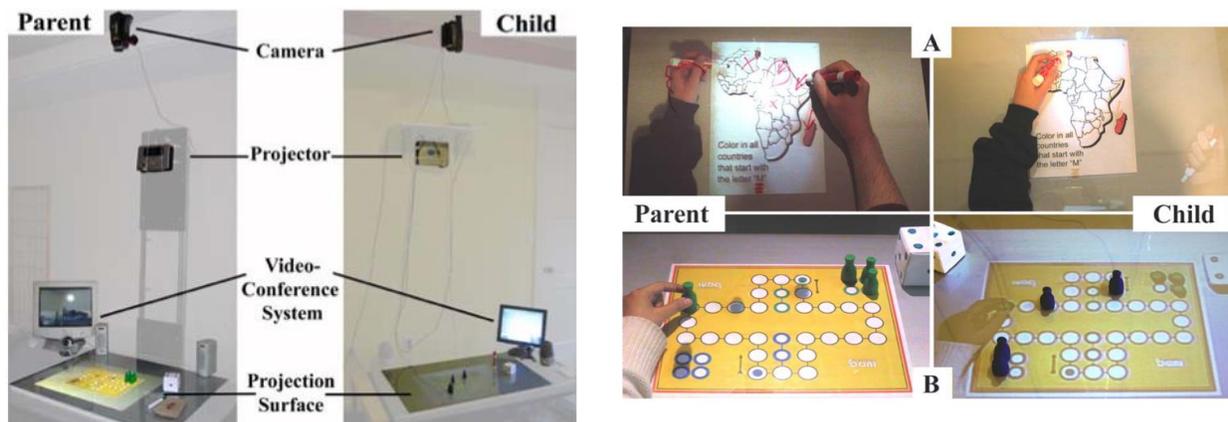


Figure 67. Le Dispositif ShareTable (Yarosh et al. 2009).

Strong et Gaver (1996) ont proposé trois concepts de dispositifs minimalistes visant le support d'une communication implicite, personnelle et expressive. Le premier système, appelé Feather (Plume), est conçu pour les situations où l'un des interactants est en voyage pendant que l'autre reste à la maison. Le voyageur emporte un petit objet lui permettant de signaler qu'il pense à l'autre utilisateur, ce petit objet pouvant par exemple être un cadre photo. Lorsque le cadre est posé, la photo est affichée à l'envers ; elle se retourne lorsque le voyageur se saisit du cadre. Le deuxième objet est de taille plus importante, comme un meuble, et reste à la maison. Il contient un petit ventilateur électrique silencieux dans sa base, et une simple plume repose sur

une grille placée au-dessus du ventilateur. Lorsque le partenaire en déplacement saisit son cadre photo pour bien en voir l'image, le ventilateur se met en marche, faisant flotter la plume, laquelle est prise dans un cône en plastique transparent partant de la base du meuble (voir figure 68). Ainsi, lorsque le voyageur se saisit du cadre, il peut correctement voir la photo puisque celle-ci se retourne, et son activité est signalée par le biais de la plume à la maison.

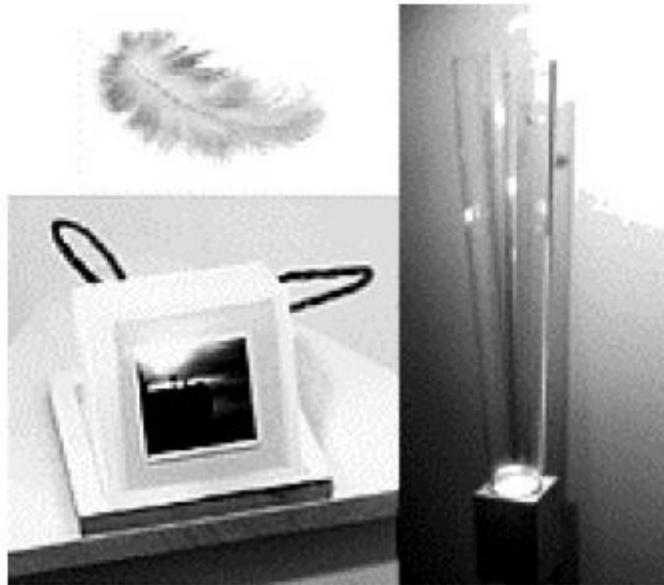


Figure 68. Feather, par Strong et Gaver (1996).

Le deuxième concept, Scent (Senteur), est un dérivé de Feather. Cette fois-ci, lorsque le voyageur se saisit du cadre photo, cela déclenche dans la maison un système chauffant se trouvant au fond d'un bol en métal qui contient des huiles essentielles (voir figure 69). Ainsi, la maison se remplit peu à peu de parfum, ce qui symbolise les pensées du voyageur.

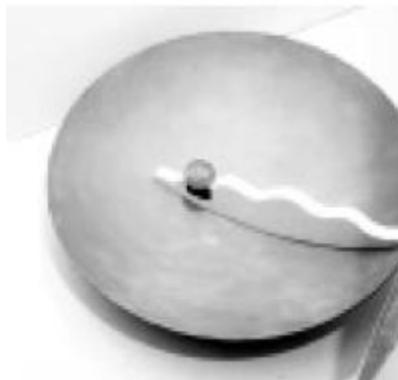


Figure 69. Scent, par Strong et Gaver (1996).

Le troisième concept, Shake (Secouer), est destiné à des relations moins intimes et à une communication plus symétrique. Le système est composé de deux paires de dispositifs, chacune des paires étant portée par l'un des interactants (voir figure 70). Chaque paire d'objets contient une bobine magnétique, composée d'une tige métallique entourée d'un fil de fer. C'est grâce à cette bobine que l'on envoie un message ; lorsqu'on la secoue, le mouvement de la tige induit un courant qui traverse le fil de fer. Ce courant est généré dans l'objet de réception du partenaire, le faisant s'agiter de manière proportionnelle à la façon dont la bobine est secouée.



Figure 70. Shake, par Strong et Gaver (1996).

Chung, Lee et Selker (2006) proposent Lover's Cups (voir figure 71), une interface permettant de partager l'instant de boire une boisson avec quelqu'un se trouvant à distance. L'interface se présente sous la forme de deux verres (un pour chacun des interactants) reliés par une connexion sans fil, qui communiquent donc entre eux et amplifient l'aspect social du fait de boire. Lorsqu'un verre s'illumine, cela signifie que l'autre interactant est en train de boire.

Schmeer et Baffi (2011) proposent un système de communication asynchrone appelé Touch trace mirror. Le système de communication est basé sur des messages asynchrones, volatiles et collaboratifs, et est prévu pour le contexte d'usage d'interaction pour un couple à distance. Il est composé de deux miroirs (un à chaque domicile) qui permettent de laisser un message comme on écrit sur la buée d'un miroir de salle de bain. En écrivant sur l'un des miroirs, on envoie le message écrit à l'autre miroir, qui s'éclaire alors. Pour pouvoir lire le message, le partenaire doit placer sa main sur la lumière, qui va alors se déplacer ; en suivant la lumière du doigt, le partenaire va alors retracer le message initialement envoyé (voir figure 72).



Figure 71. Le verre de l'homme s'illumine du fait que la femme boive (Chung, Lee et Selker, 2006).



Figure 72. Une utilisatrice devant son Touch trace mirror suit du doigt la lumière émise, retraçant ainsi le message de l'autre interactant (Schmeer et Baffi, 2011).

Tous ces systèmes permettent d'enrichir la communication à distance, en se concentrant sur l'aspect émotionnel de l'échange.

2.2.1.3.2 Le toucher dans les environnements virtuels partagés

L'article [An Experimental Study on the Role of Touch in Shared Virtual Environments](#) présente une expérience concernant le rôle du toucher dans les environnements virtuels partagés (Basdogan et al., 2000). Les auteurs se demandent dans quelle mesure l'ajout de communication haptique entre les individus dans un environnement virtuel contribuerait à l'expérience partagée. Leur but est d'étudier le rôle du feedback haptique dans les tâches collaboratives, et de voir si la communication haptique par le biais d'un retour de force peut accentuer le sentiment d'être et de collaborer avec un partenaire distant.

L'étude cherche à évaluer l'impact d'un retour de force sur les performances de tâche, à mieux comprendre le rôle de la communication haptique dans les interactions humain-humain, à étudier l'impact du toucher sur le sentiment subjectif de collaboration avec un humain rapporté par les sujets sur la base de ce qu'ils pouvaient voir et sentir, et à examiner si le genre, la personnalité ou les expériences émotionnelles des utilisateurs peuvent affecter la communication haptique dans les environnements virtuels partagés (ibid).

L'expérience vise à faire collaborer deux individus à distance dans un environnement virtuel partagé. La tâche consiste à déplacer un anneau le long d'un fil (voir figure 73), sans toucher ce dernier ; l'action des deux personnes est nécessaire pour faire avancer l'anneau et accomplir la tâche. Dans l'une des conditions, les individus n'ont qu'un retour visuel puis un retour visuel et tactile, dans l'autre ils ont d'abord un retour visuel et un retour de force, puis seulement un retour visuel.

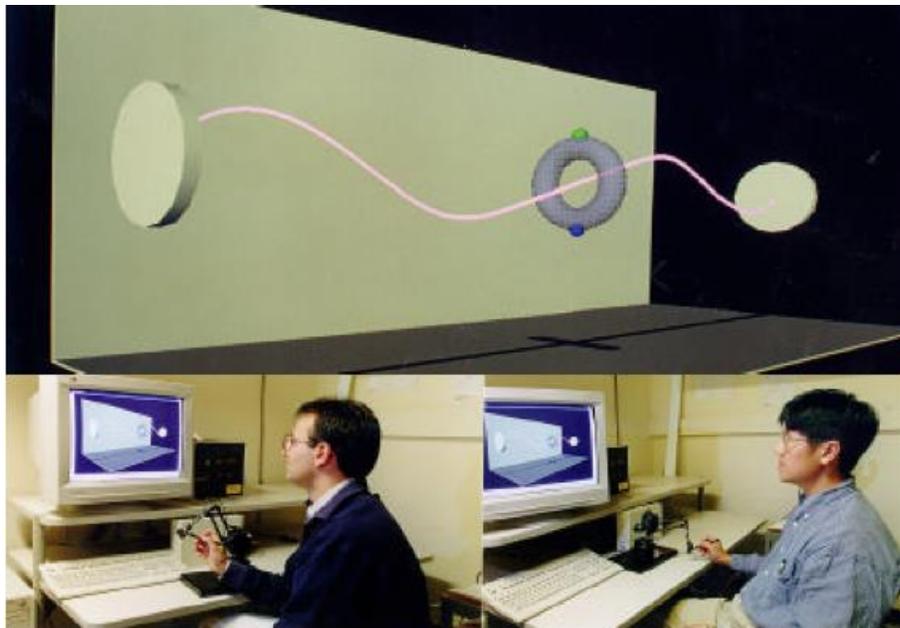


Figure 73. La tâche proposée dans l'étude de Basdogan et al (2000).

Les résultats suggèrent que le feedback haptique améliore significativement les performances et contribue à un sentiment d'« être ensemble » dans l'environnement virtuel partagé (ibid). De plus, les résultats montrent que l'expérience du feedback visuel seul au départ, auquel on ajoute un retour haptique par la suite permet de meilleures performances que de donner d'abord un retour à la fois visuel et haptique suivi par un retour visuel seul (ibid).

La réalité virtuelle induit la création d'expériences multi-sensorielles par le biais de moyens artificiels. Être capable d'explorer et de manipuler des objets par le biais d'interfaces haptiques dans des environnements virtuels se décline en plusieurs applications dans de nombreux champs, parmi lesquels la médecine, l'éducation, le loisir,... Dans le domaine des téléopérateurs, les performances des opérateurs augmentent significativement dans la manipulation d'objets lointains quand il y a un feedback haptique. D'autres expériences sur l'automatisation et la planification de tâches complexes par le biais d'un contrôle et d'une supervision humaine montrent que les temps de réalisation des tâches peuvent être réduits quand il y a un retour haptique. De plus il apparaît que le toucher et la manipulation d'objets améliorent l'expérience subjective (ibid). Le toucher, en comparaison des autres sens, est plus localisé et est bidirectionnel, ce qui est lié à l'intimité.

L'article supporte ces théories et pose l'impact positif de l'haptique sur la performance de tâche et le sentiment subjectif d'être ensemble dans un environnement virtuel partagé.

2.2.1.3.3 Synthèse

Les dispositifs de communication à distance présentés visent une communication à distance relationnelle, riche sur le plan émotionnel, soit en intégrant le sens du toucher (ou une forme de sens du toucher), soit en passant par d'autres biais (en vue de communiquer la présence, l'awareness, la relatedness, l'intimité,...).

On peut distinguer plusieurs façons de classer ces dispositifs. En ce qui concerne ceux qui proposent un toucher à distance, deux catégories principales se distinguent ; il y a d'une part les systèmes tactiles, qui stimulent les récepteurs cutanés de la peau, et d'autre part les systèmes kinesthésiques (ou systèmes à retour d'effort), basés sur des retours de force. Certains systèmes combinent les deux, en vue de stimuler à la fois le sens cutané et le sens kinesthésique. L'utilisation de vibreurs est la méthode la plus courante pour transmettre des sensations tactiles, et on trouve des systèmes à retour de force pour la stimulation du sens kinesthésique, mais on notera que d'autres possibilités existent et sont adoptées : stimulations thermiques, systèmes gonflables, ou encore utilisant l'humidité.

On peut aussi identifier le classement en dispositifs basés sur l'échange de signaux codés, comme par exemple des émoticônes tactiles (qui peuvent être proposées par les concepteurs ou définis par les interactants), et dispositifs permettant des interactions directes. Les icônes haptiques sont donc un système de représentation symbolique, toujours plus ou moins conventionnel (de même que les smileys), tandis que les systèmes d'interaction haptique directe délivrent des stimuli liés au sens du toucher en fonction des actions du partenaire. On remarquera que les interfaces d'interaction tactile directe sont développées dans le souci non pas de transmettre une information, mais de permettre une expérience émotionnelle ; c'est la démarche pour laquelle nous allons opter.

On notera que l'une des méthodes les plus adoptées pour proposer des interfaces haptiques consiste à enrichir les conversations déjà possibles, que ce soit par le téléphone ou la visiophonie. De nombreux dispositifs présentés ci-dessus consistent ainsi à ajouter un retour de nature tactile afin d'enrichir l'audio ou la vidéo. Nous choisissons de ne pas adopter cette voie, afin de pouvoir mettre le tactile au centre de l'interaction proposée, et non comme un enrichissement d'une interaction déjà mise en place et centrée sur la communication informative.

On notera enfin que les dispositifs proposés sont conçus pour une situation d'usage particulière ; il s'agit, sauf exception, d'interactants à distance, avec un niveau d'intimité supposé particulier (en général, des couples, membres d'une même famille ou amis proches). Il est important pour le concepteur d'envisager la situation d'usage dans laquelle son dispositif s'inscrit, afin de répondre au mieux aux besoins spécifiques du contexte ayant été identifié. La situation d'usage que nous envisageons est délibérément large : nous souhaitons en effet pouvoir proposer une expérience accessible au plus grand nombre, aussi nous délimitons notre contexte à celui « d'interactants se trouvant à distance, se connaissant ou non ». Une grande variété d'interactions et de paires d'interactants est ainsi possible.

Suite à cet état de l'art, nous souhaitons plus particulièrement souligner la notion d'interfaces tangibles, liée au dispositif inTouch, et mise en jeu dans d'autres des dispositifs présentés. Ishii (2008) décrit de façon générale les interfaces tangibles comme proposant « une correspondance entre une entrée haptique et des représentations haptiques à distance. Aussi appelé téléprésence tangible, le mécanisme sous-jacent consiste en la synchronisation d'objets distribués et la simulation dans les gestes de dispositifs présents, par le biais par exemple du mouvement ou de la vibration, ce qui permet aux interactants à distance de transmettre les manipulations haptiques qu'ils effectuent sur les objets physiques distribués. Cela a pour

conséquence de donner aux interactants distants un sentiment de présence fantomatique, comme si une personne invisible manipulait un objet partagé. » (ibid, p. xix, traduction personnelle)²⁵.

La notion de la tangibilité est clé dans notre rapport au monde et aux autres (Declerck, 2010), et chercher à l'intégrer dans des dispositifs de communication permet l'expérience d'interactions riches, ainsi qu'un sentiment de présence accru. C'est un point-clé à considérer dans la conception de dispositifs traitant avec des données numériques, immatérielles, mais aussi dans la conception de dispositifs chargés de permettre l'interaction à distance.

La communication à distance émotionnelle est un enjeu important, notamment du fait des déplacements que les individus sont amenés à faire de plus en plus fréquemment, se retrouvant isolés de leurs famille et amis. On constate que c'est un enjeu de recherche et de conception qui concerne de plus en plus de chercheurs et de designers, pour des raisons diverses, mais toutes motivées par l'idée de palier à l'absence des proches. On constate que le tangible et le toucher constituent des pistes de choix pour répondre à cette problématique.

2.2.2 Suppléance perceptive du toucher

Nous avons traité de la question de la suppléance perceptive dans la partie 1.1.2. La plupart des dispositifs de suppléance perceptive consistent à transmettre des informations liées au sens de la vue (en passant soit par du tactile, soit par de l'audio). Nous allons prendre le parti, pour notre interface de communication à distance riche sur le plan émotionnel, de proposer aux utilisateurs une suppléance perceptive du toucher. Étant donné que les utilisateurs à distance ne peuvent par définition pas se toucher, nous souhaitons leur proposer une alternative fonctionnant sur le principe de la suppléance perceptive.

Plusieurs propriétés fonctionnelles du sens du toucher peuvent être prises en considération pour en proposer une suppléance perceptive. Comprendre la suppléance perceptive nécessite de faire la différence entre modalité sensorielle et modalité perceptive (Pacherie, 1997). La modalité sensorielle concerne le type de sens qui se retrouve stimulé, tandis que la modalité perceptive concerne le type de couplage avec le monde. Ainsi, pour faire la suppléance perceptive du toucher, on pourra réaliser une modalité perceptive tactile à l'aide d'autres types de modalité sensorielle.

²⁵ « [...] mapping haptic input to haptic representations over a distance. Also called tangible telepresence, the underlying mechanism is the synchronization of distributed objects and the gestural simulation of presence artifacts, such as movement or vibration, allowing remote participants to convey their haptic manipulations of distributed physical objects. The effect is to give a remote user the sense of ghostly presence, as if an invisible person was manipulating a shared object. » (Ishii, 2008, p. xix).

Sur le plan de l'objectivité, le toucher correspond au contact physique d'une chose avec une autre. Sur le plan psycho-physiologique, sa caractéristique générale est que le champ récepteur (que ce soit un point ou une surface) qui permet la perception de l'objet coïncide spatialement avec cet objet perçu. Ceci est une caractéristique de la modalité perceptive tactile : il y a coïncidence spatiale entre capteur et objet perçu. Pour un toucher interindividuel, cela signifie que chacun des interactants ne reçoit un retour sensoriel que lorsque les interactants sont au même endroit en même temps. Nous appelons cette caractéristique du sens du toucher le fait qu'il soit proximal, coïncidental²⁶.

En plus de la coïncidence, on retrouve bien sûr la propriété liée au fait que la perception tactile peut être une perception active. Elle implique alors le déploiement spatio-temporel d'un geste d'exploration nécessaire à la perception. On parle dans ce cas de perception haptique, ou tactilo-kinesthésique (Hatwell, 2003). L'utilisateur devra par conséquent être capable de faire varier les retours sensoriels qu'il reçoit en fonction de ses actions au travers du dispositif proposé.

Il découle de la recherche en suppléance perceptive que l'on peut produire une forme de perception tactile sans avoir à mobiliser d'input sensoriel tactile. Ainsi, notre proposition va consister à fournir des informations de nature tactile, sans nécessairement passer par un retour de nature tactile, mais par exemple par un retour visuel, ou encore auditif.

2.3 Design d'interface : Touch Through

2.3.1 L'application Touch Through : introduction

Notre première proposition d'interface pour la communication émotionnelle à distance est une application smartphone (ou tablette), appelée Touch Through. Elle a été développée sur la base des propriétés du croisement perceptif (voir partie 1.1.1.1.3) et du sens du toucher (plus particulièrement son importance pour la communication des émotions), ainsi que sur la recherche en suppléance perceptive. En effet, comme nous l'avons vu dans la partie 2.2.1.2, le toucher interpersonnel influe nos interactions sociales de tous les jours et nous donne accès aux expériences tactiles les plus émotionnelles (Gallace et Spence, 2010) ; ce fait nous pousse fortement dans la direction d'un dispositif incluant du tactile pour la transmission d'émotions à distance. Étant donné qu'un « toucher à distance » est par définition impossible, nous optons pour la focalisation sur certaines caractéristiques fonctionnelles du toucher afin d'en proposer

²⁶ Coïncidental est ici à prendre dans le sens de « qui se produit en même temps », et non pas dans celui de « qui se produit par hasard ».

une suppléance perceptive. Ainsi, notre application smartphone donne la possibilité de transmettre un « toucher » d'un point de vue fonctionnel.

Cette application permet la mise en relation haptique directe entre individus, ce qui bien sûr ne les empêche pas de mettre au point un système porteur de significations déterminées en utilisant le dispositif (nous esquisserons ce développement dans la partie 4). Avec un mode d'interaction direct, le sens que les utilisateurs associent au dispositif résulte d'un construit social : le sens émerge en effet alors dans l'interaction (et n'est donc pas défini au préalable par des concepteurs). La perception d'images d'autrui ou d'images de l'état émotionnel d'autrui (par le biais de smiley ou d'émoticônes haptiques) n'est pas toujours suffisante pour constituer autrui comme un être intentionnel ni pour associer un contenu de nature émotionnel à cette constitution. C'est surtout par la perception de l'activité perceptive d'autrui dirigée vers soi que nous pensons que ces émotions peuvent émerger.

Notons que l'application fonctionne seule : elle ne vise pas à enrichir une communication téléphonique ou visiophonique pré-existante. Elle est destinée à l'interaction entre deux personnes distantes, et offre la possibilité d'interagir avec des inconnus. Le contexte d'usage est donc très large et souple ; nous souhaitons en effet que le dispositif soit utilisable par le plus grand nombre.

Touch Through a été codée en java, et fonctionne sous un système d'exploitation Android.

2.3.2 L'application Touch Through : description

Pour percevoir de manière tactile, le point « de vue » doit être à la même position que l'objet à percevoir. Dans l'espace, je suis à l'endroit où je touche. Pour permettre l'impression d'un « toucher distal », cette structure spécifique du tactile doit être respectée : l'utilisateur n'aura un retour que s'il est là où il y a quelque chose ou quelqu'un à percevoir tactilement. Pour pouvoir retrouver cette propriété de la coïncidence alors même que l'on se trouve à distance, il faut donc avant tout que les interactants puissent partager un même espace. Le point à partir duquel on perçoit tactilement (le bout du doigt par exemple) doit être au même endroit que la chose qui est perçue. Pour que deux personnes interagissent de façon 'tactile', elles doivent donc partager un même espace virtuel, et elles ne se toucheront que quand elles seront exactement à la même position dans cet espace. Il va donc s'agir non pas que les personnes soient à la même place (ce qui est impossible si elles se trouvent à distance), mais que leurs activités perceptives le soient.

Touch Through étant une application smartphone, l'espace d'interaction va être projeté sur l'écran. Les écrans des dispositifs (smartphone ou tablette) des interactants vont donc définir un même espace d'interaction partagé.

Touch Through permet une forme de toucher à distance respectant les propriétés fonctionnelles du toucher de la coïncidence et de la perception active. L'application propose un espace d'interaction où la stimulation survient lorsque les foci perceptuels des interactants se trouvent au même endroit. L'interaction tactile considérée est donc minimale. Ce n'est bien sûr pas l'idée de simuler le toucher de la peau qui est en jeu, mais la propriété fonctionnelle de la coïncidence du point de perception (là où je touche) et de ce qui est perçu (le partenaire).

Nous allons à présent exposer les différentes modalités selon lesquelles il est possible d'effectuer la suppléance perceptive du toucher par le biais de l'application. Les deux interactants doivent tout d'abord lancer l'application et se connecter l'un avec l'autre, afin de pouvoir démarrer une interaction. Une fois cette dernière lancée, les écrans des dispositifs deviennent noirs jusqu'à ce que l'un des interactants touche son écran.

Modalité visuelle

Cette modalité est la seule à être déclinable en deux modes : distal et proximal (ou coïncidental).

** Modalité visuelle en mode distal*

Le mode distal permet de fournir à tout moment des informations sur la localisation du partenaire dans l'espace partagé sans qu'il y ait contact, ce qui ne respecte donc pas la propriété de la coïncidence. Le fait que la coïncidence ne soit pas obligatoire dans ce mode permet de faciliter la compréhension du fonctionnement de l'application, lors de sa découverte. Quand l'un des utilisateurs touche la surface de l'écran de son dispositif, un disque de couleur apparaît sous son doigt (voir figure 74) et accompagne le doigt dans ses déplacements, pour autant qu'il reste en contact avec l'écran (le disque disparaît dès que le doigt quitte la surface tactile de l'écran). La taille du disque est par défaut légèrement plus grande que la surface de contact entre le doigt et l'écran, de façon à ce que l'utilisateur puisse le voir sous son doigt ; cependant on peut régler la taille pour qu'elle soit plus petite. Le disque de couleur apparaît sur les écrans des deux interactants, et se déplace de la même façon sur les deux.



Figure 74. Lorsque l'un des interactants touche son écran en modalité visuelle distale, un disque de couleur apparaît sur les deux écrans.

Quand l'autre interactant pose à son tour son doigt sur son écran, un second disque qui le représente apparaît sur les deux écrans (voir figure 75). Les disques sont de deux couleurs différentes, et ils accompagnent le doigt de l'utilisateur auquel ils sont respectivement associés.

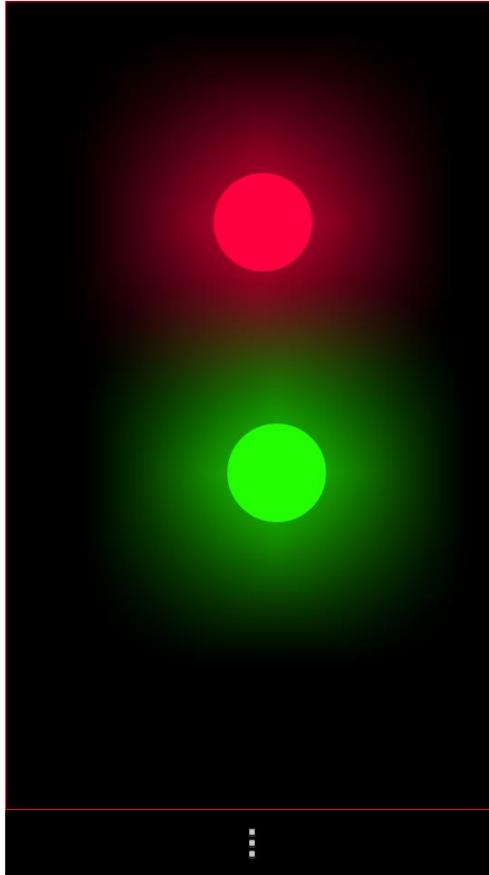


Figure 75. Lorsque le deuxième interactant touche à son tour son écran, un deuxième disque de couleur apparaît sur les deux écrans.

Quand les doigts des deux utilisateurs atteignent le même endroit dans l'espace d'interaction, les deux disques fusionnent en un disque d'une troisième couleur (voir figure 76). Les doigts des utilisateurs touchent alors « le même endroit » (de l'espace d'interaction). Les partenaires de l'interaction font alors l'expérience d'une forme de toucher à distance.

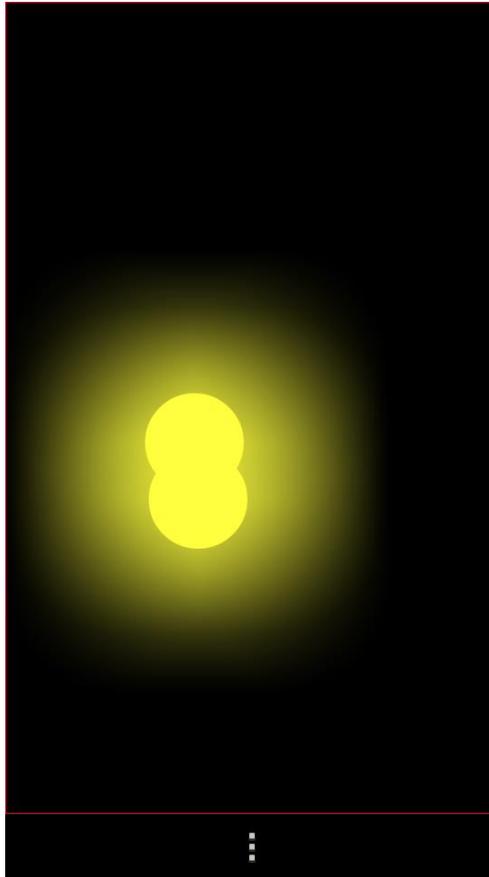


Figure 76. Lorsque les deux interactants touchent le même endroit de l'espace partagé, leurs disques fusionnent en adoptant une troisième couleur.

** Modalité visuelle en mode proximal*

En mode proximal, la modalité visuelle ne permet de voir apparaître un disque de couleur que lorsqu'il y a coïncidence, c'est-à-dire lorsque les doigts des partenaires sont au même endroit. Les utilisateurs ne peuvent alors avoir d'information sur la position de leur partenaire avant de l'avoir rencontré. Ils ne connaissent la position de leur partenaire que quand ils partagent cette position avec lui, ce qui est conforme à la façon dont fonctionne le sens du toucher.

Le mode proximal est disponible dans deux autres modalités : audio, et vibratoire.

Modalité audio

Dans cette modalité, la rencontre des doigts des partenaires au même endroit de l'espace d'interaction virtuel représenté sur l'écran des téléphones engendre un retour audio. Un son plus

faible peut aussi être émis lorsque les partenaires sont proches, afin de les aider à se trouver (ce qui se rapproche d'une modalité distale).

Modalité vibratoire

Dans cette modalité, c'est un retour vibratoire que les sujets reçoivent lorsqu'ils se « touchent ». Une vibration plus faible peut être émise lorsque les partenaires sont proches, afin de les aider à se trouver (ce qui là aussi se rapproche d'une modalité distale). Notons que la modalité vibratoire propose une suppléance perceptive du toucher en donnant un retour qui est de nature tactile ; étant donné que la suppléance est proposée pour compenser l'impossibilité d'un toucher à distance, ce n'est donc pas incompatible de faire passer les propriétés fonctionnelles du toucher que nous avons mises en avant par le biais d'un retour sensoriel tactile.

Réglages

Une fois l'application lancée et l'utilisateur connecté, ce dernier se retrouve dans la salle d'attente de l'application, où il peut configurer les réglages comme il le souhaite (partenaire pour la prochaine interaction (dont possibilité d'interagir au hasard et d'interagir avec quelqu'un ne faisant pas partie de la liste de contacts), couleur et taille du disque,...) ainsi que la ou les modalité(s) qu'il souhaite utiliser pour l'interaction à venir (voir figure 77). Toutes les modalités sont en effet combinables.

Notons que les disques de couleur propres à chaque utilisateur constituent en fait leur avatar dans l'espace d'interaction proposé. Ils sont composés d'un corps propre (ce champ récepteur qui est sensible à ce qui se trouve dans l'espace virtuel partagé) et d'un corps-image (que le champ récepteur de l'autre utilisateur peut percevoir). En modalité visuelle distale, les avatars sont visibles dès que les utilisateurs touchent leur écran. En modalité visuelle proximale, les avatars ne sont visibles que lorsqu'il y a « toucher », et ils ne sont jamais visibles en modalité audio ou vibratoire seules.

Touch Through propose au total deux modes d'interaction ; distal (disponible uniquement pour la modalité visuelle) et proximal. Dans le mode proximal, il n'est pas aisé pour les partenaires de se trouver, étant donné qu'ils n'ont au début de l'interaction aucune indication sur la localisation de leur partenaire. Le suivi nécessite de l'attention ; il est délicat, incertain, hésitant. La réussite du suivi est en fait le signe le plus certain de l'attention perceptive du partenaire. Le sentiment de présence est alors porté par la rencontre des deux activités perceptives.



Figure 77. La salle d'attente de l'application, où l'utilisateur peut voir les autres personnes connectées et choisir parmi elles son partenaire pour la prochaine interaction (ici, aléatoire), lancer la prochaine interaction (démarrer), configurer les modalités (vibratoire, audio et visuelle), choisir la taille de son avatar ainsi que sa couleur, et enfin choisir entre mode distal (voire l'autre activé) et le mode proximal (voir l'autre désactivé).

2.3.3 Premiers retours utilisateurs

Notre application a tout d'abord été proposée à six paires d'utilisateurs en situation d'usage réelle, afin d'évaluer leur réaction ainsi que leur plaisir à utiliser l'application, pour améliorer cette dernière.

Les utilisateurs, placés dans des pièces différentes, se voient présenter chacun un téléphone portable sur lequel l'application a été réglée en mode distal (modalité visuelle) et où

une interaction a été lancée. L'écran des téléphones est donc noir quand ils se saisissent du téléphone.

La moitié des participants touchent d'eux-mêmes la surface du téléphone, tandis que les autres se montrent perplexes et demandent ce qu'il convient de faire, étant donné que rien ne se passe à l'écran (l'expérimentateur les enjoint alors à toucher l'écran). Les utilisateurs se montrent curieux vis-à-vis des disques de couleur (que ce soit celui qui suit leur doigt ou celui qui se déplace « tout seul »). Dans la majorité des cas, les utilisateurs sont surpris lorsque le premier « toucher » avec le partenaire se produit et que les disques changent de couleur ; ils cherchent alors tout de suite à reproduire l'expérience et se mettent à se suivre l'un l'autre en une danse du croisement perceptif.

Après cette phase d'introduction, les utilisateurs découvrent les modalités audio et vibratoires en les combinant au mode distal visuel. Ils essayent ensuite le mode proximal seul, sans le retour distal visuel ; ils n'ont alors de retour que lorsque la coïncidence de leur disque se produit, exactement comme c'est le cas pour le sens du toucher. Ils estiment alors qu'il est plus difficile de se trouver l'un l'autre, mais ils persistent à se chercher et tentent de maintenir leur croisement perceptif le plus longtemps possible. Après un moment, les utilisateurs les plus joueurs s'échappent du contact de leur partenaire, afin que ce dernier les poursuive.

En laissant la main aux utilisateurs sur les réglages, les utilisateurs commencent tout d'abord par activer toutes les modalités, en mettant le visuel en distal. Ainsi, ils voient leur partenaire à distance, et lors du contact, ils ont un retour tactile, un retour audio et un retour visuel (les disques fusionnent et changent de couleur). Mais rapidement, les utilisateurs changent les réglages afin de retirer le mode distal du visuel ; ils trouvent alors l'expérience plus excitante, plus « touchante », plus intéressante car elle présente un défi. Le mode proximal, qui est moins riche du point de vue informationnel, permet donc de stimuler l'intérêt des utilisateurs vis-à-vis de la situation d'interaction proposée, et nos utilisateurs rapportent qu'il permet un échange plus riche sur le plan émotionnel. Ce retour des utilisateurs en situation d'usage non contrôlée a été très enrichissant et a permis de justifier certains partis pris de notre conception.

Certains utilisateurs nous ont signalé que l'application leur donnait plus le sentiment d'être touché que de toucher, notamment en mode proximal vibratoire. Gibson (1962) souligne cette distinction entre le toucher actif, qui désigne ordinairement le fait de toucher, et le toucher passif ; dans ce dernier cas, les stimulations sur la peau ne sont pas du fait de celui qui perçoit, mais d'un agent extérieur.

D'autres utilisateurs remarquent que par le biais de l'application, il n'est pas possible de toucher si l'autre interagissant retire son doigt de la surface et n'est donc pas engagé dans l'interaction ; il n'est donc pas possible d'attirer l'attention par cette suppléance perceptive du

toucher, comme on pourrait par contre le faire en présence (par exemple en tapant sur l'épaule de l'autre interactant).

Globalement, les utilisateurs se sont montrés enthousiastes par l'application, mais ont questionné son attrait sur le long terme.

2.3.4 L'application Touch Through : enjeux

Touch Through permet donc une perception réciproque de l'activité perceptive d'autrui en réalisant une sorte de « caresse distale ». Elle permet une forme de toucher à distance qui respecte les propriétés de coïncidence et de perception active du toucher. Cette forme de toucher a pour but de réaliser un échange à distance émotionnel, en donnant à ressentir la présence d'autrui à distance.

Rappelons que lorsque deux individus se trouvent à distance, ils sont à la fois intouchables et « intouchants » l'un pour l'autre. Lorsque deux individus sont en interaction dans l'espace d'interaction proposé par l'application, ils deviennent à la fois touchables et « touchants » l'un pour l'autre. Touch Through permet en effet une forme de croisement perceptif. Rappelons que Lenay (2010) pose que la première valeur du contact est la reconnaissance de la présence d'autrui : je perçois autrui comme activité percevante. La médiation technique qui met les individus en contact ne peut être porteuse d'émotion, de signification, que si la face visible d'autrui qu'elle nous donne est directement liée à son activité perceptive ; c'est bien le cas avec Touch Through. Elle permet en effet un croisement perceptif : je te touche toi qui me touches, et tu me touches moi qui te touche.

Lors de l'utilisation du mode proximal de l'application (et ce quelle que soit la modalité), il est difficile pour les partenaires de se trouver, car ils n'ont au début aucune indication quant à l'endroit où se trouve leur partenaire. Mais dès qu'ils réussissent, ils peuvent se suivre bien que de façon hésitante ; ils doivent faire preuve de délicatesse et d'attention pour réussir. L'interaction permise par l'application est fragile et peut être interrompue très facilement, par exemple si l'un des participants se déplace trop vite pour que l'autre arrive à le suivre. La réussite à suivre le partenaire est la preuve de son attention. Suivre l'autre interactant est difficile, mais c'est ce qui participe d'un échange émotionnel. En effet, la réussite, toujours partielle et incertaine, de ce suivi permet à chacun de sentir l'attention perceptive de son partenaire et c'est cette rencontre entre deux activités perceptives attentives qui est porteuse du sentiment de présence, et qui est typiquement tactile.

Les difficultés et les limitations de l'application sont pensées dans la perspective de proposer un échange émotionnel. L'intérêt de Touch Through est donc, d'après nous,

paradoxalement dans sa simplicité et dans les limites qu'elle impose à l'interaction. Grâce à la situation minimale proposée, la situation de communication se trouve enrichie. Les utilisateurs ne peuvent percevoir qu'ils sont avec quelqu'un d'autre dans l'espace d'interaction que lorsqu'ils sont avec eux : lorsqu'ils se trouvent au même endroit de l'espace. Dans la mesure où je ne peux savoir que je suis avec autrui que lorsqu'il est au même endroit que moi, cette connaissance est toujours incertaine, jamais complètement dépendante de moi, et c'est sans doute pourquoi l'interaction permise par Touch Through est aussi intéressante. Nous posons l'hypothèse que c'est cet amoindrissement qui est porteur d'une émotion plus grande.

La direction adoptée au cours de l'interaction entre les individus émerge de la dynamique d'interaction. Il est donc intéressant de comparer les modes d'utilisation possible de Touch Through ; avec la modalité visuelle distale, on peut dire « il est là-bas, je suis ici, je peux aller le rejoindre ». Le vécu et le sentiment de présence ne seront pas les mêmes avec les modalités proximales, où l'on est dans le noir et n'avons un retour que lorsque l'on a « touché » l'autre participant. Touch Through représente donc un outil qui va permettre de comparer l'expérience et le sentiment de présence qui sont permis dans les deux modes d'utilisation (distal et proximal). Cette comparaison va nous permettre de tester formellement notre hypothèse d'un espace perceptif appauvri qui permet une expérience riche et touchante. Pour l'effectuer, nous avons réalisé une expérience que nous présentons dans la partie suivante.

2.4 Étude expérimentale : toucher à distance avec Touch Through (Distal vs Proximal)²⁷

De façon à évaluer le sentiment de présence (comme premier pas possible vers un échange émotionnel) permis par notre dispositif Touch Through, nous avons élaboré une étude expérimentale basée sur le célèbre « jeu d'imitation » d'Alan Turing (Turing, 1950), plus connu en intelligence artificielle sous le terme de « test de Turing ». Il s'agissait à l'origine d'une

²⁷ Cette étude a donné lieu à une publication avec poster et à deux communications orales (dont une avec actes) :

- Le Bihan, G., Lenay, C., Tixier, M., Mara, J. (2013). Touch Through: Experiencing Remote Touch Across Different Modalities, Proceedings of CHI, Paris, France. April 27 - May 2, 2013

- Le Bihan, G. (2013). Touch at a Distance for the Remote Communication of Emotions, Proceedings of TEI, Barcelona, Spain. February 9 - 13, 2013

- Le Bihan, G. (2013). Touch Through: Distal Touch vs Proximal Touch. *Séminaire Phiteco (Philosophie, Technologie, Cognition), From Perceptual Interactions to Extended Cognition*, Université de Technologie de Compiègne, Compiègne, France, 21-25 janvier 2013

épreuve destinée à décider si une machine, suffisamment complexe, pouvait se faire passer pour un humain sans être démasquée par un autre humain. Dans le test imaginé et décrit par Turing, un sujet s'engage dans une conversation écrite à l'aide d'un clavier avec deux interlocuteurs placés dans deux pièces séparées de la sienne. L'un des interlocuteurs est un autre sujet humain, tandis que le second est un ordinateur. La tâche pour le premier humain consiste à essayer de savoir, uniquement par le biais d'un dialogue par écrit, dans quelle pièce est la machine, sachant que cette dernière (ainsi que le sujet humain) fait tout pour être reconnue comme étant un humain. Si le sujet humain se trouve dans l'impossibilité de désigner qui est la machine et qui est l'humain, alors la machine aura réussi le test, et on pourra considérer qu'elle est intelligente, ou du moins qu'elle est capable de simuler l'intelligence (Searle, 1980). Nous voulons ici proposer un test de Turing non-verbal en utilisant Touch Through, afin de pouvoir mettre en lumière que notre application permet, malgré sa simplicité, de faire la différence entre un autre humain et un robot du fait de la dynamique d'interaction interpersonnelle en tant qu'engagement mutuel. Ainsi, elle serait un pas dans la direction de la communication émotionnelle à distance, en permettant d'avoir le sentiment de la présence d'autrui (car si je peux faire la distinction entre un robot et un autre, c'est que je sais quand un autre est avec moi).

Contrairement à l'expérience princeps sur le croisement perceptif minimaliste (Auvray, Lenay et Stewart, 2009) que nous avons déjà présentée (voir partie 1.1.1.1.3), les sujets ne pourront rencontrer qu'une seule entité dans un environnement bidimensionnel : un autre sujet humain ou un robot. Nous nous situons donc dans le cadre d'interactions strictement dyadiques. Ce sont les robots qui joueront le rôle de « leurres mobiles », mais cette fois, leur comportement ne sera pas le reflet à distance du comportement d'un sujet humain dans l'espace, mais dans le temps. En effet, les interactions avec un robot consisteront à interagir avec un enregistrement de ce que le partenaire a effectué lors d'une interaction précédente. Le robot aura donc bien un comportement aussi complexe que celui du partenaire (étant donné qu'il rejoue ce que le partenaire a effectivement fait), mais il n'est pas sensible au sujet et ne réagit pas en fonction de lui. Notons que ce n'est pas toujours l'enregistrement de la même interaction qui est rejouée ; le robot ne se déplace donc pas toujours de la même façon à chaque essai où il intervient. Dans ce cadre, nous postulons que les sujets humains parviendront à mutuellement se reconnaître du fait d'une co-régulation dynamique de leurs activités perceptives. Nous faisons aussi l'hypothèse que les sujets amélioreront leurs performances au fur et à mesure de l'expérience. En effet, la répétition des essais se traduira par une stabilisation progressive de la dynamique interpersonnelle, qui peut donc servir de support de plus en plus pertinent pour la reconnaissance mutuelle des interactants.

2.4.1 Population

24 sujets, répartis en 4 groupes expérimentaux de 6 sujets chacun (soit 3 binômes dans chaque groupe), ont participé à cette expérience. Les sujets sont des étudiants de l'Université de Technologie de Compiègne âgés de 18 à 28 ans. Ils n'ont pas de connaissance de l'application Touch Through et n'ont pas connaissance des éléments et hypothèses théoriques mobilisés dans l'expérience.

Les groupes expérimentaux sont définis en fonction de l'ordre de passation des conditions expérimentales décrites ci-après.

2.4.2 Matériel

Pour cette expérience, nous avons fait utiliser aux sujets l'application Touch Through, décrite dans la partie précédente. Chaque participant est donc installé devant un poste de travail qui comprend un smartphone où l'application est installée.

Le dispositif expérimental comprend un ordinateur et deux smartphones : l'ordinateur constitue le serveur, qui va gérer les connexions-clients et enregistrer les données, tandis que les téléphones constituent les clients et sont destinés à chacun des deux sujets de chaque paire. Les smartphones sont des HTC One S.

L'espace sur lequel se déplacent les sujets est un espace rectangulaire en 2-D de 888 par 540 pixels qui correspond géométriquement à la surface de l'écran des smartphones diminuée d'un espace destiné aux commandes (une bande en bas de l'écran où se trouvent la touche retour et les icônes du téléphone).

En plaçant leur doigt sur l'écran tactile, les sujets déplacent un avatar circulaire de 108 pixels de diamètre (le corps-image), auquel correspond topologiquement un corps percevant circulaire de 108 pixels de diamètre (soit 1,1 cm) dont la configuration et la sensibilité varient en fonction des conditions expérimentales. La surface de contact est donc un cercle de 216 pixels de diamètre autour de chacun des centres des avatars ; il y a stimulation quand la distance entre les deux centres des avatars est de moins de 108 pixels. La surface d'interaction est de 888 x 540 pixels (soit 8,8 x 5,4 cm), et la diagonale (soit la distance maximale entre les deux avatars) de 1039 pixels (soit 10,3 cm).

Le corps percevant est sensible soit à l'avatar du partenaire, soit à l'avatar du robot, qui est en fait un enregistrement de ce qu'a fait le partenaire durant une interaction précédente.

2.4.3 Procédure

Dans chacun des quatre groupes expérimentaux, 6 sujets sont répartis en 3 binômes. Chaque binôme passe 2 conditions expérimentales, chacune des conditions expérimentales étant définie par un mode (distal ou proximal) et une modalité (visuelle ou vibratoire) de l'application Touch Through, ce qui engendre une configuration différente des corps-images et des corps-percevants et de leur sensibilité associée.

Lorsque le champ récepteur élémentaire croise au moins un pixel (en l'occurrence de l'avatar de l'autre sujet ou d'un enregistrement préalable de cet avatar), cela active une stimulation par l'intermédiaire d'un calcul de collision exécuté par l'application Touch Through.

Dans la condition Distale, Touch Through est réglé en mode distal et en modalité visuelle. Le corps-image circulaire d'un sujet apparaît sous forme de point sur les deux écrans lorsque le sujet pose son doigt sur l'écran ; le corps-image se déplace en même temps que le doigt évolue sur l'écran tactile. Un deuxième corps-image circulaire en forme de point d'une autre couleur est associé à l'autre sujet. Lorsque les corps percevants (confondus avec les corps-images) entrent en contact, les corps-images adoptent une troisième couleur.

Dans la condition Proximale Visuelle, Touch Through est réglé en mode proximal et en modalité visuelle. Le corps-image des sujets n'apparaît alors plus que lorsque les corps percevants sont en contact ; ils sont circulaires, et de la même couleur.

Dans la condition Proximale Vibratoire, Touch Through est réglé en mode proximal et en modalité vibratoire. Le corps-image des sujets n'est alors jamais visible à l'écran, mais lorsque les corps percevants sont en contact, les deux téléphones se mettent à vibrer. Notons que nous aurions pu aussi utiliser le retour sonore au lieu du retour vibratoire dans cette condition ; ces deux types de retours étant équivalents, nous n'avons pas jugé utile de faire une condition Proximale Sonore.

Cependant, il est à noter que les deux conditions proximales adoptées ici ne sont pas équivalentes. Le mode Proximal Visuel est en effet différent du mode Proximal Vibratoire, pas simplement parce que le retour sensoriel diffère, mais aussi parce que dans le mode visuel, le sujet a accès à la position relative du partenaire au moment du contact. On dispose donc d'une information locale concernant la position relative du partenaire au moment du contact. Par opposition, la seule information fournie dans le mode Proximal Vibratoire est donnée au moment du contact.

Chaque condition expérimentale est constituée de 16 essais, dont 4 essais de familiarisation et 12 essais expérimentaux. Les deux sujets font l'expérience de la même condition en même temps. Chaque paire de sujets fait l'expérience de deux des trois modalités ; tous les sujets passent la condition Distale, mais seule la moitié d'entre eux fait l'expérience en

condition Proximale Visuelle, tandis que l'autre moitié passe la condition Proximale Vibratoire. Les binômes passent donc successivement deux des trois conditions expérimentales.

Les groupes expérimentaux, tous composés de trois binômes, sont constitués en fonction des conditions passées et de l'ordre dans lequel elles sont passées. Les quatre groupes indépendants définis en fonction de l'ordre de passation des conditions expérimentales sont donc :

- groupe 1 : condition Distale, puis condition Proximale Visuelle ;
- groupe 2 : condition Distale, puis condition Proximale Vibratoire ;
- groupe 3 : condition Proximale Visuelle, puis condition Distale ;
- groupe 4 : condition Proximale Vibratoire, puis condition Distale.

Les passations se font en binôme. Les sujets sont accueillis dans une salle, et l'application leur est très brièvement présentée, ainsi qu'une consigne globale décrivant l'expérience à laquelle ils vont participer. Dans cette consigne, les sujets sont informés que l'expérience va consister à interagir par le biais de l'application dans un espace partagé avec soit l'autre participant, soit un objet mobile, que l'on appelle le robot, se déplaçant de manière aussi complexe que le partenaire. Il est également annoncé que l'expérience est divisée en deux phases, chacune de ces phases étant décomposée en deux parties distinctes : une phase de familiarisation et une phase de test. Après cette consigne, les participants sont installés dans deux salles adjacentes, de façon à les séparer et à éviter tout indice perceptif autre que les stimulations délivrées par l'application résultant de la rencontre entre les corps-images dans l'environnement numérique (comme par exemple les réactions des participants lors de la rencontre). Les sujets sont assis devant le smartphone qui leur est respectivement attribué. Lorsque les sujets sont prêts le premier essai de la première condition de chaque groupe expérimental commence. Notons que durant tous les essais, l'activité des sujets dans l'application est enregistrée sous forme numérique.

Le premier essai est un essai de familiarisation avec Touch Through et avec la condition expérimentale en cours (mode et modalité de l'application). Les sujets sont donc connectés dans un espace partagé. La consigne pour cet essai indique au sujet de poser son doigt sur l'écran tactile, de le déplacer pour chercher son partenaire et de suivre ce dernier le plus longtemps possible. Cet essai n'est pas chronométré, et dure tant que les sujets le jugent utile. L'expérimentateur s'assure que les sujets ont bien compris, et leur indique le cas échéant que l'essai suivant va commencer.

Le deuxième et le troisième essais consistent à nouveau pour les sujets à se chercher et se suivre dans l'espace, excepté que cette fois, ils disposent de 30 secondes pour interagir, ce qui sera le temps alloué pour la tâche de test.

Le quatrième essai n'est plus en interaction : chaque sujet se retrouve dans un espace différent, partagé avec un objet mobile. Le robot rejoue les mouvements effectués par le partenaire lors de la première interaction de 30 secondes (ce qui correspondait au deuxième essai passé). Cet essai dure lui aussi 30 secondes. Il consiste pour les sujets à pouvoir essayer l'interaction avec le robot, après avoir interagi uniquement avec le partenaire jusqu'ici. Cet essai clôt la phase de familiarisation.

Le cinquième essai est le premier essai de test. Les sujets sont informés que l'expérience à proprement parler commence. L'objectif consiste alors à explorer l'espace pendant 30 secondes, et à déterminer s'ils partagent l'espace avec le partenaire, ou avec le robot. À la fin de l'essai, les sujets disposent de 10 secondes pour indiquer leur réponse sur une feuille prévue à cet effet. À la fin de ces 10 secondes, l'expérimentateur lance l'essai suivant, et ce jusqu'à l'effectuation de 12 essais consécutifs. Aucune information supplémentaire quant à la nature du robot ou à la façon dont il se déplace n'est délivrée. De plus, aucun feedback n'est donné aux sujets pendant la durée de l'expérience.

Au cours de ces 12 essais, les sujets se retrouvent 6 fois face à face, et 6 fois face au robot, dans un ordre fixé préalablement.

À la fin des 12 premiers essais expérimentaux, une pause de deux minutes est proposée aux participants, pendant laquelle ils ne peuvent pas interagir entre eux. Ensuite, la deuxième phase de l'expérience peut commencer. La deuxième condition expérimentale commence alors, suivant le même scénario que la première condition passée (soit 4 essais de familiarisation suivis de 12 essais expérimentaux). Les sujets sont invités à faire attention aux modifications du type de stimulations reçues lors des essais de familiarisation, de façon à s'assurer que la différence de condition expérimentale est comprise par les deux sujets.

À la fin de l'expérience, les sujets sont invités à faire un retour sur leur expérience en répondant à un questionnaire. Nous réunissons ensuite les sujets de façon à procéder à un débriefing, où les sujets peuvent nous faire part de leurs difficultés éventuelles ou expliciter leur stratégie.

2.4.4 Hypothèses opérationnelles

Les sujets parviendront à distinguer leur partenaire et le robot, en vertu de l'émergence d'un processus dynamique de co-ajustements et de co-régulation de l'interaction, processus qui est un indicateur du sentiment de présence.

Le taux de succès des sujets sera différent pour les différentes conditions. Nous pensons que les conditions proximales seront, malgré leur plus grande simplicité, plus à même de permettre aux sujets de se reconnaître.

Du fait d'un effet d'apprentissage, les sujets feront mieux cette distinction lors des six derniers essais que lors des six premiers. De même, les sujets se reconnaîtront mieux lors de la deuxième condition passée que lors de la première condition passée.

Les sujets se reconnaîtront davantage de façon mutuelle que de façon unilatérale, la reconnaissance elle-même émergeant du processus de l'interaction, et non d'un jugement individuel unilatéral.

2.4.5 Résultats

2.4.5.1 Analyses des réponses individuelles²⁸

Nous nous intéressons tout d'abord aux réponses individuelles globales indépendamment des conditions expérimentales et de l'ordre des passations. Il existe deux types de réponses : « partenaire » et « robot ». Une bonne réponse consiste à répondre « partenaire » lors d'une interaction avec le partenaire et « robot » lors d'une interaction avec le robot. Il y a alors reconnaissance. À l'inverse, une mauvaise réponse consiste à répondre « partenaire » suite à une interaction avec le robot, et « robot » suite à une interaction avec le partenaire. Il n'y a alors pas reconnaissance. Globalement, nous observons que le nombre moyen de bonnes réponses individuelles est supérieur au nombre moyen de mauvaises réponses individuelles ($d = 6,75$ réponses) (voir tableau 2). La différence est significative (test de Student : $t_{(46)} = 7.5649$; $p = 1.3042 \times 10^{-9}$).

²⁸ Par « Analyse des réponses individuelles », nous n'entendons pas une analyse de profils individuels de 24 réponses, mais une analyse des réponses fournies par les sujets indépendamment des réponses de leur partenaire.

Tableau 2. Nombre et fréquence moyens de bonnes et mauvaises réponses fournies par les 24 sujets, de manière globale, selon chaque condition et selon le passage de la première ou de la deuxième condition.

	Reconnaissance	Non-reconnaissance
Sujet total	15,375 (64%)	8,625 (36%)
Condition Distale	8,00 (67%)	4,00 (33%)
Condition Proximale Visuelle	4,00 (67%)	2,00 (33%)
Condition Proximale Vibratoire	3,375 (56%)	2,625 (44%)
Première condition passée	7,083 (59%)	4,917 (41%)
Deuxième condition passée	8,292 (69%)	3,708 (31%)

La différence globale entre le nombre moyen de bonnes réponses et de mauvaises réponses suit la même évolution dans les trois conditions expérimentales (à savoir plus de bonnes réponses que de mauvaises réponses), cependant elle se manifeste dans des proportions différentes : la différence observée est faible pour la condition Proximale Vibratoire ($d = 0,75$ réponse), alors qu'elle est plus importante pour les deux autres conditions ($d = 2$ réponses) (voir tableau 2 et figure 78). Suite à l'analyse statistique, il est établi que la différence est significative pour la condition Distale (test de Student : $t_{(46)} = 5.5377$; $p = 1.4190 \times 10^{-6}$) et pour la condition Proximale Visuelle (test de Student : $t_{(22)} = 4.7844$; $p = 8.8910 \times 10^{-5}$), mais qu'elle ne l'est pas pour la condition Proximale Vibratoire (test de Student : $t_{(22)} = 1.3763$; $p = 0.1826$).

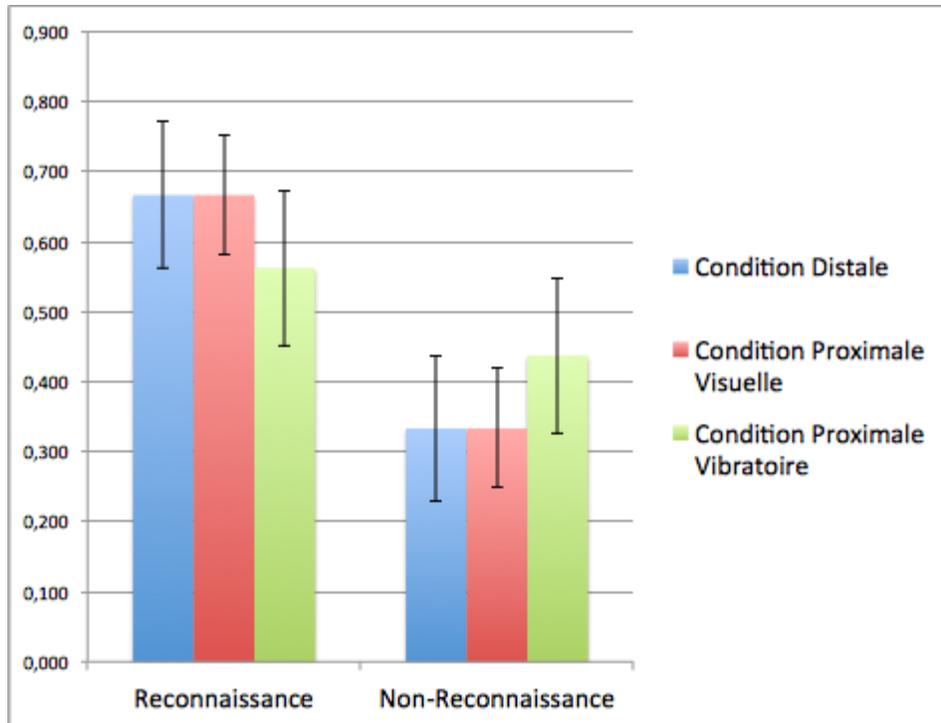


Figure 78. Fréquences moyennes de bonnes et mauvaises réponses en fonction des conditions expérimentales.

On notera que le fait que la condition Proximale Vibratoire se démarque des deux autres ne va pas dans le sens de ce que nous attendions ; nous avons l'hypothèse que les conditions proximales seraient plus proches entre elles que de la condition Distale, mais il semblerait que ces deux conditions présentent d'importantes différences, bien qu'elles aient toutes deux été conçues sur un mode proximal.

D'une façon générale, nous n'observons pas d'effet d'apprentissage à l'intérieur des conditions expérimentales. En effet, toutes conditions expérimentales confondues, le nombre moyen de bonnes réponses est comparable au cours des 6 derniers essais comparativement aux 6 premiers ($d = 0,29$ réponse) ; parallèlement, le nombre moyen de mauvaises réponses est légèrement plus faible au cours des 6 derniers essais (voir figure 79). Les différences ne sont pas significatives (test de Student : $t_{(46)} = -0.2737$; $p = 0.7856$).

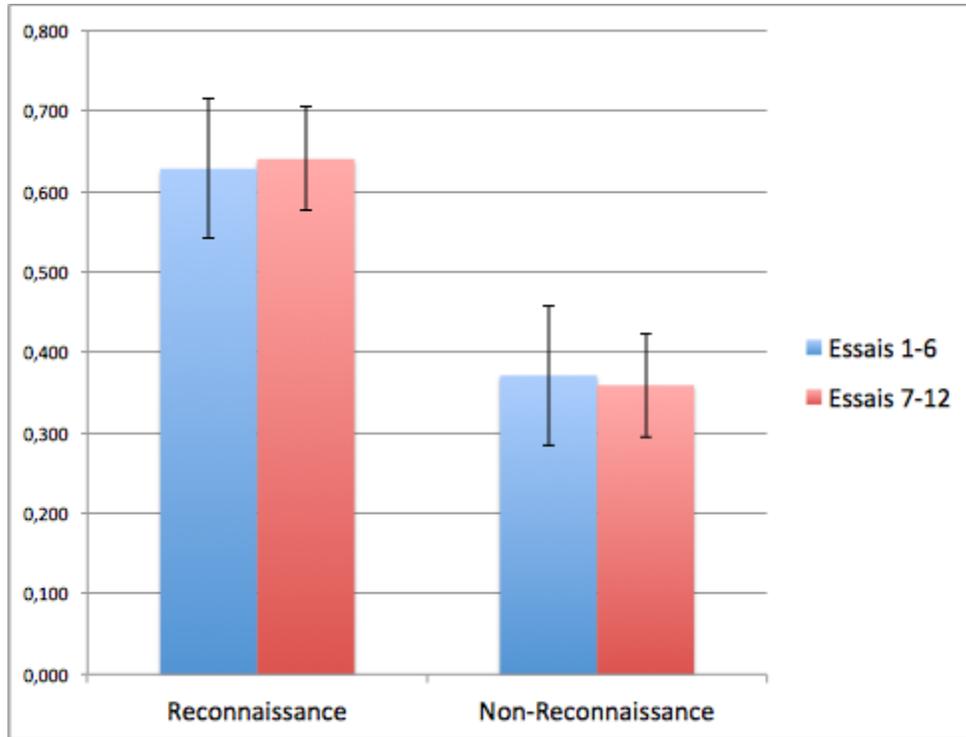


Figure 79. Fréquence moyenne de bonnes et mauvaises réponses en fonction des deux blocs d'essais (essais 1 à 6 vs essais 7 à 12).

De façon descriptive, l'effet de l'apprentissage par condition est très faible en condition Distale ($d = 0,04$ réponse), ainsi qu'en condition Proximale Vibratoire ($d = -0,04$ réponse) (voir figure 80). Cependant, il est plus important en condition Proximale Visuelle ($d = 0,25$ réponse). Aucune de ces différences n'est significative (condition Distale, test de Student : $t_{(46)} = -0.1028$; $p = 0.9186$; condition Proximale Visuelle, test de Student : $t_{(22)} = -0.4925$; $p = 0.6277$ et condition Proximale Vibratoire, test de Student : $t_{(22)} = 0.0679$; $p = 0.9465$).

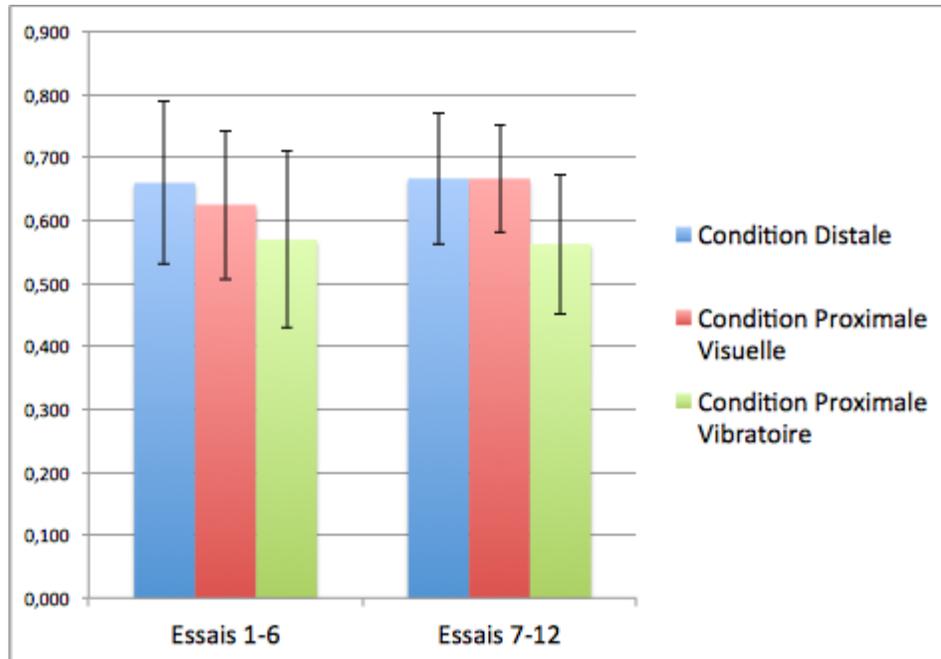


Figure 80. Fréquence moyenne de bonnes réponses en fonction des conditions expérimentales.

L'effet d'apprentissage global se constate en fait mieux en observant l'ordre dans lequel les conditions sont passées : le nombre moyen de bonnes réponses est supérieur lors de la deuxième condition passée par rapport à la première ($d = 1,21$ réponse) (voir figure 81). Cependant, cette différence n'est pas significative pour un seuil de 5% (test de Student : $t_{(46)} = -1.7434$; $p = 0.0883$).

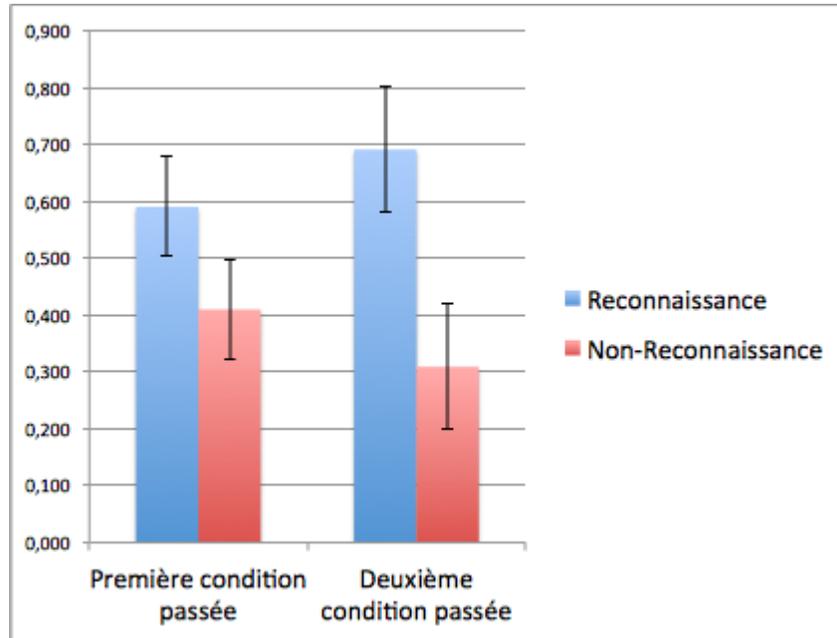


Figure 81. Fréquence moyenne de bonnes et mauvaises réponses pour la première et la deuxième condition passée.

En s'intéressant au nombre de réponses « Partenaire », on constate que de manière globale, les sujets répondent plus fréquemment « Partenaire » que « Robot » ($d = 9,38\%$) (voir tableau 3). Cette différence est significative (test de Student : $t_{(46)} = 4.0162$; $p = 0.0002$). Les sujets ont donc tendance à prêter des intentions humaines à l'interactant. Par condition, la différence est presque nulle en condition Distale ($d = -1,38\%$) et est plus marquée dans les conditions Proximales Visuelle ($d = 22,22\%$) et Vibratoire ($d = 18,06\%$). Les différences dans ces deux dernières conditions sont significatives (respectivement, test de Student : $t_{(22)} = 4.3620$; $p = 0.0002$ et test de Student : $t_{(22)} = 5.3274$; $p = 2.3930 \times 10^{-5}$). La différence du type de réponse entre les trois conditions est significative (test de Fischer : $F_{(2, 45)} = 2.6085$; $p = 0.0108$). En comparant les conditions deux à deux, on trouve des différences significatives entre la condition Distale et la condition Proximale Visuelle (test de Student : $t_{(34)} = -2.6397$; $p = 0.0147$) et entre la condition Distale et la condition Proximale Vibratoire (test de Student : $t_{(34)} = -2.7209$; $p = 0.0105$). La condition Distale est donc bien la seule où la différence entre le nombre de réponses « Partenaire » et « Robot » n'est pas significative. En conditions Proximales, les sujets répondent donc plus fréquemment « Partenaire » : ils ont donc plus l'impression d'être face à un humain que dans les autres conditions. Ces conditions sont donc particulièrement intéressantes pour permettre le sentiment de présence, que ce soit à tort ou à raison.

Tableau 3. Fréquence des réponses « Partenaire » et « Robot » de manière globale et en fonction des conditions expérimentales quelque soit l'interactant.

	Réponses « Partenaire »	Réponses « Robot »
Général	54,69 %	45,31 %
Condition Distale	49,31 %	50,69 %
Condition Proximale Visuelle	61,11 %	38,89 %
Condition Proximale Vibratoire	59,03 %	40,97 %

Nous allons à présent nous intéresser aux réponses en fonction du type d'interactant. Pour rappel, chaque sujet est confronté à deux types d'interactants, rencontrés chacun 6 fois au cours des 12 essais expérimentaux : l'autre sujet humain, et un robot (en fait un enregistrement d'une interaction précédente avec le partenaire). Nous regardons tout d'abord la fréquence de la réponse « Partenaire » en fonction du type d'interactant (Partenaire vs Robot). De façon générale, nous observons que la fréquence de réponses « Partenaire » est supérieure lorsque le sujet se trouve face à l'autre humain que lorsqu'il se trouve face au robot ($d = 28\%$) (voir figure 82). Cette différence est significative (test de Student : $t_{(46)} = 6.4080$; $p = 7.3792 \times 10^{-8}$).

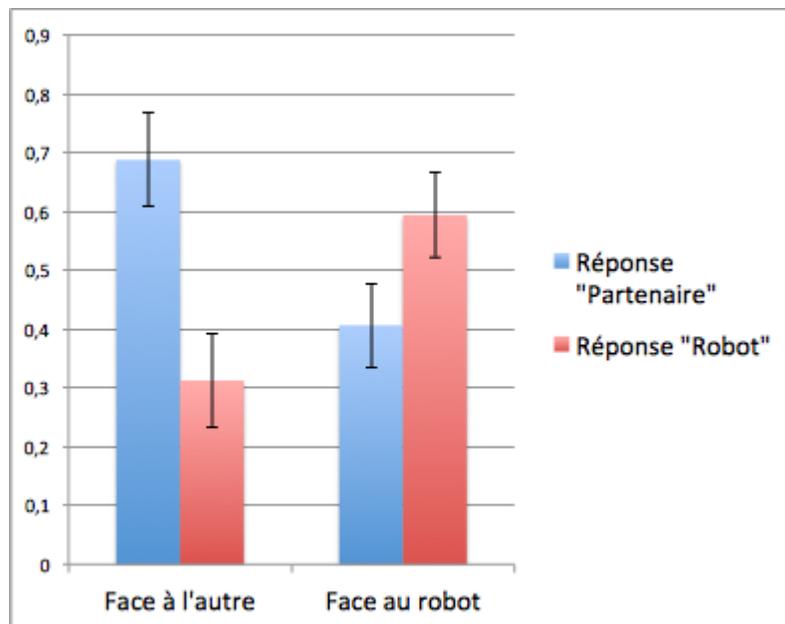


Figure 82. Fréquences des deux types de réponses (« Partenaire » ou « Robot ») en fonction des deux types d'interactants (face à l'autre sujet ou face au robot).

On constate que l'on retrouve la différence entre les fréquences de la réponse « Partenaire » face à l'autre sujet et face au robot dans toutes les conditions expérimentales (voir figure 83). À chaque fois, cette différence est significative (pour la condition Distale, test de Student : $t_{(46)} = 4.7009$; $p = 2.3953 \times 10^{-5}$, pour la condition Proximale Visuelle, test de Student : $t_{(22)} = 2.5771$; $p = 0.0174$, et pour la condition Proximale Vibratoire, test de Student : $t_{(22)} = 2.7835$; $p = 0.0112$).

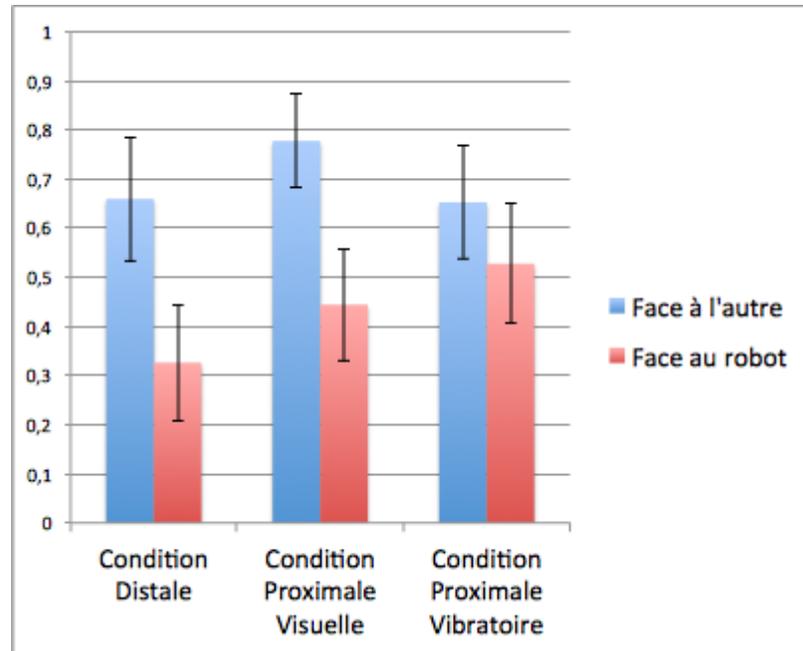


Figure 83. Fréquences de réponses « Partenaire » en fonction du type d'interactant (interaction avec le partenaire ou avec le robot) et en fonction des conditions expérimentales.

En d'autres termes, les sujets répondent plus fréquemment « Partenaire » dans l'absolu, mais aussi plus fréquemment lorsqu'ils ont effectivement interagi avec l'autre sujet que lorsqu'ils ont interagi avec le robot. Le robot est plus souvent identifié par erreur comme étant un humain que l'humain comme étant un robot, mais les sujets répondent tout de même bien plus souvent « Partenaire » lorsqu'ils sont face à autrui que lorsqu'ils sont face au robot.

Il est intéressant d'observer le nombre moyen de bonnes réponses selon le type d'interactant. Face à l'autre participant, les sujets donnent plus souvent la bonne réponse (à savoir « Partenaire ») que face au robot (où la bonne réponse est alors « Robot ») ($d = 9\%$) (voir figure 84). Cette différence est significative ($t_{(46)} = 2.1360$; $p = 0.0381$). Autrui est donc plus facile à reconnaître comme étant un humain que le robot comme étant un robot ; ainsi, le fait d'être face à face aide de façon significative les sujets à trouver la bonne réponse.

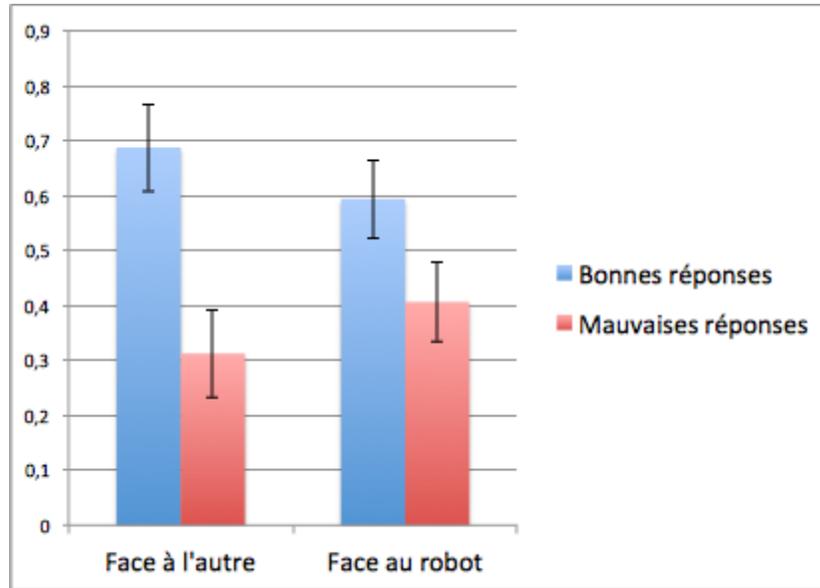


Figure 84. Fréquences des bonnes et mauvaises réponses en fonction du type d'interactant.

Par condition, on constate que dans la condition Distale, la différence entre le nombre de bonnes réponses face à l'autre sujet et face au robot est quasiment nulle ($d = -1\%$). Elle est par contre plus marquée pour les deux conditions Proximales ; en Proximale Visuelle $d = 22\%$, et en Proximale Vibratoire $d = 18\%$ (voir figure 85). Cependant, seule la différence en condition Proximale Visuelle est significative ($t_{(22)} = 2.5771$; $p = 0.0174$). Il y a donc là une différence entre la condition Distale et les deux conditions Proximales. Dans la première, le nombre de bonnes réponses est semblable quelque soit l'interactant ; les bonnes réponses dans la condition Distale sont réparties également entre les interactions face au partenaire et celles face au robot. À l'inverse, les bonnes réponses dans les conditions Proximales sont plus élevées face à l'autre sujet que face au robot. Ce résultat va dans le sens de notre hypothèse posant les conditions Proximales comme étant plus adaptées pour le sentiment de présence ; lorsqu'il y a reconnaissance, elle a plus souvent lieu face à l'autre sujet que face au robot. Il y a dans ces conditions une tendance générale à reconnaître plutôt la présence de l'humain que le robot. Notons que c'est dans la condition Proximale Visuelle que la reconnaissance de l'autre sujet lorsque les partenaires sont en interaction est la plus élevée, les bonnes réponses s'élevant alors à 78% (voir figure 85).

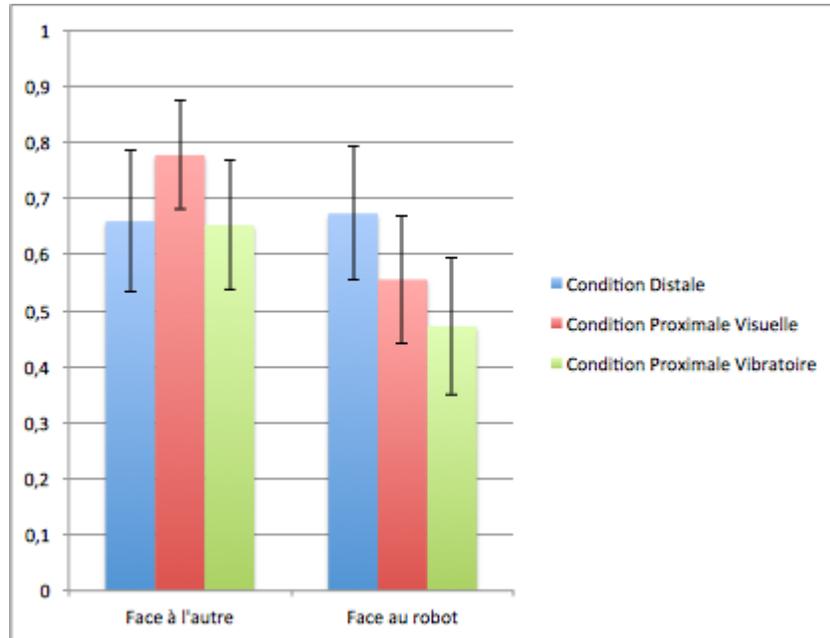


Figure 85. Fréquences de bonnes réponses en fonction des trois conditions expérimentales et en fonction du type d'interactant.

2.4.5.2 Analyses des réponses collectives

Après avoir analysé les réponses individuelles des sujets indépendamment des réponses de leur partenaire, nous nous intéressons maintenant aux données collectives, c'est-à-dire à la qualité de la réponse du binôme pour chaque essai. Dans ce cadre, nous n'allons dans un premier temps prendre en compte pour les bonnes et mauvaises réponses que les cas où les sujets se retrouvaient face à face. En effet, les essais où les sujets étaient tous deux en interaction avec un robot ne permettent pas d'accéder à l'exactitude d'une réponse collective, mais plutôt à deux réponses individuelles. Nous considérons donc trois cas de figure, correspondant à trois modalités de réponse par binôme en face à face : les bonnes réponses mutuelles (c'est-à-dire les cas où les deux sujets se sont mutuellement identifiés comme des humains), les bonnes réponses solitaires (c'est-à-dire les cas où seul un des deux sujets a reconnu qu'il interagissait avec son partenaire), et les mauvaises réponses mutuelles (c'est-à-dire les cas où aucun des deux sujets n'a identifié son partenaire comme étant humain). Dans les cas où nous considérerons les réponses des binômes sans qu'ils soient nécessairement face à face, nous utiliserons le terme « double bonne réponse » pour signifier deux bonnes réponses, et « simple bonne réponse » pour signifier une bonne et une mauvaise réponse.

Globalement, nous observons que la fréquence de bonnes réponses mutuelles est supérieure à la fréquence de bonnes réponses solitaires ($d = 25\%$) et à la fréquence de mauvaises réponses mutuelles ($d = 38\%$) (voir figure 86). Une analyse statistique confirme que les trois modalités de réponses par binôme sont statistiquement différentes (test de Fischer : $F_{(2, 33)} = 23.6129$; $p = 4.3083 \times 10^{-7}$). En les comparant deux à deux, les différences sont significatives entre les bonnes réponses mutuelles et les bonnes réponses solitaires (test de Student : $t_{(22)} = 4.1781$; $p = 0.0004$), entre les bonnes réponses solitaires et les mauvaises réponses mutuelles (test de Student : $t_{(22)} = -2.5701$; $p = 0.0176$) et entre les bonnes réponses mutuelles et les mauvaises réponses mutuelles (test de Student : $t_{(22)} = 6.5123$; $p = 2.7614 \times 10^{-6}$). Ainsi, les mauvaises réponses mutuelles sont le type de réponse le moins statistiquement fréquent, alors que les bonnes réponses mutuelles sont données à plus de 50% lorsque les sujets sont face à face.

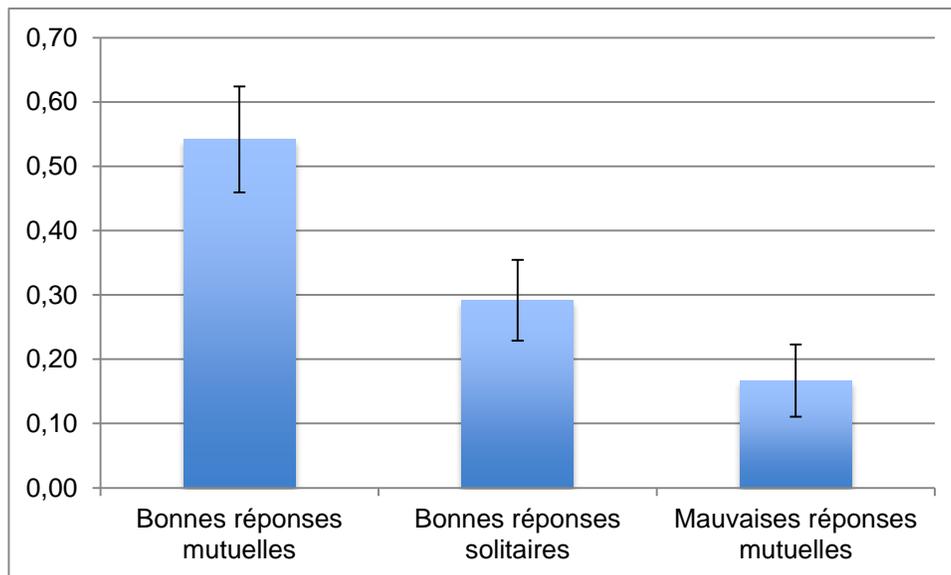


Figure 86. Fréquence des trois modalités de réponses par binôme face à face pour l'ensemble des binômes.

De façon descriptive, nous observons que les différences entre les trois modalités de réponse par binôme s'expriment de façon différente selon les conditions expérimentales (voir figure 87).

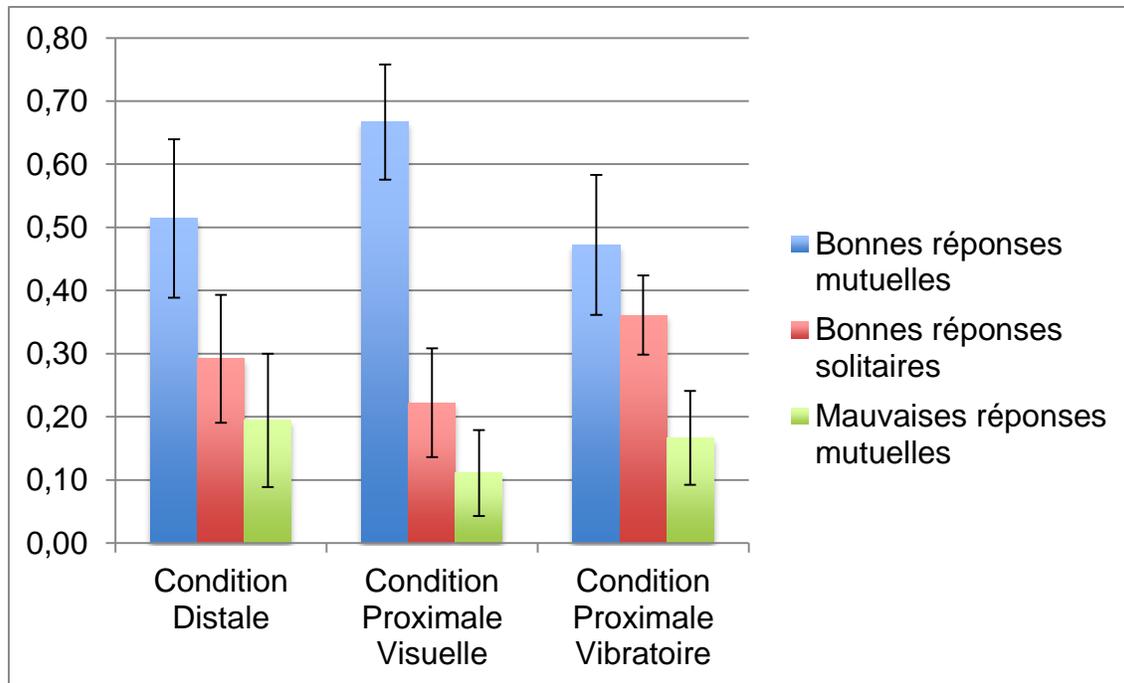


Figure 87. Fréquence des trois modalités de réponse par binôme en face à face en fonction des conditions expérimentales.

En ce qui concerne la condition Distale, nous observons que la fréquence de bonnes réponses mutuelles est supérieure à la fréquence de bonnes réponses solitaires ($d = 22\%$), elle-même supérieure à la fréquence de mauvaises réponses mutuelles ($d = 10\%$). Les bonnes réponses mutuelles constituent 51% des réponses à deux dans cette condition. Les différences entre les trois modalités de réponse sont statistiquement significatives (test de Fischer : $F_{(2, 33)} = 6.4906$; $p = 0.0042$). Ce sont les bonnes réponses mutuelles qui sont statistiquement plus fréquentes que les bonnes réponses solitaires (test de Student : $t_{(22)} = 2.3872$; $p = 0.0264$) et que les mauvaises réponses mutuelles (test de Student : $t_{(22)} = 3.3721$; $p = 0.0028$).

Concernant la condition Proximale Visuelle, on retrouve la même évolution : la fréquence de bonnes réponses mutuelles est supérieure à la fréquence de bonnes réponses solitaires ($d = 44\%$), elle-même supérieure à la fréquence de mauvaises réponses mutuelles ($d = 11\%$). C'est dans cette condition que la fréquence de bonnes réponses mutuelles est la plus élevée, et est élevée de la façon la plus importante par rapport à la deuxième modalité de réponse la plus fréquente : les bonnes réponses mutuelles constituent 67% des réponses à deux dans la condition Proximale Visuelle. Les différences entre les trois modalités de réponse sont statistiquement significatives (test de Fischer : $F_{(2, 15)} = 19.0975$; $p = 7.5274 \times 10^{-5}$). Ce sont les bonnes réponses mutuelles qui sont statistiquement plus fréquentes que les bonnes réponses

solitaires (test de Student : $t_{(10)} = 4.3388$; $p = 0.0015$) et les mauvaises réponses mutuelles (test de Student : $t_{(10)} = 5.9764$; $p = 0.0002$).

Enfin, en ce qui concerne la condition Proximale Vibratoire, la fréquence de bonnes réponses mutuelles est aussi plus élevée que la fréquence de bonnes réponses solitaires ($d = 11\%$), qui est elle-même plus élevée que la fréquence de mauvaises réponses mutuelles ($d = 19\%$). Les bonnes réponses solitaires constituent 47% des réponses à deux dans cette condition. Les différences sont statistiquement significatives (test de Fischer : $F_{(2, 15)} = 4.9485$; $p = 0.0223$). Cette fois, ce sont les mauvaises réponses mutuelles qui sont statistiquement moins fréquentes que les bonnes réponses mutuelles (test de Student : $t_{(10)} = 2.8035$; $p = 0.0212$) et que les bonnes réponses solitaires (test de Student : $t_{(10)} = 2.4450$; $p = 0.0352$).

Si nous nous centrons sur la fréquence de bonnes réponses mutuelles en fonction des conditions expérimentales, nous observons que la fréquence de bonnes réponses mutuelles est plus importante en condition Proximale Visuelle qu'en condition Distale ($d = 15\%$) et en condition Proximale Vibratoire ($d = 19\%$) (voir figure 87). Cependant, la différence n'est pas significative (test de Fischer : $F_{(2, 9)} = 1.2524$; $p = 0.3063$).

En détaillant les réponses données par binôme selon le type d'interactant (soit le partenaire, soit le robot), les doubles bonnes réponses sont en fait une reconnaissance mutuelle face au partenaire, et deux bonnes réponses face au robot. En comparant les fréquences de doubles bonnes réponses face à l'autre sujet et face au robot, on constate que c'est bien face à l'autre participant que les deux sujets trouvent le plus souvent la bonne réponse de concert ($d = 18\%$) (voir figure 88). La différence est significative (test de Student : $t_{(22)} = 2.6899$; $p = 0.0134$). Pour les bonnes réponses solitaires, c'est face au robot qu'elles sont le plus fréquentes ($d = 17\%$). Enfin, les fréquences de doubles mauvaises réponses sont comparables entre les deux types d'interactant ($d = 0,7\%$). Les bonnes réponses mutuelles sont la réponse à deux la plus fréquente face à l'autre sujet (54%), tandis que face au robot, ce sont les simples bonnes réponses qui se produisent le plus souvent (47%).

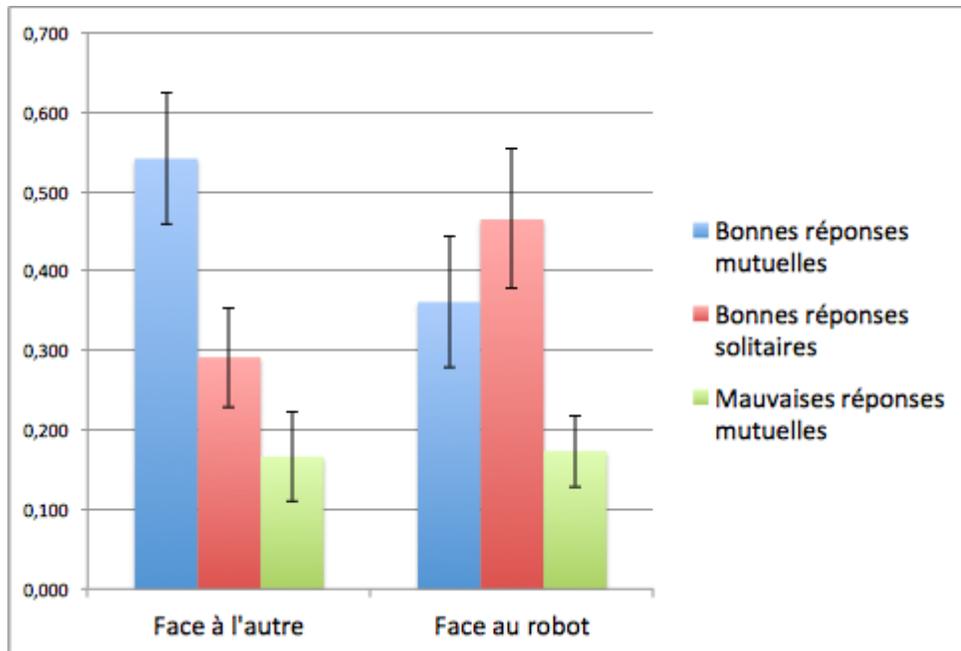


Figure 88. Fréquence des trois modalités de réponses par binôme en fonction du type d'interactant.

À présent, nous souhaitons nous intéresser aux réponses sous l'angle de la nature de la réponse (« Partenaire » ou « Robot »). Nous avons dans ce cas considéré toutes les réponses données par les binômes, qu'ils soient face à face ou face au robot ; nous détaillerons dans nos résultats la différence entre les deux cas. Soit les deux sujets répondent « Partenaire », soit seul l'un des deux le fait, soit ils répondent tous deux « Robot ». On constate que de manière générale, les sujets en binôme répondent plus souvent (à 37,85 %) l'un « Partenaire » et l'autre « Robot » (voir tableau 4). Ils répondent tous deux « Partenaire » à 35,76 %, et tous deux « Robot » à 26,39 %. La différence entre deux réponses « Partenaire » et deux réponses « Robot » est significative (test de Student : $t_{(22)} = -2.5206$; $p = 0.0195$). En observant les réponses « Partenaire » et « Robot » données par condition, on retrouve que les sujets répondent plus souvent différemment en conditions Distale (37,50 %) et Proximale Vibratoire (43,06 %). En condition Proximale Visuelle, ils répondent plus souvent tous les deux « Partenaire » (à 44,44 %). La différence entre deux réponses « Partenaire » et deux réponses « Robot » est significative pour les deux conditions Proximales (en Visuel, test de Student : $t_{(10)} = 3.5777$; $p = 0.0083$, et en Vibratoire, test de Student : $t_{(10)} = 7.0502$; $p = 3.5606$), mais pas pour la condition Distale.

Tableau 4. Nombre moyen de réponses « Partenaire » ou « Robot » par binôme de manière générale et par condition.

	Deux réponses « Partenaire »	Une réponse « Partenaire », une réponse « Robot »	Deux réponses « Robot »
Général	35,76 %	37,85 %	26,39 %
Condition Distale	30,56 %	37,50 %	31,94 %
Condition Proximale Visuelle	44,44 %	33,33 %	22,22 %
Condition Proximale Vibratoire	37,50 %	43,06 %	19,44 %

Les différences entre les modalités de réponses « Partenaire » ou « Robot » par binôme sont statistiquement significatives ($F_{(2, 33)} = 6.2253$; $p = 0.0051$). En comparant les modalités de réponses deux à deux, on constate que la différence entre les réponses par binôme deux réponses « Partenaire » et aucune réponse « Partenaire » est significative (test de Student : $t_{(22)} = 2.7943$; $p = 0.0109$), ainsi que la différence entre les réponses par binôme une réponse « Partenaire » et aucune réponse « Partenaire » (test de Student : $t_{(22)} = 3.1044$; $p = 0.0052$).

Dans la condition Distale, les sujets répondent en binôme plus souvent différemment que deux fois « Partenaire » ($d = 6,94$ %) et que deux fois « Robot » ($d = 5,56$ %), mais ces différences ne sont pas significatives.

Dans la condition Proximale Visuelle, les sujets répondent en binôme plus souvent deux fois « Partenaire » que différemment ($d = 11,11$ %) et que deux fois « Robot » ($d = 22,2$ %). La différence globale entre ces trois types de réponses est significative (test de Fischer : $F_{(2, 15)} = 4.6154$; $p = 0.0274$). En les comparant deux à deux, la seule différence significative est celle entre deux réponses « Partenaire » et aucune réponse « Partenaire » (test de Student : $t_{(10)} = 3.5777$; $p = 0.0083$).

Dans la condition Proximale Vibratoire, les sujets répondent en binôme plus souvent différemment que deux fois « Partenaire » ($d = 5,56$ %) et que deux fois « Robot » ($d = 23,61$ %). La différence globale entre ces trois types de réponses est significative (test de Fischer : $F_{(2, 15)} = 25.7609$; $p = 1.4075 \times 10^{-5}$). En les comparant deux à deux, les différences significatives sont celles entre deux réponses « Partenaire » et aucune réponse « Partenaire » (test de Student : $t_{(10)} = 7.0502$; $p = 3.5606 \times 10^{-5}$) et celle entre une réponse « Partenaire » et aucune réponse « Partenaire » (test de Student : $t_{(10)} = 6.2493$; $p = 0.0003$).

En ce qui concerne les résultats obtenus par binôme en fonction du type d'interactant selon les conditions, nous nous intéressons aux fréquences des doubles bonnes réponses (voir figure 89). Pour toutes les conditions, les doubles bonnes réponses sont plus fréquentes face à l'autre sujet (reconnaissance mutuelle) que face au robot ; pour la condition Distale $d = 7\%$, pour la condition Proximale Visuelle $d = 33\%$ et pour la condition Proximale Vibratoire $d = 25\%$. Les différences sont les plus importantes pour les conditions Proximales, ce qui indique que ces conditions sont donc plus efficaces pour réussir la tâche face au partenaire que face au robot. En valeur globale, c'est la condition Proximale Visuelle qui permet la plus haute fréquence de reconnaissance mutuelle (bonnes réponses mutuelles face à face), avec une valeur de 67%.

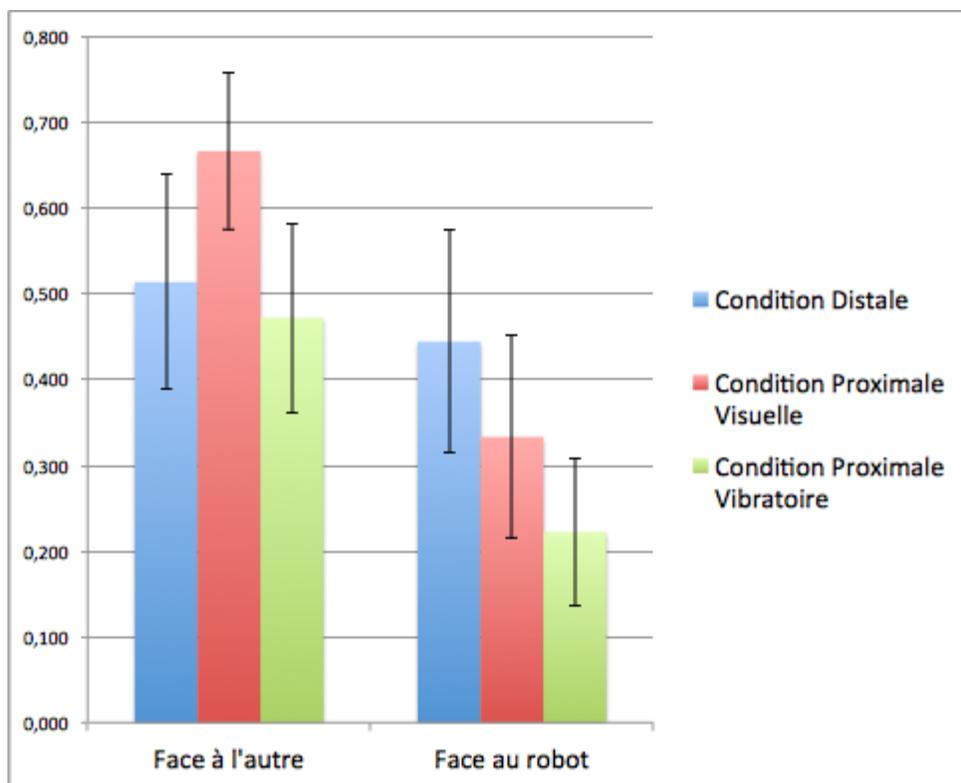


Figure 89. Fréquence de doubles bonnes réponses en fonction du type d'interactant et des conditions.

Nous allons à présent nous intéresser aux réponses données par le deuxième sujet sachant la réponse du premier sujet. Nous allons prendre en considération d'une part si les réponses données sont bonnes ou mauvaises, et si ce sont des réponses « Partenaire » ou « Robot ». Nous allons de plus faire la différence entre la situation de face à face, et celle face au robot. On constate que, de manière globale, lorsque le sujet A a donné une bonne réponse, dans 70,65 %

des cas, le sujet B aussi fournit une bonne réponse (voir tableau 5.a). Dans la condition Distale, cela arrive dans 72,83 % des cas (voir tableau 5.a), contre 75,13 % pour la condition Proximale Visuelle (voir tableau 5.b) et 62,19 % pour la condition Proximale Vibratoire (voir tableau 5.b). Lorsque le sujet A répond mal, le sujet B ne fournira une bonne réponse que dans 52,25 % des cas (voir tableau 5.a). Dans la condition Distale, cela arrivera à 53,83 % des cas (voir tableau 5.a) ; dans la condition Proximale Visuelle, 49,65 % des cas (voir tableau 5.b), et enfin 48,57 % des cas pour la condition Proximale Vibratoire (voir tableau 5.b). Face à face, une bonne réponse du sujet A entraîne une bonne réponse du sujet B dans 78,82 % des cas (contre 61,29 % lorsque les sujets sont face au robot, et donc n'interagissent pas entre eux ; voir tableau 5.a). Face à face, une mauvaise réponse du sujet A n'entraîne une bonne réponse du sujet B que dans 46,56 % des cas (contre 56,55 % des cas face au robot ; voir tableau 5.a). On constate donc une différence notable des réponses du deuxième sujet en fonction de la réponse du premier, en particulier lorsque les deux sujets sont face à face, ce qui appuie le rôle de la dynamique de co-ajustement mis en place par les sujets (les sujets se cherchent mutuellement, là où le robot ne réagit pas en temps réel à ce que fait le sujet).

Tableau 5.a. Proportions des types de réponses collectives en fonction de la valeur bonne ou mauvaise de la réponse, de la valeur donnée (« Partenaire » ou « Robot »), selon le type d'interactant, de manière générale et pour chaque interactant (global et condition Distale).

		Global			Condition Distale		
Réponse du sujet A	Réponse du sujet B	Total	Face à face	Face au robot	Total	Face à face	Face au robot
Reconnaissance	Reconnaissance	70,65%	78,82%	61,29%	72,83%	78,60%	67,19%
	Non-reconnaissance	29,35%	21,18%	38,71%	27,17%	21,40%	32,81%
Non-reconnaissance	Reconnaissance	52,25%	46,56%	56,55%	53,83%	40,86%	67,75%
	Non-reconnaissance	47,75%	53,44%	43,45%	46,17%	59,14%	32,25%
« Partenaire »	« Partenaire »	65,48%	78,82%	43,45%	62,02%	78,60%	32,25%
	« Robot »	34,52%	21,18%	56,55%	37,98%	21,40%	67,75%
« Robot »	« Partenaire »	41,66%	46,56%	38,71%	36,94%	40,86%	32,81%
	« Robot »	58,34%	53,44%	61,29%	63,06%	59,14%	67,19%

Tableau 5.b. Proportions des types de réponses collectives en fonction de la valeur bonne ou mauvaise de la réponse, de la valeur donnée (« Partenaire » ou « Robot »), selon le type d'interactant, de manière générale et pour chaque interactant (conditions Proximale Visuelle et Proximale Vibratoire).

Réponse du sujet A	Réponse du sujet B	Condition Proximale Visuelle			Condition Proximale Vibratoire		
		Total	Face à face	Face au robot	Total	Face à face	Face au robot
Reconnaissance	Reconnaissance	75,13%	85,82%	61,38%	62,19%	72,64%	47,72%
	Non-reconnaissance	24,87%	14,18%	38,62%	37,81%	27,36%	52,28%
Non-reconnaissance	Reconnaissance	49,65%	49,21%	48,18%	48,57%	51,30%	46,78%
	Non-reconnaissance	50,35%	50,79%	51,82%	51,43%	48,70%	53,22%
« Partenaire »	« Partenaire »	73,33%	85,82%	51,82%	63,54%	72,64%	53,22%
	« Robot »	26,67%	14,18%	48,18%	36,46%	27,36%	46,78%
« Robot »	« Partenaire »	41,67%	49,21%	38,62%	52,53%	51,30%	52,28%
	« Robot »	58,33%	50,79%	61,38%	47,47%	48,70%	47,72%

Lorsque l'on compare les deux blocs d'essais en fonction du type d'interactant, on constate que, face à l'autre sujet, les fréquences des réponses par binômes font état d'un effet d'apprentissage à deux au cours du passage de la condition. En effet, la fréquence des bonnes réponses mutuelles augmente ($d = 0,14$), atteignant les 61% dans le deuxième bloc d'essais, tandis que les fréquences des bonnes réponses solitaires et des mauvaises réponses mutuelles diminuent (respectivement, $d = 0,11$ et $d = 0,03$) (voir figure 90).

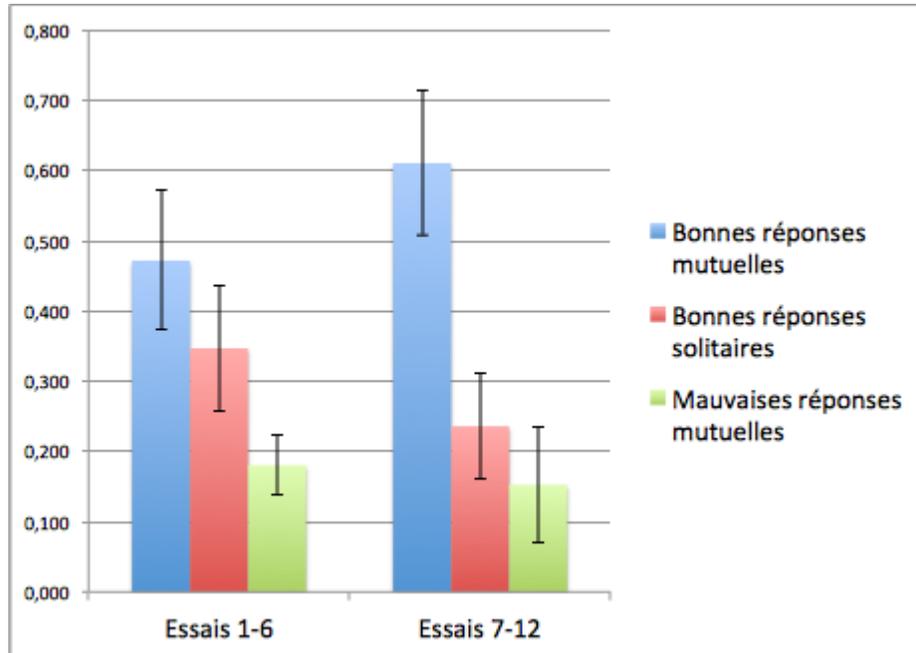


Figure 90. Fréquence des trois modalités de réponses entre les deux blocs d'essais face au partenaire.

Face au robot cependant, on constate que l'on ne retrouve pas cet effet d'apprentissage. En effet, la fréquence des doubles bonnes réponses reste constante, tandis que la fréquence des simples bonnes réponses diminue ($d = 0,07$) et que celle des doubles mauvaises réponses augmente ($d = 0,07$) (voir figure 91).

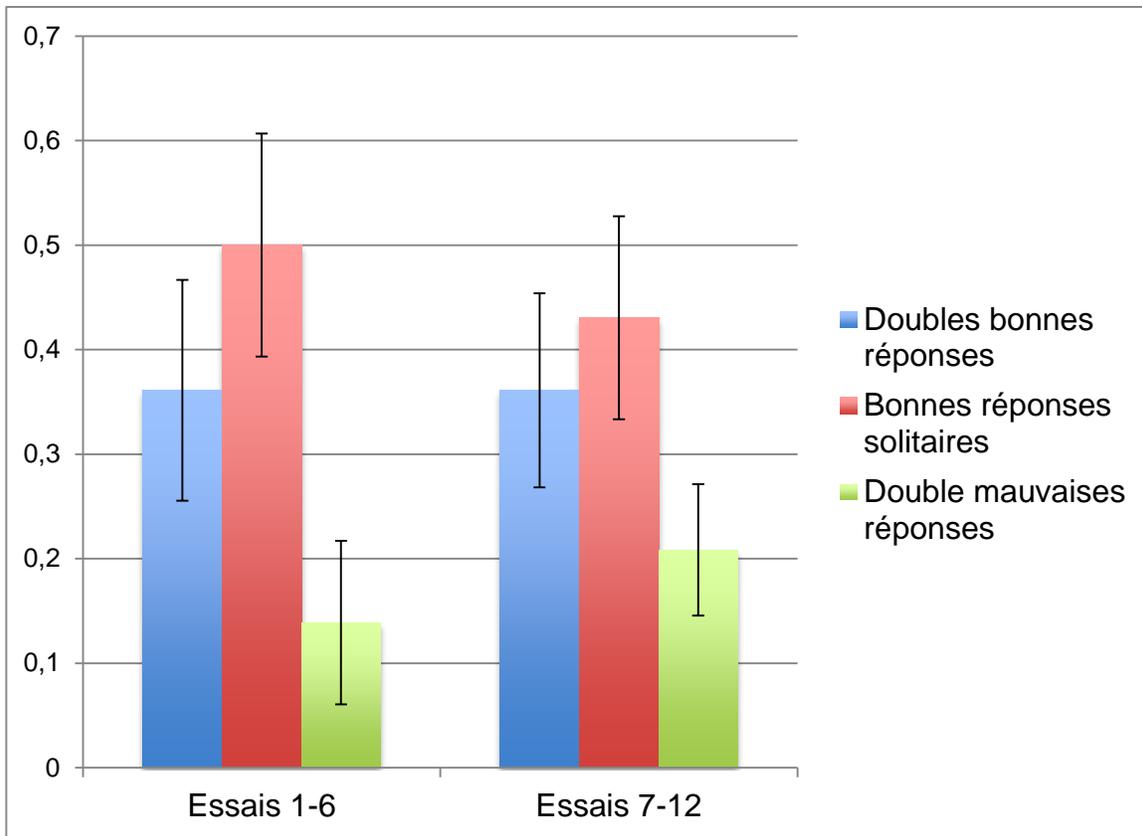


Figure 91. Fréquence des trois modalités de réponses entre les deux blocs d’essais face au robot.

Enfin, en ce qui concerne les réponses par binôme en fonction du type d’interactant, nous nous intéressons aux résultats par condition expérimentale. Face au partenaire, toutes les conditions voient la fréquence de bonnes réponses mutuelles augmenter (voir figure 92). Cette augmentation est la plus forte en condition Proximale Visuelle ($d = 0,22$) ; la fréquence des bonnes réponses mutuelles atteint les 78% dans cette condition pendant le deuxième bloc d’essais. En conditions Distale et Proximale Vibratoire, l’augmentation est plus faible (respectivement, $d = 0,14$ et $d = 0,06$). C’est de plus dans la condition Proximale Visuelle que la fréquence de bonnes réponses mutuelles est la plus élevée dans les deux blocs d’essais. C’est la condition dans laquelle les sujets apprennent le mieux à se reconnaître, et se reconnaissent le mieux.

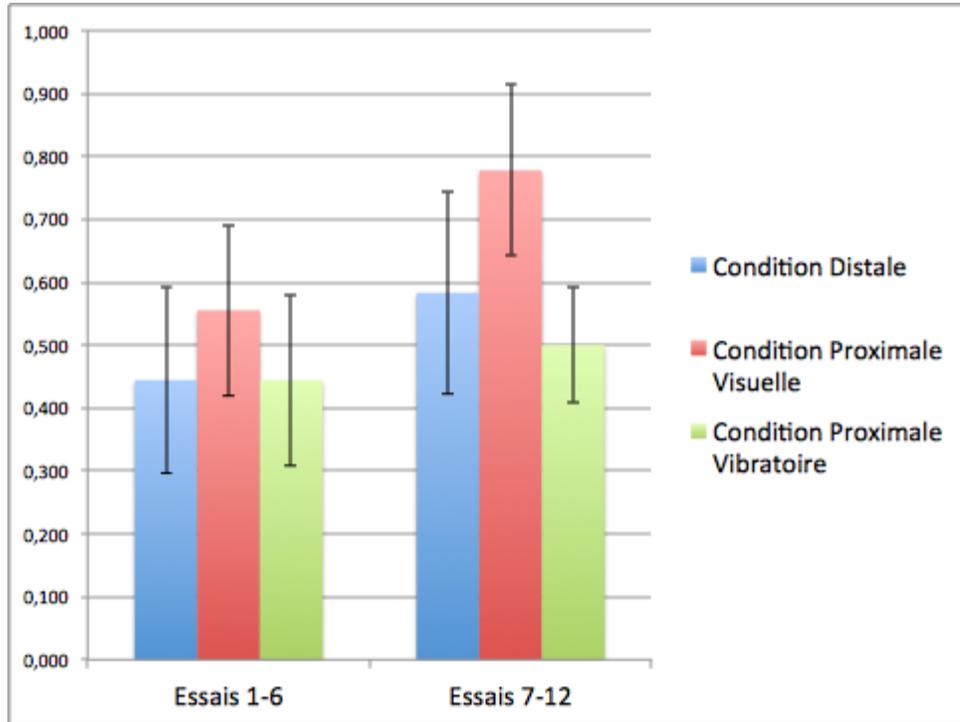


Figure 92. Fréquence de bonnes réponses mutuelles par condition face au partenaire pour les deux blocs d’essais.

Face au robot, on ne constate aucun effet d’apprentissage entre les deux blocs d’essais par condition. En effet, la fréquence des doubles bonnes réponses reste constante entre les deux blocs d’essais pour toutes les conditions (voir figure 93). Les doubles bonnes réponses sont les plus fréquentes en condition Distale.

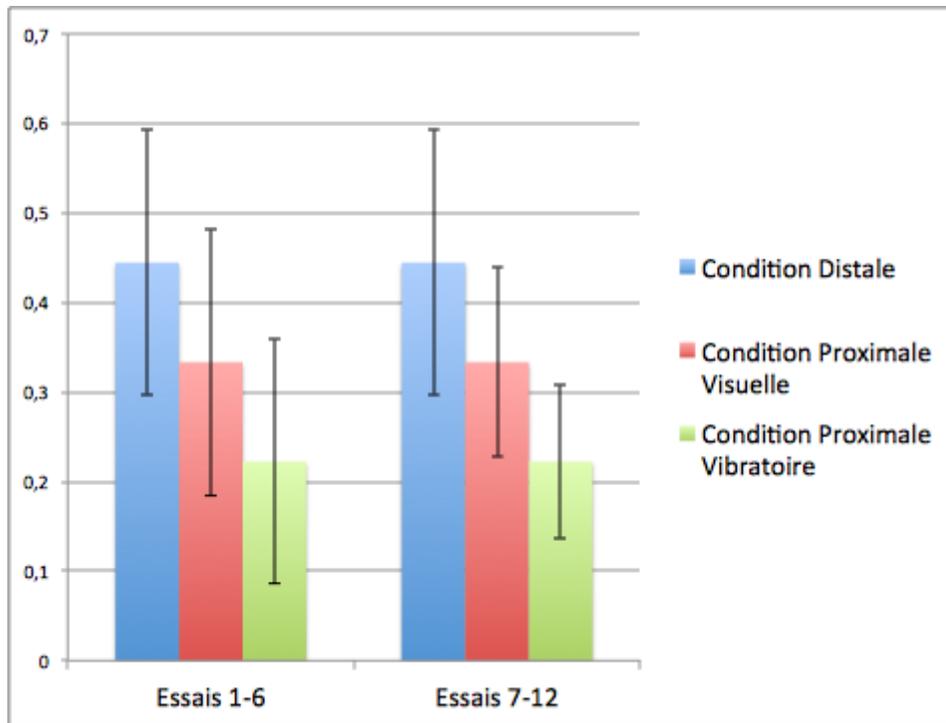


Figure 93. Fréquence de doubles bonnes réponses par condition face au robot pour les deux blocs d'essais.

2.4.5.3 Analyse des trajectoires

Parallèlement à l'analyse des réponses des sujets, nous avons procédé à une analyse des trajectoires perceptives déployées par les sujets lors de chaque essai. Ces analyses sont rendues possibles grâce à la méthodologie minimaliste que nous avons adoptée, ainsi qu'à la médiation technique (ici l'application Touch Through) qui a permis l'enregistrement de ces trajectoires. Ce type d'analyse va nous permettre d'expliquer les résultats obtenus, en mettant en évidence les critères qui ont amené les sujets à répondre « Partenaire » ou à répondre « Robot », que ce soit à tort ou à raison. Il s'agit de comprendre ce qui motive les réponses, bonnes ou mauvaises, des sujets. Le dispositif présente donc l'intérêt de pouvoir rentrer dans l'intimité perceptive des sujets pour comprendre ce qui les pousse à répondre ce qu'ils répondent. Nous allons pouvoir ainsi proposer quelques éléments d'explication.

D'une façon générale, nous avons analysé plusieurs indicateurs, de façon à tenter de comprendre les réponses données par les sujets et de caractériser leurs trajectoires perceptives. Il s'agit de faire le lien entre la valeur des indicateurs et la nature de la réponse donnée par le sujet. D'un point de vue méthodologique, nous avons procédé d'une part à une prise en compte des retours donnés par les utilisateurs pendant la séquence de débriefing de l'expérience, et d'autre

part à une analyse qualitative des trajectoires à partir de leur représentation graphique²⁹. Elles nous ont permis de poser un certain nombre d'hypothèses, de façon à guider une analyse quantitative des indicateurs qui nous semblaient pertinents. Le choix de différents indicateurs permet alors de procéder à une analyse quantitative systématique. La richesse des données a conduit à analyser de très nombreux indicateurs, mais par souci de clarté, nous ne relaterons ici que les résultats principaux.

Deux axes distincts ont guidé notre analyse : d'une part l'interaction des sujets avec les deux types d'interactants (Partenaire vs Robot), et d'autre part l'interaction entre les sujets en fonction de la nature de leurs réponses individuelles (Reconnaissance vs Non-Reconnaissance et « Partenaire » vs « Robot ») et collectives (doubles bonnes réponses, simple bonne réponse, doubles mauvaises réponses, à distinguer dans les cas où les sujets étaient face à face et les cas où ils étaient face au robot).

1.1.1.1.1 Interaction des sujets avec les deux types d'interactant

Tout d'abord, nous nous sommes intéressés aux périodes que les sujets ont passées en contact avec l'interactant (soit l'autre sujet, soit le robot). Pour chaque essai, nous avons considéré que les séquences d'interaction étaient définies par les séquences pendant lesquelles les écarts entre les stimulations sont inférieurs à deux secondes. Les durées sont exprimées en millisecondes, et les distances en pourcentages (100% correspond à la diagonale de l'écran, soit 1101 pixels ; cela permet de rapporter la distance parcourue à la taille de l'espace, même si évidemment les sujets n'ont que très rarement suivi le parcours de la diagonale).

Nous avons tout d'abord observé la distance parcourue à deux pendant l'essai ; c'est la somme totale des distances parcourues lors des séquences d'interaction par les deux interactants. De manière individuelle, les sujets ont parcouru plus de distance lors des séquences d'interaction avec l'autre interactant lorsqu'ils étaient face à face que lorsqu'ils étaient face au robot ($d = 87,57$) (voir figure 94). Cette différence est significative (test de Student : $t_{(46)} = 2.8403$; $p = 0.0067$). Cette différence se retrouve dans les trois conditions expérimentales. C'est dans la condition Proximale Visuelle que l'écart est le plus marqué ($d = 141,71$) et c'est la seule condition où la différence est significative (test de Student : $t_{(22)} = 3.3845$; $p = 0.0035$) ; l'écart est moins important dans la condition Distale ($d = 84,11$) et dans la condition Proximale Vibratoire ($d = 40,89$). En valeur totale, c'est aussi en condition Proximale Visuelle que les sujets parcourent le plus de distance lors des séquences d'interaction avec le partenaire (plus de 4

²⁹ Pour plus de détails concernant la méthode d'analyse utilisée pour l'étude des trajectoires, veuillez vous reporter à l'Annexe B.

fois la diagonale de l'écran au total). On notera de plus que la distance parcourue lors des séquences d'interaction avec le robot est toujours sensiblement la même (de manière globale, 254,62%, avec 245,83% pour la condition Distale, 266,52% pour la condition Proximale Visuelle et 260,32% pour la condition Proximale Vibratoire).

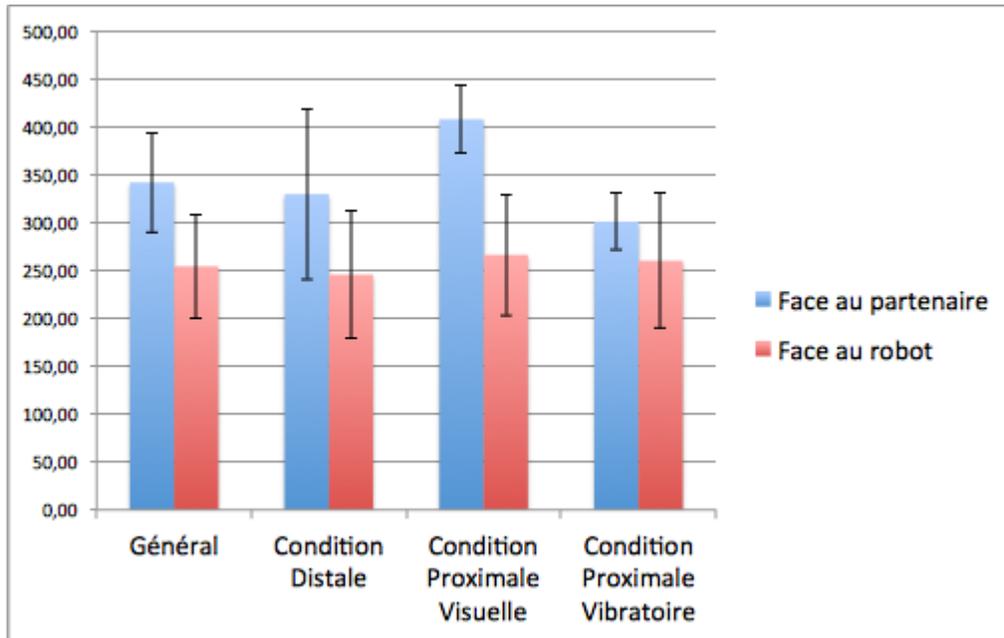


Figure 94. Moyennes de la distance totale parcourue par un sujet avec l'autre interactant en fonction des conditions expérimentales et du type d'interactant.

Nous allons regarder cet indicateur en fonction de la réponse donnée (« Partenaire » ou « Robot ») par les sujets, pour voir s'ils se basent dessus pour fournir leur réponse. En comparant la valeur de cet indicateur selon que la réponse donnée soit « Partenaire » ou « Robot », on constate que de manière globale, lorsque les sujets répondent « Partenaire », la distance totale parcourue est supérieure comparée à quand ils répondent « Robot » ($d = 74,49\%$) (voir tableau 6). On retrouve ce résultat dans toutes les conditions. La différence globale est significative (test de Student : $t_{(46)} = 2.3176$; $p = 0.0253$), ainsi que celle dans la condition Distale, qui vaut $96,23\%$ (test de Student : $t_{(46)} = 2.0414$; $p = 0.0471$).

Tableau 6. Valeur de l'indicateur Distance totale parcourue avec l'interactant selon que la réponse d'un sujet soit « Partenaire » ou « Robot », de manière générale et pour chaque condition.

	Réponse « Partenaire »	Réponse « Robot »
Général	337,39	262,90
Condition Distale	332,01	235,79
Condition Proximale Visuelle	363,75	318,70
Condition Proximale Vibratoire	279,44	267,40

En mettant en évidence la corrélation entre le pourcentage de réponses « Partenaire » et l'indicateur Distance totale parcourue avec l'interactant, on constate que cette corrélation avec la réponse « Partenaire » est la plus forte pour la condition Distale (voir figure 95.a), puis pour la condition Proximale Vibratoire (voir figure 95.b), puis enfin pour la condition Proximale Visuelle (voir figure 95.c).

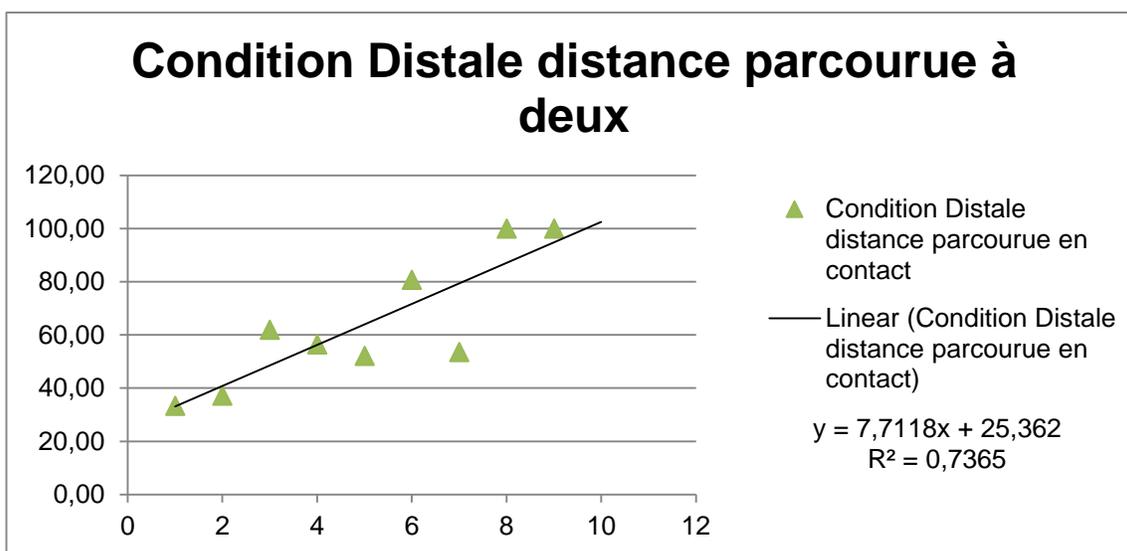


Figure 95.a. Corrélation entre le pourcentage de réponses « Partenaire » et la distance totale parcourue avec l'interactant : condition Distale.

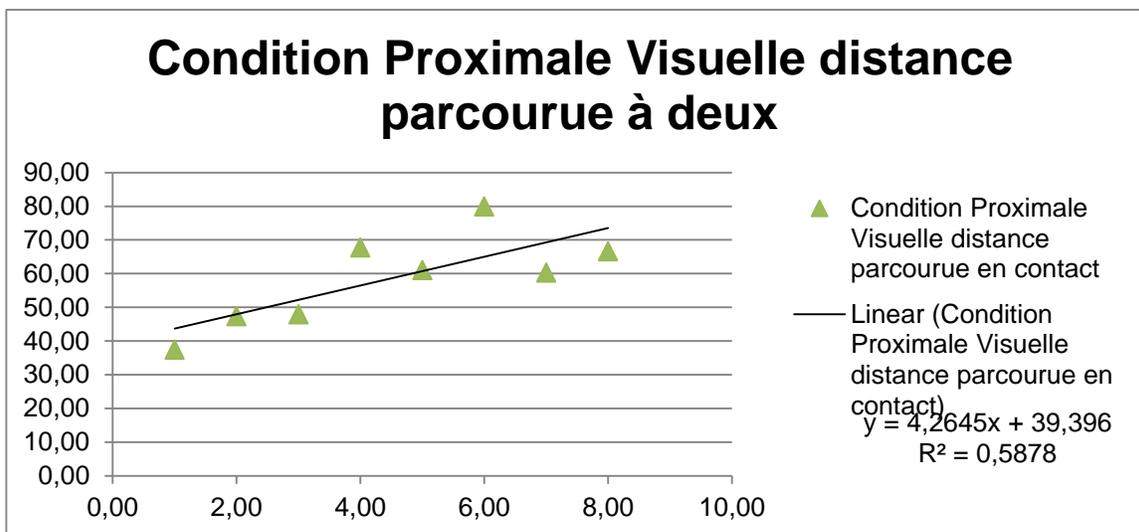


Figure 95.b. Corrélation entre le pourcentage de réponses « Partenaire » et la distance totale parcourue avec l'interactant : condition Proximale Visuelle.

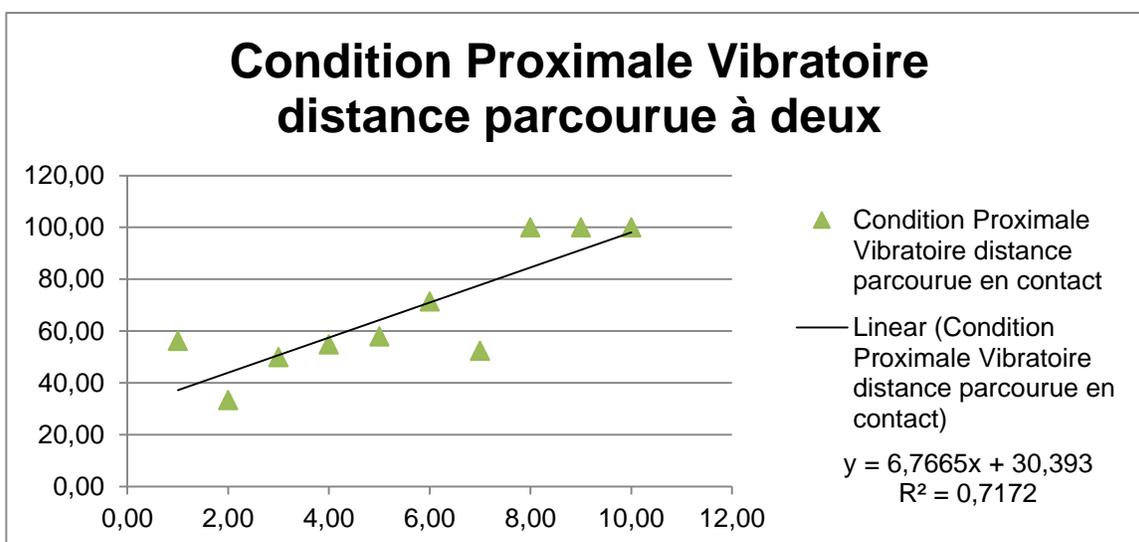


Figure 95.c. Corrélation entre le pourcentage de réponses « Partenaire » et la distance totale parcourue avec l'interactant : condition Proximale Vibratoire.

Ainsi, de manière générale les sujets associent une grande distance parcourue à deux à la présence de leur partenaire, et ils l'associent plus particulièrement en condition Distale, bien que cette dernière soit la condition où il est le plus facile de parcourir une grande distance totale (étant donné que l'on sait à chaque instant où se trouve l'interactant), et que ce soit dans la condition Proximale Visuelle qu'une grande distance totale parcourue en séquence d'interaction est synonyme du partenaire dans les faits.

Le deuxième indicateur que nous avons pris en considération est le temps total des séquences d'interaction pendant l'essai. Nous avons donc mesuré combien de temps, sur les 30 secondes de chaque essai, les sujets se trouvaient en séquence d'interaction avec l'interactant. De manière générale, les sujets restent en moyenne plus longtemps en séquence d'interaction avec le partenaire qu'avec le robot ($d = 3926$ ms) (voir figure 96). Ce résultat se retrouve à nouveau dans les trois conditions expérimentales. À nouveau, la différence la plus grande entre partenaire et robot se retrouve dans la condition Proximale Visuelle ($d = 5999$ ms), suivie par celle dans la condition Distale ($d = 2823$ ms) puis celle dans la condition Proximale Vibratoire ($d = 3934$ ms). La valeur la plus forte se retrouve à nouveau dans la condition Proximale Visuelle (22556 ms, soit plus des deux tiers de l'essai). Les différences sont significatives de manière globale (test de Student : $t_{(46)} = 2.8058$; $p = 0.0073$) et dans la condition Proximale Visuelle (test de Student : $t_{(22)} = 3.5113$; $p = 0.0028$).

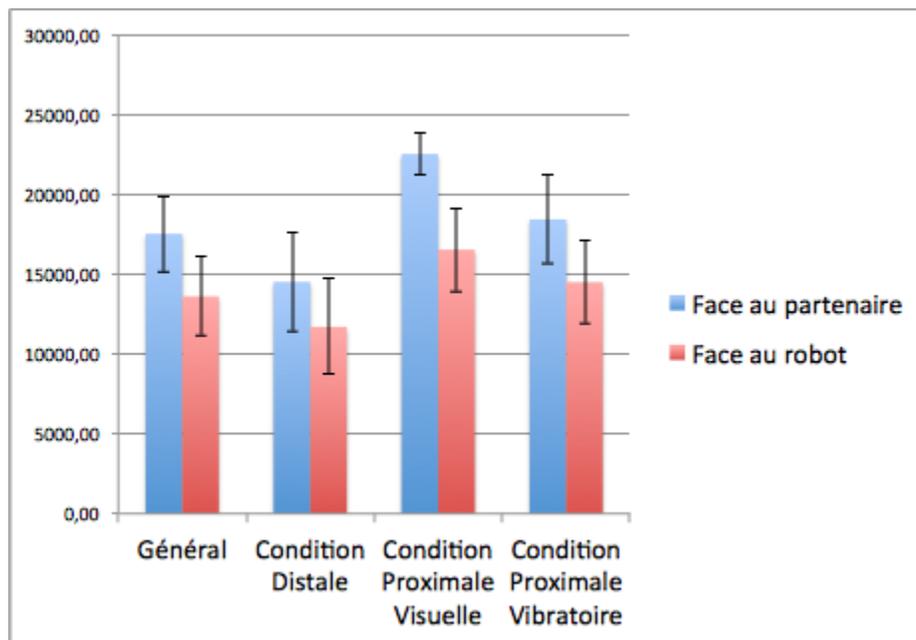


Figure 96. Moyennes en millisecondes du temps total passé pour un sujet en séquence d'interaction avec l'autre interactant en fonction des conditions expérimentales et du type d'interactant.

L'indicateur du temps total passé en séquence d'interaction est donc un indicateur de la présence du partenaire : globalement, plus ce temps est long, plus l'interactant est probablement l'autre sujet et non le robot. Cela est particulièrement vrai pour la condition Proximale Visuelle. Regardons à présent si les sujets s'appuient sur cet indicateur pour fournir leur réponse ; si c'est le cas, on observera une différence de valeur de l'indicateur selon que les sujets répondent

« Partenaire » ou « Robot ». De manière générale, la valeur de l'indicateur est supérieure lorsque les sujets répondent « Partenaire » par rapport à quand ils répondent « Robot » ($d = 3471,41$ ms) (voir tableau 7). La différence est significative (test de Student : $t_{(46)} = 2.2786$; $p = 0.0274$). Les sujets utilisent donc de manière générale cet indicateur pour donner leur réponse (un temps long entraînant une réponse « Partenaire »). On observe que l'on retrouve une valeur de l'indicateur supérieure en cas de réponse « Partenaire » par rapport au cas de réponse « Robot », ce qui indique une tendance des sujets à prendre en compte cet indicateur pour répondre ; cependant, la différence n'est significative dans aucune des conditions.

Tableau 7. Valeur de l'indicateur Temps total passé en séquence d'interaction selon que la réponse d'un sujet soit « Partenaire » ou « Robot », de manière générale et pour chaque condition.

	Réponse « Partenaire »	Réponse « Robot »
Général	17226,98	13809,57345
Condition Distale	14733,41	11259,67873
Condition Proximale Visuelle	20547,33	18251,65417
Condition Proximale Vibratoire	17091,88	16079,27254

La corrélation entre le pourcentage de réponses « Partenaire » et l'indicateur temps total passé en séquence d'interaction se révèle être la plus forte pour la condition Proximale Visuelle (voir figure 97.b), puis pour la condition Proximale Vibratoire (voir figure 97.c), puis enfin pour la condition Distale (voir figure 97.a.).

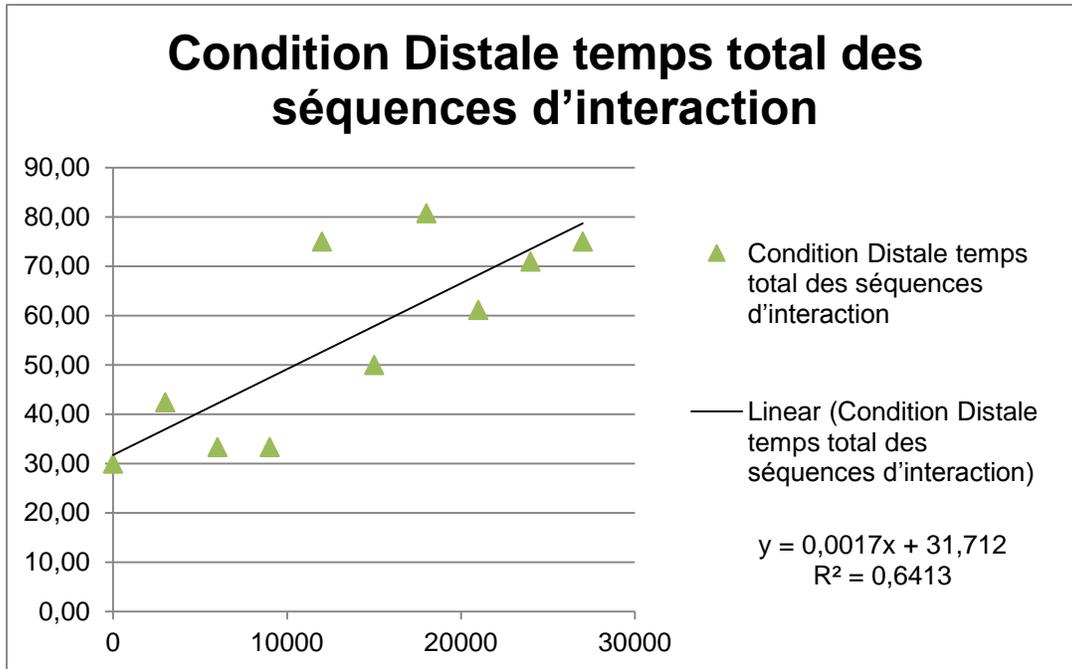


Figure 97.a. Corrélation entre le pourcentage de réponses « Partenaire » et le temps total passé en séquence d'interaction : condition Distale.

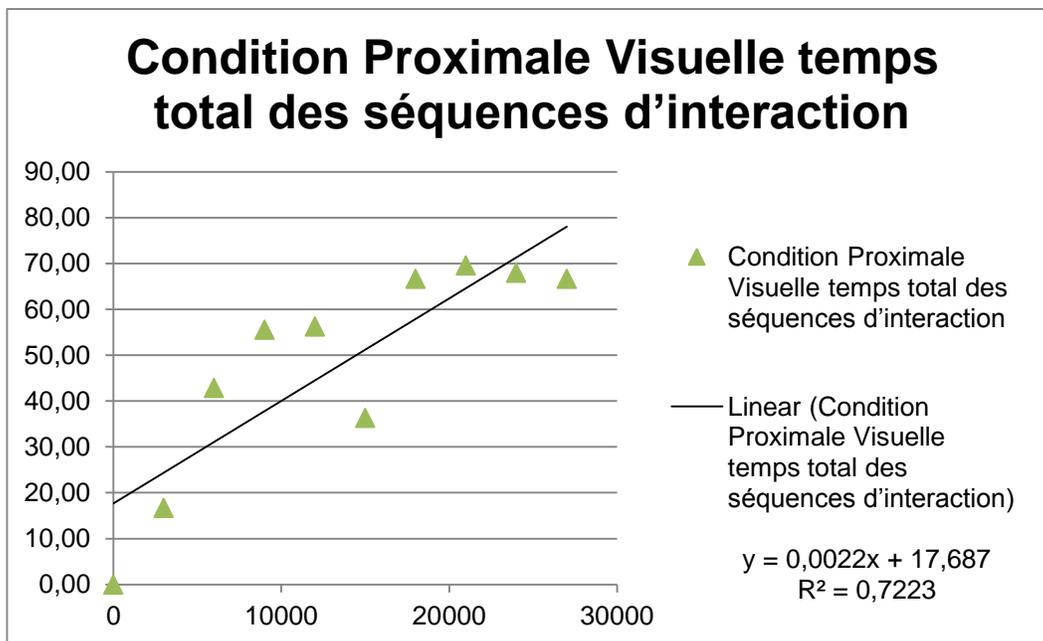


Figure 97.b. Corrélation entre le pourcentage de réponses « Partenaire » et le temps total passé en séquence d'interaction : condition Proximale Visuelle.

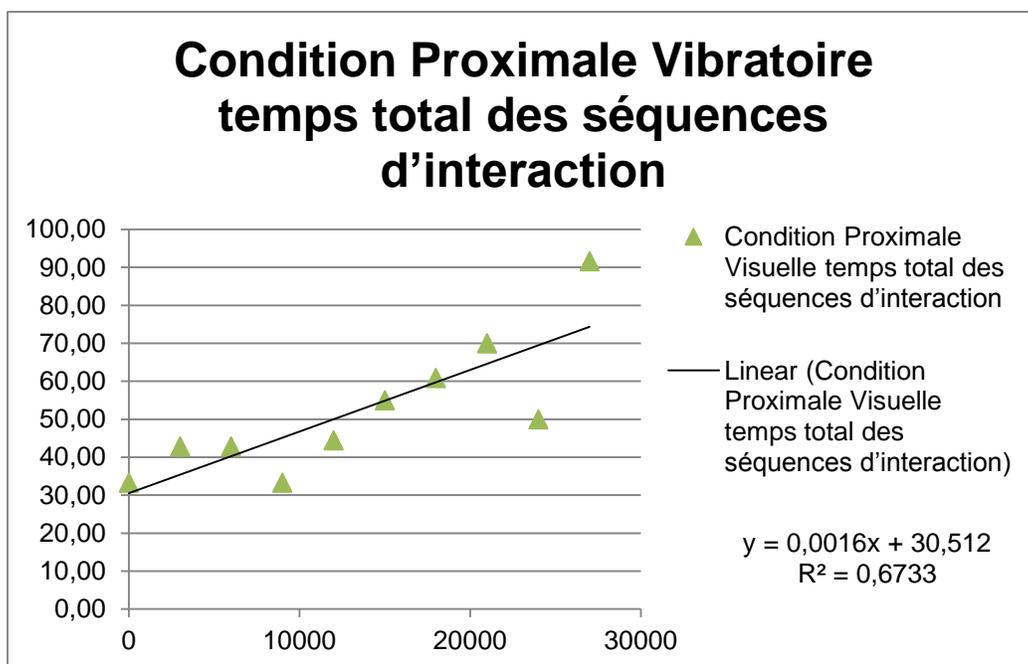


Figure 97.c. Corrélation entre le pourcentage de réponses « Partenaire » et le temps total passé en séquence d'interaction : condition Proximale Vibratoire.

Le troisième indicateur retenu concernait le nombre de séquences d'interaction ayant eu lieu pendant un essai. De manière générale, il y avait plus de séquences d'interaction lors des essais face au robot que lors de ceux face au partenaire ($d = 0,54$) (voir figure 98). La différence est significative (test de Student : $t_{(46)} = -2.1962$; $p = 0.0332$). Elle se retrouve dans les trois conditions expérimentales, avec un écart plus marqué dans la condition Proximale Visuelle ($d = 0,79$) et dans la condition Proximale Vibratoire ($d = 0,71$) que dans la condition Distale ($d = 0,33$). Les différences ne sont cependant pas significatives. Il est à noter que le plus petit nombre de séquences d'interaction, que ce soit face au partenaire ou au robot, a lieu en condition Proximale Visuelle (en moyenne, respectivement 2,08 et 2,88 séquences d'interaction), tandis que le plus grand nombre de séquences d'interaction a lieu en condition Distale (3,44 et 3,76 séquences d'interaction respectivement).

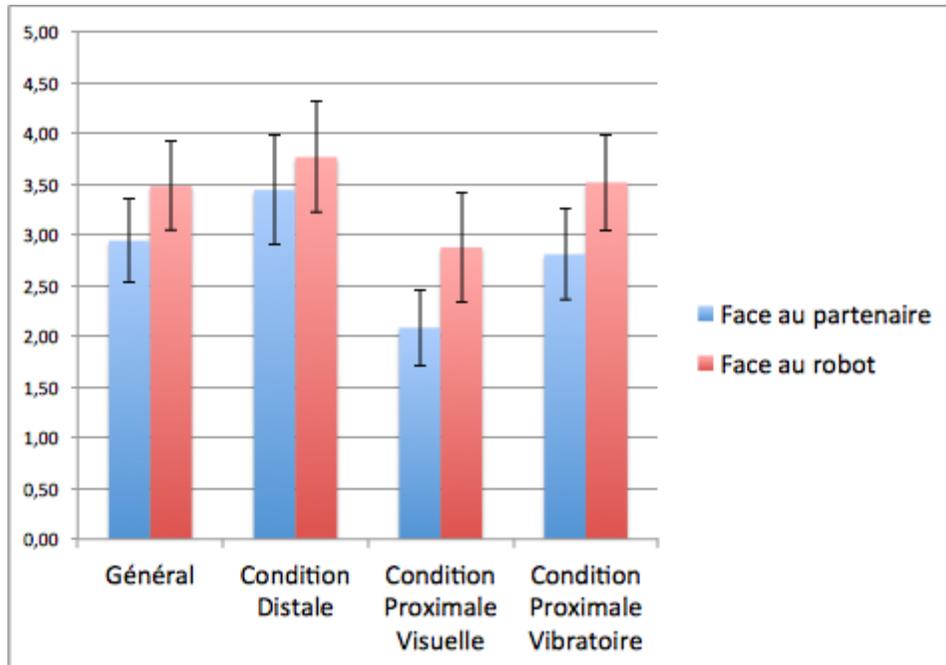


Figure 98. Moyennes du nombre de séquences d'interaction en fonction des conditions expérimentales et du type d'interactant.

Le nombre de séquences d'interaction est donc un indicateur à considérer ; plus le nombre est grand, plus le sujet a de chances d'interagir avec le robot. En effet, face au partenaire, les deux sujets cherchent à maintenir le contact, ce qui entraîne un nombre de séquences d'interaction plus petit : en se cherchant mutuellement, on peut facilement se retrouver en moins de deux secondes. Pour vérifier si les sujets prennent en compte cet indicateur pour répondre, il s'agit de vérifier les réponses données (« Partenaire » ou « Robot ») par les sujets. De manière générale, la valeur du nombre de séquences d'interaction est inférieure lorsque les sujets répondent « Partenaire » par rapport à quand ils répondent « Robot » ($d = 0,52$) (voir tableau 8). La différence n'est cependant pas significative pour un seuil de 5 % (test de Student : $t_{(46)} = -1.9548$; $p = 0.0568$). La différence se retrouve dans toutes les conditions ; pour la condition Distale, $d = 0,19$; pour la condition Proximale Visuelle, $d = 0,70$; enfin, pour la condition Proximale Vibratoire, $d = 0,59$. On observe que l'on retrouve une valeur de l'indicateur inférieure en cas de réponse « Partenaire » par rapport au cas de réponse « Robot », ce qui indique une tendance des sujets à prendre en compte cet indicateur pour répondre ; cependant, la différence n'est significative dans aucune des conditions.

Tableau 8. Valeur de l'indicateur Nombre total de séquences d'interaction selon que la réponse d'un sujet soit « Partenaire » ou « Robot », de manière générale et pour chaque condition.

	Réponse « Partenaire »	Réponse « Robot »
Général	3,02	3,54
Condition Distale	3,59	3,78
Condition Proximale Visuelle	2,32	3,02
Condition Proximale Vibratoire	2,90	3,49

La corrélation entre le pourcentage de réponses « Partenaire » et l'indicateur nombre total de séquences d'interaction ne se révèle intéressante qu'en condition Distale (voir figure 99.a) et en condition Proximale Vibratoire (voir figure 99.b). Si l'indicateur se révèle très important pour la réponse en condition Distale, il est intéressant de voir que son effet, bien que présent, est bien plus faible en condition Proximale Vibratoire, et absent en condition Proximale Visuelle.

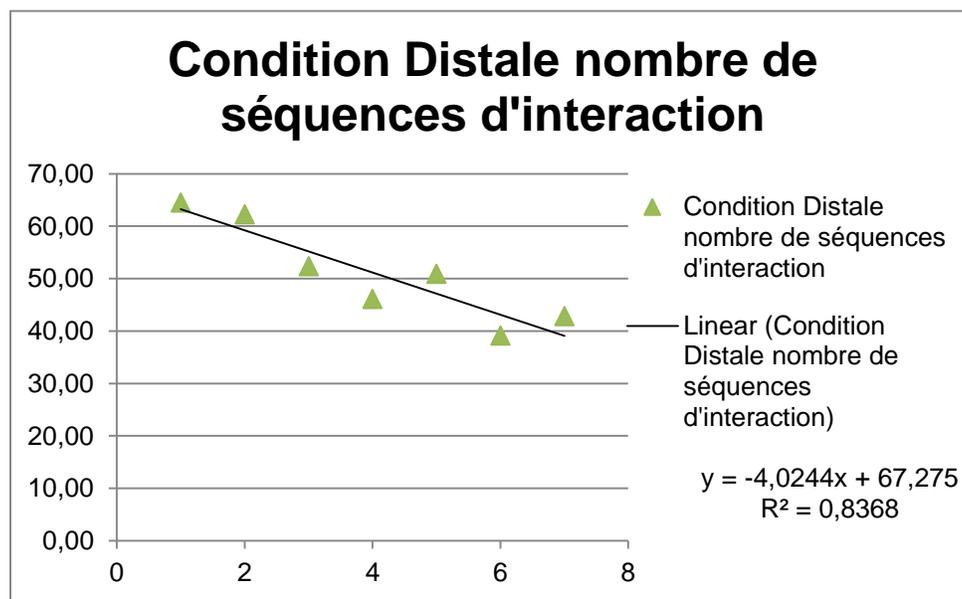


Figure 99.a. Corrélation entre le pourcentage de réponses « Partenaire » et le nombre total de séquences d'interaction : condition Distale.

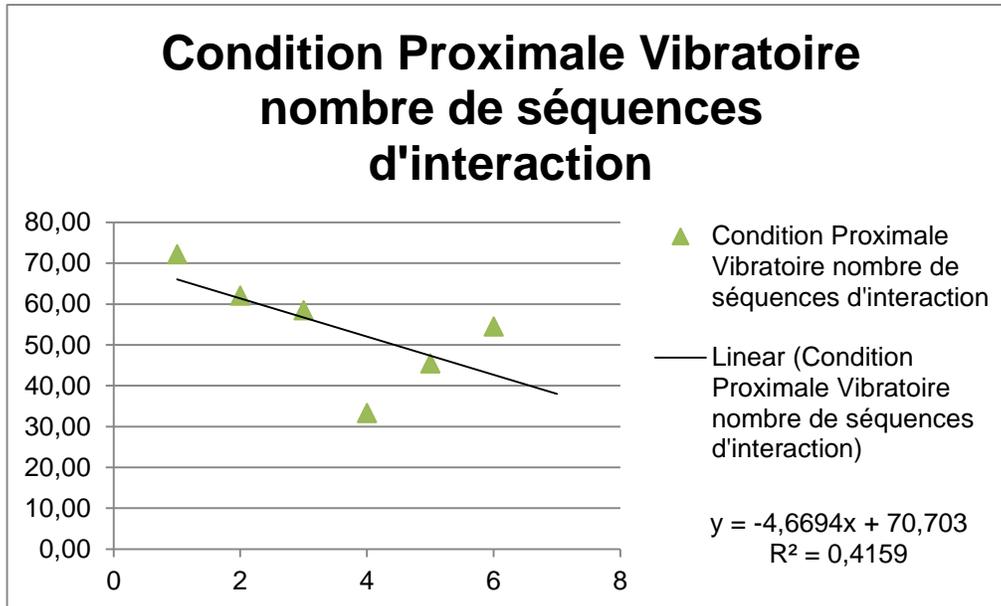


Figure 99.b. Corrélation entre le pourcentage de réponses « Partenaire » et le nombre total de séquences d'interaction : condition Proximale Vibratoire.

Enfin, le quatrième indicateur permet une nouvelle approche des deux précédents ; il s'agit de la durée moyenne d'une séquence d'interaction. On constate que le temps moyen d'une séquence d'interaction est plus long face au partenaire que face au robot ($d = 4161$ ms) (voir figure 100). La différence est significative (test de Student : $t_{(46)} = 2.8812$; $p = 0.0061$). Elle se retrouve dans toutes les conditions expérimentales : c'est dans la condition Proximale Visuelle que la différence est la plus grande ($d = 8457$ ms) (elle est significative, test de Student : $t_{(22)} = 3.1135$; $p = 0.0051$), suivi de la condition Proximale Vibratoire ($d = 4013$ ms) et enfin de la condition Distale ($d = 2038$ ms). En valeur totale face au partenaire, c'est en condition Proximale Visuelle qu'une séquence d'interaction est la plus longue : 16892 ms, contre 10151 en condition Proximale Vibratoire et 6379 en condition Distale.

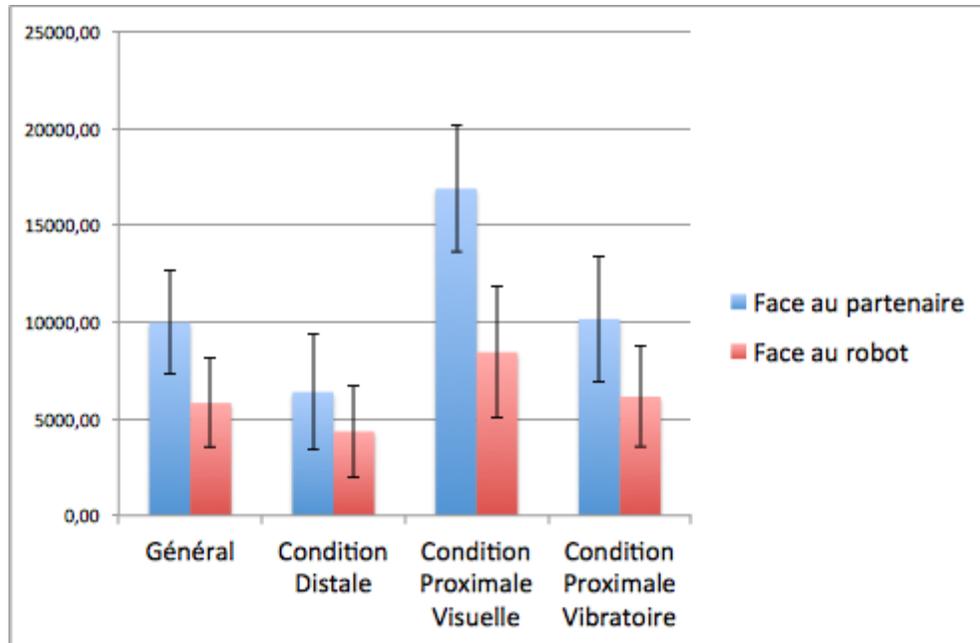


Figure 100. Moyennes du temps moyen d'une séquence d'interaction en fonction des conditions expérimentales et du type d'interactant.

Le temps moyen d'une séquence d'interaction est donc un indicateur pertinent : plus ce temps est long, plus la probabilité de se trouver face au partenaire est grande. Pour savoir si les sujets s'appuient sur cet indicateur pour donner leur réponse, nous allons observer sa valeur selon que les sujets répondent « Partenaire » ou « Robot ». De manière générale, on observe que le temps moyen d'une séquence d'interaction est plus long lorsque les sujets répondent « Partenaire » que lorsqu'ils répondent « Robot » ($d = 2686,13$ ms) (voir tableau 9). À nouveau, la différence n'est pas significative pour un seuil de 5 % (test de Student : $t_{(46)} = 1.8357$; $p = 0.0730$). On retrouve la différence dans toutes les conditions ; pour la condition Distale, $d = 1916,03$ ms ; pour la condition Proximale Visuelle, $d = 2555,16$ ms ; et enfin, pour la condition Proximale Vibratoire, $d = 1736,80$ ms. La différence n'est significative pour aucune des conditions. Cet indicateur n'est donc pas non plus significativement utilisé par les sujets pour répondre, mais on constate malgré tout une tendance à le prendre en considération.

Tableau 9. Valeur de l'indicateur Temps moyen d'une séquence d'interaction selon que la réponse d'un sujet soit « Partenaire » ou « Robot », de manière globale et pour chaque condition.

	Réponse « Partenaire »	Réponse « Robot »
Général	9096,03	6409,91
Condition Distale	6129,42	4213,39
Condition Proximale Visuelle	13100,75	10545,59
Condition Proximale Vibratoire	9334,35	7597,55

La corrélation entre le pourcentage de réponses « Partenaire » et l'indicateur temps moyen d'une séquence d'interaction ne se retrouve donc pas vraiment ; c'est dans la condition Proximale Visuelle qu'il est le plus marqué (voir figure 101). Cependant il n'est pas étonnant que cet indicateur soit invisible, car il nécessite un calcul et n'est pas directement accessible pour les sujets.

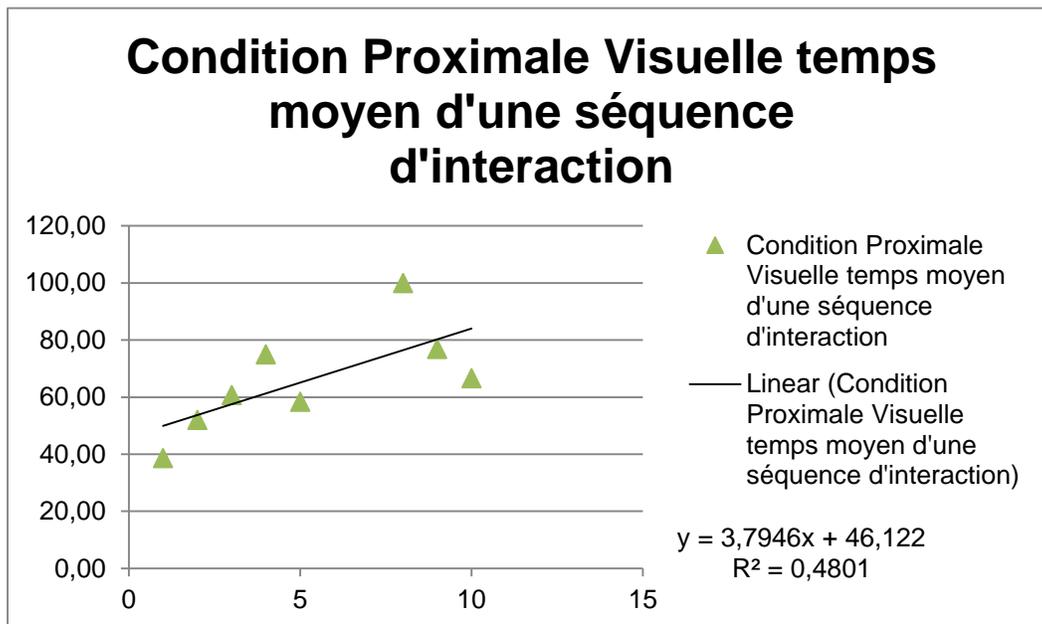


Figure 101. Corrélation entre le pourcentage de réponses « Partenaire » et le temps moyen d'une séquence d'interaction : condition Proximale Visuelle.

En conclusion, les différents indicateurs retenus sont tous un indice de la présence du partenaire, mais ce sont la distance totale parcourue à deux et le temps total des séquences d'interaction qui influent le plus sur la réponse des sujets.

2.4.5.3.1 Interaction entre les sujets en fonction de la qualité des réponses

Nous allons à présent considérer les indicateurs sous le jour de la qualité des réponses données par les utilisateurs, tout d'abord à un niveau individuel (Reconnaissance vs Non-reconnaissance) puis à un niveau collectif (doubles bonnes réponses, simples bonnes réponses, doubles mauvaises réponses).

2.4.5.3.1.1 Au niveau individuel

Commençons par les résultats individuels. Tout d'abord, la distance parcourue lors des séquences d'interaction avec l'autre interactant varie peu selon la qualité des réponses individuelles. En effet, de manière générale la différence entre la distance parcourue lors d'une bonne réponse et lors d'une mauvaise réponse est faible ($d = 1,04$) (voir figure 102). Ce résultat se retrouve dans les trois conditions expérimentales ; pour la condition Distale, $d = 24,62$, tandis que pour les conditions Proximales, $d = -13,56$ en Visuel, et $d = -10,52$ en Vibratoire. Il est intéressant de noter que dans les conditions Proximales, la distance parcourue lors des séquences d'interaction avec l'autre interactant est plus longue dans le cas d'une mauvaise réponse ; cela s'explique par le fait que la durée des séquences d'interaction est un critère de reconnaissance du partenaire pour les sujets.

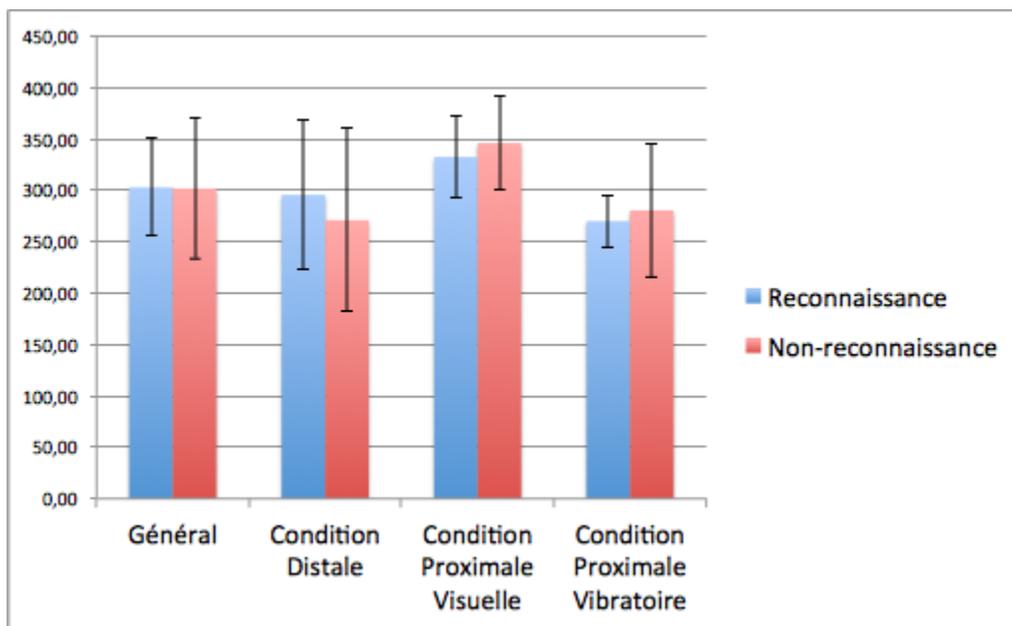


Figure 102. Moyennes de la distance parcourue lors des séquences d'interaction avec l'autre interactant en fonction de la qualité des réponses individuelles.

Ensuite, le temps total des séquences d'interaction varie lui aussi peu selon la qualité des réponses individuelles. En effet, de manière générale, la différence de temps total des séquences d'interaction selon une bonne ou une mauvaise réponse individuelle est de 518 ms (voir figure 103). Pour la condition Distale, cette différence est de 797 ms, quand elle est de 155 ms pour la condition Proximale Visuelle et de 1395 ms pour la condition Proximale Vibratoire.

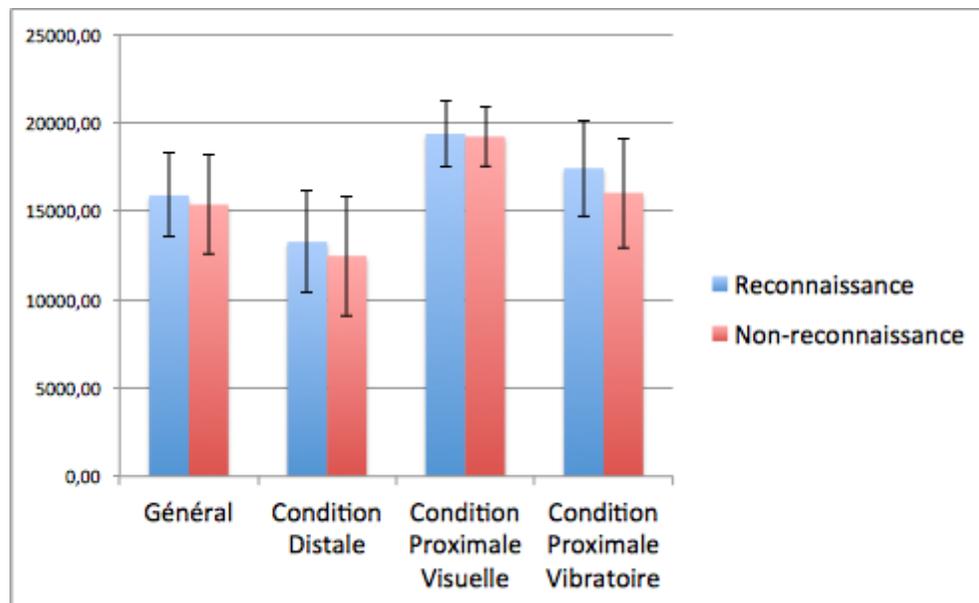


Figure 103. Moyennes du temps total des séquences d'interaction avec l'autre interactant en fonction de la qualité des réponses individuelles.

En ce qui concerne le nombre de séquences d'interaction, on constate qu'il est plus élevé dans le cas d'une mauvaise réponse individuelle que dans le cas d'une bonne réponse individuelle ($d = 0,19$) (voir figure 104). Ce résultat se retrouve dans les trois conditions expérimentales : $d = 0,11$ en condition Distale, $d = 0,44$ en condition Proximale Visuelle et $d = 0,12$ en condition Proximale Vibratoire.

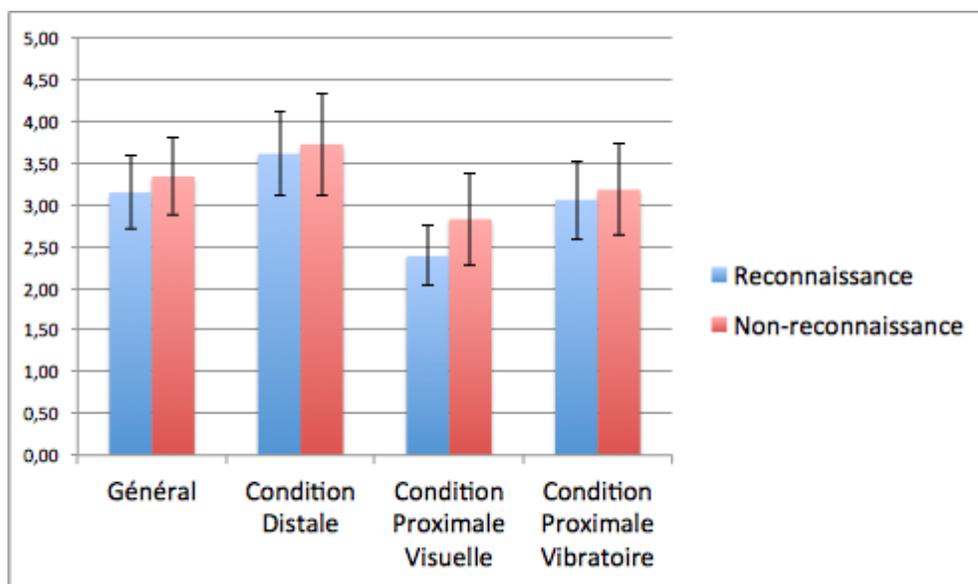


Figure 104. Moyennes du nombre de séquences d'interaction avec l'autre interactant en fonction de la qualité des réponses individuelles.

Enfin, le temps moyen d'une séquence d'interaction en fonction des réponses individuelles est généralement plus long en cas de reconnaissance qu'en cas de non-reconnaissance ($d = 918$ ms) (voir figure 105). Ce résultat se retrouve dans toutes les conditions expérimentales : si la différence n'est pas très importante en condition Distale ($d = 477$) et en condition Proximale Vibratoire ($d = 450$), elle l'est plus en condition Proximale Visuelle ($d = 2526$). On observe que c'est en condition Proximale Visuelle que cette durée est la plus importante : en cas de reconnaissance, elle s'élève à 13043 ms, contre 8783 ms en condition Proximale Vibratoire et 5309 ms en condition Distale.

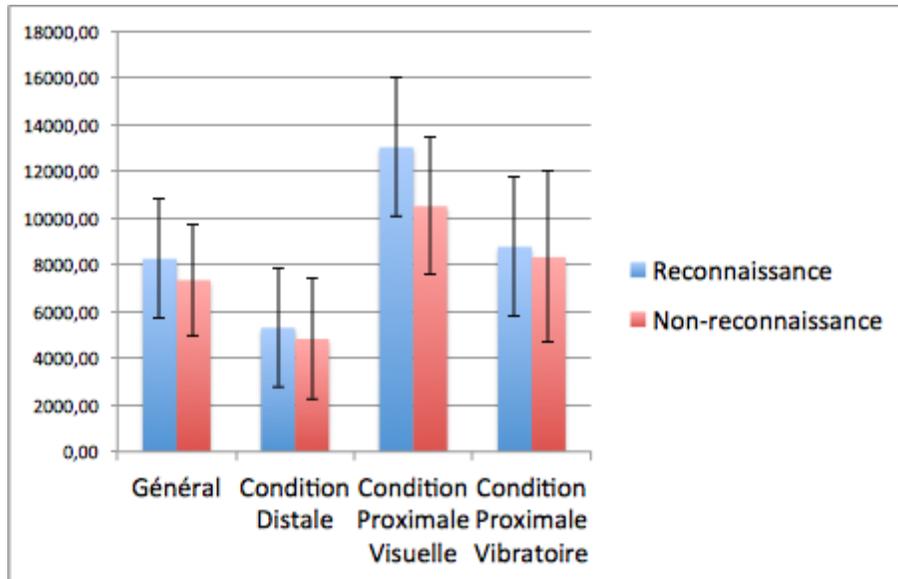


Figure 105. Moyennes du temps moyen d'une séquence d'interaction entre le sujet et l'autre interactant en fonction de la qualité des réponses individuelles.

2.4.5.3.1.2 Au niveau collectif

À présent, nous nous intéressons aux indicateurs en fonction des deux réponses des binômes. Pour la distance totale parcourue lors des séquences d'interaction, on constate que de manière générale, elle est plus élevée dans le cas de deux bonnes réponses que dans le cas d'une seule bonne réponse ($d = 40,73$) et de deux mauvaises réponses ($d = 40,16$) (voir figure 106). Ce résultat se trouve au niveau des conditions expérimentales, excepté pour la condition Proximale Visuelle, où la plus grande distance parcourue lors des séquences d'interaction est trouvée dans le cas des doubles mauvaises réponses, avant les doubles bonnes réponses ($d = 23,15$) et des simples bonnes réponses ($d = 102,05$).

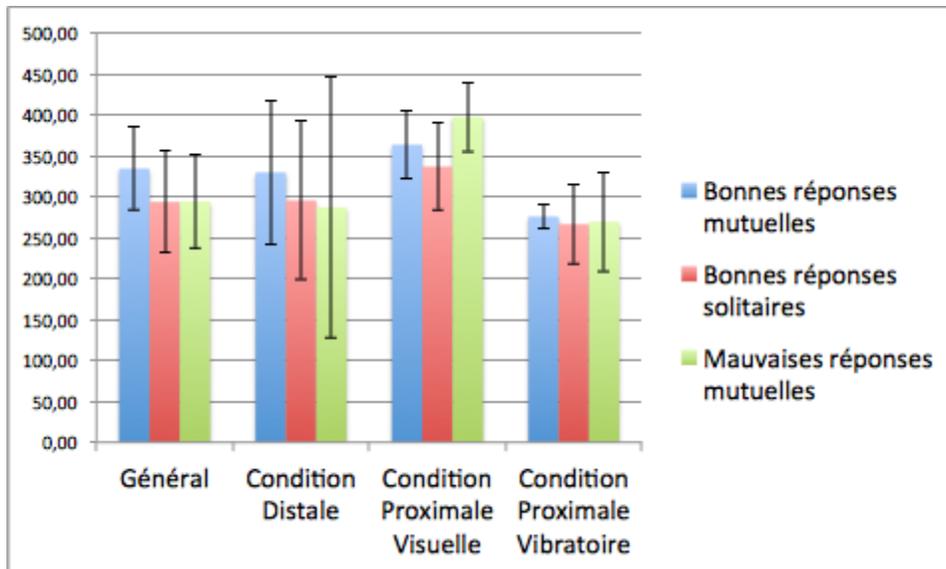


Figure 106. Moyennes de la distance parcourue lors des séquences d'interaction avec l'autre interactant en fonction des conditions expérimentales et du type de réponse collective.

Lorsque l'on regarde les résultats en fonction du type de réponse et du type d'interactant, on constate que dans tous les cas de réponses, la distance totale parcourue lors des séquences d'interaction avec le partenaire est plus élevée que la distance totale parcourue lors des séquences d'interaction avec le robot (voir figure 107). Pour les doubles bonnes réponses, $d = 152,48$ (la différence est significative, test de Student : $t_{(22)} = 3.0161$; $p = 0.0065$), pour les simples bonnes réponses $d = 91,21$ et pour les doubles mauvaises réponses $d = 48,06$. En valeur totale, la plus longue distance parcourue lors des séquences d'interaction avec le partenaire a lieu dans le cas des bonnes réponses mutuelles (375,32%).

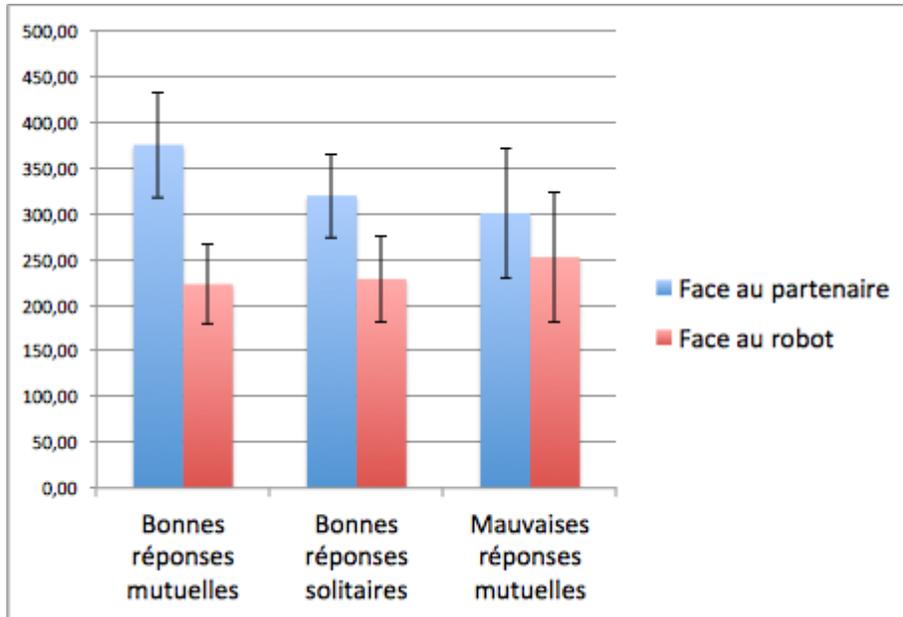


Figure 107. Moyennes de la distance parcourue lors des séquences d'interaction avec l'autre interactant en fonction du type d'interactant et de la qualité des réponses collectives.

En ce qui concerne le temps total des séquences d'interaction, on constate que le type de réponses le plus propice à une plus longue durée des séquences d'interaction est celui des doubles bonnes réponses (voir figure 108). Cela se retrouve dans toutes les conditions expérimentales. En valeur totale, c'est en condition proximale visuelle que les sujets passent le plus de temps en séquence d'interaction (20675 ms dans le cas des doubles bonnes réponses).

Lorsque l'on s'intéresse aux résultats en fonction du type d'interactant, on trouve des résultats très intéressants. Dans le cas des doubles bonnes réponses, la durée des séquences d'interaction face au partenaire est bien plus longue que face au robot ($d = 7285$ ms) (voir figure 109). La différence est significative (test de Student : $t_{(22)} = 3.4936$; $p = 0.0021$). Dans le cas des simples bonnes réponses, cette différence est moins marquée ($d = 4081$ ms), jusqu'à pratiquement disparaître dans le cas des doubles mauvaises réponses ($d = 22$ ms).

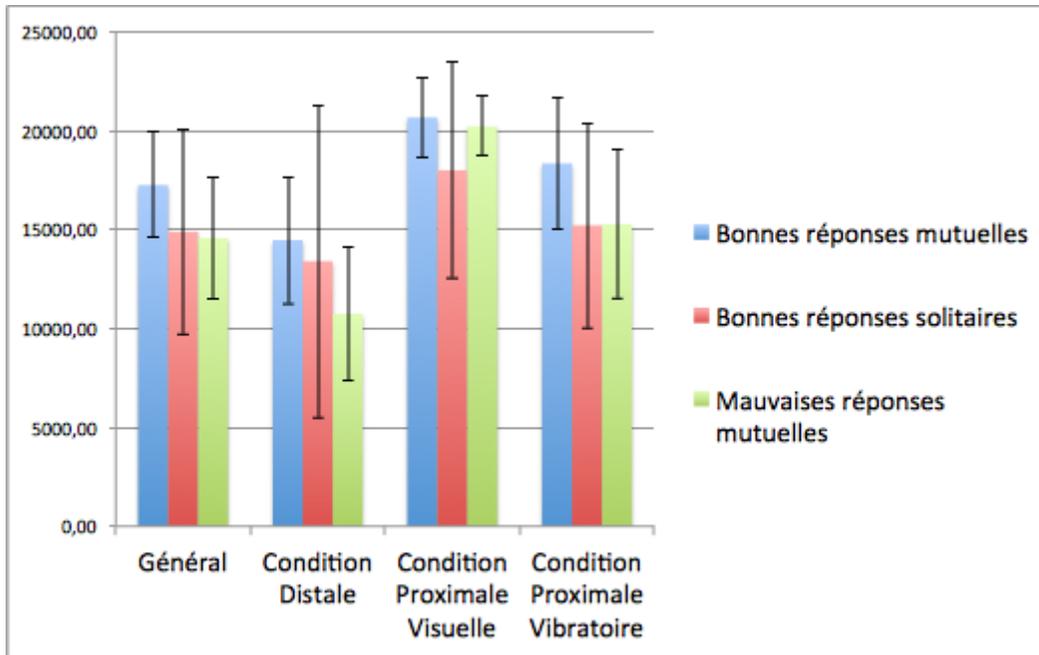


Figure 108. Moyennes du temps total des séquences d'interaction avec l'autre interactant en fonction des conditions expérimentales et du type de réponse collective.

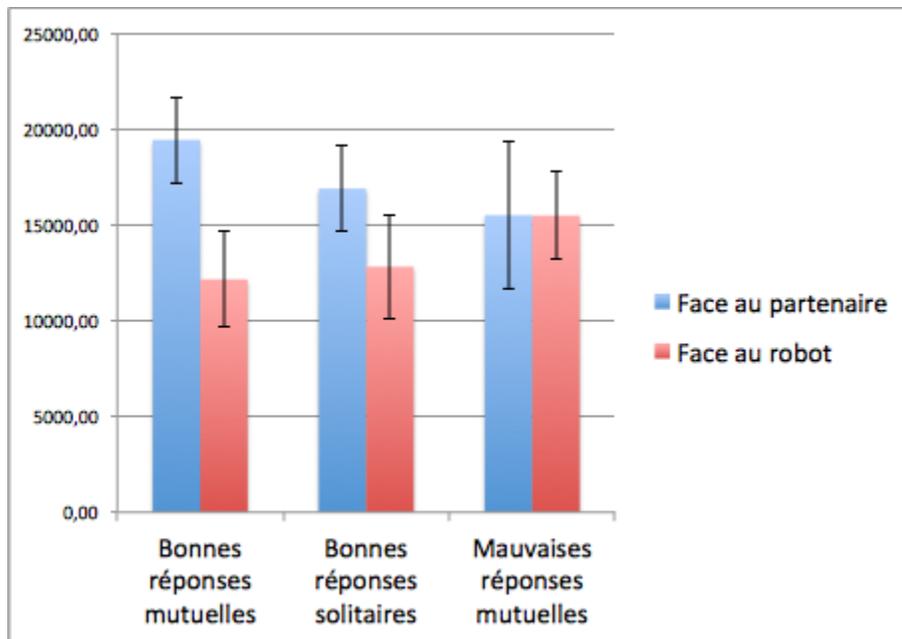


Figure 109. Moyennes du temps total des séquences d'interaction avec l'autre interactant en fonction du type d'interactant et de la qualité des réponses collectives.

Le nombre de séquences d'interaction avec l'autre interactant est, de manière générale, plus faible dans le cas des doubles bonnes réponses que dans les autres cas de réponses (voir figure 110). Ce résultat se retrouve dans toutes les conditions expérimentales, et l'on notera aussi que le nombre de séquences d'interaction est pratiquement équivalent dans les cas de simples bonnes réponses et de doubles mauvaises réponses (en global, $d = 0,04$, en condition Distale $d = -0,03$, en condition Proximale Visuelle $d = 0,04$ et en condition Proximale Vibratoire $d = 0,19$).

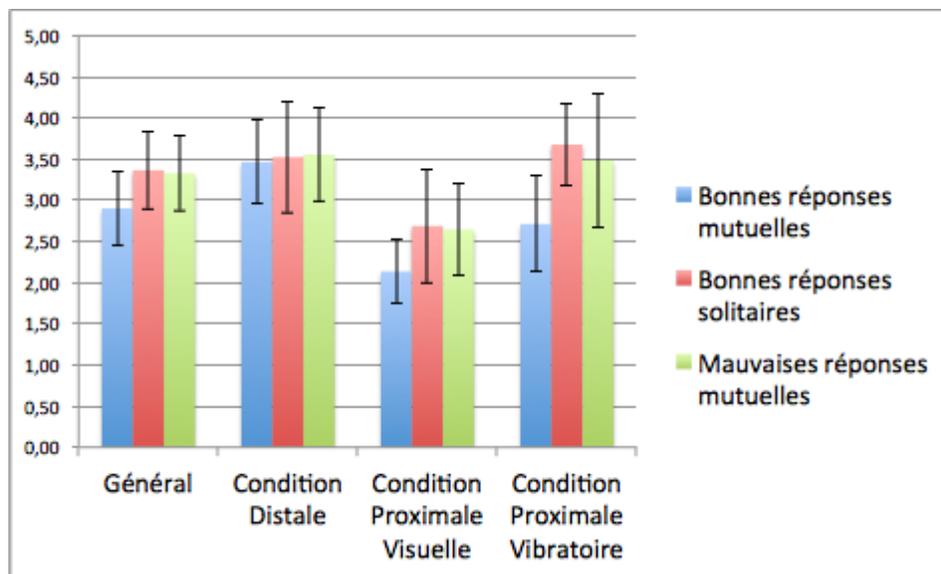


Figure 110. Moyennes du nombre de séquences d'interaction avec l'autre interactant en fonction des conditions expérimentales et du type de réponse collective.

Lorsque l'on s'intéresse aussi au type d'interactant, on constate que dans le cas des doubles bonnes réponses, le nombre de séquences d'interaction est plus faible face au partenaire que face au robot ($d = 1,37$) (voir figure 111). La différence est significative ($t_{(22)} = -3.1157$; $p = 0.0058$). La différence est bien moins marquée dans le cas des simples bonnes réponses ($d = 0,22$), et est même inversée dans le cas des doubles mauvaises réponses ($d = -0,62$).

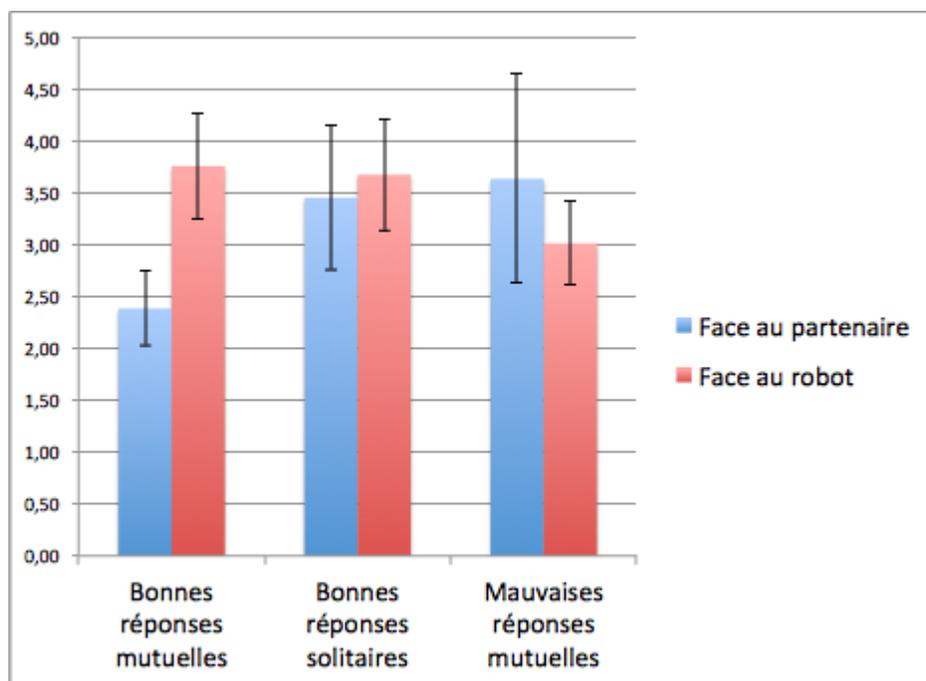


Figure 111. Moyennes du nombre de séquences d'interaction avec l'autre interactant en fonction du type d'interactant et de la qualité des réponses collectives.

Enfin, le temps moyen d'une séquence d'interaction en fonction des réponses collectives est plus important dans le cas des doubles bonnes réponses que lors de simples bonnes réponses ($d = 3034$ ms) et de doubles mauvaises réponses ($d = 2876$ ms) (voir figure 112). Ce résultat se retrouve dans les deux conditions Proximales : en Proximale Visuelle, on obtient respectivement $d = 4197$ ms et $d = 3801$ ms, et en condition Proximale Vibratoire $d = 6256$ ms et $d = 2532$ ms. En condition Distale cependant, la durée moyenne d'une séquence d'interaction est comparable en cas de doubles bonnes réponses et de simples bonnes réponses ($d = -258$ ms). En cas de doubles mauvaises réponses, il y a alors une durée moindre (avec les bonnes réponses mutuelles, $d = 2695$ ms).

En considérant le type d'interactant, on trouve que le temps moyen d'une séquence d'interaction est plus long face au partenaire que face au robot dans toutes les modalités de réponses (voir figure 113). Cependant cette différence est bien plus marquée en cas de doubles bonnes réponses ($d = 8647$ ms ; la différence est significative, test de Student : $t_{(22)} = 4.1323$; $p = 0.0005$) qu'en cas de simples bonnes réponses ($d = 2854$ ms) et de doubles mauvaises réponses ($d = 2313$ ms).

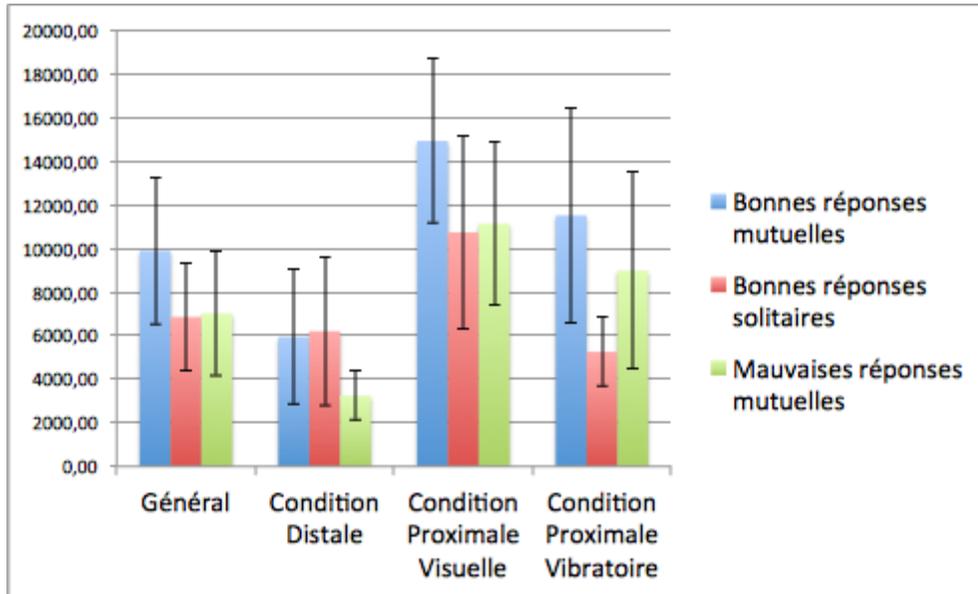


Figure 112. Moyenne de la durée d'une séquence d'interaction en fonction des conditions expérimentales et du type de réponse collective.

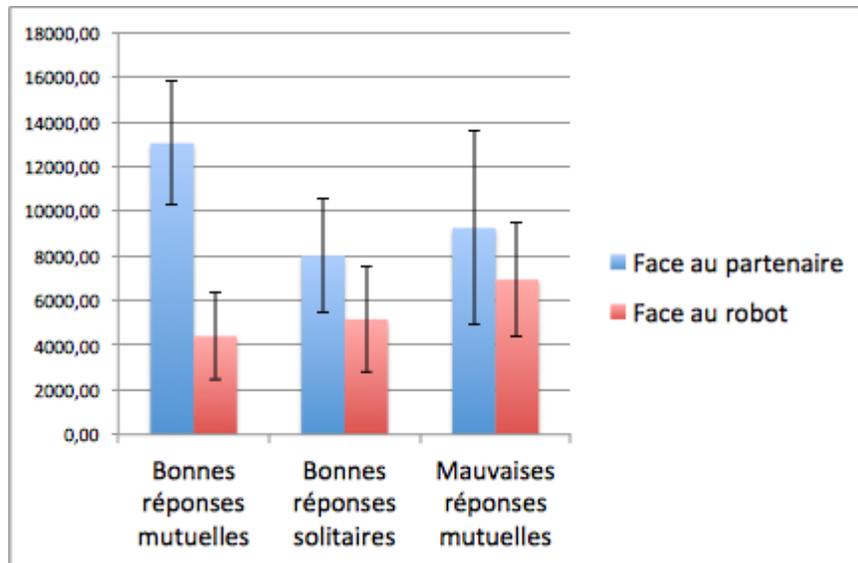


Figure 113. Moyenne de la durée d'une séquence d'interaction en fonction du type d'interactant et de la qualité des réponses collectives.

2.4.6 Discussion

Notre expérience avait pour vocation de tester la possibilité pour deux sujets engagés dans l'environnement minimaliste proposé par notre application de faire la différence entre la présence de l'autre sujet et une trace de l'autre sujet (le robot). En effet, nous avons posé l'hypothèse que le sentiment de présence d'autrui constitue un premier pas possible vers une expérience émotionnelle ; notre expérience nous permet de comparer une présence et une absence. Nous avons de plus posé l'hypothèse que les interactions entre les sujets humains font émerger une dynamique particulière, cette dynamique pouvant servir de support aux participants pour la discrimination entre partenaire et robot. Dans les conditions de la situation mise en place, et conformément à notre position théorique concernant l'émergence d'une signification partagée au cœur d'une coordination des activités perceptives, nous postulons que l'organisation du processus d'interaction entre les sujets (la dynamique des interactions humaines présentant une organisation spécifique) pouvait conduire à une fréquence de reconnaissance plus élevée que la fréquence de non-reconnaissance, ce qui témoignerait du sentiment de présence de l'autre sujet. Par ailleurs, la coordination des foci attentionnels se réalisant par un engagement mutuel des participants, la fréquence de bonnes réponses mutuelles devait être plus importante que la fréquence de bonnes réponses solitaires. Pour autant, les conditions expérimentales devaient avoir un effet sur la réussite de la tâche, puisque la nature du dispositif de couplage entre les actions et les sensations individuelles devait conduire à des modes d'organisation de l'interaction différents, notamment entre la condition Distale et les conditions Proximales. Il s'avère en fait que les trois conditions mènent à des résultats différents ; la condition Distale se distingue en effet des conditions Proximales, mais ces deux dernières ne sont pas pour autant assimilables. En outre, nous nous attendions à ce que la réussite soit meilleure dans le deuxième bloc d'essais de chaque condition expérimentale et dans la deuxième condition passée par rapport à la première, témoignant d'une stabilisation de la dynamique collective et d'une appropriation individuelle du dispositif de couplage.

D'une façon globale, nous pouvons conclure à un succès de la tâche, que ce soit à un niveau individuel (reconnaissance de l'interactant), ou à un niveau collectif (bonnes réponses mutuelles). En effet, de manière individuelle et toutes conditions confondues, le nombre moyen de bonnes réponses est supérieur au nombre moyen de mauvaises réponses ; ce résultat se retrouve dans toutes les conditions. Le fait que cela soit particulièrement vrai face au partenaire est un indicateur que l'on peut donc reconnaître ce dernier au cours d'une interaction par le biais de Touch Through, ce qui pose notre application comme permettant un sentiment de présence. De plus, nous pouvons conclure à un effet d'apprentissage individuel global, qui se traduit par une augmentation des bonnes réponses entre les deux conditions passées. On n'observe par

contre peu de différence entre les deux blocs d'essais à un niveau individuel. Dans les conditions de l'expérience, les sujets ont donc pu percevoir s'ils interagissaient avec leur partenaire ou avec un enregistrement de ce dernier.

D'un point de vue collectif et de manière globale, la fréquence des bonnes réponses mutuelles est supérieure à la fréquence des bonnes réponses solitaires et des mauvaises réponses mutuelles. Ce résultat semble montrer que la réussite de la tâche est attribuable à un engagement interpersonnel (Reddy, 2011), plus qu'à une simple « lecture » unilatérale de la dynamique du mouvement de l'autre participant (Moore, Angelopoulos et Bennett, 1997). Nous pouvons de plus conclure à un effet d'apprentissage collectif global ; il y a une augmentation des bonnes réponses mutuelles et une diminution des bonnes réponses solitaires entre les deux blocs d'essais. Dans la deuxième condition passée, le nombre de bonnes réponses mutuelles augmente, mais c'est plus du fait d'une diminution des mauvaises réponses mutuelles. On notera par ailleurs que le nombre de doubles bonnes réponses est bien plus important face au partenaire que face au robot ; ainsi donc les sujets se reconnaissent bien plus entre eux qu'ils ne reconnaissent le robot.

En outre, si nous pouvons attester d'une réussite globale, il convient de nuancer ces résultats selon les conditions expérimentales. À un niveau individuel, si les trois conditions amènent plus de bonnes réponses que de mauvaises réponses, on constate que les résultats sont moins bons pour la condition Proximale Vibratoire que pour les deux autres conditions. À un niveau collectif, ce résultat se retrouve : la fréquence de doubles bonnes réponses est la plus importante par rapport aux autres réponses dans les conditions Distale et Proximale Visuelle, mais pas dans la condition Proximale Vibratoire, où la fréquence de simples bonnes réponses est supérieure à la fréquence de doubles bonnes réponses, qui est supérieure à la fréquence de doubles mauvaises réponses. La condition Proximale Vibratoire est donc la condition où les sujets réussissent le moins bien la tâche, tant à un niveau individuel qu'à un niveau collectif. Ce résultat nous a tout d'abord semblé ne pas être cohérent ; en effet, les conditions Proximales se ressemblent énormément, fonctionnant sur le même principe de coïncidence (on n'a de stimulation que lorsque l'on est au même endroit). Cependant, il est vrai qu'elles ne délivrent pas tout à fait la même information : lors du contact en modalité visuelle, les sujets peuvent *voir* la position des disques, ce qui leur permet d'avoir accès à une information quant à la direction adoptée par le partenaire, ce qui n'est pas le cas dans la condition Proximale Vibratoire, où le suivi de l'autre sujet est donc encore plus difficile. Cette différence a été exploitée par les sujets et leur a permis de réussir la tâche de manière remarquable en condition Proximale Visuelle, là où les résultats en condition Proximale Vibratoires sont moyens.

En regardant les résultats selon le type d'interactant, on constate à un niveau individuel que les sujets fournissent plus de bonnes réponses face au partenaire que face au robot, et ce dans toutes les conditions expérimentales. C'est dans la condition Proximale Visuelle que le plus

grand nombre de bonnes réponses est atteint (face au partenaire). C'est aussi dans cette condition qu'il existe le plus grand écart entre le nombre de bonnes et de mauvaises réponses selon le type d'interactant. De manière collective, les scores sont à nouveaux meilleurs face au partenaire que face au robot. C'est face au partenaire qu'il y a le plus de bonnes réponses mutuelles, bien avant les bonnes réponses solitaires et les mauvaises réponses mutuelles. Face au robot, le nombre de doubles mauvaises réponses est comparable, mais il y a beaucoup plus de simples bonnes réponses, qui sont le cas de réponses le plus fréquent. Dans toutes les conditions, il y a plus de doubles bonnes réponses face au partenaire que face au robot, et c'est dans la condition Proximale Visuelle que les doubles bonnes réponses sont le plus fréquentes. On constate par ailleurs un effet d'apprentissage face au partenaire : en effet, le nombre de doubles bonnes réponses augmente entre les deux blocs d'essais face au partenaire (alors qu'il stagne face au robot). Nous avons constaté que le partenaire est plus souvent identifié comme humain que ne l'est le robot. L'interaction perceptive entre deux sujets humains semble donc qualitativement différente, et conduit à une meilleure identification.

L'analyse des trajectoires perceptives nous apporte des éléments déterminants pour tenter de comprendre ce qui amène les sujets à se reconnaître. Nous nous intéressons ici précisément au déploiement des activités perceptives et à la coordination interpersonnelle, telles qu'elles sont rendues accessibles à l'analyse via la méthode minimaliste (Lenay, 2006). Tout d'abord, on constate de nettes différences dans les trajectoires perceptives selon le type d'interactant. D'une part, les sujets ont parcouru plus de distance avec le partenaire qu'avec le robot, et ce dans toutes les conditions expérimentales. Ce résultat est tout à fait cohérent avec le second, qui stipule que les sujets ont passé plus de temps face au partenaire que face au robot (à nouveau, ceci est valable dans toutes les conditions) : en étant plus longtemps face à face, ils se sont donc plus déplacés à deux. Ces deux indicateurs ont globalement été utilisés par les sujets pour fournir leur réponse ; plus les indicateurs étaient élevés, plus les sujets avaient tendance à répondre qu'ils étaient face au partenaire. C'est dans la condition Proximale Visuelle que les sujets se basaient de la manière la plus marquée sur ces indicateurs pour répondre. Le troisième indicateur concernait le nombre de séquences d'interaction, qui s'avère être plus faible face au partenaire que face au robot, et ce dans toutes les conditions. Rappelons qu'une nouvelle séquence d'interaction n'est comptabilisée que si la durée séparant deux retours sensoriels est supérieure à deux secondes. Cela indique que le partenaire est plus facile à suivre, et à retrouver en moins de deux secondes si le contact a été perdu, que le robot. Cela est lié au fait que le partenaire est lui aussi à ma recherche et veut maintenir le contact avec moi afin de m'identifier. Des deux résultats précédents, il résulte que la durée d'une séquence d'interaction est supérieure face au partenaire par rapport à face au robot, et ce dans toutes les conditions. On notera que la condition Proximale Visuelle s'avère être celle où la durée moyenne d'une séquence d'interaction est la

plus longue : en effet, le temps total des séquences d'interaction y est le plus long et le nombre de séquences d'interaction le plus faible. C'est donc dans cette condition que les partenaires arrivent le mieux à rester en interaction, et à ainsi s'identifier. En effet, il est facile de rester en contact avec l'interactant en condition Distale (étant donné que l'on voit là où l'autre sujet se trouve même à distance), mais cela n'apporte pas d'indication quant au fait que l'autre participant soit lui aussi en train d'essayer de maintenir le contact. Les stratégies doivent donc être différentes : si en condition Proximale une durée moyenne d'une séquence d'interaction est un indicateur du partenaire, en condition Distale il faudra chercher d'autres moyens pour réussir à distinguer le partenaire du robot.

On trouve des résultats similaires en s'intéressant à la qualité des réponses pour analyser les trajectoires perceptives, en particulier à un niveau collectif. À un niveau individuel, on n'observe en effet que peu de différence entre les niveaux des indicateurs suite à une bonne réponse individuelle ou suite à une mauvaise réponse individuelle. Nous pouvons par contre observer des différences entre les conditions expérimentales : si la distance parcourue lors des séquences d'interaction n'est pas si remarquable, le temps total des séquences d'interaction est lui plus important dans les conditions Proximales, surtout Visuelle. Les sujets passent donc moins de temps en contact avec l'interactant en condition Distale, mais ils effectuent des déplacements plus grands dans cette condition. Cela confirme l'idée qu'il n'est pas nécessaire de prêter grande attention aux déplacements pendant les séquences d'interaction en condition Distale (là où cette attention est un indicateur de la présence de l'autre sujet en conditions Proximales).

À un niveau collectif, on constate que les différences de la qualité des réponses impactent les indicateurs. S'il y a peu de différence pour la distance parcourue à deux de manière globale, en s'intéressant au type d'interactant on constate que quel que soit le type de réponse, les sujets se déplacent plus en interaction avec le partenaire qu'avec le robot. La plus grande distance parcourue à deux face au partenaire est atteinte dans le cas des bonnes réponses mutuelles, tandis que face au robot, elle l'est dans le cas des doubles mauvaises réponses. On retrouve l'idée que qu'une interaction plus longue est synonyme pour les sujets de la présence du partenaire.

En ce qui concerne le temps total des séquences d'interaction, on constate de manière générale qu'il est supérieur dans le cas des doubles bonnes réponses. Ce résultat se retrouve dans toutes les conditions, bien qu'en condition Proximale Visuelle, le temps total des séquences d'interaction dans les cas de doubles bonnes réponses et de doubles mauvaises réponses sont comparables. Cela s'explique à nouveau par le fait que lorsque les sujets arrivent à maintenir un long contact avec l'interactant, ils vont avoir tendance à penser se trouver face au partenaire dans le cas des conditions Proximales. Étant donné qu'ils arrivent mieux à maintenir le contact en condition Proximale Visuelle qu'en condition Proximale Vibratoire, on comprend que le temps

total des séquences d'interaction avec le robot puisse être élevé dans le cas des mauvaises réponses mutuelles. En s'intéressant à l'indicateur en considérant le type de réponse et le type d'interactant, on trouve un temps plus long face au partenaire, avec des résultats comparables dans le cas des doubles mauvaises réponses. Le temps le plus long passé face au robot se retrouve à nouveau dans le cas des doubles mauvaises réponses.

Pour le nombre de séquences d'interaction, il est de manière générale plus faible dans le cas des doubles bonnes réponses, ce qui se situe dans la continuité des résultats précédents. Ce résultat se retrouve dans toutes les conditions expérimentales, bien qu'il soit plus marqué dans le cas des conditions Proximales. En considérant aussi le type d'interactant, on constate que le plus faible nombre de séquences d'interaction se produit dans le cas des bonnes réponses mutuelles face au partenaire (et le plus important face au robot se produit dans le cas des doubles bonnes réponses).

Enfin, en ce qui concerne le temps moyen d'une séquence d'interaction, il est plus important dans le cas des doubles bonnes réponses. Ce résultat se retrouve de manière très marquée dans les conditions Proximales, mais pas dans la condition Distale. De manière globale, ce temps est plus important dans les conditions Proximales que dans la condition Distale (sauf dans le cas des simples bonnes réponses en condition Proximale Vibratoire). Selon le type d'interactant, on constate que dans le cas des doubles bonnes réponses, le temps d'une séquence d'interaction face au partenaire est trois fois plus grand que face au robot. Dans les autres cas de réponses, il est toujours supérieur, de manière moins marquée cependant.

Les sujets mettent en effet en place un jeu de co-ajustements dynamiques, qui résulte de l'engagement mutuel des sujets dans les séquences d'interaction (De Jaegher et Di Paolo, 2007 ; Fogel, 1993) et qui permet aux sujets de se reconnaître. Lors du croisement perceptif entre deux sujets, chacun des partenaires est engagé dans la régulation de son propre mouvement d'exploration en fonction des retours sensoriels reçus. Étant donné que les retours sensoriels résultent de la rencontre de l'activité perceptive de l'autre participant, l'engagement de l'un des sujets régule l'engagement de l'autre. Cette co-régulation dynamique produit une structure temporelle, qui est propre au processus interactif, ce qui témoigne de l'émergence d'une dynamique relationnelle autonome (De Jaegher, Di Paolo et Gallagher, 2010). Il n'y a donc pas une simple production de contingences qui seraient détectées par l'autre participant : dans la mesure où les deux sujets sont activement engagés dans l'interaction, ils s'entraînent mutuellement dans une accroche perceptive, accroche qui émerge de leur volonté réciproque à établir un contact, à le maintenir et à le retrouver lorsque celui-ci est perdu. Ainsi, ils participent activement à la co-régulation de l'interaction, en saisissant les opportunités d'action disponibles localement (Krueger, 2011 ; Lenay et Stewart, 2012).

Les résultats mettent finalement en avant le fait que la condition Proximale Visuelle se révèle la plus efficace pour la réalisation de la tâche. En effet, au niveau des binômes, c'est la condition entraînant le plus de doubles bonnes réponses, le plus fort apprentissage entre les deux blocs d'essais, ainsi que le plus fort taux de bonnes réponses mutuelles lorsque les interactants étaient face à face. On notera de plus que les conditions Proximales, et en particulier la condition Proximale Visuelle, sont plus propices à voir apparaître une réponse « Partenaire » qu'une réponse « Robot », alors qu'en condition Distale, les sujets répondent « Partenaire » la moitié du temps seulement. Cela indique aussi que les conditions Proximales sont plus aptes à donner le sentiment de la présence de l'autre sujet que la condition Distale.

En vue de maintenir le croisement perceptif permis par l'application, qui n'est possible que grâce aux propriétés fonctionnelles du toucher que sont la coïncidence et la perception active implémentées dans l'application, les sujets mettent en place des stratégies, telles que rester au plus près du partenaire, le suivre, le chercher en se rendant notamment au dernier endroit où on l'a croisé,... Grâce au questionnaire, les sujets ont pu nous faire part de leurs stratégies ; leurs retours se recoupent avec nos analyses. Les sujets ont cherché avant tout à établir un contact, et se sont basés sur le nombre de stimulations et la réaction de l'interactant pour pouvoir l'identifier. Plus de stimulations étaient synonymes du partenaire en conditions Proximales. Si l'interactant essayait de les suivre, cela était aussi synonyme du partenaire. Les stratégies en condition Distale étaient plus difficiles pour les sujets à expliciter ; quelques sujets reconnaissaient que le parcours fait par le robot ressemblait à un parcours précédent. Seuls quatre sujets ont identifié que le robot refaisait ce que le partenaire avait déjà fait.

De façon minimale, nous pourrions caractériser l'interaction humaine par l'émergence d'une recherche perceptive mutuelle. C'est en effet cette recherche qui semble caractériser la reconnaissance d'autrui dans les croisements perceptifs minimalistes, qui est alors décrite comme le produit d'un attracteur de la dynamique collective (Auvray, Lenay et Stewart, 2009 ; Lenay, 2010).

2.4.7 Conclusion

Outre ses résultats d'analyse qui montrent que Touch Through permet un sentiment de présence, cette expérience nous a permis de procéder à l'évaluation de notre application, dont l'objectif est d'enrichir la relation à distance en permettant une suppléance perceptive du toucher. Grâce aux questionnaires et aux retours des utilisateurs suite à l'expérience, nous avons pu accéder à leur évaluation de l'application.

Pour décrire l'expérience vécue lors de l'utilisation de l'application, il a été demandé aux utilisateurs d'utiliser 5 mots-clés de leur choix. Le terme le plus fréquent a été « intéressant » (8 occurrences) ; cela est encourageant pour le développement de l'application. Ensuite, les sujets ont répondu « difficile » (6 réponses) ; cela n'est pas étonnant étant donné le caractère minimaliste du dispositif, qui nécessite d'une part la compréhension du système et d'autre part le déroulement de la tâche en lui-même. La troisième réponse la plus fréquente est « toucher » (5 occurrences), ce qui est extrêmement positif ; en effet, il n'a jamais été mentionné que l'application permettait une suppléance perceptive du toucher. C'est donc bien au cours de l'utilisation que les utilisateurs ont ressenti la mise en jeu de ce sens. Il est à noter que ce terme n'a pas été répondu uniquement par les sujets ayant passé la condition Proximale Vibratoire, mais aussi par des sujets n'ayant passé que les deux conditions en modalité Visuelle. Ensuite, ont été cités à égalité (4 occurrences) les termes « intrigant », « interaction » et « amusant ». Une fois de plus, c'est encourageant quant à l'accueil que les utilisateurs réservent à l'application. 54 autres termes ont aussi été cités, une ou deux fois chacun.

À la question « vous êtes-vous senti proche de votre partenaire ? », les sujets ont répondu oui à 54%, et non à 46%. Par condition, on retrouve le même score pour la condition Distale, un score 50%-50% pour la condition Proximale Visuelle, contre un 58%-42% pour la condition Proximale Vibratoire. Ainsi, c'est la condition de suppléance perceptive du toucher à distance par une stimulation tactile qui permet aux sujets de ressentir le plus facilement un sentiment de présence ; il est intéressant de noter que le ressenti des sujets diffère des résultats de l'expérience, où ils ont été plus performants en terme de résultats dans les autres conditions. Notons que la condition Proximale Visuelle n'est pas aussi bien évaluée que les deux autres pour le critère de la performance, alors qu'elle délivre d'une part pratiquement les mêmes informations que la condition Proximale Vibratoire et que d'autre part elle est celle où les sujets ont le mieux réussi la tâche de reconnaissance. Il est à noter que beaucoup de sujets précisent avoir répondu « non » du fait qu'à cause de la tâche même, ils étaient dans le doute : comment se sentir proche de son partenaire alors que l'on est par définition pas sûr que ce soit lui ? Cela étant rappelé, les réponses données sont encourageantes.

Au niveau de la stratégie, les sujets ont plusieurs fois rapporté avoir des stratégies différentes en fonction de la condition passée. Par exemple, un sujet déclare : « La première condition [condition Proximale Visuelle] peut se faire sous la forme d'un suiveur et d'un leader avec échange des rôles de temps en temps. Deuxième condition [condition Distale] : jeu de cache-cache ; poser le doigt et voir si le partenaire s'y dirige puis lever son doigt pour le poser un peu plus tard autre part. » Un autre pose : « En condition tactile, j'ai essayé de rester statique puis de faire de grands mouvements, afin d'ensuite chercher contact. En condition [distale] visuelle, j'ai effectué des mouvements plus fluides avec une allure variable. » Au niveau

stratégique, les sujets ont donc fait la différence entre la condition Distale et celle des conditions Proximales qu'ils ont effectuée.

Les sujets ont souvent précisé avoir trouvé la condition Distale plus facile (alors que l'analyse détaillée des résultats montre qu'en fait ce n'est pas la condition où les sujets réussissent le mieux la tâche), mais ont aussi trouvé plus d'intérêt aux conditions Proximales : plus de challenge, plus de besoin d'interprétation et d'imagination, plus de contact,...

En ce qui concerne l'évaluation globale de l'application, les retours utilisateurs sont très positifs et encourageants. La plupart trouvent le fonctionnement intuitif et agréable. On notera particulièrement deux remarques. La première insiste sur l'idée de proximité potentielle avec le partenaire : « L'application a sans doute plus de sens quand on connaît le partenaire. Elle a surtout un côté ludique. Au bout des essais on a l'impression que l'on pourrait caractériser l'interaction avec le partenaire, avec tout un style (par exemple le fait qu'il cherche le contact ou non). C'est rigolo pour jouer avec quelqu'un de proche. Je l'imagine bien pour l'utiliser avec quelqu'un d'éloigné, mais proche dans le cœur. Avec quelqu'un avec qui on n'est qu'un simple pote ça perdrait son intérêt. » Nous avons développé l'application dans l'idée qu'elle puisse être utilisée à deux par des individus plus ou moins proches (voire des inconnus), mais il est très intéressant de noter que spontanément, certains utilisateurs identifient un potentiel caractère intime et relationnel de l'échange. La deuxième remarque reçue que nous souhaitons mettre en avant insiste sur les différences entre condition Distale et condition Proximale, et sur le fait que la première, bien qu'apportant plus d'informations, n'est pas forcément la meilleure ou la plus agréable : « L'expérience s'est déroulée en deux étapes. Pendant la première on voyait le déplacement de son partenaire/du robot sur l'écran [condition Distale]. Pendant la seconde, on ne sentait plus qu'une vibration [condition Proximale Vibratoire]. Contrairement à ce que je pensais, la deuxième condition était bien plus intéressante. Je m'y suis senti proche de ma partenaire alors que j'avais moins d'information sur sa localisation. Et je pense avoir alors mieux discerné le robot de ma partenaire. »

Globalement et suite à ces premières évaluations, notre première proposition d'interface pour la communication émotionnelle à distance par le biais d'une suppléance perceptive du toucher tient ses objectifs et permet aux utilisateurs de faire l'expérience du sentiment de présence en l'utilisant.

En ce qui concerne les perspectives, plusieurs pistes sont à envisager. D'une part continuer l'évaluation et l'amélioration de l'application grâce à des retours utilisateurs, et d'autre part continuer à comparer les différences entre conditions et leur impact sur une communication riche sur le plan émotionnel. Une ouverture de Touch Through consisterait à proposer son utilisation à plus de deux utilisateurs à la fois.

2.5 L'articulation Théorie-Design-Expérimentation : premiers résultats

Dans cette partie, nous avons effectué une première itération de la boucle Théorie-Design-Expérimentation (voir figure 114). Partant du socle théorique présenté en partie 1 et de recherches concernant le sens du toucher, nous avons tout d'abord établi que ce dernier était pertinent dans la communication émotionnelle à distance. Il rend en effet possible un type de communication efficace sur le plan émotionnel, et permet aux individus de se sentir en contact. Les caractéristiques du toucher sont identifiées comme efficaces pour la communication à distance basée sur le relationnel, et non sur l'informationnel comme le sont les interfaces les plus répandues aujourd'hui. Nous avons présenté les diverses interfaces proposées par des chercheurs et des designers pour permettre l'échange relationnel et émotionnel à distance, en regroupant celles qui permettent une sorte de toucher à distance. Afin de proposer notre propre interface, nous avons conçu une application smartphone permettant une suppléance perceptive du toucher, rendant ainsi possible une forme de toucher à distance. Nous avons ensuite utilisé cette application comme dispositif minimaliste dans une expérimentation mettant en jeu un test de Turing. Suite à l'analyse des résultats et aux retours des utilisateurs, nous avons établi que Touch Through, malgré sa simplicité, permet un sentiment de présence pour les utilisateurs s'en servant. La suppléance perceptive du toucher que nous proposons s'inscrit donc dans les interfaces au centre de notre recherche. L'expérimentation nous a de plus permis d'étudier les différences entre les deux modes de toucher permis par l'application, distal et proximal. Le mode proximal, bien que moins riche sur le plan informationnel, semble plus adapté pour la communication émotionnelle et relationnelle.

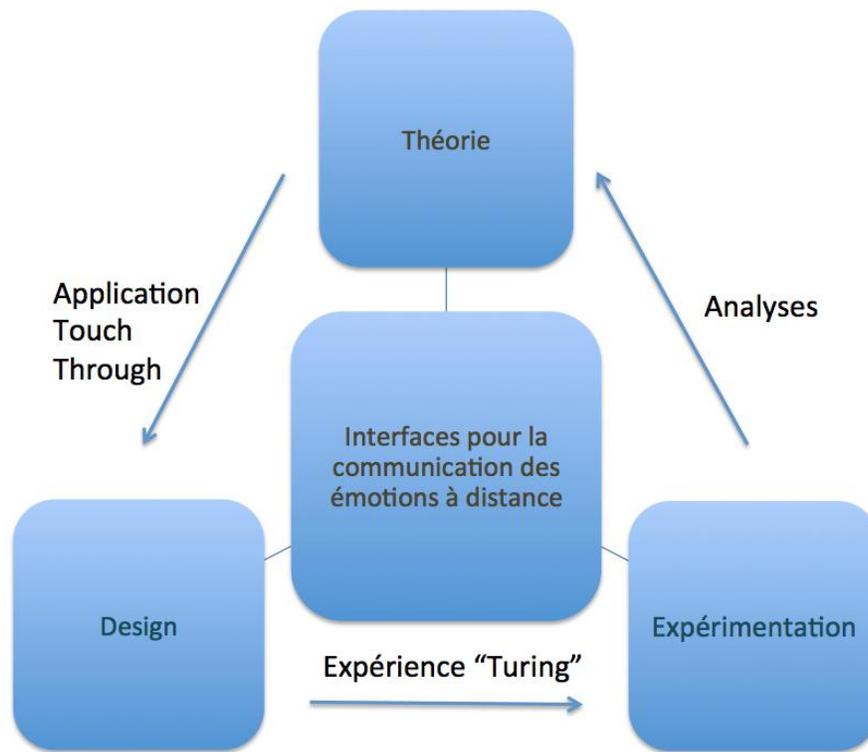


Figure 114. La première itération de la boucle Théorie-Design-Expérimentation.

3 L'attention partagée : du théorique vers le design en passant par l'expérimentation

3.1 Introduction

Dans la partie 2, nous avons vu une première articulation entre théorie, design et expérimentation : de la théorie découle une solution de design, à partir de laquelle il est possible de mettre en place une expérimentation. C'est la démarche que nous avons adoptée lors du développement de l'application Touch Through. Ce faisant, nous avons réussi, dans une certaine mesure, à éclaircir ce qui était essentiel dans la rencontre dyadique. De là, nous avons dégagé des principes de design que nous avons appliqués et testés.

Nous allons à présent considérer une nouvelle approche ; partant de la théorie, nous allons imaginer une expérimentation, dont les résultats vont nous servir comme base pour un nouveau design (voir figure 115). Il s'agit à présent d'ouvrir la relation dyadique à la perception commune d'un monde partagé, afin de pouvoir proposer des relations plus riches. La nouvelle étape essentielle pour créer un espace d'interaction émotionnel consiste alors pour les interactants à mettre en place une attention conjointe autour d'objets communs. De même que lors du travail sur l'interaction dyadique effectué précédemment, nous voulons là encore dégager des principes essentiels de la dynamique d'interaction pour coordonner des activités perceptives autour d'un objet partagé ; c'est là la question classique de l'attention conjointe décrite en cognition sociale.

Dans le cadre de notre travail, la question de la perception mutuelle d'objets communs se révèle comme étant tout à fait pertinente. L'interaction à distance peut en effet se faire autour d'objets distants mis en commun, dans des espaces numériques partagés. Ces espaces font entrer en jeu non seulement des rencontres interpersonnelles, mais aussi une coordination des activités perceptives, que ce soit pour l'interaction ou pour l'exploration de contenus partagés (autour desquels peuvent s'organiser des interactions significatives). Quels mécanismes permettent aux utilisateurs de savoir qu'ils sont dans un monde commun ? Notamment, comment organisent-ils et coordonnent-ils alors leurs activités perceptives ? Le cadre théorique de l'attention conjointe permet de se saisir de ces questions.

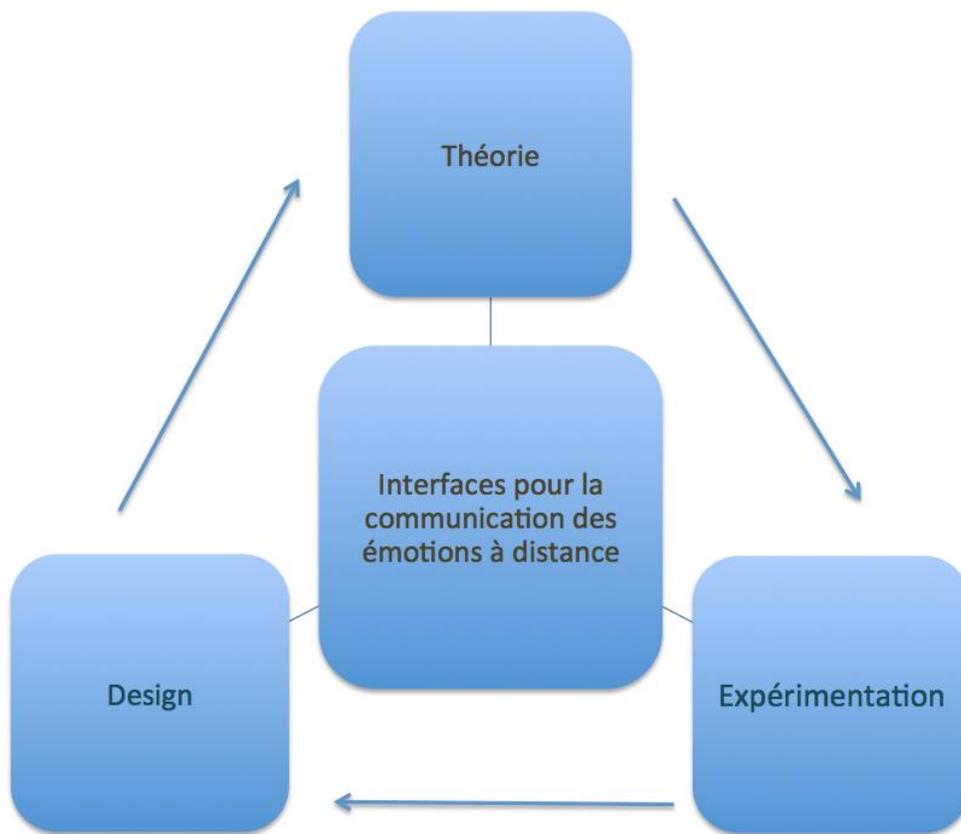


Figure 115. Un second cycle : théorie-expérimentation-design.

Les espaces numériques partagés offrent de nombreuses opportunités de collaboration, qui impliquent une coordination interpersonnelle des activités perceptives individuelles autour d'objets communs. Il s'agit donc d'une part de pouvoir percevoir les autres utilisateurs ainsi que le focus de leur attention, mais aussi d'autre part de percevoir les objets présents dans les espaces numériques partagés. Il est donc capital que les utilisateurs puissent identifier les objets comme étant partagés, et par conséquent perceptibles par les autres utilisateurs. On peut considérer par exemple la confrontation aux mediaspaces et aux situations de collaboration à distance (Fraser et al., 2006, Luff et al., 2003). On notera que ces solutions sont basées sur des informations sonores et visuelles (écrans, audioconférence, visioconférence). Il s'avère que ces situations se révèlent peu propice à la reconnaissance des objets partagés par les interactants à distance et à l'organisation d'une attention conjointe autour de ces objets (Heath et Luff, 1991, 1992, Heath, Luff et Sellen, 1977 ; De Fornel 1992, 1988, 1994).

Dans le cadre de la collaboration à distance, Hindmarsh et Heath (2000) s'intéressent à la façon dont les collègues travaillant ensemble établissent, ne serait-ce que momentanément, une orientation mutuelle dirigée vers des « objets », tels que des documents ou des écrans. Ils

identifient trois thèmes clés quant à ce type d'interaction. Le premier concerne la façon dont le visuel, couplé au discours, est utilisé pour faire des références démonstratives. Le deuxième aspect est celui de la production en interaction de références démonstratives, accompagnée d'une coordination dynamique des actions au cours de la réussite collaborative de l'orientation mutuelle. Le troisième enfin concerne les façons dont la production et l'intelligibilité des actions référentielles trouvent leur fondement dans les activités communes, dans lesquelles, par lesquelles et pour lesquelles elles émergent.

Il apparaît comme évident que l'attention conjointe est un mécanisme complexe, faisant entrer en jeu deux interactants et un objet sur lequel tous deux portent leur attention de manière conjointe (en ayant conscience que l'attention de l'autre participant se porte aussi sur ce même objet ; l'attention est donc à la fois dirigée vers un objet commun, ainsi que vers le partenaire et là où se situe son attention). Elle implique, pour chaque partenaire de l'interaction, une interprétation de la situation, qui doit passer par une interprétation des postures et mouvements du partenaire, ainsi que par une interprétation ou compréhension du focus du partenaire, et de l'objet perçu par ce partenaire. Comment ce mécanisme compliqué se met-il en place, dans quelles conditions peut-il avoir lieu ? Comment l'enfant apprend-il l'attention conjointe ? Comment peut-on proposer une solution d'interface au problème de l'attention conjointe ?

Nous allons dans cette partie dans un premier temps exposer la partie théorique prise en considération (celle de l'attention conjointe). Nous allons présenter cette opération complexe plus en détails en nous appuyant sur plusieurs définitions. Nous allons exposer les limites d'une approche classique pour considérer ce mécanisme, et expliciter notre posture. Nous allons ensuite considérer l'attention conjointe comme coordination perceptive, et allons proposer un modèle théorique énoncé de cette idée. Après cela, nous allons présenter la mise en place de l'attention conjointe chez l'enfant, ce qui est éclairant dans notre travail étant donné que cela permet de se pencher sur les mécanismes de fonctionnement du processus. Nous allons donc présenter les deux étapes de développement de l'attention conjointe chez l'enfant (à savoir d'une part répondre à la sollicitation d'attention conjointe, et d'autre part initier l'attention conjointe), qui apporteront des indices d'éléments à prendre en considération. Nous verrons ensuite le rôle de la dynamique du mouvement et du contexte dans l'attention conjointe, la posant comme une pratique sociale incarnée. À ce moment, nous considérerons aussi l'influence des différents sens dans ce mécanisme.

Dans un second temps, nous allons décrire l'expérience minimaliste que nous avons mise en place autour de l'attention conjointe ; il s'agit de l'étude expérimentale « objets communs et objets privés ». Elle a pour but de nous permettre de proposer des hypothèses sur les mécanismes en jeu lors de l'attention conjointe et sur les conditions nécessaires pour la réussir. Nous

présenterons cette expérience, ses résultats, et les conclusions que nous pouvons en tirer dans le cadre de nos travaux.

Nous exposerons enfin une proposition de design s'appuyant sur les résultats présentés ; elle consiste en une solution destinée à la plateforme du réseau social facebook, et est axée sur la possibilité d'attention conjointe (participant de la visibilité entre public et privé d'une part, et proposant un nouveau type d'interaction sur cette plateforme d'autre part).

3.2 Attention conjointe, objet partagé : contexte théorique

3.2.1 Attention conjointe : définitions et enjeux

D'une façon générale, l'attention conjointe fait référence à la capacité d'un sujet donné à coordonner son attention avec un partenaire social, de manière à partager avec lui un point de référence, que ce soit un objet, un événement ou un autre sujet (Mundy et Newell, 2007). L'attention conjointe désigne donc le fait que deux personnes s'intéressent en même temps à la même chose, et soient conscientes d'être toutes les deux focalisées sur cet objet. Ainsi, chacun des individus a son attention portée à la fois sur l'objet d'une part, et sur le fait que l'autre personne porte elle aussi son attention à l'objet d'autre part. L'un des exemples les plus connus du phénomène d'attention conjointe est celui du suivi du regard d'autrui : l'objet de l'attention n'est plus le regard d'autrui, comme dans le cas des interactions en face-à-face impliquant un regard mutuel (Argyle et Cook, 1976), mais la cible du regard de l'autre agent. On parle de regard déictique (Gomez, 1991).

On se rend bien compte que le simple fait que deux personnes s'intéressent au même objet en même temps n'est pas suffisant pour qu'une attention conjointe soit en jeu. Ainsi, Call et Tomasello se basent sur la définition suivante : « Joint attention is not just two individuals looking at the same thing at the same time. Joint attention requires that each of the individuals knows that the other is attending to the same thing as they are attending to; that is what makes it a joint, rather than merely a simultaneous activity » (Call et Tomasello, 2005, p.45)³⁰. On retrouve cette idée chez Eilan : « For [...] joint attention to be in play, it is not sufficient that the

³⁰ Traduction personnelle : « L'attention conjointe ne consiste pas simplement en ce que deux individus regardent la même chose au même moment. L'attention conjointe nécessite que chacun des deux individus sache que l'autre est attentif à la même chose qu'eux ; c'est cela qui crée une activité conjointe et non une simple activité simultanée. »

infant and the adult are in fact attending to the same object, nor that the one's attention cause the other's. » (Eilan, 2005, p.1)³¹.

De même, Gomez procède à cette distinction, et précise: « I propose that JA can be analysed into two main components: Attention following (a subject attends to the same target of attention as another subject in response to the latter's attention) and Attention contact (two subjects mutually attending to each other's attention). » (Gomez, 2005, p.65)³². On notera que Gomez met aussi en avant le fait que c'est en réponse au comportement de l'un des deux individus que l'autre se met à diriger son attention sur le même objet. Par exemple, le mécanisme dans le cas du visuel consiste souvent à ce que l'un des individus pointe du doigt l'objet de l'attention ; dans le cas de l'audition, il sera plutôt indiqué de faire silence et de tendre l'oreille. Ainsi, lorsqu'une personne attire avec succès l'attention d'une autre sur un troisième objet, les deux individus sont en situation d'attention conjointe. Par exemple, la figure 116 nous montre un couple dans leur jardin. Chacun perçoit le corps image d'autrui, sa capacité à percevoir et son intentionnalité. La femme a entrepris des actions perceptives spécifiques en relation avec l'espace qu'ils partagent ; elle regarde dans la direction de l'arbre et pointe du doigt dans cette même direction. Ce dernier fait semble pouvoir permettre à l'homme de regarder dans la même direction et de focaliser son attention sur l'élément particulier qui est désigné. En réaction à l'activité perceptive de la femme, l'homme regarde dans la direction qu'elle perçoit et qu'elle désigne. C'est bien l'activité perceptive de la femme qui entraîne les actions perceptives de l'homme. Ils portent tous deux leur attention sur un objet extérieur, et sont conscients d'être tous les deux en train de s'intéresser à cet objet.

³¹ Traduction personnelle : « Pour qu'une attention conjointe soit mise en jeu, il n'est pas suffisant que le bébé et l'adulte soient effectivement en train de porter leur attention au même objet, ni que l'attention de l'un entraîne l'attention de l'autre. » En effet, l'attention de l'un peut par exemple entraîner l'attention de l'autre sans que le premier en soit conscient ; ce n'est alors pas de l'attention conjointe.

³² Traduction personnelle : « Je propose que l'on analyse l'attention conjointe comme étant constituée de deux composantes principales : le suivi d'attention (un sujet est attentif à la même cible qu'un autre sujet en réponse à l'attention de ce dernier) et le contact d'attention (deux sujets sont mutuellement attentifs à l'attention de l'autre) ».



Figure 116. Exemple d'attention conjointe ; deux individus observent avec attention un arbre de leur jardin.

Comme nous l'avons vu dans la partie 1.1.1.2, il existe pour l'étude des interactions (auxquelles appartient l'attention conjointe) deux approches principales : une approche dite classique, et une approche sensori-motrice. L'approche cognitive classique consiste à expliquer la cognition sociale en supposant l'existence d'une théorie naïve des comportements intentionnels d'autrui (la théorie de l'esprit) présente chez chaque sujet. L'interaction sociale est alors considérée comme un simple input sur lequel s'opèrent des processus internes (Tomasello, 1995 ; Tomasello, 1999). Dans l'approche classique, l'attention conjointe implique en premier lieu la conscience que les deux participants dirigent un focus attentionnel sur le monde, et, de manière cruciale, que celui-ci peut être partagé (Tomasello 1995). La perception mutuelle d'un objet serait alors la résultante d'une capacité réflexive des individus se trouvant engagés dans l'interaction, capacité qui rend possible la compréhension du fait qu'autrui a un focus attentionnel dirigé sur le même objet que soi (ibid).

Cependant, en la posant ainsi, l'attention conjointe désigne un mécanisme individuel qui est destiné à la compréhension de l'état attentionnel d'autrui, et qui rend possible le partage d'une représentation de l'objet de son attention (ce partage étant lui-même représenté chez l'individu). L'attention conjointe repose alors sur une conception mentaliste de l'esprit (Baron-Cohen, 1997 ; Tomasello, 1995 ; Tomasello, 1999) : un monde interne privé se trouve nécessairement à l'origine du comportement, et la compréhension du comportement d'autrui se ferait par une psychologie de sens commun, classiquement organisée en une théorie de l'esprit (Astington et Gopnik, 1991 ; Premack et Woodruff, 1978). Dans cette approche, l'attention conjointe serait alors la résultante d'une coordination d'états mentaux internes précédant l'action conjointe. De manière minimale, l'action conjointe englobe toute forme d'interaction sociale où au moins deux individus coordonnent leurs actions dans l'espace et dans le temps afin de produire un changement dans l'environnement (Sebanz, Bekkering et Knoblich, 2006). Ici, la coordination des actions ne pourrait se réaliser que si les interactants partagent à priori un même référent – et ont conscience qu'ils le partagent. Il s'agit donc dans cette approche de l'attention conjointe d'appliquer une théorie des actions intentionnelles par des sujets rationnels (voir Tomasello et al. 2005) ; c'est une perspective représentationaliste.

Cependant, dans notre travail, nous cherchons à créer une situation conjointe complètement nouvelle dans une relation médiatisée à distance. Il s'agit donc de situations d'interaction inédites. Le cadre théorique de l'approche cognitive classique se révèle alors peu pratique : il y aurait en effet besoin pour les interactants, en plus de la théorie de l'esprit, d'une théorie de l'usage de l'interface pour pouvoir conduire leur raisonnement. Ce besoin d'une théorie de l'interface impliquerait une solution compliquée du point de vue du design, étant donné que nous cherchons à provoquer un sentiment de présence à distance et de partage, et non seulement un raisonnement abstrait et hypothétique. Il semble donc plus approprié d'adopter d'autres approches qui prennent au maximum en compte l'activité perceptive elle-même, ainsi que la coordination des actions, comme par exemple celle de Gallagher (Gallagher, 2004 ; Gallagher, 2011 ; Fiebich et Gallagher, 2012).

Grâce à l'approche sensori-motrice adoptée dans notre travail, nous allons replacer l'attention conjointe dans le cadre d'une pratique sociale incarnée (voir 4.2.2 et 4.2.4). D'autre part, nous allons essayer par une approche minimaliste (pratiquée jusque là dans le cadre de cette thèse) de comprendre cette dynamique d'interaction qui fait passer de la relation dyadique face-à-face à une relation triadique organisée autour d'une perception commune d'un objet tiers. Cette approche est pour nous intéressante car grâce au minimalisme, nous allons pouvoir regarder précisément les conditions nécessaires à l'attention conjointe d'une part, et d'autre part les opérations et la dynamique d'interaction qui réalisent l'attention conjointe. On cherche à tirer de

cette analyse et de l'expérience associée des bases pour nous guider dans notre design d'interaction.

Cependant, notre travail nous amène à tout d'abord nous interroger sur l'interaction spécifique mise en jeu lors d'une attention conjointe. Par quels mécanismes deux personnes se rendent-elles compte que leur attention est focalisée sur le même objet ? Comment peut-on attirer l'attention d'autrui vers un objet, autrement qu'en le pointant du doigt ? Quelle sorte d'interaction l'attention conjointe nous permet-elle d'avoir ?

Nous souhaitons en particulier nous intéresser à la question de ce qui est public, et de ce qui est privé ; de ce qui est partagé et de ce qui ne l'est pas. Comment se rend-on compte qu'autrui n'a pas accès à un objet que l'on perçoit ? À l'inverse, comment sait-on que notre partenaire peut percevoir un certain objet auquel nous avons nous-même accès ? Cette question est cruciale dans le cadre de notre recherche. Elle permet une réflexion sur ce qui se joue dans l'interaction et la collaboration dans les espaces numériques. L'idée que l'on puisse évoluer dans un espace numérique et y partager des objets avec autrui soulève la question des conditions de possibilité de la collaboration et de son expérience vécue. Un échange de ce type riche sur le plan émotionnel nécessite une collaboration efficace et un sentiment de proximité avec la personne avec laquelle on interagit. Il est donc important de s'intéresser au statut partagé ou non des objets dans l'espace numérique, et d'étudier la manière dont les utilisateurs identifient ce statut. L'interaction s'organisant autour des objets communs constitue aussi un point d'intérêt important. Le rôle de l'interface y est en effet déterminant, et impacte la qualité de l'interaction proposée ainsi que l'expérience vécue par les utilisateurs.

3.2.2 L'attention conjointe comme coordination perceptive

Nous faisons le choix, dans ce travail concernant notamment la conception d'interfaces, de nous placer dans une perspective de perception active. Nous adoptons une perspective interactionniste, éactive, et constructionniste. Dans l'approche incarnée et située de la cognition, permettant notamment de réhabiliter l'action et l'engagement mutuel pour rendre compte de la compréhension sociale (Gallagher, 2004 ; Reddy et Morris, 2004 ; De Jaegher et Di Paolo 2007), la coordination qui a lieu dans le phénomène d'attention conjointe serait d'abord perceptive, et donc sensori-motrice (Gallagher, 2011), plus qu'elle ne serait une coordination interne d'états mentaux (comme cela est considéré dans l'approche classique). L'attention conjointe serait dans cette perspective au cœur d'une activité perceptive partagée, permettant la co-construction d'un cadre de référence pragmatique. On considère alors l'attention non comme un *état* rendant possible l'engagement de l'individu dans le monde, mais comme un *processus* qui découle de cet

engagement (Reddy, 2011). Dans cette perspective, l'attention conjointe est dirigée dans le monde à travers l'engagement actif du sujet. Être conscient de l'attention d'autrui, c'est donc être conscient de son engagement dans le monde, celui-ci étant directement perceptible au travers de ses actions (Gallagher, 2004 ; Reddy, 2011).

Dans le cadre d'une approche éactive, on peut construire un modèle de croisement perceptif de l'attention conjointe en considérant d'une part la façon dont deux individus perçoivent l'environnement, et d'autre part la situation du croisement perceptif. En ce qui concerne la perception du monde, il s'agit de dupliquer le schéma de la boucle sensori-motrice d'un individu ; deux individus perçoivent alors le monde (voir figure 117.a).

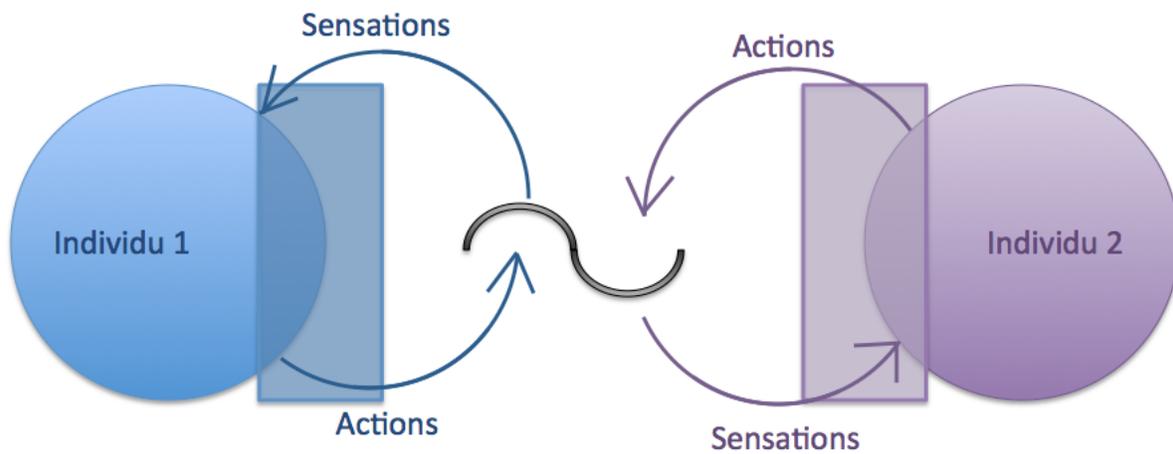


Figure 117.a. Modèle théorique de la boucle-sensori-motrice pour deux individus percevant le monde.

D'autre part, le croisement perceptif peut se ramener à l'idée que chaque individu devient ce qui est perçu par l'autre. Ainsi, les actions de l'un agissent sur les sensations reçues par l'autre, et réciproquement (voir figure 117.b).

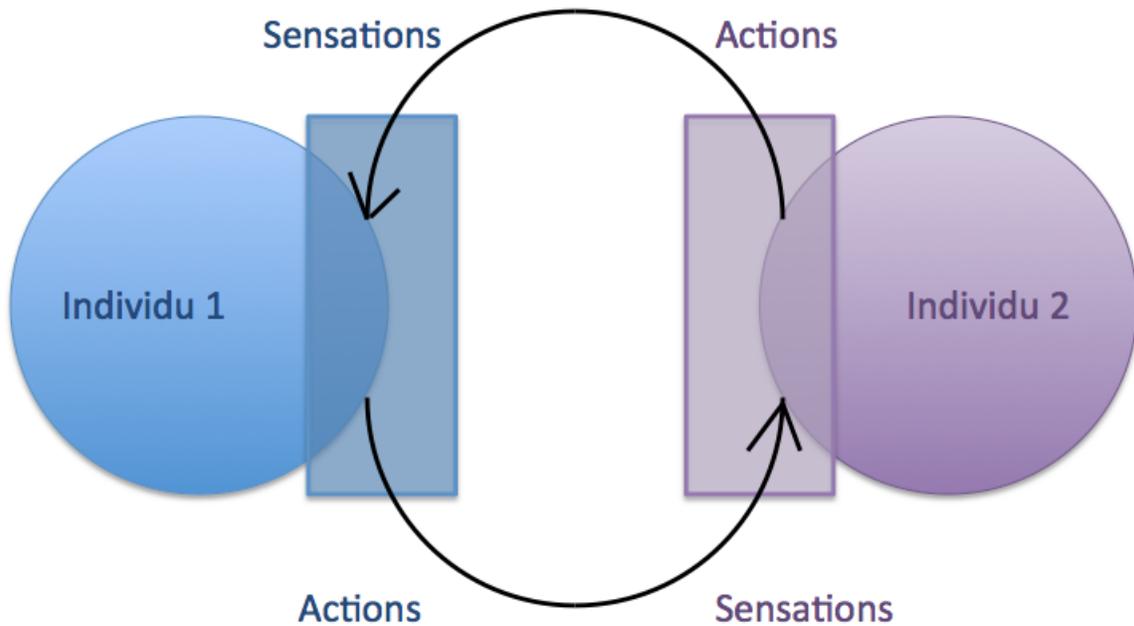


Figure 117.b. Modèle théorique du croisement perceptif ayant lieu entre deux individus.

Dans une situation d'attention conjointe, d'une part les deux individus sont focalisés sur la même partie du monde (ce que l'on peut ramener à la figure 117.a) et d'autre part ils sont conscients de là où se focalise l'attention de l'autre participant (ce qui évoque une des caractéristiques du croisement perceptif de la figure 117.b). Ainsi, lorsque l'on ramène la situation d'attention conjointe dans le cadre d'une approche sensori-motrice de la perception, on obtient donc le modèle de croisement perceptif suivant (voir figure 117.c) :

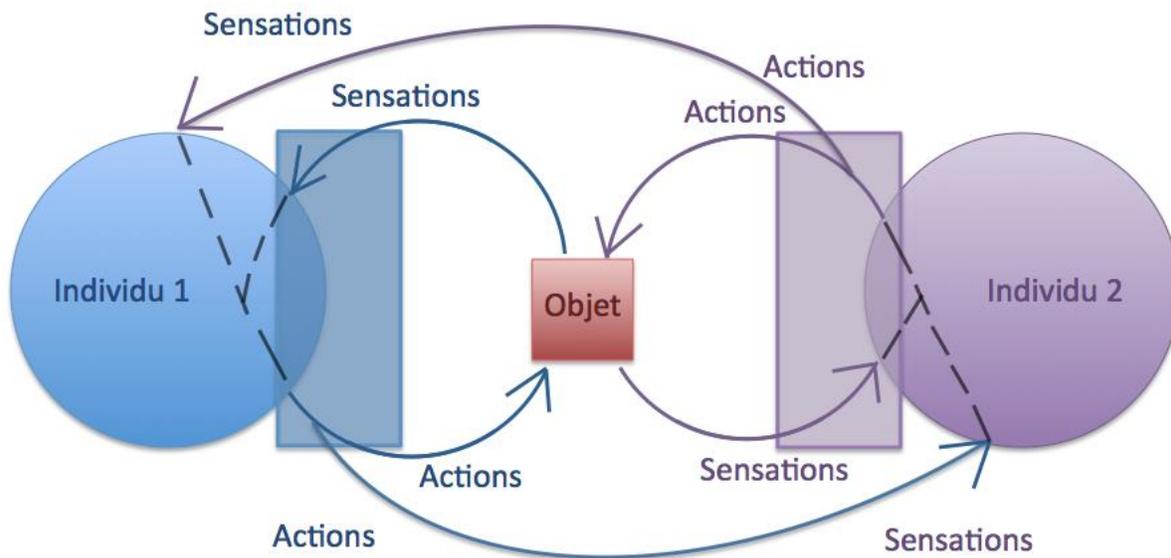


Figure 117.c. Modèle théorique du croisement perceptif ayant lieu entre deux individus qui partagent la perception d'un même objet de manière conjointe.

Les actions perceptives s'accrochent l'une l'autre ; ce n'est pas seulement que l'on perçoit l'activité perceptive d'autrui. La perception d'un objet extérieur est à présent incluse dans notre modèle théorique du croisement perceptif entre deux individus. Dans ce modèle, les deux cercles les plus externes de la figure indiquent la perception réciproque entre les deux sujets, perception qui permet de se focaliser conjointement sur un même objet.

Notre exemple de la figure 116 correspond à ce modèle : l'individu 1 est la femme, qui regarde et qui pointe quelque chose sur l'arbre (qui est donc l'objet de son attention). Lorsque les perceptions réciproques des individus se coordonnent, l'homme (individu 2) entreprend des actions en relation avec celles de la femme et regarde lui aussi en direction de l'arbre. La perception est toujours dans la boucle actions-sensations. Les actions perceptives dirigées vers l'objet indiquent ce que perçoit le partenaire, ce qui permet le partage d'un espace d'attention commun. Les actions entreprises par le partenaire envers l'objet sont perçues comme faisant partie de l'activité perceptive d'autrui. Ainsi, je me concentre sur l'objet, mon partenaire se concentre sur l'objet, et nous savons tous deux que chacun est concentré sur l'objet (et sait que l'autre l'est aussi).

Ce qui est capital dans ce type d'interaction, c'est bien que les deux attentions s'entraînent réciproquement, de manière assez semblable à la façon dont deux activités perceptives s'attirent lors d'un croisement perceptif : je regarde l'objet, mon partenaire regarde

l'objet, et je sais que mon partenaire regarde l'objet, tout comme il sait que je regarde l'objet (et ainsi à l'infini : je sais que mon partenaire sait que je regarde l'objet, il sait que je sais qu'il regarde l'objet, etc.) « Your perceptual experience is just the same whether you are engaged in solitary attention to [an object] or in joint attention to it. What is distinctive of the case in which you and another are jointly attending is that (a) you are monitoring the direction of the other person's attention, and (b) one of the factors controlling the direction of your own attention is the direction of the other person's attention » (Campbell, 2005, p.287)³³.

L'interaction sociale impliquant un objet est donc un mécanisme attentionnel mettant en jeu une capacité de coordination complexe, où autrui n'est plus une cible (contrairement à l'interaction dyadique) mais une source perceptuelle qui permet d'organiser son propre comportement (Gomez, 1991). Les signaux émis par autrui (comme les mouvements des yeux et de la tête) sont informatifs, étant donné qu'ils signalent la présence potentielle d'un objet intéressant dans l'espace. Ils agissent donc comme des affordances sociales, qui offrent des possibilités d'action dans le contexte de l'interaction (Krueger, 2011). On peut définir l'attention conjointe comme un mécanisme fondamentalement perceptif (Butterworth et Jarrett, 1991), qui acquiert du sens au fur et à mesure des expériences sociales où il est mis en jeu (Racine et Carpendale, 2007). La perception mutuelle d'un objet, et la conscience de cette perception seraient le produit de la rencontre des activités perceptives, qui établissent un *participatory sense-making* dans lequel la connaissance serait co-construite (De Jaegher et Di Paolo, 2007).

Ramener l'attention conjointe dans le cadre du croisement perceptif nous permet de comprendre les mécanismes mis en jeu lors de ce type particulier d'interaction autour d'un objet, mécanismes au nombre desquels on trouve la coordination perceptive. Dans une perspective énonciationniste, la situation d'interaction est comprise par les interactants par le biais des activités perceptives, donc dans la situation elle-même. Mais comment cela se réalise-t-il ? S'intéresser, chez l'enfant, à la mise en place du mécanisme de l'attention conjointe va nous offrir des éléments afin de répondre en partie à cette question (voir Butterworth et Jarrett, 1991 ; Reddy, 2003).

³³ Traduction personnelle « L'expérience perceptive est exactement la même que l'on porte son attention sur un objet de manière solitaire ou de manière conjointe. Ce qui est distinctif dans le cas où l'on porte son attention de manière conjointe est que (a) on oriente la direction de l'attention de l'autre personne et que (b) un des facteurs orientant la direction de sa propre attention est la direction de l'attention de l'autre personne. »

3.2.3 Les étapes de développement de l'attention conjointe

Pour comprendre les étapes de l'adoption d'une nouvelle interface et des nouvelles possibilités d'interaction sociale chez les adultes, il peut être très intéressant de s'intéresser aux étapes du développement où ces capacités de cognitions sociales se mettent en place chez l'enfant. Il s'avère en effet éclairant de regarder dans ce cadre le développement et la mise en place de l'attention conjointe. Il s'agit de décomposer le problème en en regardant la genèse. Cela se révèle d'autant plus pertinent que nous voulons mettre en place des situations d'interaction nouvelles, qui nécessiteront d'apprendre à mettre en place une forme nouvelle d'attention conjointe. Butterworth et Jarrett (Butterworth et Jarrett, 1991) par exemple décomposent ainsi l'attention conjointe : il s'agit de passer du croisement perceptif (entre les deux interactants) à un suivi de mouvement dans la direction du regard d'autrui, puis d'effectuer une bascule vers l'objet. On se rapproche des mécanismes possibles qui seraient le résultat de la dynamique d'interaction, et non le résultat d'un raisonnement.

Comment le complexe mécanisme de l'attention conjointe se met-il en place ? Les recherches concernant l'attention conjointe portent en majorité sur l'apparition de cette dernière au cours du développement de l'enfant, généralement vers 12 mois (« Infants follow the direction of other people's gaze in the first year of life » (Scaife & Bruner, 1975).³⁴). Sont aussi mises en avant les conséquences de l'apparition de l'attention conjointe, notamment dans l'acquisition du langage et dans le développement de capacités d'interaction (« In this chapter I want to examine one specific aspect of joint attention: that is the production of the pointing gesture by infants. In particular, I will report my own work on this topic and discuss it in a broader joint attention context. I will try to argue that pointing, like Harlequin, serves two masters: social cognition and language development. » (Franco, 2005, p.129)³⁵). Ainsi, la plupart des études présentent des expériences mettant en jeu des bambins ou des enfants autistes (on observe en effet un déficit en attention conjointe chez les personnes autistes). Racine et Carpendale suggèrent de plutôt considérer les diverses formes d'interactions pour lesquelles l'attention conjointe est constitutive de divers degrés de compréhension (Racine et Carpendale, 2007).

³⁴ Traduction personnelle : « Les bébés commencent à suivre la direction du regard d'autres personnes au cours de la première année de vie ».

³⁵ Traduction personnelle : « Dans ce chapitre je souhaite prendre en considération un aspect spécifique de l'attention conjointe, à savoir le fait que les bébés se mettent à pointer du doigt. En particulier, je vais rapporter mon propre travail sur ce sujet et le mettre en perspective dans le contexte plus large de l'attention conjointe. Je vais essayer de démontrer que le fait de pointer du doigt répond à deux objectifs : la cognition sociale et le développement du langage. »

Lors du développement de l'attention conjointe chez l'enfant, deux étapes principales sont mises en place : d'une part tous les mécanismes qui permettent à l'enfant de focaliser son attention sur un élément désigné par l'adulte (comme par exemple suivre le regard de l'adulte), et d'autre part tous les mécanismes permettant à l'enfant d'attirer l'attention de l'adulte sur un élément (par exemple en pointant du doigt). Ces deux étapes sont autant d'indices de mécanismes mis en jeu lors d'une interaction conjointe autour d'un objet.

La première étape lors de l'acquisition de l'attention conjointe est donc le suivi de direction d'attention ; le bébé se met par exemple à suivre la direction du regard de l'adulte. « Butterworth and colleagues (e.g., Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991; see also Morissette, Ricard, & Gouin-Decarie, 1995) have shown that, initially, when following gaze, infants do not follow the gaze to the particular object that the other is looking at. Rather, they turn to the appropriate side, but then focus their own gaze on the first object that they see, whether or not it is the one at which the adult is looking. Only in the 2nd year will infants locate the particular object to which the other has turned. » (Moore, Angelopoulos et Bennett, 1997, p.84)³⁶. On comprend bien que le suivi de regard ne mène pas encore forcément à l'attention conjointe, car l'enfant ne focalise pas systématiquement son attention sur l'objet sur lequel l'adulte porte son attention. L'enfant cherche au début de l'acquisition un objet dans son propre champ de vision, dans la direction du regard de l'adulte ; c'est tout d'abord la détection du changement de cette direction qui est clé. Des mécanismes de plus en plus complexes vont être progressivement mis en place par l'enfant pour déterminer avec plus de précision quel est l'objet sur lequel l'adulte porte son attention ; s'aidant du mouvement et de la position de la tête et des yeux de l'adulte, il va chercher l'objet pouvant attirer l'attention dans cette direction, tout d'abord en se restreignant à son propre champ de vision, puis en réussissant à en sortir (Butterworth et Jarrett, 1991). L'enfant utilise donc les indices comportementaux d'autrui afin de détecter la direction de son focus attentionnel (Mundy, 1995).

La première étape consiste donc à ce que l'enfant cherche à fixer son attention sur le même objet que l'adulte, à l'initiative de ce dernier. Mundy et Newell considèrent que ce n'est qu'un des pendants de l'attention conjointe ; on peut en effet soit répondre à l'attention d'un autre, soit chercher à *initier* son attention. Ainsi, la deuxième étape consiste à diriger activement l'attention d'autrui, par exemple en regardant ou en pointant du doigt un objet sur lequel on souhaite attirer l'attention. « Responding to joint attention (RJA) refers to infants' ability to follow the direction of the gaze and gestures of others in order to share a common point of

³⁶ Traduction personnelle : « Butterworth et ses collègues ont montré que, initialement, quand ils suivent un regard, les bébés ne suivent pas le regard jusqu'à l'objet spécifique que l'autre regarde. C'est plutôt qu'ils se tournent du bon côté, mais qu'ensuite ils concentrent leur propre regard sur le premier objet qu'ils voient, que ce soit ou non celui que l'adulte est en train de regarder. »

reference. Alternatively, initiating joint attention (IJA) involves infants' use of gestures and eye contact to direct others' attention to objects, to events, and to themselves. The function of IJA is to show or spontaneously seek to share interests or pleasurable experience with others. » (Mundy et Newell, 2007, p. 269)³⁷. L'enfant utilise donc intentionnellement l'orientation de son propre focus attentionnel ou des gestes de pointage dans le but de diriger le focus attentionnel d'autrui (Mundy, 1995).

Ces deux étapes sont l'illustration de la différence à faire entre ce qui est de l'ordre de la *réponse* à l'attention conjointe, et ce qui est de l'ordre de l'*initiation* de l'attention conjointe (Mundy et Acra, 2006). Ces deux ensembles de comportements présentent des trajectoires développementales différentes (Mundy et al., 2007), mais pourtant complémentaires.

Heal (2005) récapitule ainsi les étapes d'acquisition de l'attention conjointe : suite à l'attention que l'enfant et la mère se portent de façon intense (yeux dans les yeux, sourires, imitation,...) au cours des premières semaines et des premiers mois de la vie, à partir du sixième mois environ apparaissent des épisodes d'attention dirigée vers des objets extérieurs à la dyade³⁸. Ces épisodes font donc entrer en jeu trois éléments : l'enfant, l'adulte et la partie du monde sur laquelle ils se concentrent tous deux. En premier, vers 12 mois, se met en place la tendance de l'enfant à suivre le regard de l'adulte vers l'objet extérieur. Les variantes de cette capacité se développent ensuite peu à peu, comme par exemple la vérification par le regard dans le cas où l'enfant n'arrive pas à identifier un objet auquel s'intéresser après avoir suivi le regard de la mère. Des mécanismes de plus en plus sophistiqués se mettent en place de manière progressive.

³⁷ Traduction personnelle : « La réponse à l'attention conjointe fait référence à la capacité des enfants à suivre la direction du regard et des gestes des autres dans le but de partager un point de référence commun. Alternativement, l'initiation d'une attention conjointe implique l'utilisation par les enfants de gestes et d'un contact du regard en vue de diriger l'attention de l'autre sur certains objets, certains événements, et sur soi-même. La fonction de l'initiation d'attention conjointe est de montrer ou de chercher spontanément à partager de l'intérêt ou des expériences agréables avec les autres. »

³⁸ Reddy soutient à ce sujet : « L'attention conjointe dirigée vers un objet mise en jeu à la fin de la première année est généralement considérée comme le témoin de la découverte par l'enfant de l'attention des autres. Je soutiens que l'attention mutuelle mise en jeu au cours des premiers mois de la vie fait déjà entrer en jeu chez l'enfant une conscience de la direction de l'attention de l'autre. Pour l'enfant, c'est lui-même qui est le premier objet de cette focalisation ; elle s'étend ensuite graduellement à des 'objets' de plus en plus distants. » (Reddy, 2003, p.397) Traduction personnelle : « Joint attention to an external object at the end of the first year is typically believed to herald the infant's discovery of other people's attention. I will argue that mutual attention in the first months of life already involves an awareness of the directedness of attention. The self is experienced as the first object of this directedness followed by gradually more distal 'objects'. »

L'enfant se met aussi à chercher à diriger l'attention de l'adulte vers quelque chose, par exemple en mettant en place une manœuvre de pointage.

L'étude de l'acquisition de l'attention conjointe par l'enfant met en évidence divers mécanismes intéressants à prendre en considération lors de la conception d'interfaces permettant d'agir ensemble à distance sur du contenu numérique partagé. En effet, savoir quand autrui a accès à la même chose que nous, et pouvoir attirer son attention sur un objet en particulier, sont des aspects clés pour que ce type d'interaction soit efficace.

3.2.4 Inscription dans le mouvement et dans le contexte

Nos actions et interactions de tous les jours se réalisent comme mouvement spatial, et se produisent dans un contexte social. Il y a un mouvement local d'une part, et d'autre part le sens, qui se rapporte au contexte en tant qu'environnement. La majorité des actions et interactions de tous les jours se réalisent en effet comme mouvement corporel spatial et en parallèle toujours dans un environnement (technique et social), un contexte. Le mécanisme de l'attention conjointe répond à cette idée. Il ne peut en effet avoir lieu sans mouvement (il me faut agir pour percevoir, et je dois décaler mon point de vue entre l'objet à percevoir et autrui) et est porteur de sens (nous nous mettons d'accord sur l'objet à explorer à deux et faisons sens de cette exploration). Le mouvement et le sens sont partie intégrante de nos interactions. La réciprocité et la synchronisation mises en place dans nos coordinations interactionnelles (comme par exemple dans le cas de l'attention conjointe) peuvent en effet se comprendre au travers d'une théorie dynamique du comportement (Schmidt et Richardson, 2008). On constate aussi, en ce qui concerne les émotions, que la perspective des systèmes dynamiques peut s'appliquer au développement émotionnel (Fogel et al, 1992) : les émotions ne sont alors pas vues comme des états mais comme des processus dynamiques auto-organisés, intimement liés au cours des activités individuelles en contexte. D'autre part, les processus émotionnels sont intégrés dans les systèmes socioculturels.

Intéressons-nous de plus près à l'attention conjointe comme étant inscrite dans le mouvement et dans le contexte. En ce qui concerne le mouvement, la détection du focus attentionnel d'autrui dérive de l'intégration de la direction du regard et de l'orientation corporelle, parallèlement à différents indices verbaux, plus que la direction du regard en soi (Vaughan Van Hecke et Mundy, 2007). L'enfant peut donc s'engager dans le suivi de regard via les aspects dynamiques du mouvement d'orientation attentionnelle (Moore, Angelopoulos et Bennett, 1997). Le mouvement joue un rôle crucial dans la capacité à suivre le regard d'un partenaire ; Moore et al (1997) ont déterminé que les caractéristiques dynamiques de la tête de l'adulte jouaient un rôle

plus important dans le suivi de regard par l'enfant que l'orientation statique finale de la tête. La perception de l'ensemble des comportements de l'agent interactif, qui s'oriente dynamiquement dans un monde d'objets, permet de déterminer où se situe le focus intentionnel de l'agent. La production d'un mouvement d'orientation dans une direction donnée par l'agent est un indicateur de la présence d'un objet ou d'un événement à percevoir dans cette direction.

Il est frappant de constater que les travaux classiques menés sur l'attention conjointe sont majoritairement conduits dans la modalité visuelle ; il nous semble pourtant que se limiter à une seule modalité pour étudier le phénomène réduit l'universalité de la question. D'autant que la vue n'est pas nécessaire pour réaliser l'attention conjointe : des personnes souffrant de déficience visuelle peuvent en effet la mettre en place. Ainsi, l'attention conjointe peut être considérée en étudiant des enfants aveugles, ou des enfants et leurs parents aveugles (Preisler, 1991 ; Bigelow, 2003 ; Hobson, 1993 ; Ingsholt, 2000 ; Rattray et Zeedyk, 2005). L'attention conjointe se réfère traditionnellement à la coordination des foci attentionnels via la modalité visuelle, mais elle peut également dériver d'informations multimodales (Baron-Cohen, 1997), telles que les informations tactiles ou auditives (Sarriá, Gómes et Tamarit, 1996). Par exemple, dans un contexte d'attention conjointe, une fréquence élevée de vocalisations de la part de la mère est observée, et ce, quel que soit le statut visuel des participants (Rattray et Zeedyk, 2005). Ces indications vocales fournissent d'une part à un enfant déficient visuel une présence rassurante, et d'autre part elles encouragent ou restreignent l'action elle-même, en renseignant l'enfant sur son déroulement et sur sa signification en fonction de l'état du monde. De plus, elles rendent possible la constitution de patterns d'interaction synchronisés et réciproques (proto-conversations) ; cette constitution est une étape nécessaire à l'acquisition du langage par la consolidation de « formats » structurants et signifiants (Bruner, 1983).

En ce qui concerne le sens du toucher, son utilisation est principalement influencée par le statut visuel de l'enfant engagé dans la situation d'attention conjointe (Rattray et Zeedyk, 2005). Les mères d'enfants déficients visuels, qu'elles soient elles-mêmes déficientes visuelles ou non, utilisent en effet le toucher actif pour enrichir le processus d'interaction, et ce en particulier lors des interactions impliquant des objets. Les mères apportent dans ce cas les objets à l'enfant et elles l'assistent dans leur manipulation de manière à compenser l'absence de retour visuel. La mère « montre » ainsi à l'enfant déficient visuel ce qu'il y a à faire et à percevoir par le biais du toucher actif, quand les mères d'enfants voyants peuvent dans ce contexte donner des indications à l'enfant en passant par le regard ou le pointage. Dans tous les cas, l'adulte structure le contexte de l'interaction, en exposant à l'enfant les affordances disponibles et en assistant ses actions sur les objets. Chez le non-voyant, cette structuration de l'interaction permet de favoriser d'une part l'acquisition de connaissances spatiales grâce à l'aide de la mère (Bruner, 1983), et d'autre part la découverte du fait que l'expérience qu'ils font des objets peut être partagée (Bigelow, 2003).

Ainsi, la déficience visuelle de l'un des partenaires ou même des deux ne réduit pas intrinsèquement la qualité des échanges, mais implique la co-construction de modes de communications alternatifs (Pérez-Pereira et Conti-Ramsden, 1999).

Or ces modes d'interaction alternatifs sont révélateurs des activités perceptives et sociales. En effet, souvent paradoxalement, ils facilitent l'observation des activités perceptives. Par exemple pour les personnes aveugles, la modalité tactile étant nécessairement proximale, on constate dans ce cas un geste d'exploration que nous pouvons observer et qui est indispensable à la perception. C'est alors qu'on parle de perception haptique, ou tactilo-kinesthésique (Hatwell, 2003, voir partie 2.2.2). Dans le cadre de l'interaction, le toucher actif désigne alors la coordination de ces activités déployées dans l'espace et dans le temps. En ce qui concerne l'audition, il s'agit d'un flux d'informations sonores ou verbales, dont le déploiement est cette fois purement temporel (Gibson 1966). L'échange de ces informations nécessite alors une coordination temporelle entre les participants, par le biais de la mise en place d'une alternance de tours de parole. L'usage de ces modalités perceptives met en évidence le fait que la focalisation attentionnelle commune dérive d'une coordination spatio-temporelle des activités perceptives des interactants. Dans ce sens, la coordination qui a lieu dans l'attention conjointe est d'abord motrice et affective (Gallagher 2011), et peut même être décrite comme une action conjointe basique (Fiebich et Gallagher, 2012). L'attention conjointe est alors considérée comme étant une séquence spécifique d'activités coordonnées, qui prennent place dans le cadre d'un contexte interactif plus global (Kidwell et Zimmerman, 2007). Enfin, si le toucher et l'audition rendent plus aisément observable le déploiement de ces activités, ce déploiement est également nécessaire en ce qui concerne la vision : à chaque instant, un grand nombre d'informations spatialisées peuvent être capturées par la multiplicité des champs récepteurs de la rétine, cependant la perception visuelle ne se réalise que lorsque le sujet maîtrise les lois de couplage qui relient son action (le mouvement de l'œil, de la tête ou du corps dans l'espace) avec la variation des stimulations visuelles captées par la rétine (O'Regan et Noë, 2001).

Il est frappant que le visuel soit aussi privilégié pour les études concernant l'attention conjointe, même si on le comprend. En effet, le caractère distal de la modalité visuelle permet une séparation entre le point de vue, l'objet, et le partenaire. Cependant, nous nous concentrons sur la recherche des conditions générales permettant l'attention conjointe, lesquelles ne sont pas forcément liées à une modalité perceptive de type visuelle (où sont séparés un point de vue et un objet perçu). En continuité avec le cadre théorique adopté dans notre travail, nous allons travailler ici en nous basant d'une part sur une approche de type tactile, qui nous semble pour des raisons déjà présentées dans la partie 2 porteuse sur le plan émotionnel, et d'autre part sur une approche minimaliste afin de dégager au mieux les conditions nécessaires et les mécanismes mis en jeu lors de l'attention conjointe.

Le dispositif Tactos développé dans notre laboratoire (voir partie 1.1.2.2) permet la mise en place d'une coordination perceptive autour d'objets communs, bien qu'il ne permette pas de perception visuelle, et n'en soit pas non plus un substitut. L'utilisateur de Tactos a accès à des informations proprioceptives (à savoir les informations kinesthésiques de son propre mouvement d'exploration, voir Auvray et al, 2005), mais il n'a pas accès aux informations kinesthésiques liées au mouvement de l'autre utilisateur. L'utilisateur doit relier son propre mouvement aux retours tactiles séquentiels qu'il reçoit, lesquels résultent de la rencontre de son corps-percevant avec un objet de l'espace, qu'il soit fixe ou mobile (par exemple un objet en mouvement ou le corps-image d'un autre utilisateur). Ainsi, le dispositif ne permet pas la perception directe ou immédiate du mouvement : l'utilisateur n'accède qu'aux positions successives des objets rencontrés, en fonction de son propre mouvement d'exploration. Malgré la prégnance de la modalité visuelle dans la compréhension du mécanisme de l'attention conjointe, il nous semble qu'un dispositif comme Tactos peut être utile pour l'étudier, notamment du fait qu'il permette l'étude des interactions et de leur organisation. Ainsi, on pourra étudier avec Tactos la façon dont l'interaction s'organise et dont les partenaires arrivent à coordonner leurs perspectives attentionnelles autour d'objets partagés.

L'attention conjointe fait typiquement référence à la mise en place d'une focalisation attentionnelle mutuelle pour l'engagement interpersonnel dans une séquence d'interaction. Il est à noter que pour que cette interaction puisse se maintenir dans le temps, la coordination des foci attentionnels est également nécessaire. Le maintien de l'engagement mutuel dépend d'un ensemble de coordinations attentionnelles locales, lesquelles sont inscrites dans une routine d'interaction (Bruner, 1983). L'attention conjointe n'est donc pas qu'une simple procédure de désignation ; c'est une structure d'échange permettant d'une part de coordonner les foci attentionnels autour d'un même objet ou d'une même activité, mais aussi et surtout d'autre part de maintenir cette structure dans le temps (ibid). L'attention conjointe se réfère typiquement à la coordination des foci attentionnels via la modalité visuelle, mais elle peut également dériver d'informations multimodales (Baron-Cohen, 1997). L'attention conjointe visuelle est établie lorsque les partenaires impliqués alternent leur regard entre l'objet de l'attention et autrui (Mundy et al., 2007). L'engagement attentionnel des individus est pris dans un jeu de co-ajustements sensori-moteurs qui dépendent des modalités perceptives disponibles (soit des retours sensoriels couplés à des possibilités d'action spécifiques), et qui définissent la signification même des comportements en fonction du contexte et des objets présents. L'attention conjointe est une pratique sociale incarnée.

La coordination des foci attentionnels ne fait donc d'autre part sens que dans le cadre d'un contexte pragmatique qui donne une structure à l'orientation attentionnelle de l'individu et à celle du partenaire. L'attention conjointe se trouve alors enracinée dans une pratique partagée,

pratique qui fait sens pour les participants en fonction du contexte dans lequel l'interaction se déroule (Carpendale et Lewis, 2004).

Les capacités d'engagement des individus dans le cadre d'une interaction impliquant un objet, qu'elles conduisent ou non à un suivi de regard, se manifestent dans des contextes pragmatiques, concrets où la présence même d'objets donne du sens à l'engagement de chacun des interactants. L'objet pour l'individu est une affordance qui attire le focus attentionnel et organise l'action. Dans le contexte d'une interaction, l'objet présent dans le monde partagé peut être vu comme une *affordance simultanée* (Marsh et al, 2006) : il attire réciproquement les foci attentionnels, et fournit une opportunité d'action identique pour les interactants. L'objet commun lui-même peut donc être conçu comme l'attracteur des activités perceptives, organisant alors la rencontre et donnant lieu à une coordination interpersonnelle.

Nous postulons que la mise en place de cette coordination se réalise par le biais de l'émergence d'une dynamique relationnelle autonome (De Jaegher, Di Paolo et Gallagher, 2010 ; Lenay et Stewart, 2012). La *conscience* ou la *connaissance* du fait que l'objet de l'attention est partagé découle de cette dynamique, car c'est à travers elle que peut se co-construire le sens via un *participatory sense-making* (De Jaegher et Di Paolo, 2007). Ce qui permet de donner du sens aux contributions respectives des interactants, c'est la rencontre de leurs engagements dans le monde. Dans le même temps, cette rencontre structure ces mêmes engagements. La coordination interpersonnelle ne concerne pas seulement un co-ajustement des activités individuelles dirigées l'une vers l'autre, mais elle concerne un co-ajustement des activités individuelles dirigées successivement l'une vers l'autre et vers un objet partagé. Pour mettre l'accent sur la dynamique de la rencontre des activités perceptives, il est donc nécessaire d'inclure l'influence de l'objet lui-même sur le processus de coordination interpersonnelle. L'attention conjointe va alors au-delà des compétences individuelles et d'une simple relation en face-à-face. Il faut élargir le cadre conceptuel de l'interaction afin d'y inclure dès le départ un monde de relations interpersonnelles ainsi que des objets partagés. Chapman propose d'adopter le « triangle épistémique » en tant que cadre d'analyse des interactions (Chapman, 1991, p. 211). Il s'agit alors de considérer les relations mutuelles entre le sujet, l'objet de sa connaissance et un autre sujet, qu'il soit présent ou non (ibid.). La compréhension de l'attention conjointe nécessite la pleine prise en compte du contexte de l'interaction, des actions individuelles respectives se déployant dans ce contexte, et du processus d'interaction qui émerge de la rencontre des activités individuelles.

Dans ce cadre de l'attention conjointe inscrite dans le mouvement et le contexte, nous proposons une expérimentation qui vise à étudier l'engagement interpersonnel dans un espace numérique partagé. Cette expérimentation a plusieurs objectifs complémentaires.

Premièrement, il s'agit de tenter de démontrer expérimentalement que la coordination interpersonnelle des activités perceptives (dirigées vers autrui et vers les objets) est une condition

nécessaire et suffisante pour l'attention conjointe. Nous entendons ici l'attention conjointe en tant que processus permettant d'aboutir à une connaissance du fait que les foci attentionnels des interactants sont dirigés vers le même objet. Cette connaissance émerge de la rencontre interpersonnelle autour d'objets communs, qui fait apparaître une dynamique relationnelle entre les sujets et les objets, et qui sert de support à la co-construction d'un cadre de référence partagé. Pour forcer le déploiement de cette coordination, nous partirons d'un cadre minimaliste strict. Notons qu'il peut paraître étrange d'utiliser Tactos pour notre expérience, qui est un dispositif destiné aux aveugles. Cependant, nous avons vu que si les théories classiques se concentrent sur des études faisant entrer la vue en jeu, il est pourtant possible pour des personnes aveugles de mettre en place l'attention conjointe. La vue n'est donc pas nécessaire pour ce mécanisme. De plus, nous pensons qu'en ne se focalisant pas sur la vue, des éléments révélateurs de l'activité perceptive vont pouvoir être mis en évidence (de même que nos expériences minimalistes sont révélatrices). Ainsi, nous optons pour la stratégie de passer par une interface destinée aux personnes aveugles afin de pouvoir forcer le déploiement perceptif.

Deuxièmement, il va s'agir de comparer différentes configurations de corps-image/corps-percevant, afin de pouvoir observer l'effet du dispositif de couplage lui-même sur l'émergence d'une telle dynamique. Si cet axe peut se révéler pertinent d'un point de vue théorique pour tenter d'isoler les éléments fondamentaux de la coordination interpersonnelle autour d'objets partagés, il permettra également d'explorer des solutions techniques afin de mettre à disposition des utilisateurs un dispositif qui favorise au mieux l'interaction dans les espaces numériques partagés. Enfin, cette expérience va avoir pour but de mettre en avant des mécanismes mis en jeu lors de l'attention conjointe ainsi que les conditions nécessaires à sa réussite, en vue de permettre le développement d'une solution de design pour le type d'interaction qui nous intéresse (interaction à distance riche sur le plan émotionnel). Il s'agit de mettre en évidence des éléments à intégrer pour permettre l'attention conjointe, et notamment l'idée que les deux individus ont conscience que l'attention d'autrui se porte sur le même objet qu'eux.

3.3 Étude expérimentale objets communs et objets privés : une expérience minimaliste mettant en jeu l'attention conjointe³⁹⁴⁰

3.3.1 Population

24 sujets, répartis en 3 groupes expérimentaux de 8 sujets chacun (soit 4 binômes dans chaque groupe), ont participé à cette expérience. Les sujets sont des étudiants de l'Université de Technologie de Compiègne âgés de 18 à 25 ans. Ils n'ont pas de connaissance particulière du dispositif et n'ont pas connaissance des éléments théoriques mobilisés dans l'expérience.

Les groupes expérimentaux sont définis en fonction de l'ordre de passation des conditions expérimentales décrites ci-après.

3.3.2 Matériel

L'expérience se déroule dans deux salles adjacentes, ce qui nécessite la présence de deux expérimentateurs. Chaque salle contient un poste de travail destiné aux sujets. Dans une des deux salles, un ordinateur serveur permet de connecter les sujets dans un espace partagé, et de piloter les différentes phases de l'expérience.

Chaque poste de travail destiné aux sujets comprend :

- un ordinateur portable sur lequel est installé le logiciel Tactos ;
- un effecteur souris pour contrôler le déplacement du curseur. L'accélération logicielle est contrôlée et identique pour tous les sujets ;
- un stimulateur tactile composé de deux cellules Braille piézoélectriques (soit 16 picots), configurés adéquatement selon les conditions expérimentales décrites ci-après ;
- un casque audio, qui permet de délivrer des consignes simultanément aux deux sujets participant à une passation expérimentale.

Le dispositif expérimental comporte trois ordinateurs : un serveur qui va gérer les connexions et deux ordinateurs-clients destinés à chacun des deux sujets.

³⁹ Nous remercions vivement Loïc Deschamps pour sa collaboration au cours de la réalisation de cette expérience, et en particulier pour son aide inestimable concernant l'analyse statistique.

⁴⁰ Cette étude a mené à une communication orale :

- Le Bihan, G., Deschamps, L., Lenay, C. & Aubert, D. (2012). A tactile device for remote communication of emotions. Consortium of European Research on Emotion (CERE), University of Kent, Canterbury, Great Britain

L'espace partagé dans lequel se déplacent les sujets est un espace horizontal unidimensionnel de 400 pixels qui se boucle sur lui-même. Il s'agit en réalité d'un tore géométrique dans lequel le sujet n'a pas la conscience de bords ou de limites dans l'espace d'action : le franchissement d'un des bords de l'espace entraîne la réapparition simultanée au bord opposé. Du point de vue d'un observateur, le franchissement du bord gauche par un curseur se traduit par une réapparition simultanée à droite, et inversement.

À l'aide de leur main dominante placée sur l'effecteur souris, les sujets déplacent un avatar de 2 pixels de long (le corps-image), situé au centre d'un corps-percevant d'une longueur de 16 pixels dont la sensibilité et la configuration varient en fonction des conditions expérimentales. L'index de la main non-dominante du sujet est posé sur le stimulateur tactile.

Pour chaque essai, trois objets fixes d'une longueur de 2 pixels sont disposés le long de l'espace unidimensionnel. La position de ces objets est différente pour chacun des 8 essais des 3 conditions expérimentales. Parmi ces trois objets, un seul est perceptible conjointement et simultanément par les deux sujets : il s'agit d'un objet « public », ou « commun ». Les deux autres objets sont des objets « privés », c'est-à-dire que chacun d'eux n'est perceptible respectivement que par un des sujets évoluant dans l'espace partagé. En d'autres termes, pour chaque sujet, deux objets fixes sont présents à chaque essai : un objet commun et un objet privé (voir figure 118).

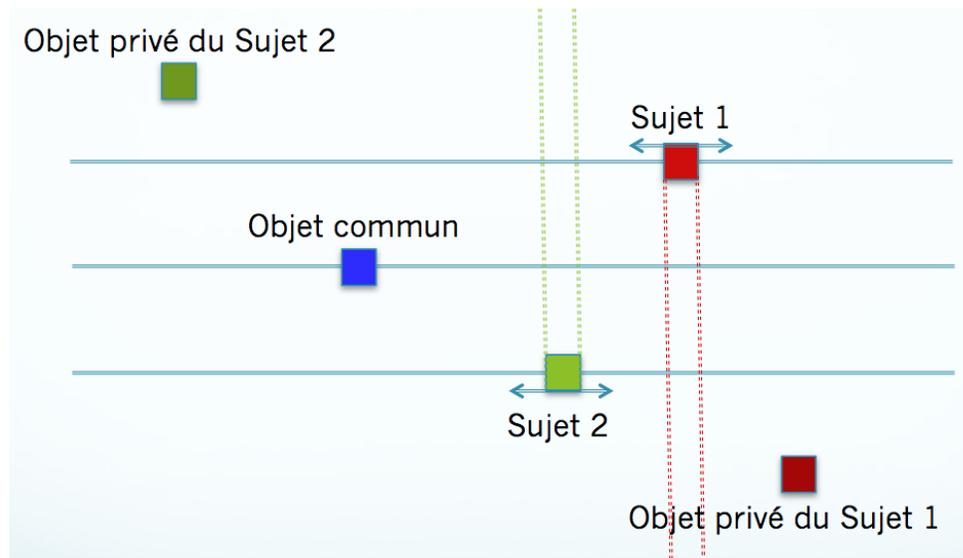


Figure 118. Schéma de la situation expérimentale pour objets communs et objets privés.

3.3.3 Procédure

Dans chaque groupe expérimental, 8 sujets sont répartis en 4 binômes. Chaque binôme passe les 3 conditions expérimentales, chacune des conditions expérimentales étant définie par une configuration différente des corps-percevants et de leur sensibilité associée.

Dans la condition « Mono », le corps-percevant est composé d'un unique champ récepteur élémentaire d'une longueur de 16 pixels. La rencontre de ce champ récepteur avec les objets ou le corps-image de l'autre participant active la montée simultanée des 16 picots de la matrice de stimulateurs tactiles (voir figure 119.a). En ce sens, la perception des objets et la perception du corps-image du partenaire conduisent à des stimulations tactiles indifférenciées.

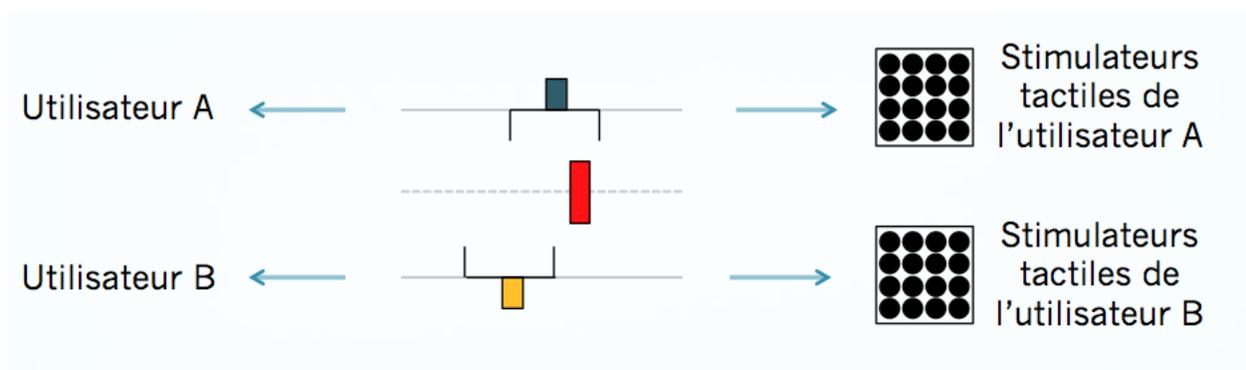


Figure 119.a. Schéma des différentes configurations des corps-percevants et de leur sensibilité associée en fonction des conditions expérimentales : condition « Mono ».

Dans la condition « Différenciée », le corps-percevant est composé d'un unique champ récepteur élémentaire d'une longueur de 16 pixels, comme pour la condition « Mono ». En revanche, la rencontre de ce champ récepteur avec les objets ou avec le corps-image de l'autre participant induit une activation différenciée des stimulateurs tactiles. En effet, la rencontre de ce champ récepteur avec un objet (qu'il soit commun ou privé) active la montée simultanée des 4 picots supérieurs de la matrice de 16 stimulateurs. À l'inverse, la rencontre de ce champ récepteur avec le corps-image de l'autre participant active la montée simultanée des 4 picots inférieurs de la matrice de 16 stimulateurs. De cette façon, il est possible de percevoir simultanément un objet et le corps-image de l'autre participant : dans ce cas, les 4 picots supérieurs et les 4 picots inférieurs sont activés simultanément (voir figure 119.b).

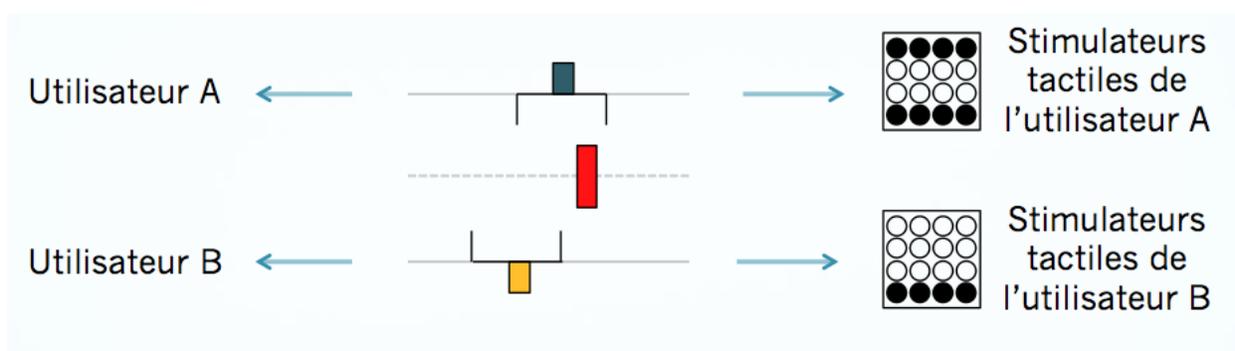


Figure 119.b. Schéma des différentes configurations des corps-percevant et de leur sensibilité associée en fonction des conditions expérimentales : condition « Différenciée ».

Dans la condition « Parallélisme », le corps-percevant est composé de 4 champs récepteurs élémentaires de 4 pixels chacun (soit une longueur totale de 16 pixels). Chaque champ récepteur élémentaire est couplé de façon topologique à chaque colonne de stimulateurs tactiles. Ainsi, lorsque le champ récepteur situé le plus à gauche du corps-percevant rencontre un objet (qu'il soit commun ou privé) ou le corps-image de l'autre participant, les quatre stimulateurs tactiles situés le plus à gauche de la matrice de stimulateurs (soit la première colonne) s'activent simultanément. De cette façon, cette condition expérimentale implique un couplage des champs récepteurs avec la matrice de stimulateurs tactiles tel que plusieurs sensations (par exemple l'une liée à l'objet, l'autre liée au partenaire) peuvent être données simultanément, mais sans être différenciées. Ici, les stimulations ne sont donc plus différenciées, mais il est tout de même possible de percevoir simultanément un objet et le corps-image de l'autre participant, à condition que les deux sources de stimulation ne se trouvent pas au même endroit et que la distance qui les sépare soit inférieure à la taille du corps-percevant (soit 16 pixels) (voir figure 119.c).

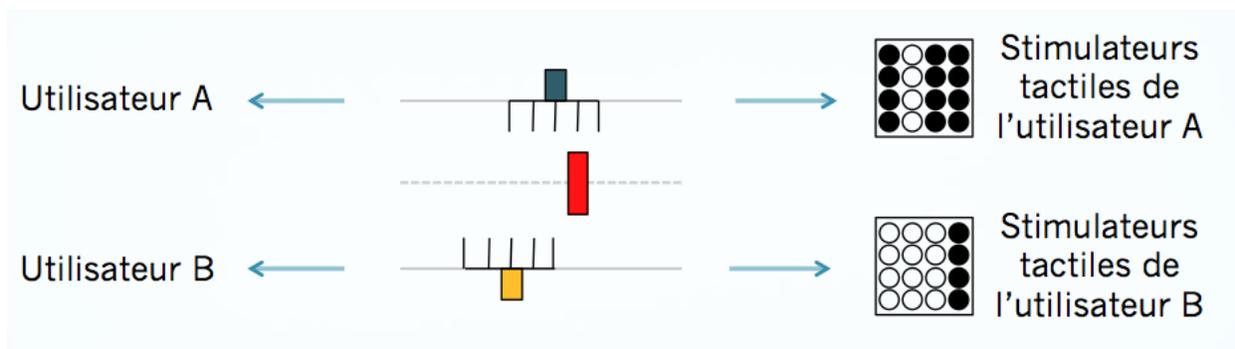


Figure 119.c. Schéma des différentes configurations des corps-percevant et de leur sensibilité associée en fonction des conditions expérimentales : condition « Parallélisme ».

Tous les binômes passent successivement les trois conditions expérimentales, mais dans un ordre différent : les groupes expérimentaux sont ainsi constitués en fonction de l'ordre de passation des conditions. Les binômes du premier groupe expérimental commencent par la condition « Mono », puis passent la condition « Différenciée » et enfin la condition « Parallélisme ». Les binômes du second groupe expérimental commencent par la condition « Différenciée », puis passent la condition « Parallélisme » et enfin la condition « Mono ». Les binômes du troisième groupe expérimental commencent quant à eux par la condition « Parallélisme », puis passent la condition « Mono » et enfin la condition « Différenciée ».

Chaque condition expérimentale est constituée de 13 essais, dont 5 essais de familiarisation et 8 essais expérimentaux.

Les passations se font en binôme. Les sujets sont accueillis dans une salle, et le dispositif leur est brièvement présenté, ainsi qu'une consigne globale décrivant l'expérience à laquelle ils vont participer. Dans cette consigne, les sujets sont informés que l'expérience va consister à interagir tactilement avec l'autre participant dans un espace partagé où sont disposés des objets fixes. Le but est de deviner lequel des objets est un objet commun, c'est-à-dire lequel des objets perçus est également perceptible par l'autre participant. Il est également annoncé que l'expérience est divisée en trois phases (les trois conditions expérimentales), chacune de ces phases étant décomposée en deux parties distinctes : une phase de familiarisation et une phase de test. Après cette consigne, les sujets sont installés dans deux salles adjacentes, de façon à isoler les participants et à éviter tout indice perceptif autre que les stimulations tactiles résultant de la rencontre entre les corps-image dans l'environnement numérique (tels que les bruits de la souris sur la table, ceux des stimulateurs tactiles ou encore les réactions des sujets lors des rencontres). Lorsque les sujets sont assis devant leur dispositif Tactos respectif, nous leur demandons d'installer le casque audio sur leurs oreilles de façon à recevoir des informations orales concernant chaque essai (le début de l'essai, la fin de l'essai, et le moment de chaque essai expérimental où les sujets pourront cliquer de façon à désigner l'objet commun). Les sujets n'ont pas les yeux bandés pendant l'expérience, mais aucune information n'est visible sur leur écran. Lorsque les sujets sont prêts le premier essai de la première condition de chaque groupe expérimental commence.

Le premier essai est un essai de familiarisation avec le dispositif et avec la condition expérimentale en cours. Les sujets sont connectés dans des espaces privés respectifs où est disposé un objet fixe de 2 pixels de long ; l'objet est donc comparable aux objets fixes qui se trouveront dans l'espace lors des essais expérimentaux. La consigne pour cet essai indique aux sujets qu'en déplaçant la souris sur un axe horizontal, ils peuvent rencontrer un objet fixe. Cette rencontre produit l'activation des picots selon la condition expérimentale en cours. Il est

demandé aux sujets de bien faire attention à la façon dont les picots se lèvent, car ils se lèveront de façon différente lors de chaque condition. Pour le premier essai de la première condition passée par les sujets, il est également demandé que les sujets se déplacent dans une direction donnée jusqu'à obtenir deux stimulations successives. Cette étape consiste à insister sur le fait que les deux stimulations n'indiquent pas la présence de deux objets, mais qu'un seul et même objet a été croisé deux fois. Ainsi, la distance entre deux stimulations équivaut à la taille totale de l'espace. Il est enfin indiqué que cette taille sera identique pour tous les essais de l'ensemble de l'expérience. Cet essai n'est pas chronométré, et dure tant que les sujets le jugent utile. Les expérimentateurs s'assurent que les sujets ont bien compris, et leur indiquent le cas échéant que l'essai suivant va commencer.

Le deuxième essai est un essai de familiarisation avec la possibilité d'interagir avec l'autre participant de l'expérience. Les sujets sont alors connectés dans un espace partagé où aucun objet n'est présent. La seule source de stimulation tactile résulte donc de la rencontre entre les deux participants. Les sujets sont informés que deux rôles leur seront attribués alternativement par l'intermédiaire du casque audio au début de l'essai. Si le sujet A est guide, le sujet B est suiveur et inversement. Le sujet guide est informé qu'un événement sonore est disposé à un endroit de l'espace et que lui seul peut l'entendre. Son objectif est de localiser cet événement, et de tenter de guider l'autre participant à cet endroit. Le sujet suiveur est informé que l'autre participant entend un son à un endroit de l'espace, et que son objectif est de suivre le guide jusqu'à ce son. Cet essai n'est pas chronométré, et s'arrête lorsque les expérimentateurs constatent que les deux sujets sont arrêtés sur l'événement sonore ou oscillent autour de lui. Lorsque c'est le cas, les expérimentateurs lancent l'essai suivant, qui est identique si ce n'est que l'événement sonore est déplacé et que les rôles s'inversent : le guide devient suiveur et vice-versa. Cet essai est reproduit en tout quatre fois, ce qui implique que chaque sujet est alternativement deux fois guide et deux fois suiveur.

Le sixième essai est le premier essai de test. Les sujets sont alors informés que l'expérience à proprement parler commence. L'objectif consiste à explorer un espace partagé où chacun des deux sujets pourra sentir la présence de son partenaire, ainsi que la présence de deux objets fixes, suivant les mêmes conditions expérimentales que les 5 essais de familiarisation précédents. Parmi ces deux objets, un objet est commun, l'autre est privé. Le but du jeu pour chaque sujet est d'interagir avec l'autre participant pendant 1 minute, et d'essayer de deviner lequel des deux objets présents est l'objet commun. À la fin de cette minute, une consigne délivrée via le casque audio indique aux sujets qu'ils disposent de 15 secondes pour cliquer sur l'objet commun. À la fin de ces 15 secondes, les expérimentateurs lancent l'essai suivant, et ce, jusqu'à la réalisation de 8 essais consécutifs.

À la fin des 8 premiers essais expérimentaux, une pause de deux minutes est proposée aux sujets, durant laquelle ils ne peuvent pas interagir entre eux. Ensuite, la deuxième phase de l'expérience peut commencer. La deuxième condition expérimentale commence alors, suivant le même scénario que la première condition passée (soit 5 essais de familiarisation suivis de 8 essais expérimentaux). Les sujets sont invités à faire attention aux modifications du type de stimulations reçues lors des essais de familiarisation, de façon à s'assurer que la nouvelle condition expérimentale est constatée et comprise par les deux sujets.

Cette deuxième phase est suivie d'une pause de deux minutes, qui précède une troisième phase où la troisième condition expérimentale est menée de manière similaire aux deux premières. À la fin de l'expérience, nous réunissons les sujets de façon à procéder à un débriefing, où les sujets peuvent nous faire part de leurs difficultés éventuelles ou expliciter leur stratégie.

3.3.4 Hypothèses opérationnelles

Quelle que soit la configuration des corps-image/corps-percevant, les sujets cliqueront davantage sur l'objet commun que sur les objets privés. Nous nous attendons à ce que cette réussite soit rendue possible en vertu de l'émergence et de la stabilisation d'une dynamique interpersonnelle autour de l'objet commun. En ce sens, nous nous attendons à ce que la réussite soit collective : la fréquence des cas où les deux sujets cliqueront mutuellement sur l'objet commun sera plus élevée que la fréquence des cas où un seul des deux cliquera sur l'objet commun.

Malgré cette réussite globale, nous nous attendons à ce que les différentes configurations de corps-image/corps-percevant aient pour effet de faire varier la nature de cette dynamique, ce qui aura une incidence sur la réussite de la tâche. En outre, la condition « Différenciée » et la condition « Parallélisme », en tant qu'elles ont vocation à surmonter des difficultés inhérentes à la condition « Mono » (respectivement indifférenciation autre/objet et impossibilité de percevoir les deux en même temps), devront conduire à plus de clics sur l'objet commun que de mauvais clics.

Toutefois, la condition « Mono » étant la condition la plus minimaliste, elle aura pour effet de forcer un déploiement maximum des activités perceptives. On suppose que même dans cette condition minimale, les sujets pourront réussir la tâche grâce à l'accroche de leur activité perceptive mutuelle et associée à l'accroche de l'activité perceptive de chacun sur la cible commune, ce qui devrait pouvoir entraîner une majorité de clics sur l'objet commun.

En outre, l'analyse des trajectoires perceptives issues de cette condition dans les cas de réussite devrait nous permettre d'isoler des paramètres objectifs pour tenter d'expliquer la nature de la réussite en termes de coordination des activités perceptives.

Enfin, quelle que soit la condition expérimentale, nous nous attendons à ce que le nombre de clics sur l'objet commun soit supérieur dans le deuxième bloc d'essais (essais 5 à 8 de chaque condition) que dans le premier bloc d'essais (essais 1 à 4 de chaque condition). Ce résultat pourra être expliqué conjointement par la stabilisation de la dynamique interpersonnelle et par l'appropriation individuelle du système de couplage lui-même, qui permettra un meilleur engagement au fur et à mesure des essais.

3.3.5 Résultats

3.3.5.1 Analyses des réponses individuelles⁴¹

Dans un premier temps, nous nous intéressons aux types de clics indépendamment des conditions expérimentales et de l'ordre de passation. Les clics peuvent être répartis en quatre catégories : les clics sur l'objet commun, les clics sur l'objet privé du sujet, les clics sur le corps-image de l'autre sujet⁴² et les clics « vides », c'est à-dire les clics situés ni sur un objet ni sur le corps-image de l'autre sujet. Nous considérons ici le nombre moyen de clics pour chaque catégorie et pour l'ensemble des sujets.

Globalement, nous observons qu'en ramenant tous les résultats des essais expérimentaux à ceux d'un sujet moyen faisant 24 clics, le nombre moyen de clics sur l'objet commun est supérieur au nombre moyen de clics sur l'objet privé du sujet ($d = 9$), au nombre moyen de clics sur le corps-image de l'autre sujet ($d = 11.87$) et au nombre moyen de clics sur des emplacements vides ($d = 12.16$). Par ailleurs, nous observons que le nombre moyen de clics sur l'objet commun est également supérieur au nombre moyen d'occurrences où les sujets ne cliquent pas ($d = 6.91$) (voir figure 120). La différence globale est significative (test de Kruskal-Wallis : $H_{(4,120)} = 87,42915$; $p < 0.0001$).

⁴¹ Par « Analyse des réponses individuelles », nous n'entendons pas une analyse de profils individuels de réponse, mais une analyse des réponses fournies par les sujets indépendamment des réponses de leur partenaire.

⁴² D'une façon générale, ces clics sont situés aux alentours de l'objet privé de l'autre participant : par exemple, si le sujet 1 explore son objet privé et que dans le même temps, les deux sujets interagissent vivement dans cette zone, les stimulations reçues par le sujet 2 peuvent lui donner l'illusion de la présence d'un objet à cet endroit.

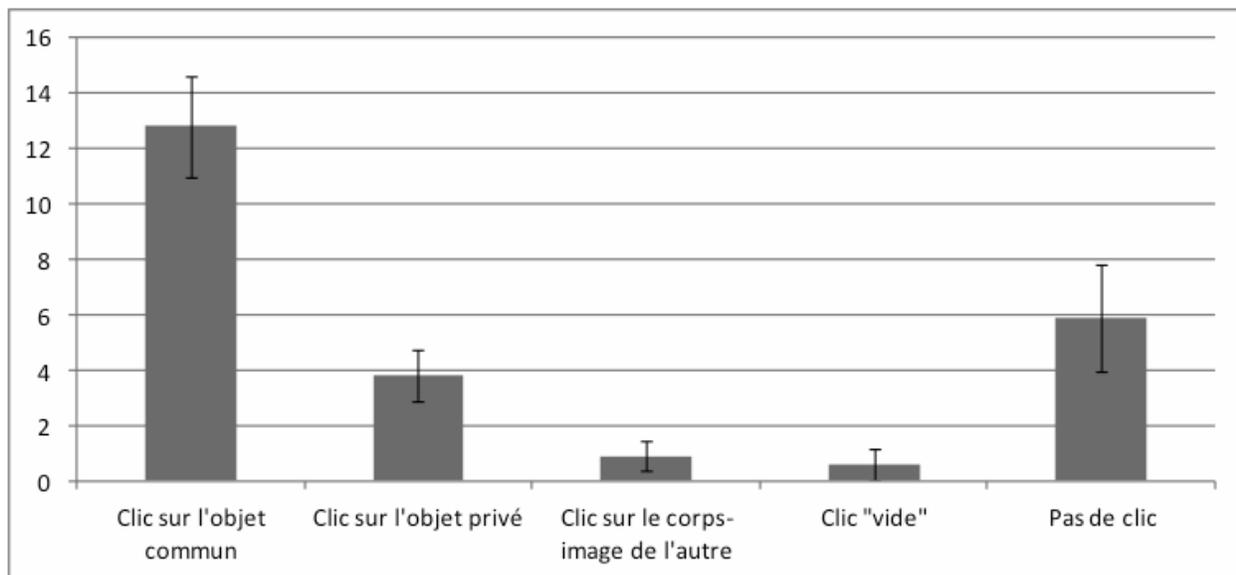


Figure 120. Nombre moyen de clics par type de clic.

De façon à simplifier les analyses, nous avons exclu les essais où les sujets n'ont pas cliqué⁴³, et nous regroupons les clics qui ne sont pas situés sur l'objet commun en une catégorie globale « Mauvais clics ». Ici encore, le nombre moyen de clics sur l'objet commun est supérieur au nombre moyen de mauvais clics ($d = 7.45$). Cette différence est significative ($t_{(23)} = 7.8369$; $p < 0.0001$).

Toutefois, nous constatons que l'effet d'interaction entre la qualité de la réponse (bons clics vs mauvais clics) et les conditions expérimentales est statistiquement significatif ($F_{(2,69)} = 31.0327$; $p < 0.0001$). En effet, si la différence globale entre le nombre moyen de clics sur l'objet commun et le nombre moyen de mauvais clics suit la même évolution dans chacune des trois conditions expérimentales, celle-ci se manifeste dans des proportions différentes : la différence observée est faible pour la condition « Mono » ($d = 0.71$) et pour la condition « Parallélisme » ($d = 1.25$), alors qu'elle est plus importante pour la condition « Différenciée » ($d = 5.5$) (voir figure 121).

⁴³ Cette exclusion nous paraît justifiée dans la mesure où une absence de clic ne nous permet pas de savoir si le sujet n'a pas souhaité fournir de réponse, ou s'il n'a tout simplement pas eu le temps de cliquer dans le temps imparti. Par ailleurs, il nous semble plus opportun de comparer les réponses effectives : soit la différence entre les clics sur l'objet commun et les mauvais clics.

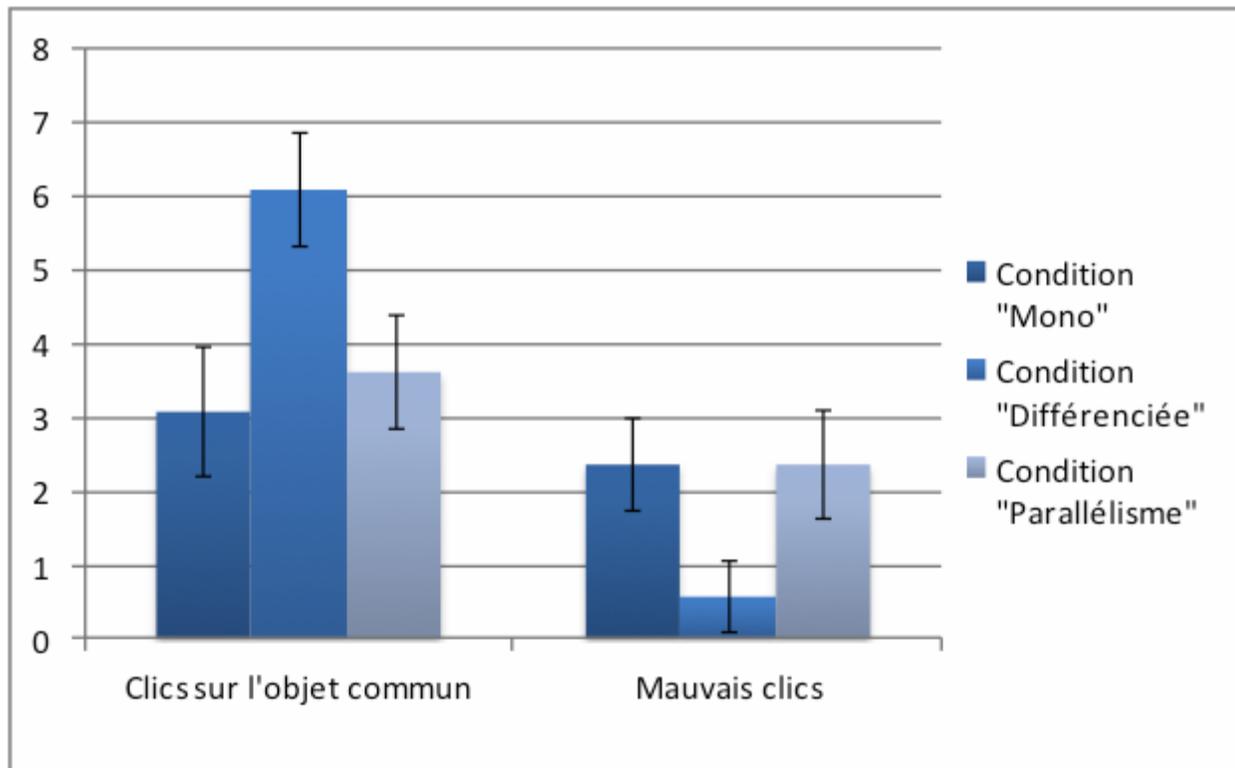


Figure 121. Nombre moyen de clics sur l'objet commun et de mauvais clics en fonction des 3 conditions expérimentales.

Si nous centrons notre analyse sur le nombre moyen de clics sur l'objet commun, nous constatons un effet global des conditions expérimentales (test de Kruskal-Wallis : $H_{(2,72)} = 28,16017$; $p < 0.0001$). De façon plus spécifique, nous observons que le nombre moyen de clics sur l'objet commun est plus faible dans la condition « Mono » que dans la condition « Différenciée » ($d = 3$) (voir figure 121). Cette différence est significative (test de Mann & Withney : $U = 60,5$; $Z_{\text{ajusté}} = 4.74722$; $p < 0.0001$). De la même façon, le nombre moyen de clics sur l'objet commun est plus faible dans la condition « Parallélisme » que dans la condition « Différenciée » ($d = 2.45$) (voir figure 121). Cette différence est significative (test de Mann & Withney : $U = 81$; $Z_{\text{ajusté}} = 4.32365$; $p = 0.00015$). Enfin, le nombre moyen de clics sur l'objet commun est plus faible dans la condition « Mono » que dans la condition « Parallélisme » ($d = 0.54$) (voir figure 121). Toutefois, cette différence n'est pas significative (test de Mann & Withney : $U = 242$; $Z_{\text{ajusté}} = 0.963283$; $p = 0.3354$). Il est à noter que cet effet des conditions expérimentales sur le nombre moyen de clics sur l'objet commun se retrouve quel que soit l'ordre de passation.

D'une façon générale, nous observons un effet d'apprentissage à l'intérieur des conditions expérimentales. En effet, toutes conditions confondues, le nombre moyen de clics sur l'objet commun est supérieur au cours des quatre derniers essais comparativement aux quatre premiers essais ($d = 1.20$). Cette différence est significative ($t_{(23)} = 2.3499$; $p = 0.0277$). Parallèlement, le nombre moyen de mauvais clics est plus important dans le premier bloc d'essais que dans le second ($d = 1.08$). Cette différence est significative ($t_{(23)} = 2.8487$; $p = 0.0091$) (voir figure 122).

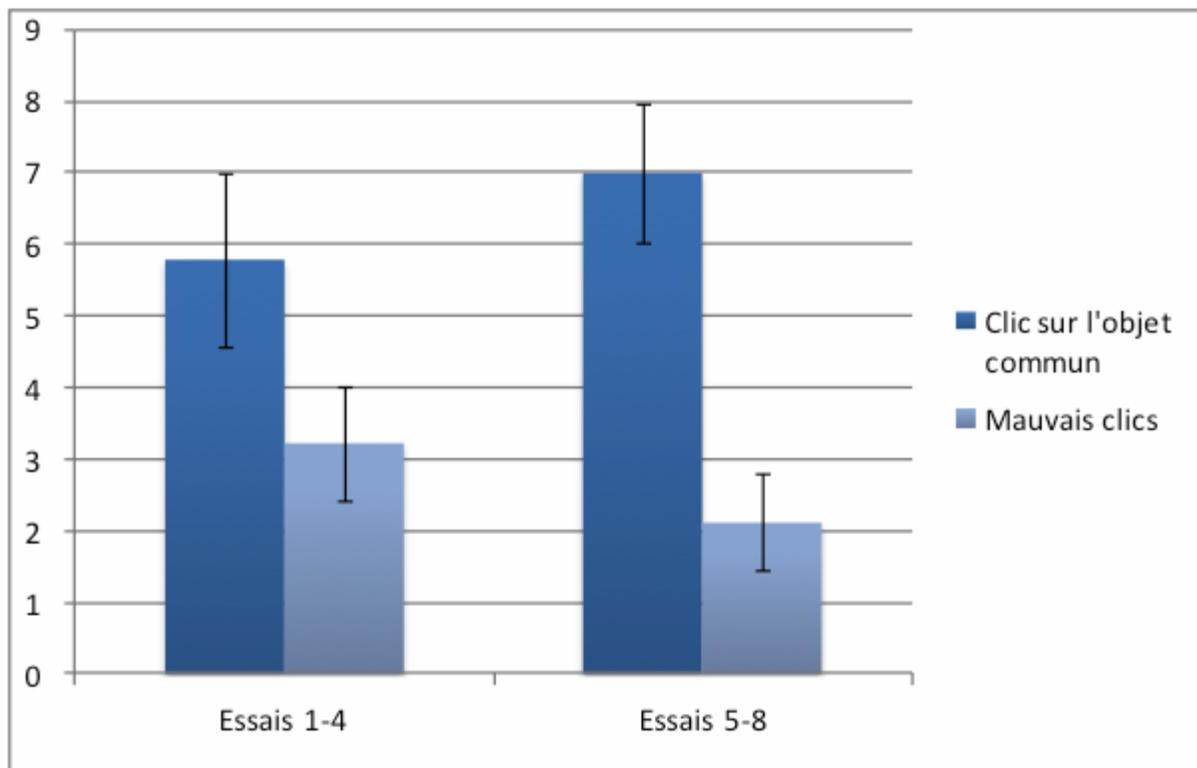


Figure 122. Nombre moyen de clics sur l'objet commun et de mauvais clics en fonction des deux blocs d'essais (Essais 1 à 4 versus Essais 5 à 8).

De façon descriptive, l'effet de l'apprentissage est très faible en condition « Différenciée » ($d = 0.08$), plus important pour la condition « Mono » ($d = 0.33$) ainsi que pour la condition « Parallélisme » ($d = 0.79$). Toutefois, nous ne pouvons pas conclure à un effet d'interaction entre le facteur « Apprentissage » et le facteur « Conditions expérimentales » ($F_{(2,69)} = 1.4932$; $p = 0.2318$).

Enfin, nous avons procédé à l'analyse des réponses d'un sujet en fonction de la réponse donnée par son partenaire. Dans toutes les conditions, on retrouve le fait que lorsque le sujet A a

bien cliqué, la probabilité que le sujet B réponde correctement est supérieure à la probabilité qu'il clique mal ou ne clique pas. Même dans la condition « Mono », la probabilité de bonne réponse correcte est alors supérieure à la somme des probabilités des deux autres types de réponses (mauvais clic et pas de clic). Au contraire, lorsque le sujet A a mal répondu, on se rend compte que la probabilité que le sujet B réponde bien chute drastiquement (16,07 % pour la condition « Mono », 10 % pour la condition « Différenciée » et 12,22 % pour la condition « Parallélisme »). Pour les conditions « Mono » et « Différenciée », la réponse la plus probable est alors pas de clic, tandis que pour la condition « Parallélisme », la réponse la plus probable lorsque le sujet A a mal cliqué est un mauvais clic. Enfin, lorsque le sujet A n'a pas cliqué, on constate que dans toutes les conditions, les probabilités que le sujet B clique correctement ou clique mal sont équivalentes. Ces résultats indiquent bien l'influence du partenaire sur la réponse fournie ; c'est grâce aux indications de mon partenaire que je donne la réponse que je donne.

Tableau 10. Proportion de réponses d'un sujet B en fonction des réponses du sujet A par condition.

Réponse du sujet A	Réponse du sujet B	Condition « Mono »	Condition « Différenciée »	Condition « Parallélisme »
Clic correct	Clic correct	59,50 %	90,48 %	68,98 %
	Mauvais clic	12,14 %	0,67 %	8,03 %
	Pas de clic	28,36 %	8,85 %	22,99 %
Mauvais clic	Clic correct	16,07 %	10,00 %	12,22 %
	Mauvais clic	38,89 %	15,56 %	59,82 %
	Pas de clic	45,04 %	74,44 %	27,96 %
Pas de clic	Clic correct	45,37 %	56,25 %	55,73 %
	Mauvais clic	54,63 %	43,75 %	44,27 %

3.3.5.2 Analyses des réponses collectives

Après avoir analysé les réponses individuelles des sujets indépendamment des réponses de leurs partenaires, nous nous intéressons à présent aux données collectives, c'est-à-dire à la qualité de la réponse du binôme pour chaque essai. Dans ce cadre, nous considérons trois cas de figure, correspondant à trois types de réponses collectives par binôme : les clics mutuels sur l'objet commun (c'est-à-dire les cas où les deux sujets ont cliqué sur l'objet commun), les clics solitaires (c'est-à-dire les cas où seul un des deux sujets a cliqué sur l'objet commun, l'autre ayant cliqué ailleurs ou n'ayant pas cliqué), et les autres cas (c'est-à-dire les cas où les deux

sujets ont cliqué ailleurs que sur l'objet commun, ou un des deux a cliqué ailleurs, l'autre n'ayant pas cliqué). Notons que dans ce dernier cas de figure, les cas où les deux sujets n'ont pas cliqué n'ont pas été pris en compte. Nous avons exclu ce cas de figure pour les mêmes raisons que lors de l'analyse des réponses individuelles.

Globalement, nous observons que la fréquence de clics mutuels est supérieure à la fréquence de clics solitaires ($d = 0.16$), et à la fréquence des autres cas ($d = 0.11$). De plus, la fréquence de clics solitaires est inférieure à la fréquence des autres cas ($d = 0.04$) (voir figure 123). La différence entre les trois cas de figure est significative ($F_{(2,33)} = 6.7480$; $p = 0.0035$).

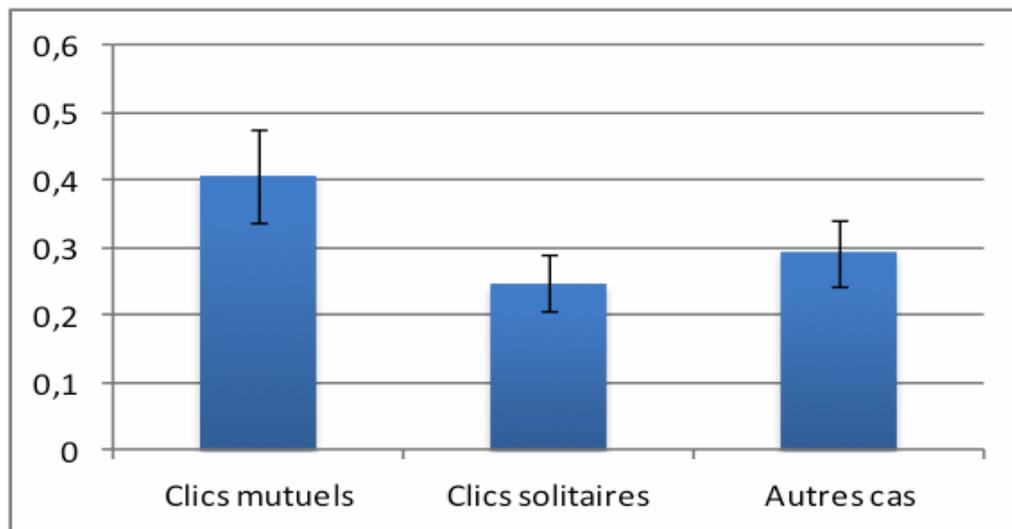


Figure 123. Fréquence des trois types de réponses collectives pour l'ensemble des binômes.

De façon descriptive, nous observons que les différences entre les trois types de réponses collectives s'expriment de façon différente selon les conditions expérimentales (voir figure 124). En effet, l'effet d'interaction entre les types de réponses collectives des binômes et le facteur « Conditions expérimentales » est statistiquement significatif ($F_{(4,66)} = 12.9981$; $p < 0.0001$).

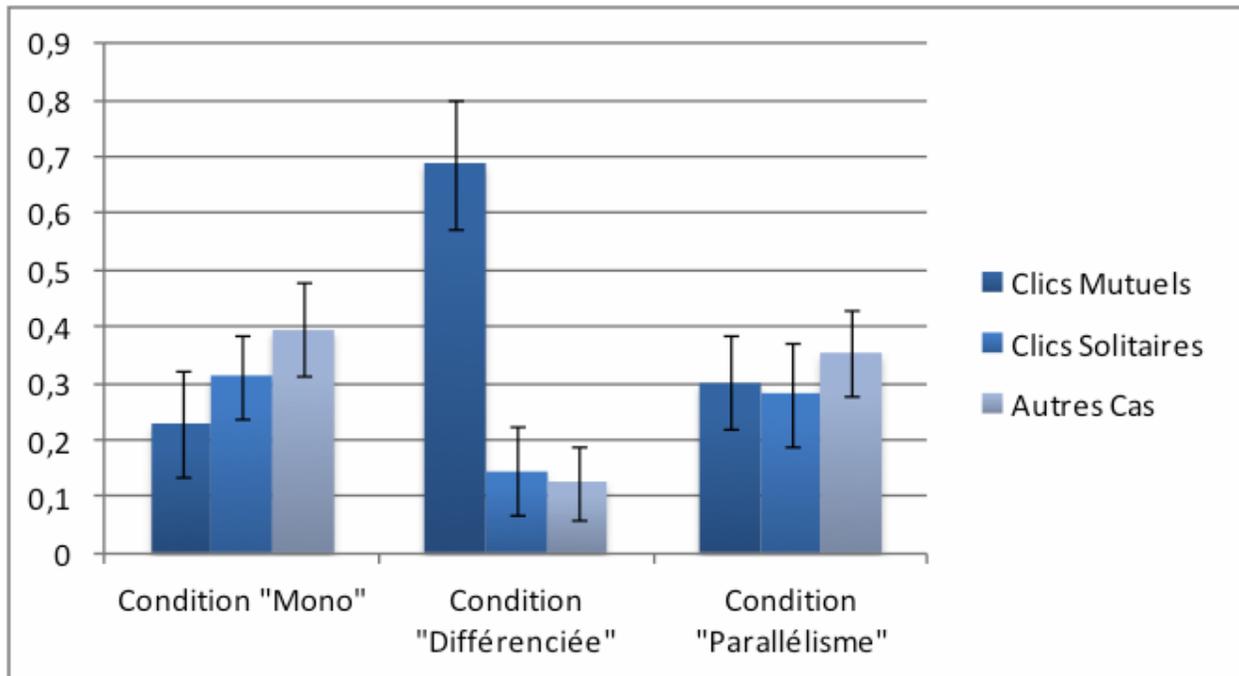


Figure 124. Fréquence des trois types de réponse en fonction des conditions expérimentales.

En ce qui concerne la condition « Mono », nous observons que la fréquence de clics mutuels est inférieure à la fréquence de clics solitaires ($d = 0.08$), elle-même inférieure à la fréquence des autres cas ($d = 0.08$) (voir figure 124). Toutefois, la différence globale n'est pas significative ($F_{(2,22)} = 2.0230$; $p = 0.1561$).

Concernant la condition « Différenciée », la fréquence de clics mutuels est supérieure à la fréquence de clics solitaires ($d = 0.54$), et à la fréquence des autres cas ($d = 0.56$). Enfin, la fréquence de clics solitaires est supérieure à la fréquence des autres cas ($d = 0.02$). La différence globale est significative ($F_{(2,22)} = 26.1250$; $p < 0.0001$). Enfin, en ce qui concerne la condition « Parallélisme », la fréquence de clics mutuels est supérieure à la fréquence de clics solitaires ($d = 0.02$), mais est inférieure à la fréquence des autres cas ($d = 0.05$). Enfin, la fréquence de clics solitaires est inférieure à la fréquence des autres cas ($d = 0.07$). Toutefois, la différence globale n'est pas significative ($F_{(2,22)} = 0.4658$; $p = 0.6337$).

Si nous nous centrons sur la fréquence de clics mutuels en fonction des conditions expérimentales, nous observons que la fréquence de clics mutuels est plus importante en condition « Différenciée » qu'en condition « Mono » ($d = 0.46$), et « Parallélisme » ($d = 0.38$). Par ailleurs, la fréquence de clics mutuels est plus importante en condition « Parallélisme » qu'en condition « Mono » ($d = 0.07$) (voir figure 124). D'une façon statistique, nous pouvons conclure

à un effet des conditions expérimentales sur la fréquence de clics mutuels ($F_{(2,33)} = 18.8427$; $p < 0.0001$).

Ces différences étant constatées, nous souhaitons ici souligner une remarque importante. Paradoxalement, les clics solitaires nous semblent en effet révélateurs d'une dynamique collective. Nous nous intéressons au type de réponse d'un sujet B lorsque le sujet A a cliqué correctement. D'une façon globale, nous observons que la proportion de mauvais clics de B est très faible (5.90%). Par conséquent, B n'a pas cliqué dans une grande majorité des cas (94.10%). En outre, en ce qui concerne les cas où ni A ni B n'ont cliqué sur l'objet commun, nous observons que lorsque A a effectué un mauvais clic, B n'a pas cliqué dans la majorité des cas (89.93% des cas). En outre, dans les 10.07% des cas où les deux sujets ont cliqué ailleurs que sur l'objet commun, ils ont cliqué au même endroit dans 75.86% des cas (soit sur l'objet privé de l'un des deux partenaires, soit dans de plus rares cas sur un emplacement vide). Ces observations se retrouvent dans chacune des trois conditions expérimentales (voir tableau 11).

Tableau 11. Proportion des types de réponse collective dans les cas des clics solitaires et des autres cas pour chaque condition expérimentale.

Cas	Réponse du sujet A	Réponse du sujet B	Condition « Mono »	Condition « Différenciée »	Condition « Parallélisme »
Clics Solitaires	Clic correct	Pas de clic	90.62%	98.96%	92.71%
		Mauvais clic	9.38%	1.04%	7.29%
Autres cas	Mauvais clic	Pas de clic	88.54%	98.96%	82.39%
		Mauvais clic	11.46%	1.04%	17.71%

Par ailleurs, d'une façon globale, nous observons que la fréquence de clics mutuels est plus élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais ($d = 0.04$) (voir figure 125). Par ailleurs, la fréquence de clics solitaires est plus élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais ($d = 0.12$). En revanche, la fréquence des autres cas est moins élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais ($d = 0.13$).

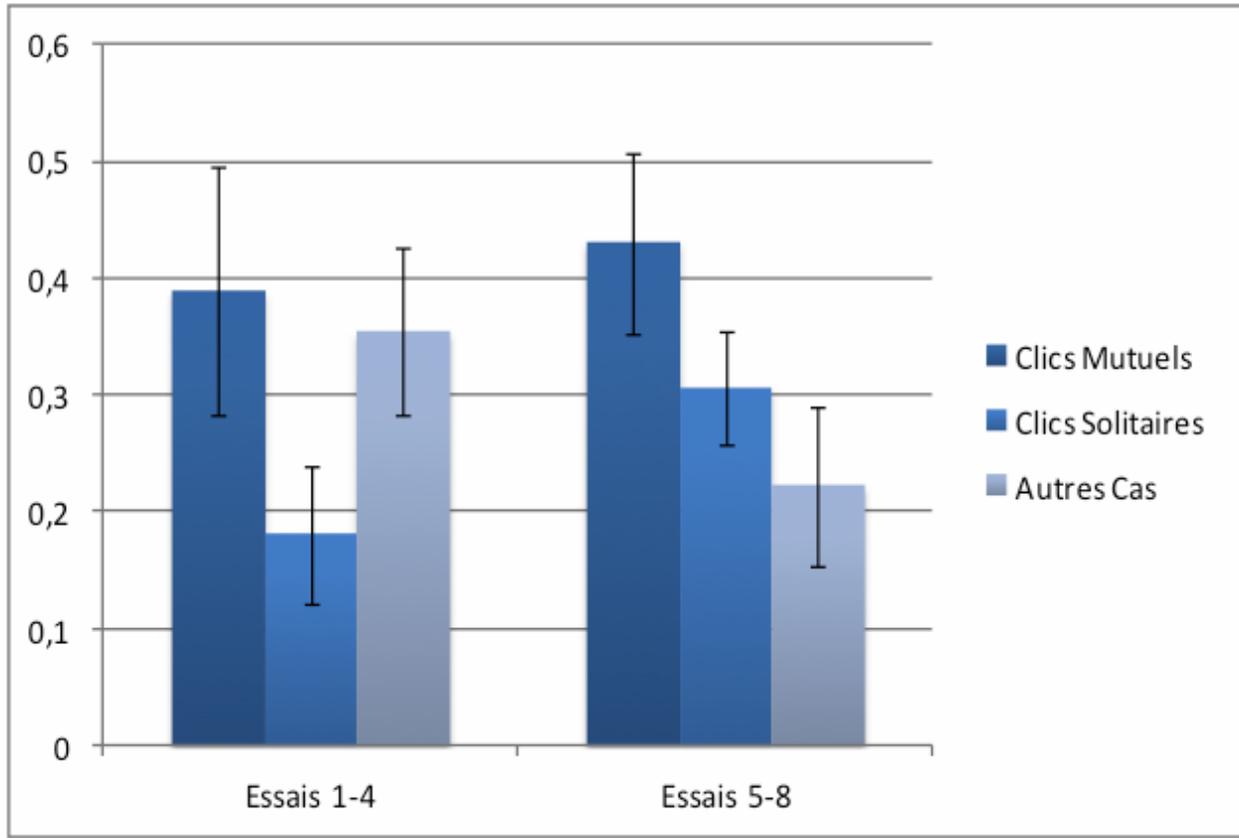


Figure 125. Fréquence des trois types de réponses collectives en fonction des deux blocs d'essais.

Pour finir, nous regardons si les conditions expérimentales ont un effet sur la fréquence de clics mutuels en fonction des deux blocs d'essais. De façon descriptive, nous observons que les fréquences de clics mutuels ne présentent que très peu de différences en ce qui concerne la condition « Mono » et la condition « Différenciée » en fonction des deux blocs d'essais (respectivement $d = 0.02$ et $d = 0.04$) (voir figure 126). En revanche, la fréquence de clics mutuels est plus élevée dans le deuxième bloc d'essai que dans le premier bloc d'essai en ce qui concerne la condition « Parallélisme » ($d = 0.10$). Toutefois, l'effet d'interaction entre les conditions expérimentales et les deux blocs d'essais n'est pas significatif ($F_{(2,33)} = 0.4304$; $p = 0.6538$).

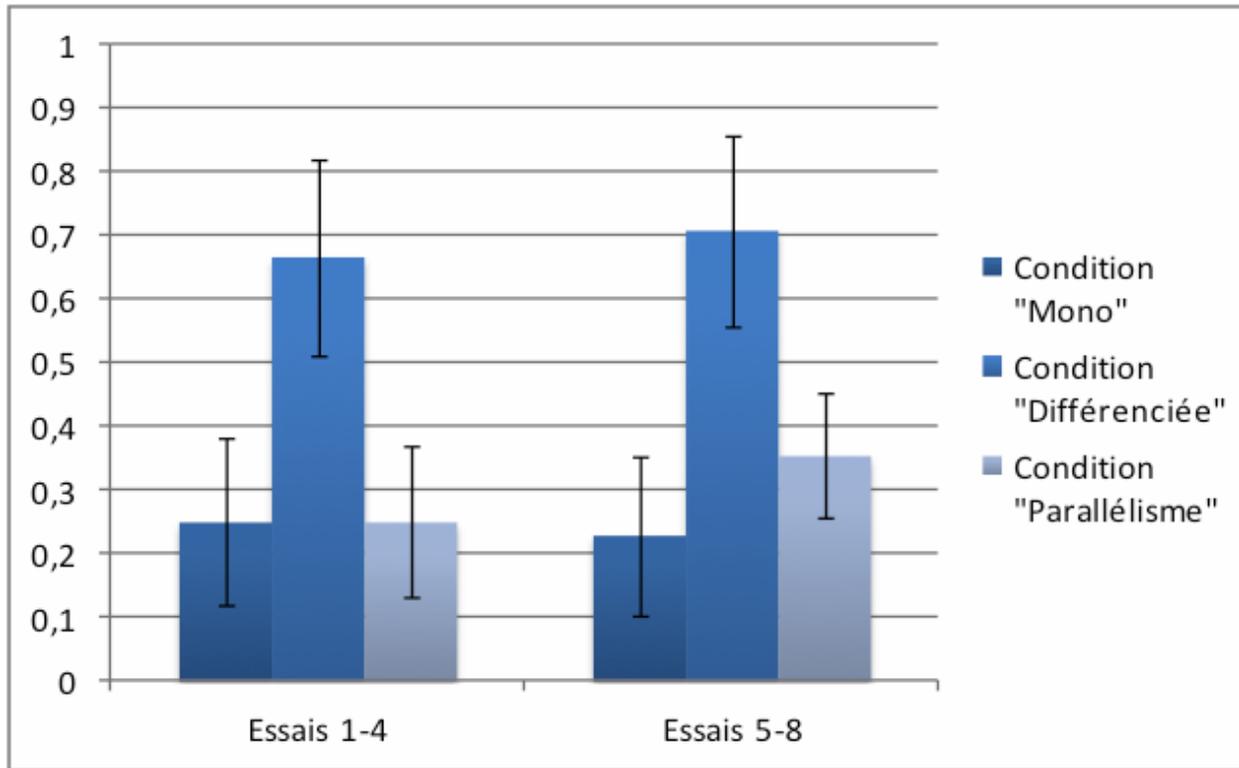


Figure 126. Fréquence de clics mutuels en fonction des deux blocs d'essais et des conditions expérimentales.

3.3.5.3 Analyses des trajectoires

Comme pour notre expérience concernant les interactions dyadiques dans le contexte de la discrimination Humain/Robot en utilisant Touch Through, une analyse des trajectoires perceptives déployées par les sujets a été menée⁴⁴. Cette analyse a pour but de permettre d'expliquer la réponse fournie par les sujets en fonction de leur façon d'interagir. Ici encore, nos choix d'indicateurs ont été guidés par les informations fournies par les sujets lors du débriefing de l'expérience, ainsi que par une analyse qualitative des graphiques représentant les différents essais (voir exemples en annexe C).

Une fois les différents indicateurs identifiés, nous avons procédé à une analyse quantitative systématique. La richesse des données nous a conduit à analyser de très nombreux indicateurs, mais par souci de clarté, nous ne relaterons ici que les résultats principaux.

⁴⁴ Pour plus de détails concernant la méthode d'analyse utilisée pour l'étude des trajectoires, veuillez vous reporter à l'Annexe C.

Trois axes distincts ont guidé notre analyse : tout d'abord l'interaction des sujets avec les deux types d'objets (Objet commun vs Objet privé), puis d'autre part l'interaction entre les sujets autour des deux types d'objets, et enfin l'interaction entre les sujets en fonction de la qualité de leur réponse individuelle (Clics sur l'objet commun vs Mauvais clics), et collective (Clics mutuels, Clics Solitaires et Autres cas).

3.3.5.3.1 Interaction des sujets avec les deux types d'objets

Dans un premier temps, nous avons isolé un premier groupe d'indicateurs, afin de pouvoir décrire l'interaction individuelle des sujets avec les deux types d'objets (Objet commun vs Objet privé), et ce, quelle que soit la réponse qu'ils aient donnée à la fin des essais. Nous nous intéressons ici à la durée totale pendant laquelle les sujets se sont trouvés dans un intervalle de 40 pixels autour de chacun des deux types d'objets, ainsi qu'au nombre de stimulations⁴⁵ dues à la rencontre de chacun d'eux.

D'une façon globale, nous observons que les sujets ont en moyenne passé plus de temps autour de l'objet commun qu'autour des objets privés ($d = 15438.81$ ms) (voir figure 127). Cette différence est significative (test de Mann & Withney : $U = 618.5$; $Z_{\text{ajusté}} = 7.885277$; $p < 0.0001$). Par ailleurs, cet effet se retrouve dans chacune des trois conditions expérimentales (voir figure 127 et tableau 12).

⁴⁵ Notons que nous nous intéressons au nombre de stimulations différentes. En effet, les stimulations sont parfois étendues dans le temps, comme par exemple dans le cas d'un arrêt sur un des objets.

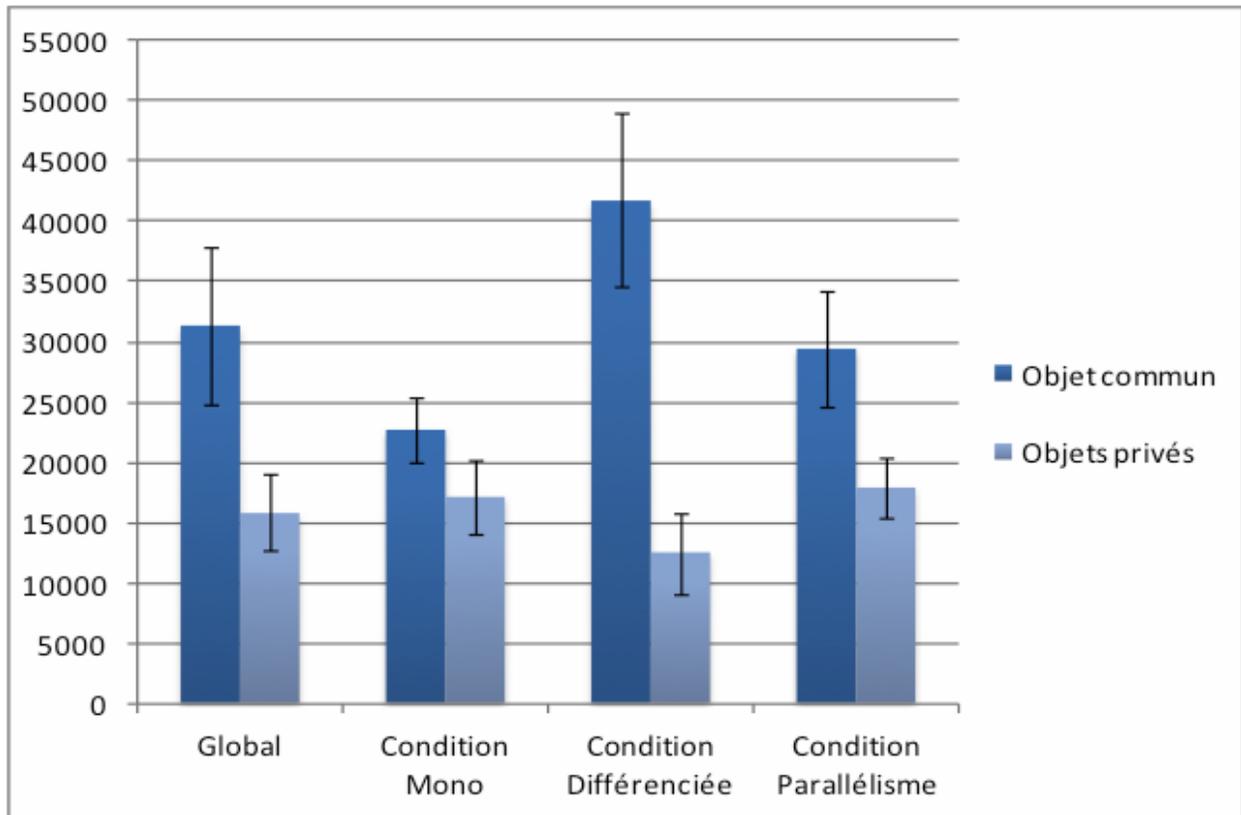


Figure 127. Durée totale moyenne (en ms) pendant laquelle les sujets se sont trouvés dans un intervalle de 40 pixels autour de chacun des deux types d'objets (Objet commun vs Objets privés) en fonction des conditions expérimentales.

Tableau 12. Résultats statistiques concernant les durées totales moyennes passées dans un intervalle de 40 pixels autour des deux types d'objets (Objet commun vs Objets privés) pour chaque condition expérimentale.

Conditions	Différence temps sur l'objet commun vs sur l'objet privé	Test	Probabilité	Taille d'effet
« Mono »	5504.06 (= 7 % du temps d'un essai)	$U = 139 ; Z_{\text{ajusté}} = 3.072328$	$p = 0.0021$	$g = 0.94$ (fort)
« Différenciée »	29326.87 (= 39 % du temps d'un l'essai)	$U = 35 ; Z_{\text{ajusté}} = 5.216772$	$p < 0.0001$	$g = 1.98$ (fort)
« Parallélisme »	11485.52 (= 15% du temps d'un essai)	$U = 72 ; Z_{\text{ajusté}} = 4.453845$	$p < 0.0001$	$g = 1.50$ (fort)

Pour autant, nous constatons des différences pour chaque type d'objet en fonction des conditions expérimentales. En effet, les sujets passent plus de temps autour de l'objet commun dans la condition « Différenciée » que dans la condition « Parallélisme » ($d = 12318.02$ ms) et dans la condition « Mono » ($d = 19069.58$ ms) (voir figure 127). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 24.76057$; $p < 0.0001$).

Parallèlement, les sujets passent plus de temps autour des objets privés en condition « Parallélisme » que dans la condition « Mono » ($d = 770.11$ ms) et dans la condition « Différenciée » ($d = 5523.34$ ms) (voir figure 127). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 24.76057$; $p < 0.0001$).

De plus, toutes conditions confondues, le nombre de stimulations résultant de la rencontre de l'objet commun est plus élevé que le nombre de stimulations résultant de la rencontre des objets privés ($d = 3.64$) (voir figure 128). Cette différence est significative (test de Mann & Withney : $U = 1924$; $Z_{\text{ajusté}} = 2.669012$; $p = 0.0076$).

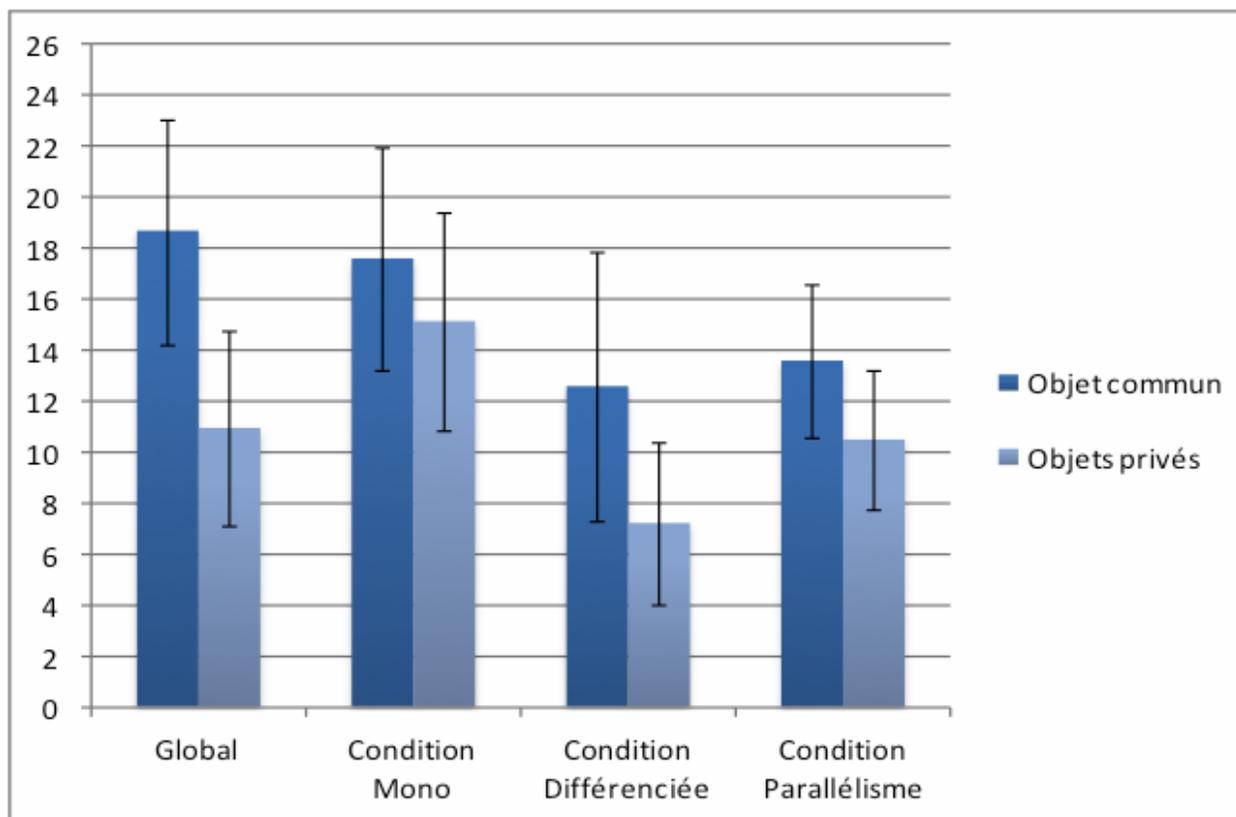


Figure 128. Nombre moyen de stimulations résultant de la rencontre des deux types d'objets (Objet commun vs Objets privés) en fonction des conditions expérimentales.

En ce qui concerne cette différence en fonction des conditions expérimentales, nous remarquons au préalable que le nombre de stimulations résultant de la rencontre des objets (qu'ils soient privés ou commun) est plus important en condition « Mono » qu'en condition « Parallélisme » ($d = 8.69$) et « Différenciée » ($d = 12.93$). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 10.65433$; $p = 0.0049$).

Nous observons que dans chacune des trois conditions expérimentales, le nombre moyen de stimulations dues à la rencontre de l'objet commun est plus élevé que celui qui résulte de la rencontre des objets privés (voir figure 128). Toutefois, nous pouvons noter que la différence n'est pas significative en ce qui concerne la condition « Mono » et la condition « Parallélisme ».

Nous constatons à nouveau des différences pour chaque type d'objet en fonction des conditions expérimentales. En effet, le nombre de stimulations résultant de la rencontre de l'objet commun est plus élevé dans la condition « Mono » que dans la condition « Parallélisme » ($d = 4.02$) et dans la condition « Différenciée » ($d = 5.02$) (voir figure 128). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 6.621082$; $p = 0.0365$).

Par ailleurs, le nombre de stimulations résultant de la rencontre des objets privés est plus élevé dans la condition « Mono » que dans la condition « Parallélisme » ($d = 4.67$) et dans la condition « Différenciée » ($d = 7.91$) (voir figure 128). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 14.95307$; $p = 0.0006$). Lorsque nous comparons les conditions expérimentales deux à deux, nous constatons que la différence réside entre la condition « Différenciée » et les deux autres conditions.

3.3.5.3.2 Interaction entre les sujets autour des deux types d'objets

Dans un second temps, nous avons isolé une série d'indicateurs concernant l'interaction entre les deux sujets autour des deux types d'objets, et ce indépendamment de la réponse que les sujets ont fourni à la fin des essais. Nous avons choisi de délimiter un intervalle de 50 pixels autour de chacun des deux types d'objets. Nous nous intéressons dans ces intervalles au nombre de stimulations résultant de la rencontre des deux sujets, ainsi qu'au nombre d'arrêts mutuels (un arrêt mutuel est défini par un arrêt simultané des deux sujets alors qu'ils sont en contact) et à leur durée totale.

Globalement, nous observons que le nombre moyen de stimulations interpersonnelles autour de l'objet commun est supérieur à celui obtenu autour des objets privés ($d = 7,90$) (voir figure 129). Cette différence est significative (test de Mann & Withney : $U = 509$; $Z_{\text{ajusté}} = 5.130746$; $p < 0.0001$). Par ailleurs, cet effet se retrouve dans chacune des trois conditions expérimentales (voir figure 129).

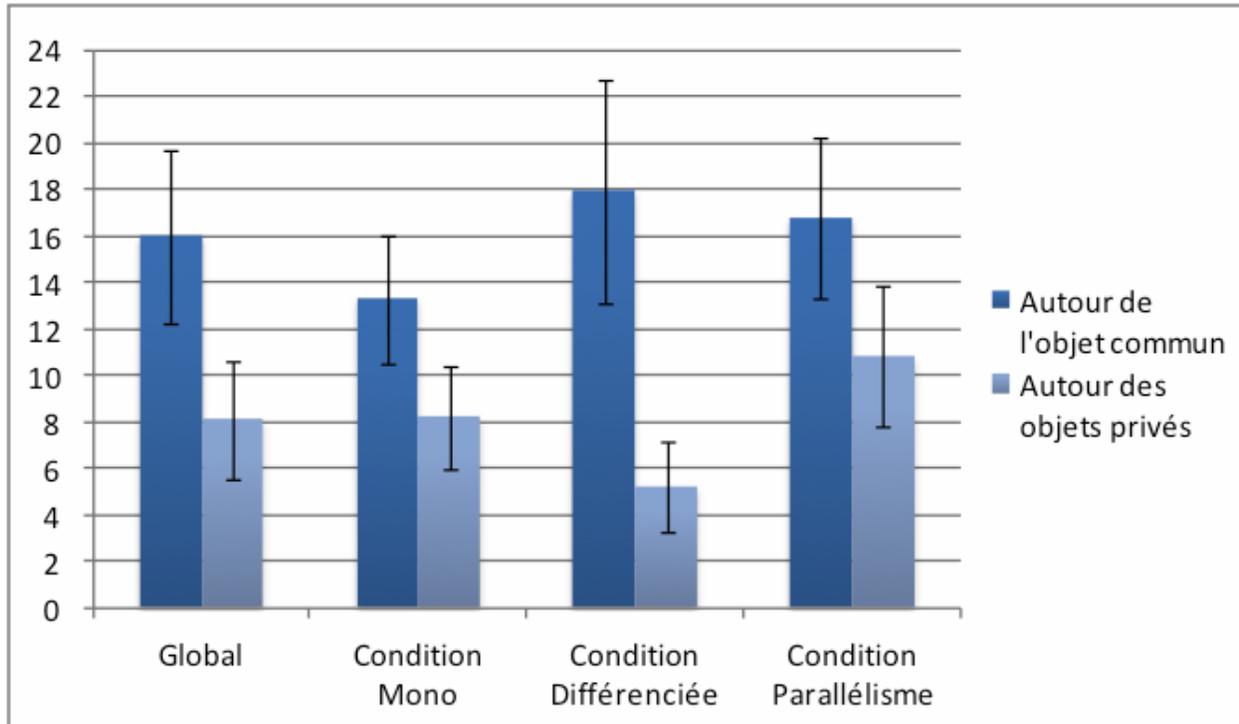


Figure 129. Nombre moyen de stimulations interpersonnelles autour des deux types d'objet (Autour de l'objet commun vs Autour des objets privés) en fonction des conditions expérimentales.

De plus, nous constatons que le nombre de stimulations interpersonnelles obtenues autour de l'objet commun est moins élevé en condition « Mono » qu'en condition « Parallélisme » ($d = 3.51$) et en condition « Différenciée » ($d = 4.65$). Mais la différence globale n'est pas significative (test de Kruskal-Wallis : $H_{(2,36)} = 0.9774775$; $p = 0.6134$).

Par ailleurs, le nombre de stimulations interpersonnelles obtenues autour des objets privés est moins élevé en condition « Différenciée » qu'en condition « Mono » ($d = 0.84$) et en condition « Parallélisme » ($d = 3.19$). Cette fois, la différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 9.769922$; $p = 0.0076$).

Globalement, nous observons que le nombre moyen d'arrêts mutuels est plus élevé autour de l'objet commun qu'autour des objets privés ($d = 7.71$) (voir figure 130). Cette différence est significative (test de Mann & Withney : $U = 392$; $Z_{\text{ajusté}} = 5.891533$; $p < 0.0001$).

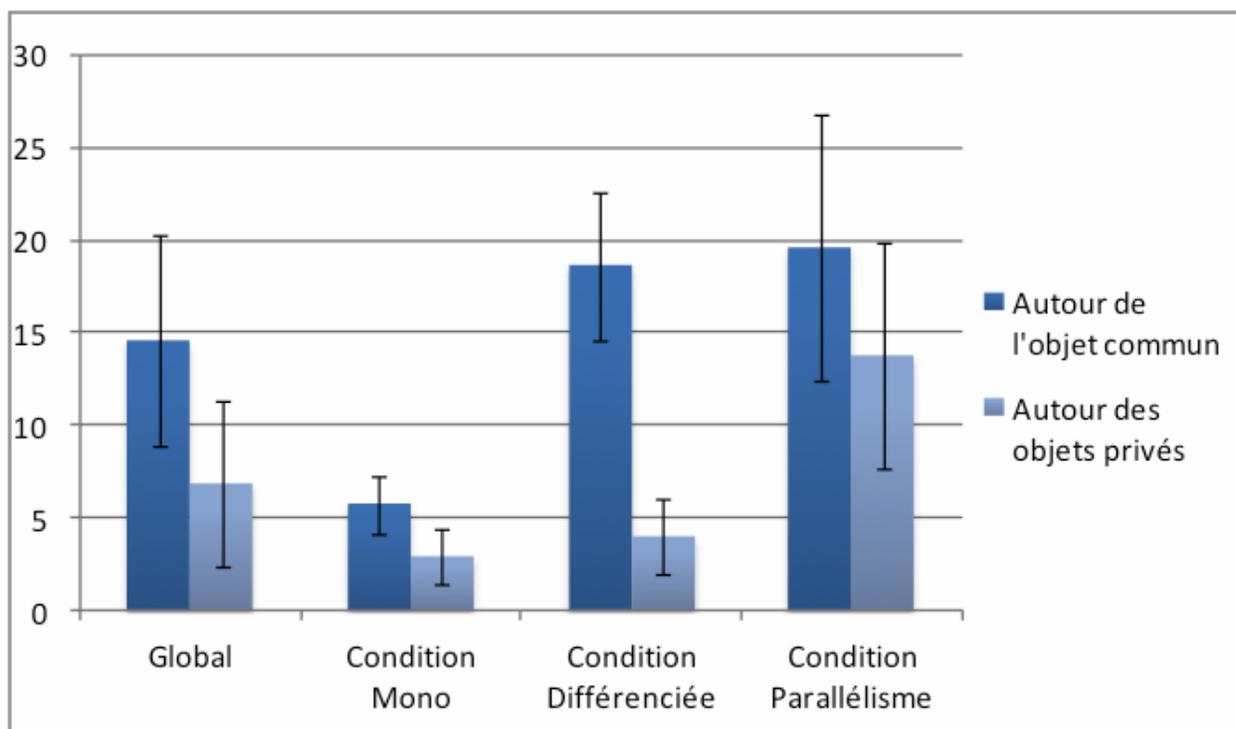


Figure 130. Nombre moyen d'arrêts mutuels autour des deux types d'objet (Autour de l'objet commun vs Autour des objets privés) en fonction des conditions expérimentales.

En ce qui concerne cette différence en fonction des conditions expérimentales, nous remarquons au préalable que le nombre d'arrêts mutuels autour des objets (qu'ils soient privés ou commun) est plus important en condition « Parallélisme » qu'en condition « Différenciée » ($d = 10.42$) et « Mono » ($d = 24.07$). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,36)} = 16.57808$; $p = 0.0003$).

Quoi qu'il en soit, nous observons que dans chacune des trois conditions expérimentales, le nombre moyen d'arrêts mutuels autour de l'objet commun est plus élevé que le nombre moyen d'arrêts mutuels autour des objets privés (voir figure 130). De plus, nous observons que le nombre moyen d'arrêts mutuels autour de l'objet commun est nettement moins élevé en condition « Mono » qu'en condition « Différenciée » ($d = 12.85$) et « Parallélisme » ($d = 13.82$). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,36)} = 15.85388$; $p = 0.0004$). Mais plus précisément, la différence n'est significative qu'entre la condition « Mono » et « Différenciée » (test de Mann & Withney : $U = 7.5$; $Z_{\text{ajusté}} = 3.72634$; $p < 0.0001$), et entre la condition « Mono » et « Parallélisme » (test de Mann & Withney : $U = 18.5$; $Z_{\text{ajusté}} = 3.0895$; $p = 0.0020$).

Par ailleurs, le nombre moyen d'arrêts mutuels autour des objets privés est nettement plus élevé en condition « Parallélisme » qu'en condition « Différenciée » ($d = 9.71$) et « Mono » ($d = 10.81$). Cette différence globale est significative (test de Kruskal-Wallis : $H_{(2,72)} = 24.37691$; $p < 0.0001$).

Enfin, la durée totale moyenne de ces arrêts mutuels est plus élevée autour de l'objet commun qu'autour des objets privés ($d = 10424.55$ ms) (voir figure 131). Cette différence est significative (test de Mann & Withney : $U = 417$; $Z_{\text{ajusté}} = 5.457165$; $p < 0.0001$).

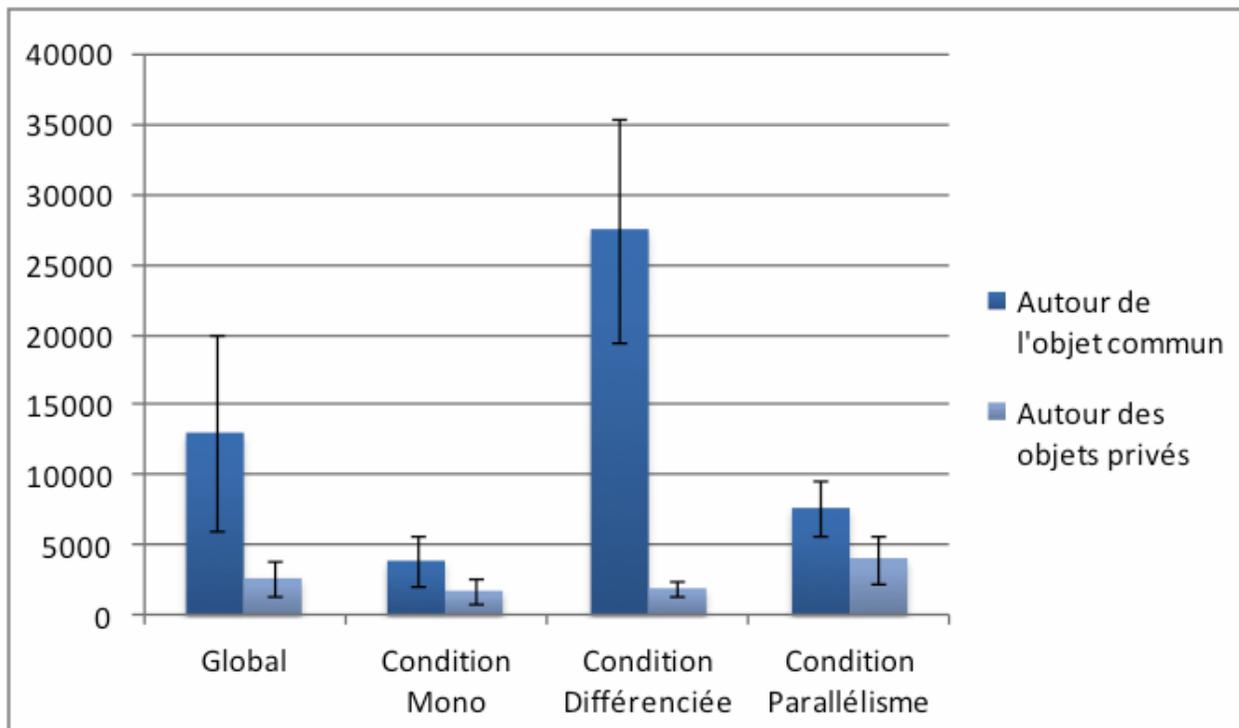


Figure 131. Durée totale moyenne des arrêts mutuels autour des deux types d'objet (Autour de l'objet commun vs Autour des objets privés) en fonction des conditions expérimentales.

Concernant la différence entre les conditions expérimentales, nous remarquons tout d'abord que la durée totale des arrêts mutuels autour des objets (qu'ils soient privés ou commun) est plus importante en condition « Différenciée » qu'en condition « Parallélisme » ($d = 17647.26$ ms) et « Mono » ($d = 23536.11$ ms). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,36)} = 20.40691$; $p < 0.0001$). Malgré tout, la durée totale moyenne des arrêts mutuels autour de l'objet commun est significativement plus élevée que la durée totale moyenne des arrêts mutuels autour des objets privés pour chacune des trois conditions expérimentales (voir figure 131).

Nous observons en outre que la durée totale des arrêts mutuels autour de l'objet commun est nettement plus élevée en condition « Différenciée » qu'en conditions « Parallélisme » ($d = 19758.61$ ms) et « Mono » ($d = 23546.40$ ms). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,36)} = 21.2177$; $p < 0.0001$).

Par ailleurs, la durée totale des arrêts mutuels autour des objets privés est plus élevée en condition « Parallélisme » qu'en conditions « Différenciée » ($d = 2072.38$ ms) et « Mono » ($d = 2249.04$ ms). La différence globale est également significative (test de Kruskal-Wallis : $H_{(2,67)} = 9,47412$; $p = 0.0088$).

3.3.5.3.3 Interaction entre les sujets en fonction de la qualité des clics

Finalemnt, nous nous intéressons à la qualité de l'interaction entre les sujets en fonction de la qualité de leurs clics individuels (Clics sur l'objet commun vs Mauvais clics), puis en fonction de leur type de réponse collective (Clics mutuels vs Clics solitaires vs Autres cas).

Pour cela, nous avons calculé la proportion d'accroche mutuelle. Nous avons alors isolé les séquences d'interaction pendant lesquelles les écarts entre les stimulations sont proches. Pour calculer la proportion d'accroche mutuelle, nous avons défini le contact comme un moment où les stimulations résultant des rencontres sont espacées de moins de deux secondes. La proportion d'accroche mutuelle se calcule en additionnant les durées totales des contacts, et en divisant cette somme par la durée totale d'un essai. Notons que nous ne prenons pas en compte la durée des arrêts mutuels lorsque ceux-ci sont supérieurs à deux secondes.

Nous observons que globalement, la proportion d'accroche mutuelle est plus élevée lors des essais qui ont conduit à un clic sur l'objet commun par rapport aux essais ayant conduit à un mauvais clic ($d = 5.12\%$) (voir figure 132). Cette différence est significative (test de Mann & Withney : $U = 1348$; $Z_{\text{ajusté}} = 2.549637$; $p = 0.0107$). Cependant, nous observons des différences en fonction des conditions expérimentales. Dans les conditions « Mono » et « Parallélisme », la proportion d'accroche mutuelle est supérieure quand le sujet a cliqué sur l'objet commun comparée à lorsqu'il a cliqué ailleurs. Ces différences sont significatives. Dans la condition « Différenciée », la proportion d'accroche mutuelle est moins élevée lorsque le sujet a cliqué sur l'objet commun que lorsqu'il a cliqué ailleurs. Toutefois, cette différence n'est pas significative (voir figure 132).

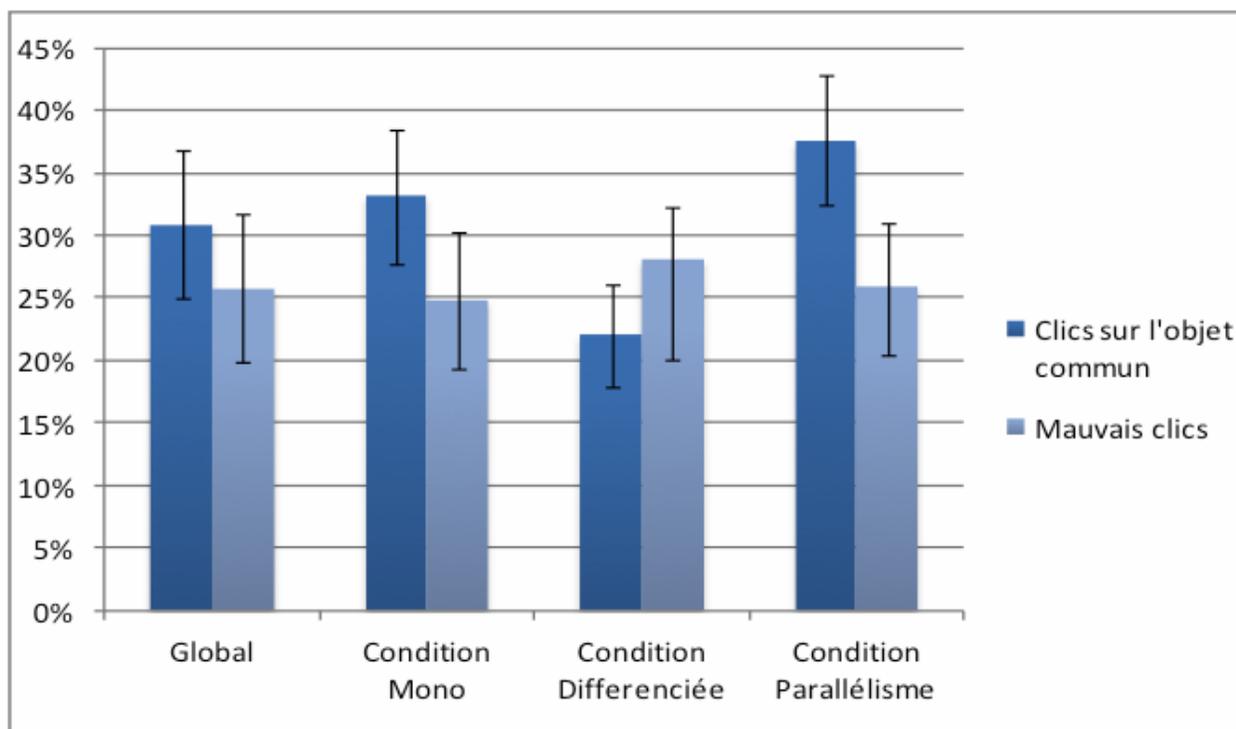


Figure 132. Proportion d'accroche mutuelle entre les deux sujets en fonction de la qualité de leurs clics (clics sur l'objet commun vs mauvais clics) en fonction des conditions expérimentales.

Par ailleurs, si nous ne considérons que les clics sur l'objet commun, nous observons que la proportion d'accroche mutuelle est plus élevée dans la condition « Parallélisme » que dans les conditions « Mono » ($d = 4.51\%$) et « Différenciée » ($d = 15.61\%$) (voir figure 132). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,71)} = 23.28953$; $p < 0.0001$).

Enfin, de façon globale, nous observons que la proportion d'accroche mutuelle est plus élevée dans le cas des clics mutuels que dans celui des clics solitaires ($d = 6.11\%$) et des autres cas ($d = 10.94\%$) (voir figure 133). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,95)} = 13.23588$; $p = 0.0013$). De façon plus spécifique, la proportion d'accroche est significativement plus élevée dans le cas des clics mutuels que dans le cas des clics unilatéraux (test de Mann & Whitney : $U = 355$; $Z_{\text{ajusté}} = 2.085286$; $p = 0.0370$) et dans les autres cas (test de Mann & Whitney : $U = 254$; $Z_{\text{ajusté}} = 3.585593$; $p = 0.0003$). En revanche, la différence n'est pas significative entre le cas des clics solitaires et celui des autres cas (test de Mann & Whitney : $U = 361$; $Z_{\text{ajusté}} = 1.500329$; $p = 0.1335$).

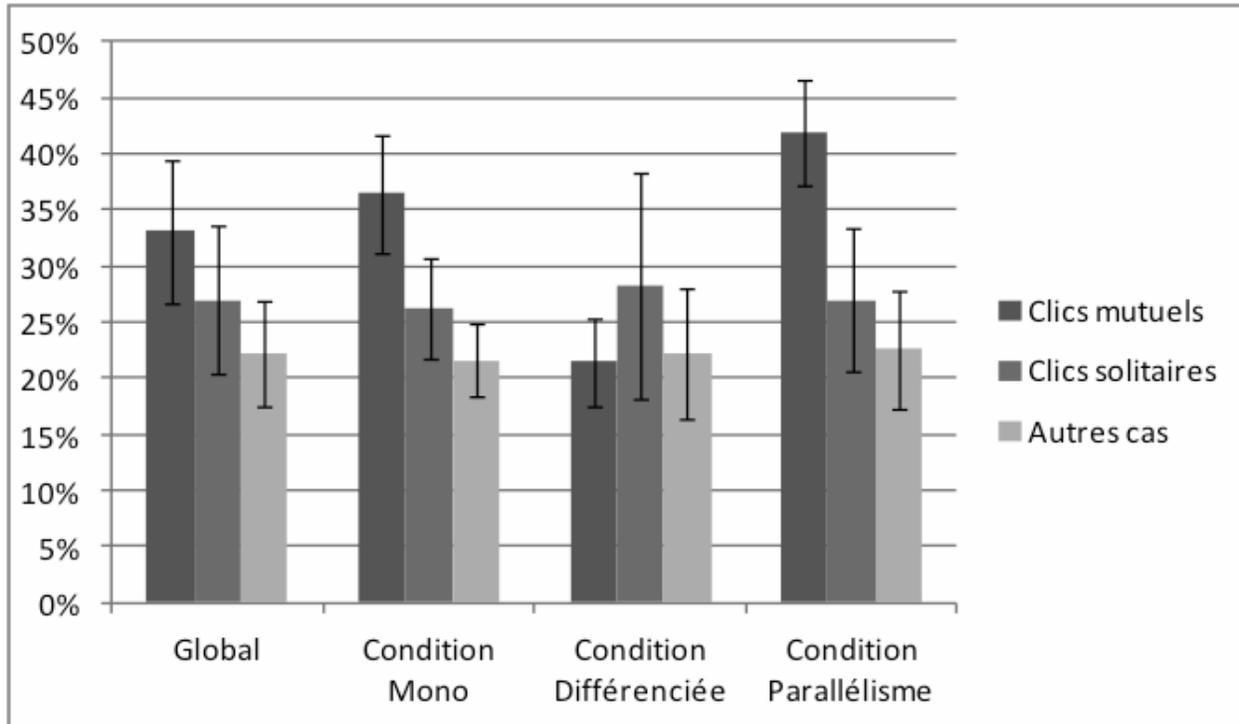


Figure 133. Proportion d'accroche mutuelle entre les deux sujets en fonction du type de réponse collective (Clics Mutuels, Clics Solitaires et Autres cas) en fonction des conditions expérimentales.

De façon plus spécifique, la différence globale entre les types de réponses collectives n'est pas significative pour la condition « Différenciée » (test de Kruskal-Wallis : $H_{(2,27)} = 0.2544407$; $p = 0.8805$). Mais en ce qui concerne les conditions « Mono » et « Parallélisme », nous retrouvons les mêmes résultats que ceux que nous avons obtenus toutes conditions confondues (voir figure 133).

Pour finir, nous comparons la proportion d'accroche mutuelle obtenue dans chaque type de réponse collective en fonction des conditions expérimentales. Concernant les cas de clics mutuels, la proportion de l'accroche est moins élevée en condition « Différenciée » qu'en condition « Mono » ($d = 15.08\%$), et « Parallélisme » ($d = 20.48\%$) (voir figure 133). La différence globale est significative (test de Kruskal-Wallis : $H_{(2,34)} = 17.79261$; $p = 0.0001$).

3.3.6 Discussion

Cette expérience avait pour vocation de rendre compte de la construction de l'attention conjointe autour d'un objet partagé, en observant la situation où les sujets doivent décider si un

objet est partagé ou non. Elle permettait donc de tester le fait qu'il soit possible ou non pour deux sujets engagés dans un environnement minimaliste de percevoir qu'un des objets présents dans cet environnement est également perceptible par l'autre sujet. En accord avec notre position théorique concernant l'émergence d'une signification partagée au cœur d'une coordination des activités perceptives, nous avons posé l'hypothèse que l'organisation du processus d'interaction entre les sujets et les objets d'une part, et entre les sujets d'autre part, pouvait conduire à une fréquence de clics sur l'objet commun plus élevée que la fréquence de mauvais clics, ce qui serait l'indice de capacités d'attention conjointe dans l'environnement minimaliste proposé. La coordination des foci attentionnels se réalisant de plus au travers d'un engagement mutuel des participants, la fréquence de clics mutuels devait être plus importante que la fréquence des clics solitaires. Par ailleurs, les conditions expérimentales devaient avoir un effet sur la réussite de la tâche, puisque la nature du dispositif de couplage entre les actions et les sensations individuelles devait conduire à des modes d'organisation de l'interaction différents. Ainsi, si la condition « Mono » devait conduire à une fréquence de clics sur l'objet commun plus faible que dans les autres conditions, cette configuration la plus minimaliste devait en même temps être particulièrement révélatrice d'une dynamique interpersonnelle, le dispositif de couplage exigeant en soi un déploiement plus important des activités perceptives et donc un niveau plus élevé de coordination de ces activités. En outre, nous nous attendions à ce que la réussite soit meilleure dans le deuxième bloc d'essais de chaque condition expérimentale, témoignant d'une stabilisation de la dynamique collective et d'une appropriation individuelle du dispositif de couplage.

Nous pouvons conclure de façon globale à un succès de la tâche, que ce soit à un niveau individuel (clics sur l'objet commun) ou à un niveau collectif (clics mutuels). Toutes conditions confondues, le nombre moyen de clics sur l'objet commun est en effet supérieur au nombre moyen de mauvais clics. En outre, nous pouvons conclure à un effet d'apprentissage individuel global, qui se traduit par une augmentation de clics sur l'objet commun et par une diminution des mauvais clics entre les deux blocs d'essais. Dans les conditions de l'expérience, les sujets ont donc pu discriminer qu'un des deux objets est également perceptible par l'autre participant : ils ont pu coordonner leurs activités perceptives de façon à établir un cadre de référence partagé, comme dans le cas du suivi de regard (Butterworth et Jarrett, 1991).

En ce qui concerne les réponses collectives, la fréquence des clics mutuels est supérieure à la fréquence des clics solitaires et des autres cas. Ce résultat semble montrer que la réussite de la tâche est attribuable à un engagement interpersonnel (Reddy, 2011), plus qu'à une simple « lecture » unilatérale de la dynamique du mouvement de l'autre sujet (Moore, Angelopoulos et Bennett, 1997), ou encore qu'à une attribution individuelle d'une intentionnalité chez autrui (Tomasello, 1995). Nous pouvons ainsi poser l'hypothèse que la connaissance du fait qu'un objet

de l'environnement est partagé émerge de la dynamique relationnelle, en tant que processus autonome qui influence les contributions individuelles et qui permet aux interactants de leur attribuer du sens (De Jaegher et Di Paolo, 2007 ; De Jaegher, Di Paolo et Gallagher, 2010). En revanche, dans la mesure où nous ne pouvons pas conclure à un effet d'apprentissage collectif global, qui se traduirait minimalement par une augmentation des clics mutuels et une diminution des clics solitaires entre les deux blocs d'essais, nous ne pouvons pas conclure à une stabilisation de cette dynamique qui aurait été profitable aux deux sujets. L'effet d'apprentissage collectif ne se traduit au contraire que par une diminution des autres cas. L'évolution de la performance ne se vérifie ainsi qu'à un niveau individuel, ce qui ne veut pour autant pas dire qu'elle ne peut pas être basée sur la dynamique collective. En effet, le fait de mettre l'accent sur le couplage interpersonnel ne revient pas à dissoudre le rôle de l'individu : celui-ci doit pouvoir re-saisir cette dynamique individuellement, de façon à construire du sens à partir d'elle (Lenay et Stewart, 2012).

De plus, si nous pouvons attester d'une réussite globale, nous pouvons conclure que celle-ci est dépendante des conditions expérimentales. À un niveau individuel, si les trois conditions amènent plus de clics sur l'objet commun que de mauvais clics, nous ne pouvons pourtant pas conclure à la réussite de la tâche dans la condition « Mono ». Par ailleurs, nous pouvons conclure que les sujets sont plus efficaces dans la condition « Différenciée » que dans les conditions « Mono » et « Parallélisme ». La condition « Différenciée » semble favoriser la perception mutuelle d'un objet commun. En revanche, à un niveau collectif, nous ne pouvons pas conclure à une fréquence de clics mutuels plus élevée que celle des autres types de réponse collective dans la condition « Mono » et dans la condition « Parallélisme ». Seule la condition « Différenciée » amène significativement plus de clics mutuels que de clics solitaires ou que des autres cas. Malgré cet apparent échec collectif, nous avons souligné que dans la majorité des cas de clics solitaires, lorsqu'un des deux sujets clique sur l'objet commun, l'autre ne clique pas. De la même façon, dans les autres cas, lorsqu'un des deux sujets clique ailleurs que sur l'objet commun, l'autre ne clique pas. Or il est difficile d'attribuer à une absence de réponse le même statut qu'une mauvaise réponse. Nous ne pouvons pas attester d'une réussite collective, mais nous ne pouvons pas pour autant affirmer que la réussite serait purement individuelle. En effet, nous avons montré que la réussite de l'un des deux sujets n'est pas liée à l'échec de l'autre, mais seulement à une absence de réponse de celui-ci.

L'analyse des trajectoires perceptives nous donne d'autre part des éléments cruciaux pour tenter de comprendre ce qui amène les sujets à cliquer sur l'objet commun. Nous nous intéressons ici au déploiement des activités perceptives et à la coordination interpersonnelle, telles qu'elles sont rendues accessibles à l'analyse via la méthode minimaliste (Lenay, 2006). Nous avons pu constater tout d'abord que les sujets passent significativement plus de temps

autour de l'objet commun qu'autour des objets privés, que ce soit à un niveau global (toutes conditions confondues), qu'à un niveau plus spécifique (dans chacune des trois conditions expérimentales). Dans le cadre de l'utilisation de Tactos, l'objet attire l'activité perceptive de l'utilisateur, car il constitue une source de stimulations qui permet d'organiser le geste d'exploration (Lenay et al, 2000). Dans notre expérience, bien que les deux types d'objet possèdent exactement les mêmes propriétés objectives (même taille, même forme), on constate que les activités perceptives déployées par les sujets sont plus attirées par l'objet commun que par l'objet privé. Cependant, nous pouvons conclure que l'objet commun attire de façon plus importante l'activité dans la condition « Différenciée » que dans les autres conditions. De plus, les sujets sont plus fortement attirés par l'objet commun en condition « Parallélisme » qu'en condition « Mono ». De manière parallèle, les objets privés attirent moins l'activité des sujets dans la condition « Différenciée » que dans les autres conditions expérimentales, alors qu'il n'existe pas de réelle différence entre les conditions « Parallélisme » et « Mono ».

Mais si cette attirance se traduit de manière globale par un nombre plus élevé de stimulations résultant de la rencontre de l'objet commun par rapport au nombre de stimulations résultant de la rencontre de l'objet privé, cette différence n'est cependant vérifiée qu'en condition « Différenciée ». Par conséquent, nous pouvons conclure qu'en fonction des conditions, les activités perceptives déployées par les sujets sont animées autour d'une zone de l'espace plus ou moins large située autour de l'objet commun. Dans la condition « Différenciée, cette zone est très étroite et est pratiquement restreinte à l'objet même, tandis qu'elle est plus large dans les conditions « Mono » et « Parallélisme ».

L'analyse des indicateurs concernant l'interaction entre les sujets autour des deux types d'objets permet d'expliquer l'attirance plus forte exercée par l'objet commun. Nous pouvons en effet conclure à un nombre plus élevé de stimulations résultant du croisement perceptif autour de l'objet commun qu'autour des objets privés, et ce quelle que soit la condition expérimentale. En d'autres termes, bien que les deux types d'objet attirent les activités perceptives individuelles, l'attirance pour l'objet commun est plus forte du fait que celui-ci attire réciproquement l'activité du partenaire. L'objet commun présente en effet une affordance simultanée (Marsh et al, 2006), laquelle fournit une opportunité d'action identique pour les deux participants. Ainsi, il est probable que la rencontre entre les deux activités perceptives ne soit pas visée pour elle-même mais qu'elle se produise autour de l'objet commun : en effet, l'objet commun conditionne un nombre élevé de rencontres autour de lui par son caractère partagé et sa disponibilité pour les deux interactants. Ici, la connaissance du fait que l'objet est partagé ne provient par conséquent pas nécessairement d'une conduite intentionnelle de la part d'un des deux sujets pour diriger activement l'attention de l'autre vers les objets présents (Mundy, 1995 ; Mundy et Acra, 2006). Bien que l'analyse qualitative des trajectoires révèle que la plupart des sujets déploient une

stratégie orientée vers les objets (ils recherchent activement les objets et cherchent à détecter la présence de l'autre sujet autour d'eux), il est à noter que dans certains cas, les sujets cherchent d'abord à se rencontrer pour explorer l'environnement de façon conjointe (la stratégie est alors orientée sujet ; voir exemples en Annexe C). Dans tous les cas, l'émergence de la connaissance du caractère commun de l'objet ne peut pas être réduite à une simple rencontre interpersonnelle inopinée autour de cet objet. Étant donné que les sujets n'accèdent pas aux stimulations reçues par leur partenaire, il faut encore qu'ils puissent se coordonner efficacement autour d'un objet pour que les actions de l'un puissent avoir du sens pour l'autre, et réciproquement. Ainsi, il faut que l'affordance simultanée de l'objet commun puisse se combiner avec une affordance sociale (Krueger, 2011), de façon à constituer une dynamique relationnelle autonome organisée autour de l'objet commun (De Jaegher et Di Paolo, 2007 ; De Jaegher, Di Paolo et Gallagher, 2010) ; c'est dans cette dynamique que peut avoir lieu un *participatory sense-making* pouvant expliquer la co-construction d'une connaissance partagée (ibid).

L'analyse de l'accroche perceptive mutuelle entre les deux sujets révèle que toutes conditions confondues, plus la coordination entre les deux sujets est grande, plus les sujets cliquent sur l'objet commun. Cette accroche ne semble en outre pas déterminer uniquement les réponses individuelles, puisque cette coordination amène les sujets à cliquer sur l'objet commun de façon mutuelle. L'accroche mutuelle est en effet significativement plus élevée dans les cas de clics mutuels que dans celui des clics solitaires et des autres cas. Par condition, ces résultats sont retrouvés dans les conditions « Mono » et « Parallélisme », mais nous ne les retrouvons pas dans la condition « Différenciée ». Par ailleurs, les cas des clics sur l'objet commun et des clics mutuels de la condition « Différenciée » présentent significativement moins d'accroche mutuelle que dans la condition « Mono » et dans la condition « Parallélisme ».

Il semblerait donc que si la condition « Différenciée » apporte les meilleurs résultats, ceux-ci ne peuvent pas être expliqués par l'accroche perceptive mutuelle, du moins pas si nous la définissons comme une coordination des activités perceptives déployées dans l'espace et dans le temps. Les résultats obtenus dans cette condition semblent par conséquent à première vue remettre en cause l'idée selon laquelle l'attention conjointe est d'abord une action conjointe simple (Fiebich et Gallagher, 2012), se définissant par une coordination interpersonnelle sensori-motrice (Gallagher, 2011). Mais le cas des arrêts mutuels nous apporte en fait des éléments de compréhension supplémentaires.

Nous avons en effet vu que d'une manière globale, les sujets s'arrêtent mutuellement plus souvent autour de l'objet commun qu'autour des objets privés. Or il s'avère que cette différence est très nette en condition « Différenciée ». En outre, la durée de ces arrêts mutuels est plus importante autour de l'objet commun qu'autour de l'objet privé. À nouveau, nous avons observé une très nette différence dans la condition « Différenciée ». Ce résultat peut s'expliquer par le

fait que dans cette condition, la différence entre le partenaire et les objets est déjà donnée au sujet par le biais de la stimulation sensorielle. En d'autres termes, contrairement aux conditions « Mono » et « Parallélisme », il n'est pas nécessaire de les discriminer à partir de l'activité perceptive. De plus, étant donné que le sujet peut percevoir en même temps l'objet et le partenaire, l'arrêt du sujet sur l'objet a plus de sens que dans les autres conditions. La stratégie adoptée majoritairement en condition « Différenciée » peut se résumer de cette manière : « Je cherche une stimulation résultant de la rencontre d'un objet. Quand j'en ai trouvé une, je suis donc sur un objet : je m'arrête alors et j'attends. Si mon partenaire vient à ma rencontre, je reçois alors deux stimulations différenciées de façon simultanée, résultant de la rencontre à la fois avec mon partenaire, et avec l'objet sur lequel je me trouve. Mais mon partenaire est peut-être alors simplement en train de me percevoir moi, et ne perçoit pas d'objet. Si mon partenaire repart, c'est qu'il n'a perçu que moi et que je suis donc en train de percevoir l'objet privé. En revanche, s'il maintient sa position, c'est que lui aussi perçoit l'objet que je suis en train de percevoir ; nous nous trouvons alors tous deux sur l'objet commun. En d'autres termes, je comprends son arrêt comme je comprends le mien : nous sommes tous les deux stimulés à la fois par un objet et par le corps-image de l'autre. »

Le maintien de l'arrêt du partenaire permet au sujet de savoir que l'objet de son attention est partagé : il fait sens du maintien de l'arrêt de l'autre sujet en fonction d'un objet que lui-même perçoit. Ainsi, si la connaissance du fait que l'objet est partagé ne résulte pas en soi d'un déploiement d'une activité perceptive conjointe autour de cet objet (telle qu'elle est entendue par notre définition d'une accroche perceptive mutuelle), elle est tout de même le résultat d'un engagement actif de la part du sujet (la recherche d'un objet). Nous pouvons par ailleurs également considérer que le fait de maintenir un arrêt est en soi la manifestation d'un engagement actif de la part du sujet, signant une accroche perceptive mutuelle autour d'un objet.

Il est possible de considérer que l'arrêt, et le fait de rester arrêté est une situation privilégiée et favorisée par le dispositif de couplage lui-même lorsqu'il est réglé dans cette condition. Ainsi, les différences observées en fonction des conditions expérimentales appuient un fait essentiel : la perception n'est pas une réception passive d'informations sensorielles mais est le résultat de l'appropriation d'un dispositif de couplage sensori-moteur. En ce sens, la médiation technique offerte par les dispositifs de suppléance perceptive rend possible la perception, en mettant à disposition de l'utilisateur un répertoire de retours sensoriels possibles qui contraint en retour les actions utiles (Lenay et al, 2000). Par exemple dans la condition « Différenciée », le système de couplage induit les arrêts mutuels, parce que les types de stimulations reçues suffisent pour comprendre l'engagement du partenaire en fonction des objets présents. Dans ce cadre, le sujet peut extraire de la situation du maintien de l'arrêt mutuel l'information selon

laquelle l'autre participant perçoit l'objet qu'il perçoit lui-même, parce que l'arrêt de l'autre sujet n'a de sens qu'en tant qu'activité perceptive dirigée simultanément vers soi et vers l'objet.

Au contraire, en condition « Mono », la différenciation entre autrui et objet n'est pas pré-donnée aux sujets, elle doit alors être constituée au travers de l'interaction. Il est de plus impossible dans cette condition de percevoir l'autre sujet et l'objet en même temps. La perception mutuelle nécessite alors une alternance dynamique des « regards » ainsi qu'un fort niveau de coordination interpersonnelle, que ce soit à un niveau microscopique (accroche mutuelle et dynamique de l'activité perceptive autour de l'objet) qu'à un niveau macroscopique (alternance structurée des regards vers l'autre sujet et vers l'objet, comme produisant un processus continu d'engagement et de désengagement interpersonnel). Nous avons par ailleurs remarqué que le nombre de stimulations résultant de la rencontre des objets est plus élevé en condition « Mono » que dans les autres conditions. Ce résultat témoigne d'un déploiement de l'activité perceptive plus élevé que dans les autres conditions.

La condition « Parallélisme » peut finalement être considérée comme une condition « Mono » augmentée. En effet, la différenciation autrui/objet n'est toujours pas pré-donnée aux sujets, mais les stimulations sensorielles présentent une spatialité intrinsèque qui permet une économie de mémoire et de mouvement pour les individus. Il est en outre possible dans cette condition de percevoir l'autre sujet et l'objet de façon simultanée si les activités perceptives s'organisent autour de l'objet. Cette condition implique alors une coordination plus fine des « regards », coordination qui est rendue possible par le biais d'un déploiement de l'activité plus restreint qu'en condition « Mono ». Cependant, nous avons également noté que cette coordination seule ne suffit pas. En effet, la condition « Parallélisme » est la condition qui présente le plus d'arrêts mutuels, même si ceux-ci sont nettement moins longs que dans la condition « Différenciée » et qu'ils ont lieu autour des deux types d'objet. Les arrêts mutuels semblent donc également être une action utile dans la condition « Parallélisme » pour faire sens de l'activité du partenaire. Ainsi, si les meilleurs résultats sont obtenus dans la condition « Différenciée », nous pensons que c'est la condition « Parallélisme » qui est la plus encourageante, puisqu'elle offre un répertoire d'actions plus complet, notamment en rendant possible le déploiement des activités perceptives nécessaire à une coordination interpersonnelle spatio-temporelle. Il s'agit également de la condition dans laquelle les résultats ont le plus progressé entre les deux blocs d'essais.

Malgré l'apparente stabilité des stratégies observées en fonction des conditions expérimentales, nous nous devons de souligner une limite importante. En effet, les sujets ont tous effectué les trois conditions expérimentales au cours de cette expérience. Si nous avons contrebalancé l'ordre de passation de ces conditions, de façon à neutraliser un éventuel effet d'ordre, il est par ailleurs possible que les stratégies déployées par les sujets dans une condition

donnée aient été influencées par les conditions expérimentales précédentes. Par exemple, le fait de commencer avec la condition « Différenciée » a pu induire un type de stratégie qui s'est avéré incompatible avec les conditions suivantes, ce qui a pu influencer les résultats. De futures analyses permettront de tester un éventuel effet d'interaction des conditions expérimentales sur les stratégies individuelles et collectives. Il serait d'autre part intéressant de séparer ces conditions expérimentales selon des groupes expérimentaux indépendants. Ainsi, nous pourrions proposer plus d'essais, de façon à favoriser l'appropriation du système de couplage par les utilisateurs, et alors permettre une meilleure stabilisation de la dynamique d'interaction pouvant conduire à de meilleurs résultats.

Les résultats soulèvent d'autre part la question de la configuration optimale des corps-image/corps-percevant pour que les utilisateurs du dispositif puissent coordonner leurs activités perceptives autour d'objets communs. Nous avons vu que dans les conditions de l'expérimentation, la différenciation des stimulations objet/autrui favorisait la reconnaissance collective qu'un objet du monde est partagé. Dans la condition « Différenciée », la résolution de la tâche est en fait plus déductive que perceptive, car elle est moins liée à l'activité perceptive. Cela tient au fait qu'il y a un indice pré-donné par le système (quant à la nature de ce qui est perçu : l'autre sujet, ou un objet fixe). Cela constitue un intérêt de la condition « Différenciée ». Mais la mise en place de cet indice pré-donné n'est pas toujours possible ni même souhaitable, et c'est alors que le parallélisme se révèle extrêmement intéressant (il est en effet excellent pour la lecture de formes par exemple), de même que le mono est intéressant pour les aspects fondamentaux. La reconnaissance dans la condition « Différenciée » est rendue possible par un arrêt de l'utilisateur sur l'objet ; il peut alors faire sens de l'arrêt du partenaire comme étant l'indice d'une activité perceptive intentionnelle dirigée vers soi et vers l'objet. Cependant, dans le cadre des espaces numériques bidimensionnels, les utilisateurs ne doivent pas seulement reconnaître que l'autre participant perçoit un objet, mais ils doivent coordonner leurs activités perceptives pour explorer ensemble un contenu partagé. Dans ce cas, la configuration « Différenciée » ne nous semble pas transposable telle quelle dans de tels espaces, dans la mesure où les stimulations, si elles sont différenciées, sont ici délivrées en tout ou rien. Ceci ne nous semble pas propice à l'exploration mutuelle de contenus spatialisés, et ce pour plusieurs raisons :

- le parallélisme du corps-percevant favorise l'exploration et la reconnaissance de formes numériques bidimensionnelles, en permettant d'accéder à des informations spatiales dès l'entrée sensorielle (Sribunruangrit, 2004).
- les stimulations en tout ou rien ne favorisent pas la coordination des activités perceptives dans les espaces numériques partagés bidimensionnels, en rendant le guidage interpersonnel difficile (Deschamps, 2013, p. 168).

Afin de tirer parti de la différenciation des stimulations objet/autrui, et du parallélisme des champs récepteurs, une solution consisterait à utiliser deux plages de deux cellules Braille piézoélectriques, chacune allouée à une perception différente (perception d'autrui pour l'une, perception des objets pour l'autre). Mais l'accès simultané à ces deux plages n'est pas forcément confortable, et l'utilisation simultanée de ces deux informations pourrait rendre la tâche du sujet encore plus compliquée, au point qu'elle deviendrait inextricable. Une deuxième solution consisterait à utiliser des modalités sensorielles différentes, en signalant par exemple la présence d'autrui par des retours sonores. Mais ces retours pourraient alors interférer avec les informations sonores ou verbales relatives au contenu partagé lui-même ; rappelons que Tactos est dédié aux personnes aveugles, et que son utilisation en couplage avec une synthèse vocale permet de décupler ses applications. La condition « Parallélisme » nous semble par conséquent être la configuration la plus indiquée pour l'exploration mutuelle de contenus bidimensionnels par le biais de Tactos.

Les résultats de cette expérience permettent d'une part d'évaluer la pertinence de la condition « Parallélisme » dans ce contexte d'exploration mutuelle d'un contenu bidimensionnel partagé (pour des analyses d'usage en situation écologique, voir Deschamps, 2013, p. 236). D'autre part, ils nous permettent une connaissance des mécanismes permettant aux utilisateurs d'identifier ce qui est commun, ou public, et ce qui est privé lors de l'exploration d'un contenu numérique. Ces mécanismes s'avèrent cruciaux pour l'interaction dans des espaces numériques partagés, qu'ils soient unidimensionnels, bidimensionnels, composés d'éléments privés et partagés. Nous avons pu comparer différentes conditions d'exploration d'un environnement numérique par le biais de Tactos. Elle a mis en évidence l'importance de la dynamique collective, de l'interaction autour des objets et entre les sujets dans le but de déterminer quels objets étaient communs, et quels objets étaient privés. Ces éléments déterminants ne se limitent pas à l'utilisation de cette interface spécifique et sont à considérer dans d'autres environnements à explorer par d'autres biais. Typiquement, la majorité des utilisateurs utilisent une souris et un écran pour naviguer dans les environnements numériques, qui peuvent être partagés entre plusieurs utilisateurs.

3.3.7 Conclusion

Cette expérience nous a permis de vérifier si les conditions techniques de l'interaction tactile sur le réseau pouvaient permettre des capacités d'attention conjointe de la part des utilisateurs. Cette question est d'une importance capitale, dans la mesure où le projet Tactos consiste à concevoir des espaces numériques partagés dans lesquels les sujets pourront

coordonner leurs activités perceptives autour d'objets communs. Elle permet aussi, en l'étendant à d'autres dispositifs, d'envisager la collaboration d'utilisateurs distants dans des espaces numériques partagés sous différentes perspectives.

Conformément à l'approche interactionniste de la cognition sociale, nous avons posé l'hypothèse que l'attention conjointe pouvait se réaliser via la coordination des activités perceptives. La dynamique relationnelle que nous avons observée dans un contexte dyadique serait ici organisée en fonction des objets présents. Ainsi, nous avons élaboré une tâche collaborative, nécessitant une coordination interpersonnelle de façon à atteindre un but prescrit par la tâche. Nous avons en ce sens pris le parti d'étudier les mécanismes de l'attention conjointe à partir d'une activité conjointe préalablement définie : les sujets sont déjà dans un même espace où ils pratiquent le croisement perceptif. Mais cela ne signifie pas que la connaissance issue de l'attention conjointe soit déjà constituée : il est crucial que les sujets s'engagent dans la tâche, et s'approprient le dispositif pour être à même de donner du sens à l'engagement de leur partenaire dans l'espace partagé. La question posée aux sujets est donc celle de faire la distinction entre l'activité conjointe autour d'un objet commun et celle autour d'un objet privé.

Les résultats principaux mis en avant suggèrent que la connaissance du fait qu'un objet de l'environnement est partagé résulte d'un engagement mutuel des participants plus que d'une simple « lecture » du comportement attentionnel de l'autre sujet (Moore, Angelopoulos et Bennett, 1997) ou d'une attribution individuelle d'intentionnalité (Tomasello, 1995). Nos résultats montrent en effet que la perception mutuelle se réalise dans le cadre du processus d'interaction (Auvray, Lenay et Stewart, 2009 ; De Jaegher, Di Paolo et Gallagher, 2010). Par conséquent, la connaissance du fait que l'objet est partagé est co-construite à travers un processus actif et dynamique de négociation qui se réalise au sein même du couplage interpersonnel (ibid.). L'attention conjointe est alors considérée comme une action conjointe basique (Fiebich et Gallagher, 2012) conduisant à une modification du statut de l'objet : ce dernier passe d'un objet privé, que l'individu rencontre dans le déploiement de son activité perceptive, à un objet commun, que les deux individus peuvent rencontrer dans le processus de leur interaction (Sebanz, Bekkering et Knoblich, 2006).

L'objet commun fournit en effet aux sujets une affordance simultanée (Marsh et al., 2006) produisant une attirance réciproque et augmentant ainsi sensiblement la probabilité de rencontrer l'autre participant autour de cet objet. Cependant, percevoir l'autre sujet et percevoir l'objet dans une zone donnée de l'espace ne signifie pas nécessairement que l'autre participant aussi perçoit cet objet. La combinaison de l'affordance résultant de la rencontre de l'objet et d'une affordance sociale (Krueger, 2011) semble donner lieu à une coordination interpersonnelle autour de l'objet permettant :

- de donner du sens à l'activité d'autrui en tant qu'il s'engage activement dans le croisement perceptif avec moi (Auvray, Lenay et Stewart, 2009 ; Lenay et Stewart, 2012) ;
- de donner du sens à cet engagement en tant qu'activité dirigée alternativement vers moi et vers l'objet (Reddy, 2011) ;
- de donner du sens à l'objet en tant que source de stimulation mutuellement perçue, ce qui conduit à la reconnaissance que l'objet de mon attention est également l'objet de l'attention d'autrui (Baldwin, 1995).

De manière parallèle, nous avons noté que la constitution de cette connaissance était influencée par les propriétés du système de couplage. Celui-ci suscite des stratégies d'actions en fonction de la nature des retours sensoriels qu'il rend possible. Ces actions sont d'une importance cruciale, dans la mesure où elles contraignent la dynamique interpersonnelle et influencent par là même ce qui fait sens pour les interactants. Lorsque les stimulations résultant du croisement perceptif et les stimulations résultant de la rencontre des objets sont différenciées, les performances sont les meilleures, et témoignent même d'un effet plafond dès le premier bloc d'essais. Les sujets peuvent alors faire sens de l'engagement de l'autre participant en fonction des objets présents, cet engagement se manifestant par un arrêt de l'individu sur l'objet qu'il perçoit. La condition « Différenciée » semble donc être une solution d'usage pertinente, qu'il convient d'envisager dans le cadre de l'exploration mutuelle de contenus numériques.

Pour autant, les meilleures performances d'exploration et de reconnaissance de formes par des utilisateurs aveugles sont obtenues avec une matrice de 4 x 4 champs récepteurs élémentaires couplés à autant de stimulateurs tactiles (Sribunruangrit, 2004). Il est à souligner que réellement explorer ensemble un contenu numérique est déjà difficile en soi (voir Deschamps, 2013). La différenciation des stimulations comme on la trouve dans la condition « Différenciée » nécessiterait l'usage de plusieurs matrices de stimulateurs (comme sur notre prototype de module d'interaction tactile), ce qui présente des limites. Le parallélisme permet l'exploration mutuelle de contenus de manière plus efficace que le simple mono, et ne présente pas les limites de la différenciation. Les utilisateurs peuvent en condition de parallélisme coordonner leurs activités perceptives afin d'organiser leur interaction en fonction des formes inscrites à l'écran. Toutefois, il peut y avoir des interférences entre les deux sources de stimulations (la forme à explorer et le corps-image d'autrui), interférences qui peuvent ponctuellement entraver la dynamique de l'interaction en masquant la forme à explorer ou le corps-image du partenaire.

L'attention conjointe est un mécanisme important à prendre en considération dans le cadre de la collaboration à distance, mais aussi dans d'autres situations. Elle permet une identification de l'autre participant et de l'objet de son attention, notions dont l'utilité s'étend au-delà de la collaboration. En effet, les espaces numériques partagés manquent de manière générale

d'indices quant au statut partagé ou privé des objets. Riches des résultats de cette expérience, nous allons à présent exposer une proposition de design mettant en jeu l'attention conjointe dans un contexte particulier.

3.4 L'attention conjointe dans facebook : expliciter le public et le privé

Dans notre expérience, nous avons proposé une tâche à réaliser dans la situation particulière, et nous l'espérons révélatrice, de Tactos afin d'étudier les mécanismes de l'attention conjointe. Nous allons à présent nous concentrer sur un autre type d'environnement d'interaction : celui du réseau social facebook. Nous nous proposons de réfléchir aux mécanismes d'attention conjointe pouvant être mis en jeu dans ce contexte, dans le but d'améliorer l'interaction permise par ce site.

Nous souhaitons souligner ici que l'étude expérimentale que nous venons de présenter mettait en jeu une attention conjointe synchrone. Cela présente un intérêt dans notre situation expérimentale : l'objet commun n'est reconnu comme commun que par la rencontre de l'activité perceptive synchrone. L'objet commun reste commun même quand l'autre participant n'est plus là, mais ce n'est pas ce que nous avons considéré dans notre expérience. Au contraire, dans facebook, quand quelque chose est commun, il l'est et ce même si les autres utilisateurs n'y sont pas. Il s'avère qu'il est donc différent de savoir si un élément est commun et si nos attentions sont conjointes sur cet objet, synchroniquement ou pas.

La distinction entre ce qui est public et privé est l'indice d'une attention conjointe possible. Il y a donc un travail à deux niveaux : celui de l'attention conjointe synchrone (soit deux personnes simultanément) et la question du public et du privé (le privé étant synonyme de l'impossibilité de l'attention conjointe, tandis que le public traduit la possibilité de l'attention conjointe). Nous allons ici considérer ces deux niveaux.

3.4.1 Facebook et le déficit en attention conjointe ; qu'est-ce qui est privé, qu'est-ce qui est public ?

Du point de vue du design, nous cherchons une manière d'exploiter la notion d'attention conjointe afin de pouvoir améliorer l'expérience proposée par le site de réseau social le plus

utilisé au monde à ce jour : facebook.com⁴⁶. Les réseaux sociaux permettent un certain contact entre leurs utilisateurs. La sociabilité sur le web propose des opportunités culturelles et sociales par le déploiement d'un cercle élargi (Cardon, 2011). Les réseaux sociaux ont donc un intérêt pour le contact et le maintien du réseau de relations de chacun.

Facebook permet à un utilisateur de se créer un compte, et de lier ce compte à celui des « amis » : toute personne dont vous acceptez la demande ou qui accepte la vôtre. Des informations relatives à chaque utilisateur se retrouvent sur la page récapitulative de son compte. Chaque utilisateur possède sa propre page (son mur), où il peut publier des statuts, des photos et partager des liens, ou recevoir des messages. L'utilisateur peut régler les paramètres de confidentialité des éléments liés à son compte ; privé (non-accessible par un tiers), accessibles à certains groupes d'amis de son choix, à tous les amis, ou par tous (utilisateurs inscrits ne faisant pas partie de la liste d'amis ou utilisateurs non-inscrits). On désigne sous le terme post tout élément publié sur un mur (statut, photo, lien, son,...). L'utilisateur dispose aussi d'un « fil d'actualité », qui est privé et où apparaissent les posts de ses contacts. Il existe d'autre part une messagerie par laquelle il est possible d'envoyer des messages à une ou à plusieurs personnes à la fois ; ces messages ne sont pas publiés sur le mur et ne sont accessibles qu'à leur(s) destinataire(s). De plus, certains autres paramètres sont réglables afin que des informations spécifiques du compte ne soient accessibles qu'à l'utilisateur. Enfin, il existe une messagerie instantanée où les utilisateurs peuvent chatter en direct avec les autres utilisateurs de leur liste d'amis qui sont simultanément connectés à ce chat. Si l'on n'y est pas connecté, on ne peut pas savoir qui est en ligne. On peut par ailleurs limiter le nombre de personnes avec lesquelles la discussion est possible en les sélectionnant ; les autres utilisateurs n'apparaîtront pas dans notre liste de personnes connectées, et nous n'apparaîtrons pas dans la leur.

Les multiples possibilités de masquage (posts, photos, informations, messages, présence, liste d'amis,...) et les changements réguliers quant à la façon de les gérer sont sources de confusion pour les utilisateurs. En effet les options de confidentialité de facebook sont régulièrement modifiées, ainsi que la façon de gérer ses amis en groupe ; l'utilisateur n'a parfois pas conscience de ces changements, et il n'est pas rare de voir apparaître des polémiques autour de ce sujet (des personnes publient à tous des statuts qu'ils ne pensaient destiner qu'à certains, par exemple).

Les réseaux sociaux comme facebook permettent le développement de nouvelles formes de communication, où des énonciations traditionnellement considérées comme privées se

⁴⁶ Près d'1,2 milliard d'utilisateurs actifs en septembre 2013. Sources : <http://www.shareholder.com/visitors/event/build3/stage/stage.cfm?mediaid=61190&mediauserid=0> et <http://investor.fb.com/releasedetail.cfm?ReleaseID=802760>

retrouvent projetées sur la scène publique (Cardon in Heaton et al, 2012). Une porosité entre la conversation et l'espace public est apparue sur les réseaux sociaux. Les conversations peuvent prendre place sur une page facebook, qui est alors vue comme un espace intermédiaire entre le privé et le public.

Quelle solution serait adaptée afin de permettre aux utilisateurs de facebook d'avoir conscience de ce qui est privé pour eux ? De ce qui est public, et de pour qui ce contenu est public ? Nous souhaitons privilégier la piste de l'attention conjointe pour apporter une solution à cette question.

Sur facebook, les notions de public et de privé peuvent s'appliquer aux posts, aux photos, aux informations du compte (informations sur l'individu mais aussi liste d'amis par exemple), à l'activité, aux messages et à la présence (en ligne/hors ligne). L'attention conjointe n'est permise sur l'interface actuelle du site qu'à un niveau très restreint. Au niveau d'un post ou d'une photo, on peut savoir si quelqu'un l'aime ou le commente soit en restant sur ou en revisitant ce post, soit en recevant une notification par email, mais il n'y a pas d'indice quant au fait que quelqu'un d'autre est ou non en train de consulter le post s'il n'y agit pas. Il n'y a pas non plus d'indice direct pointant le degré de partage d'un élément (privé, public pour certains, public pour tous). Pour les informations de notre compte, il n'y a pas non plus d'indice perceptif direct pour savoir si l'information est ouverte à tous, ouverte à certains ou privée à soi ; il faut effectuer des clics pour pouvoir accéder à cette donnée. L'activité des utilisateurs (les posts qu'ils aiment ou commentent, ou ceux qu'ils postent) est partagée avec leur liste de contacts d'une manière très peu transparente pour eux : ils n'ont pas conscience qu'elle est partagée par défaut (leur seul indice est qu'ils reçoivent l'activité de leurs amis, signe qu'en retour ces mêmes amis ont accès à leur propre activité), et la procédure pour ne pas la partager est souvent mal connue. Les messages peuvent être adressés à un ou plusieurs destinataires, dont les noms apparaissent dans l'en-tête du message. Enfin, la présence n'est partagée que par le biais de la connexion au chat (et donc uniquement aux personnes qui sont elles-mêmes connectées et avec lesquelles on a choisi de pouvoir chatter, et qui ont choisi de pouvoir chatter avec nous).

Nous allons présenter une proposition de design (encore à l'état de prototype), inspirée par nos réflexions sur les indices perceptifs de l'attention conjointe et de la distinction public/privé. Ceci devrait permettre aux utilisateurs d'avoir plus facilement et directement accès au caractère public et privé des éléments qu'ils croisent sur facebook.

3.4.2 Proposition de design⁴⁷

Nous proposons ici une interface pour les utilisateurs de facebook visant d'une part à clarifier le degré de partage (privé, semi-public, public) des posts, photos, informations, activité et messages, et d'autre part à pouvoir interagir avec les autres utilisateurs par le biais de l'explicitation de l'attention conjointe. Son fonctionnement repose sur les principes de l'attention conjointe : identifier d'une part ce qui peut être le focus de l'attention conjointe, et d'autre part les individus pouvant participer d'une attention conjointe avec soi, avant de pouvoir soit suivre l'attention d'un autre participant, soit la provoquer.

Notre proposition d'interface repose sur des indices visuels, qui sont ceux actuellement utilisés par le site ; elle n'implique donc l'ajout d'aucun matériel supplémentaire, et cherche à s'intégrer de la manière la plus discrète possible dans le système existant.

3.4.2.1 La distinction public/privé des éléments de facebook

En ce qui concerne le degré de partage des posts, photos, informations et activité, nous proposons un codage par le biais de l'interface graphique. Il s'agit de mettre en place au niveau des coins de chaque élément (post, photo, information, activité) des éléments qui symboliseront l'étendue plus ou moins restreinte du partage de l'élément pris en considération.

En ce qui concerne les éléments strictement privés (non partagés), les coins auront une forme de tête de flèche pointant vers le centre du message (voir figure 134). Les flèches seront délimitées par un trait bleu et remplies en blanc.

⁴⁷ Il s'agit bien ici d'une *proposition* ; le design envisagé n'existe qu'à l'état de concept et n'a pas été implémenté pour le moment. Nous allons ici présenter la première version du prototype et les premiers retours utilisateurs reçus, à partir desquels le développement du concept pourra, dans le futur, prendre corps.



Figure 134. Signalement d'un post, d'une photo ou d'une information sur son propre compte/mur qui n'est pas partagée.

Les éléments partagés à tous (mode public) verront dans leurs coins des têtes de flèche tournées vers l'extérieur du message (voir figure 135). Les flèches seront intégralement bleues.

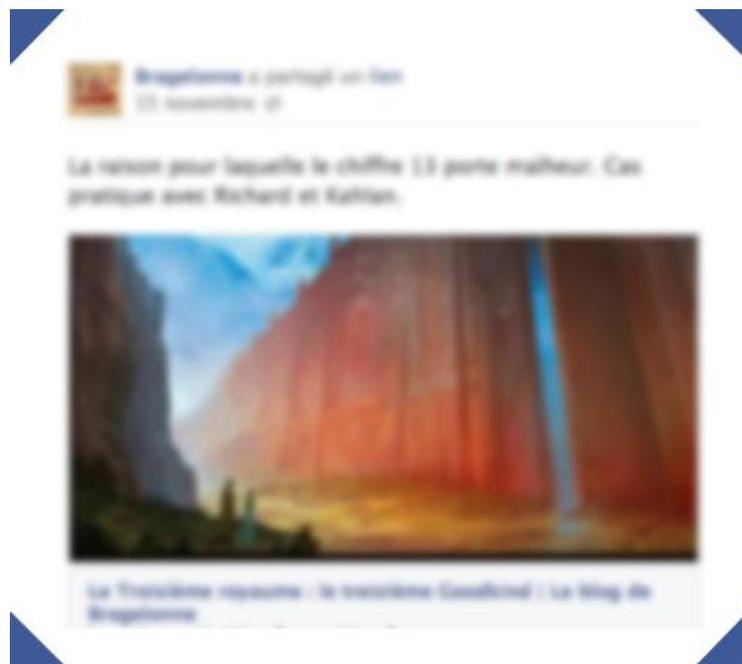


Figure 135. Signalement d'un post, d'une photo ou d'une information sur son propre compte/mur ouvert à tous.

Enfin, les messages accessibles uniquement aux amis ou à certains amis auront eux aussi des têtes de flèche tournées vers l'extérieur comme coins, mais cette fois-ci les flèches seront hachurées en bleu et blanc (voir figure 136).



Figure 136. Signalement d'un post ou d'une information sur son propre compte/mur ouvert aux amis ou à certains amis.

En ce qui concerne le partage de l'activité, nous préconisons l'ajout d'onglets sous la liste d'activité, afin de pouvoir soit accéder à l'activité des amis, soit à la nôtre (et alors pouvoir choisir le niveau de partage de l'activité, dont les éléments seront codés comme précédemment, à la différence qu'ils respecteront le code couleur de la section activité, avec des niveaux de gris au lieu du bleu). Il est à noter que les éléments codant le caractère public ou privé des éléments peut aussi s'appliquer aux boutons même de l'interface, ce qui nous donnera le sens plus ou moins partagé de nos actions ; ainsi, en 'aimant' un élément, nous pourrions savoir dans quelle mesure cette action sera partagée. Nous pourrions de plus envisager le fait de proposer aux utilisateurs de ne pas signer leurs actions ; il serait alors par exemple possible d' 'aimer' de manière anonyme.

Pour les messages enfin, nous proposons que dans le menu des messages tous les destinataires apparaissent sous la forme d'une vignette représentant leur photo de profil. La page des messages sera structurée en deux : du côté gauche, la liste des vignettes (organisée par ligne, chaque ligne présentant les destinataires d'un message particulier), et du côté droit les messages correspondants sur chaque ligne. Au delà de douze destinataires, nous proposons de regrouper

les vignettes 4 par 4 (en faisant des vignettes mosaïques de quatre destinataires chacune). En ouvrant un message, les coins changeront selon le nombre de destinataires : un coin bleu pour un (voir figure 137), deux coins bleus pour deux (voir figure 138), trois coins bleus pour trois (voir figure 139), trois coins dans un dégradé de bleus (plus foncé au plus proche du texte vers plus clair le plus loin du texte) pour quatre destinataires ou plus (voir figure 140).



Figure 137. Signalement d'un message envoyé à une personne.



Figure 139. Signalement d'un message envoyé à deux personnes.



Figure 138. Signalement d'un message envoyé à trois personnes.



Figure 140. Signalement d'un message envoyé à quatre personnes ou plus.

Le système que nous venons de présenter devrait avoir pour conséquence de rassurer les utilisateurs et de les mettre plus en confiance dans leur usage de facebook. Ils pourront en effet savoir plus facilement si leurs actions et eux-mêmes sont visibles par les autres utilisateurs, et dans quelle mesure. De plus, notre système permettra de mettre en évidence de manière plus marquée les actions des utilisateurs. L'une des leçons de nos parties théorique et expérimentale stipule que l'on reconnaît autrui à ses actions. Ce qui nous dévoile le plus n'est pas notre image, mais nos actions. En les mettant en avant, les utilisateurs pourront avoir une approche plus sereine de facebook, et en tireront une expérience interpersonnelle plus riche.

3.4.2.2 L'attention conjointe en coïncidence temporelle sur facebook

En ce qui concerne d'autre part l'interaction basée sur l'attention conjointe, il va s'agir dans un premier temps de prendre conscience des autres utilisateurs : ceux qui sont connectés bien sûr, mais aussi ceux qui explorent le même contenu que nous et ceux qui sont l'auteur de l'élément que l'on est en train d'explorer. La présence est donc le premier aspect à prendre en

considération : pour pouvoir avoir deux attentions focalisées sur un même objet partagé, il faut que les utilisateurs aient conscience l'un de l'autre.

Il est possible de mettre en place un compteur de visites sur un post dans facebook (Bernstein et al, 2013)⁴⁸. Notre proposition repose ici sur une donnée un peu plus précise : le besoin de savoir qui visite un post, une photo ou une page en même temps que nous (ou, dans le cas où l'utilisateur concerné ne souhaite pas être identifiable, si quelqu'un est en train de consulter le même contenu que nous).

Nous souhaitons en effet pouvoir signaler lorsqu'une personne est en train de consulter le même élément que nous. Nous optons pour le signalement par le biais d'un carré qui se déplacerait lentement le long du cadre de l'élément (post, photo, information, message) que l'on est en train de consulter, que celui-ci se trouve sur notre mur, sur notre compte, sur notre fil d'actualité ou sur le mur/compte d'un autre utilisateur (voir figures 141.a et 141.b). Selon que l'individu aura choisi d'être identifiable par un autre utilisateur, ce dernier verra son image de profil s'afficher ; sinon, le carré sera simplement d'une couleur unie. Si le carré présente un léger cadre extérieur, cela signifie que l'auteur de l'élément est connecté (et a accepté d'être identifié comme tel) (voir figures 141). Le carré de l'auteur tourne autour du post dans le sens contraire des carrés représentant les autres utilisateurs. Si plusieurs personnes sont en train de consulter le même élément que nous, alors plusieurs carrés correspondant chacun à une personne seront visibles (voir figure 141.a). Chacun voit se déplacer le carré qui le représente et peut se rendre anonyme en cliquant dessus. Il y a une forme de réciprocité ; on voit l'activité des autres utilisateurs, mais aussi notre propre activité telle que les autres utilisateurs la voient.

⁴⁸ Bernstein et al (2013) comptabilisent une visite sur un élément lorsque celui-ci a été affiché pendant au moins 900 millisecondes sur l'écran d'un utilisateur.



Figure 141.a. Signalement de la présence de cinq personnes (deux anonymes, deux identifiables, et l'auteur du post, entouré de bleu) sur l'un des posts, photo ou information que l'on est en train de consulter (t0).

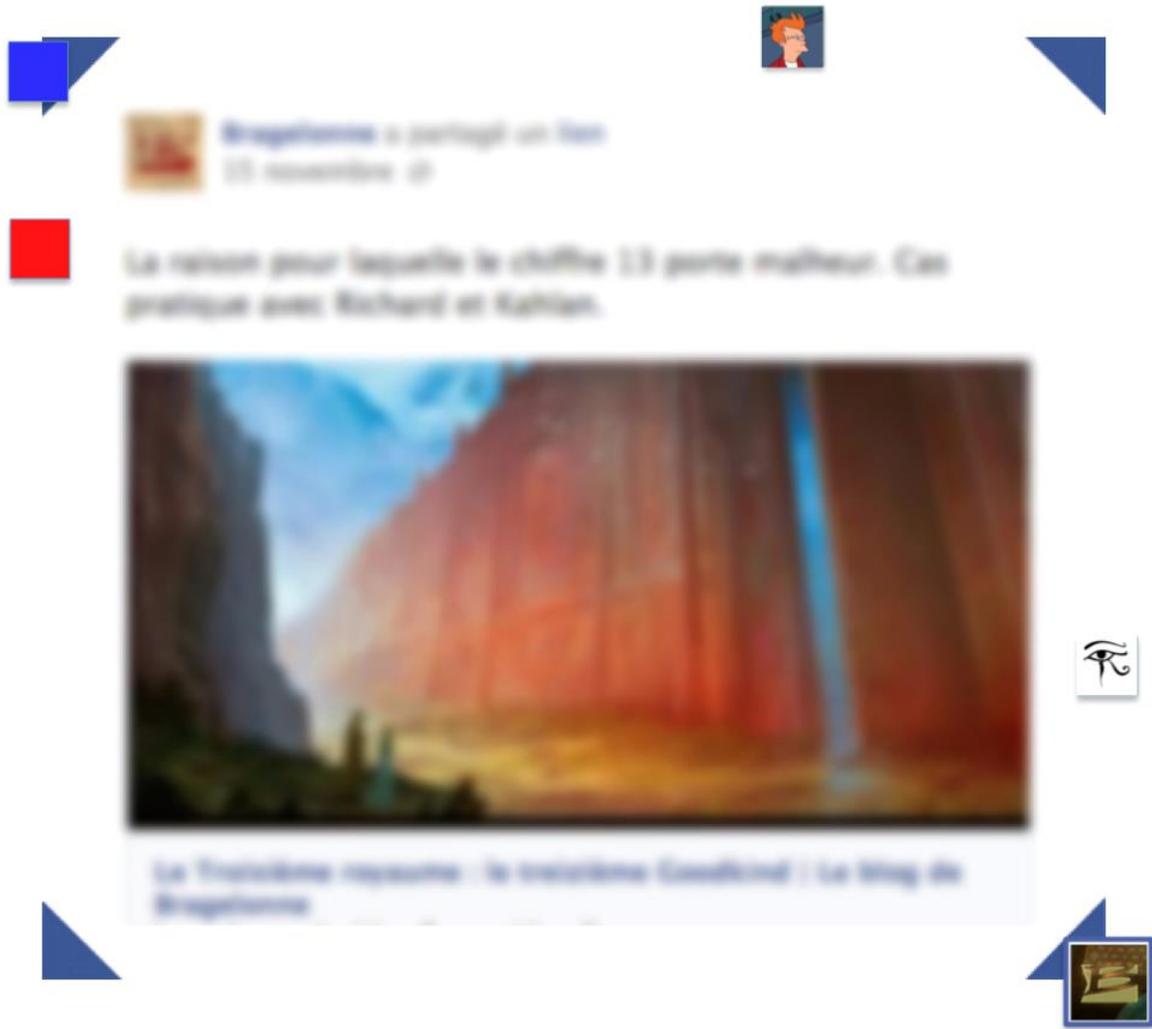


Figure 141.b. Signalement de la présence de cinq personnes (deux anonymes, deux identifiables, et l'auteur du post, entouré de bleu) sur l'un des posts, photo ou information que l'on est en train de consulter (t1).

Avec ce système, la question du respect de la vie privée entre en jeu ; tous les utilisateurs ne souhaitent pas être identifiés comme connectés, et plus particulièrement présents sur certains éléments. Nous souhaitons donc que les options soient configurables. Ainsi, l'utilisateur peut choisir de marquer son statut comme invisible (aucune identification), visible par certains amis, visible par les amis, visible par tous. Il pourra aussi régler ce statut par défaut pour certaines pages (être marqué comme toujours invisible sur cette page par exemple). Il pourra aussi choisir

le degré d'interaction qui sera possible avec lui s'il est visible : pouvoir être poké⁴⁹ sans être identifié ou pouvoir être identifié (soit pour être poké, soit pour démarrer un chat, soit pour pouvoir interagir sur l'élément actuellement visité en pointant certains éléments, qui seront alors encadrés en bleu, comme pour l'identification d'une personne sur une photo. Pour ce dernier type d'interaction, une demande sera émise et devra être acceptée par l'autre utilisateur pour que l'échange puisse avoir lieu). Toutes ces actions pourront être effectuées en cliquant sur le carré représentant la présence d'un utilisateur autour d'un post. Dans le cas du « poke aveugle », il sera bien signalé que le poke est envoyé à un utilisateur dont on ne connaît pas l'identité, et il sera possible de choisir si l'on envoie le poke en étant nous aussi masqué ou en dévoilant notre identité. Ainsi, il sera possible d'envoyer des pokes anonymes ou signés à des personnes dont on ignore l'identité, et de recevoir d'eux des pokes anonymes ou signés. Notre hypothèse est que cette interaction simple entre individus leur permettra un sentiment de présence des autres plus fort et améliorera leur expérience, et ce même lorsque l'anonymat est total.

La question de la rémanence se pose aussi, étant donné que la visite d'un élément sur facebook est en général très rapide⁵⁰ ; peut-être faudra-t-il que les utilisateurs y restent détectables pendant quelques minutes après l'avoir quitté. Dans ce cas, si une interaction autour de l'élément est lancée, il sera proposé de retourner sur l'élément en question.

Le carré en mouvement est l'écho des résultats que nous avons mis en évidence avec l'expérience objets communs et objets privés. Il est en effet une abstraction du résultat posant que l'on reconnaît l'activité perceptive d'autrui à travers ses actions. On pourrait aller encore plus loin dans une telle démarche en cherchant aussi à représenter les actions de bas niveau des utilisateurs, comme leurs mouvements de souris, ou animer par exemple le carré représentatif d'un utilisateur en le faisant se déplacer plus ou moins rapidement selon les mouvements de son curseur.

⁴⁹ Le poke est un élément fort apprécié de facebook, et nous ne manquons pas de souligner que c'est le terme le plus tactile du vocabulaire facebook.

⁵⁰ Selon le site <http://expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/>, une visite sur facebook dure en moyenne 20 minutes, et chaque utilisateur peut potentiellement être exposé à 1500 contenus différents à chaque login, soit une moyenne théorique de 800 ms par élément. Bien sûr, tous les éléments ne sont pas égaux en terme d'intérêt, mais cela témoigne d'un temps moyen théorique consacré à chaque contenu très court.

3.4.3 Retours utilisateurs

Nous avons soumis notre premier concept à une demi-douzaine d'utilisateurs par le biais d'explications et de captures d'écrans-types de ce que notre design pourrait offrir. Notre but était d'obtenir un retour quant à l'intérêt de notre proposition et aux améliorations possibles à lui apporter.

Les utilisateurs étaient tous familiarisés avec facebook, mais ne s'y rendaient pas à la même fréquence. Trois utilisateurs s'y rendaient quotidiennement, un s'y rendait environ trois fois par semaine, un une fois par semaine et un une fois par mois environ.

Les aspects du design ont été présentés de manière successive : d'abord le code autour des éléments pour le niveau de partage, ensuite le design envisagé pour l'activité, puis pour les messages, et enfin le partage de la présence et les interactions possibles qui y sont liées.

Les utilisateurs ont apprécié l'idée du codage autour des éléments (que ce soit le système post/photo/information ou le système pour les messages) ; il leur a semblé simple à comprendre, et l'idée de pouvoir connaître le degré de partage d'un élément d'un simple coup d'œil leur a paru séduisante. Certains ont noté la nécessité tout de même d'explicitier le code, en rendant par exemple une aide accessible, car si les signes choisis leur ont paru cohérents, ils ne pensaient pas pour autant qu'ils auraient pu deviner leur sens sans explication.

En ce qui concerne l'activité, les utilisateurs ont particulièrement apprécié l'idée de pouvoir avoir la main sur son aspect public ou privé. En effet, seuls deux des utilisateurs quotidiens savaient comment paramétrer l'activité et la rendre inaccessible ; les quatre autres utilisateurs ignoraient comment le faire mais ont exprimé leur envie de la masquer, ou en tous cas de ne pas la rendre accessible à tous leurs amis en permanence.

Pour l'organisation des messages, un utilisateur a exprimé son regret, avec le système que nous proposons, de ne pas pouvoir accéder à la liste des messages quand l'un d'entre eux est ouvert (ce qui est actuellement possible dans l'interface du site). En effet, notre système de présentation des destinataires implique un changement de page pour pouvoir accéder au message sélectionné ; il faut retourner en arrière pour pouvoir avoir la liste complète des messages. Les cinq autres utilisateurs ont noté que cela ne leur posait pas de problème, et qu'ils préféreraient pouvoir avoir accès en un seul coup d'œil à la liste de tous les destinataires.

Enfin, le système de présence et d'attention conjointe est celui qui a le plus intrigué et intéressé les utilisateurs. Notamment, l'idée des pokes anonymes leur a tout d'abord semblé incongrue, mais après y avoir réfléchi, tous sauf un ont déclaré que c'est une option qu'ils utiliseraient. Ils ont apprécié l'idée de pouvoir plus ou moins se « montrer », et de pouvoir avoir l'indication que l'auteur d'un élément est en ligne (si celui-ci accepte d'être identifié comme tel). L'interaction d'attention conjointe autour de l'élément visité leur a paru peut-être un peu trop

spécifique ; ils y voient une application très ponctuelle et l'utiliseraient vraisemblablement peu. Certains ont noté que pour les posts les plus populaires, il serait possible qu'il y ait trop d'utilisateurs sur un seul élément pour pouvoir faire la distinction de tous les carrés les représentant, et pouvoir interagir facilement avec l'un d'entre eux. Cet élément sera à étudier plus en profondeur avec un système fonctionnel ; en effet, il est possible d'après nous que le faible temps moyen passé sur un élément puisse compenser cet aspect, et ce même si l'on opte pour une rémanence de quelques minutes après le passage sur un élément. Mais il est vrai que les posts les plus populaires, par exemple partagés dans des groupes de plusieurs dizaines de milliers d'utilisateurs, pourront peut-être poser problème ; il faudra alors envisager une réponse de design propre à ces éléments.

Globalement, les utilisateurs ont fait preuve d'un avis positif vis-à-vis de notre proposition, appréciant son aspect simple et peu chargé ainsi que l'interaction nouvelle basée sur la présence et l'attention conjointe. Au sujet du système de présence, l'un d'entre eux a déclaré : « C'est intéressant comme système, et ça permet de réellement sentir qu'on fait partie d'un réseau social. » Un autre a souligné : « C'est plus vivant comme ça, et ça permet de mieux gérer sa visibilité. »

3.4.4 Conclusion

Comblé le déficit en attention conjointe sur le réseau social facebook est une problématique à envisager sur plusieurs plans. D'une part, cela permettrait une meilleure utilisation et navigation par les utilisateurs, pour qui il n'est pas forcément facile de faire la différence entre le contenu qu'ils partagent (ou non), et avec qui. D'autre part, cela permettrait une expérience plus riche et interactive pour les utilisateurs ; en effet, l'attention conjointe est par définition une activité qui n'est pas solitaire, et elle permet par ce biais le sentiment de présence (si je peux porter mon attention sur le même élément qu'un autre et que je le sais, alors je suis aussi conscient de l'activité de l'autre participant et du fait qu'il est « là », avec moi, en train de se concentrer sur le même élément que moi).

Aussi simple que notre proposition soit, elle met déjà en évidence des enjeux forts se déroulant dans le contexte des réseaux sociaux. En effet, les utilisateurs se montrent sensibles à la notion de public et de privé, et se montrent désireux d'avoir la main sur tous les aspects de partage de leurs informations qu'ils souhaitent pouvoir maîtriser en toute connaissance de cause. À cela s'ajoute le désir de maîtriser l'accès que les autres utilisateurs pourront avoir à leur présence (connexion, mais aussi plus précisément présence sur un post) et aux possibilités d'interaction qu'ils pourront engager avec eux. L'approche que nous proposons pourrait donc

rendre facebook plus engageant pour certaines personnes en levant les craintes qui les bloquent quant au caractère public ou privé du contenu et des actions des utilisateurs. Notre système aurait donc un rôle de rassurance pour les utilisateurs.

Le désir d'un sentiment de présence accru est lui aussi mis en avant ; les utilisateurs souhaitent plus de « vie » sur les réseaux sociaux. Notre design basé sur les concepts de l'attention conjointe est une piste de réponse à ce besoin, notamment en ce qu'il met en lumière la présence d'autrui sur un élément que je consulte.

Notons que l'interface ici proposée est basée sur les mécanismes existants de la présentation des pages web, dont l'exploration se fait de manière solitaire ; il est par conséquent particulièrement difficile de mettre en avant l'activité perceptive des individus (avant même de pouvoir faciliter leur coordination). Ainsi, l'organisation autour des objets présents ne se fait que de manière indirecte : détection de la présence de l'autre participant autour de l'élément, puis mise en place d'une interaction avec l'autre interactant, autour de l'élément (exploration conjointe) ou ailleurs (chat, poke). Cette question de la coordination perceptive est majeure dans notre contexte, et appelle d'autres solutions que celle proposée ici, qui consiste avant tout à mettre en lumière le caractère partagé des éléments ainsi que la présence d'autrui ; l'engagement dans une tâche conjointe n'est pas systématique. Contrairement à l'expérience Commun/Privé, le caractère partagé des objets n'émerge pas suite à l'interaction des participants ; elle est au contraire déjà donnée.

Notre proposition n'en est encore qu'à un stade préliminaire. Suite aux retours encourageants de nos utilisateurs, il faudra dans le futur implémenter une version fonctionnelle du prototype, afin de pouvoir véritablement la tester et l'améliorer. Les développements futurs devront porter une attention particulière à l'interaction proposée autour de la présence et de l'attention conjointe, qui est l'aspect le plus nouveau et le plus versatile de notre proposition. Les utilisateurs utilisent-ils l'interaction proposée ? Comment ? Préfèrent-ils recevoir et émettre des pokes anonymes ou identifiés ? Osent-ils signer un poke envoyé à une personne dont ils ignorent l'identité ? Qu'est-ce qui est le plus souvent utilisé : le poke, le chat ou l'exploration du contenu à deux ? Comment rendre l'exploration du contenu à deux plus riche ? Comment permettre aux utilisateurs d'accéder à cette idée de réseau vivant, engageant ? Comment développer un peu plus le sentiment de présence sur les réseaux sociaux ?

3.5 L'articulation Théorie-Design-Expérimentation : une deuxième approche

Dans cette partie, nous avons effectué une deuxième itération de la boucle Théorie-Expérimentation-Design (voir figure 142). Partant du socle théorique présenté en partie 1 et de la notion d'attention conjointe, nous avons tout d'abord mis en évidence le fonctionnement de cette dernière. L'attention conjointe fait entrer en jeu deux individus focalisés sur un objet ou événement externe, en ayant tous deux conscience que l'attention de l'autre participant se porte sur le même objet que la leur. Ce phénomène ouvre donc l'interaction dyadique à un objet tiers partagé. Afin d'étudier ce mécanisme plus en détail dans le cadre d'un environnement numérique partagé, nous avons mis en place une expérimentation utilisant le dispositif minimaliste Tactos. Les résultats de cette expérience mettent en avant le rôle du mouvement et de la dynamique d'interaction autour des objets partagés. Ils permettent aussi de réfléchir au type de retours sensoriels à mettre en place ; il ressort de notre expérience que c'est la condition « Parallélisme » qui serait la plus adaptée pour une exploration collaborative de contenu par le biais de Tactos. Dans un deuxième temps, nous avons exposé une proposition de design destinée au réseau social facebook. Elle consiste d'une part à mettre en évidence le caractère public ou privé d'un élément par le biais d'un codage graphique directement visible (ce qui s'inspire de la meilleure performance des sujets en condition « Différenciée »), et d'autre part à pouvoir savoir si d'autres individus consultent le même contenu que nous et si l'auteur de l'élément actuellement visualisé est en ligne. Il est possible d'interagir avec les individus présents sur le même contenu que nous par le biais d'un poke, d'un message ou d'une interaction (basée sur le pointage de parties spécifiques) autour de l'élément actuellement consulté. Nos premiers retours utilisateurs, positifs, ont fait état de nouvelles propositions et limites à explorer lors de l'implémentation de notre prototype. Ils ont en même temps permis de soulever une série de suggestions concernant non seulement notre proposition, mais aussi leur usage de facebook et les enjeux des réseaux sociaux en général.

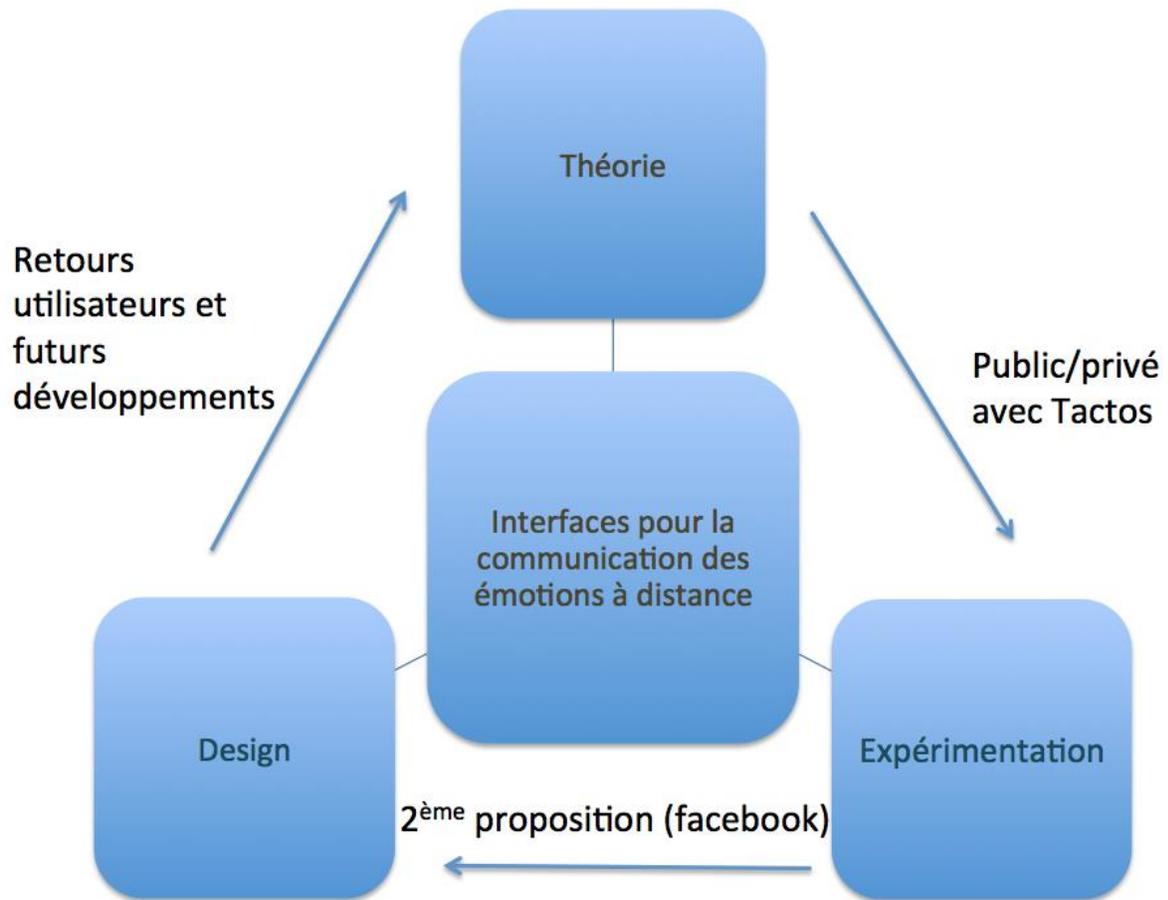


Figure 142. La deuxième itération de la boucle Théorie-Expérimentation-Design.

4 Perspectives et conclusion générale

Nous allons à présent clore notre travail, en nous penchant dans un premier temps sur les perspectives que nous envisageons pour des développements futurs possibles. Nous allons pour cela principalement présenter l'expérience préliminaire que nous avons menée pour explorer la question d'un système de signes tactiles. Dans un second temps, nous allons procéder à la conclusion générale.

4.1 Perspectives

4.1.1 Une expérience préliminaire sur un « langage » tactile

Nous avons mené une réflexion sur l'interaction tactile médiatisée sur le réseau, que ce soit par Tactos ou par Touch Through. Nous nous sommes plus précisément intéressés à la question de la valeur émotionnelle à attribuer à l'interaction tactile, dans le but de permettre une interaction porteuse de sens entre deux utilisateurs ou plus. Nous avons vu d'autre part que Barthélémy Maillet (2010) avait mis en place des émoticônes tactiles, conçues dans le cadre d'un travail exploratoire sur les interactions haptiques. Il cherchait à enrichir la dynamique interactionnelle dans le cas d'une communication asynchrone (comme celle d'un chat) grâce à la modalité tactile. Il a donc proposé une alternative aux smileys, afin de pouvoir matérialiser les indices non-verbaux et para-verbaux de la communication par le biais de la modalité tactile, pour enrichir les interactions écrites. Il a montré que les séquences de stimulations proposées pour chaque « taticône » n'ont pas de sens autre que celui que leur ont attribué ceux qui les ont produites. L'émergence de sens n'est ici possible entre n utilisateurs que dans le cadre d'une convention explicite. Plusieurs questions émergent alors, comme notamment celle des conditions d'émergence d'une signification à partir d'une stimulation tactile abstraite.

Notre propre travail sur la communication émotionnelle à distance nous a poussés à explorer la piste du toucher, et nous avons mis en place Touch Through, une application se basant sur les propriétés fonctionnelles du toucher pour en permettre la suppléance perceptive. Afin de mettre en place un échange qui serait plutôt axé sur le partage d'informations, il est intéressant de considérer l'idée d'un langage tactile minimaliste. Nous avons mis en place une étude exploratoire permettant d'examiner de plus près la construction d'un tel langage. Nous avons déjà rencontré la question du rôle du langage dans la communication émotionnelle (voir partie 1). Gumtau (2005) désigne sous le terme « sémiotique tactile » les significations que l'on attribue à différents types de sensations haptiques dans le cadre de langages basés sur le toucher.

Pour lui, ce type de langage peut être subdivisé en deux classes : la classe alphabétique et la classe symbolique.

La classe alphabétique utilise des représentations des lettres alphanumériques pour former des mots. L'écriture braille constitue un premier codage tactile des lettres pour la lecture, mais il en existe d'autres, comme par exemple le « Lorm », développé pour les individus souffrant de déficiences à la fois auditives et visuelles. Le Lorm consiste à utiliser des signaux tactiles tracés dans la main de l'interlocuteur pour représenter les lettres alphabétiques.

À l'inverse, le langage symbolique vise à représenter directement des concepts de plus haut niveau, comme des idées ou des émotions. Par exemple, on retrouve comme langage symbolique les expressions faciales, les mouvements des mains et le langage corporel, qui permettent d'exprimer l'intérêt et l'état émotionnel. Des recherches ont montré que l'utilisation des lettres alphanumériques est bien plus lente mais bien plus précise que celle d'un langage symbolique (Reed et al., 1990).

Ainsi, il existe deux approches possibles pour concevoir des systèmes haptiques dédiés aux interactions interpersonnelles. L'approche codée d'une part vise l'échange d'informations, ou l'enrichissement des informations échangées entre les utilisateurs, en passant par des signaux tactiles. Il s'agit alors de proposer des signaux haptiques aisément identifiables, différenciables et assimilables. Il n'existe pour l'instant que des propositions d'icônes haptiques, aucune n'étant mise en place à grande échelle. Il n'y a donc pas de communauté d'utilisateurs qui s'en seraient saisi et auraient établi un sens partagé.

Se distinguant de l'approche codée, l'approche symbolique ouvre d'autre part la possibilité d'une interaction directe de type tactile en tant que telle, qui ne chercherait pas à coder une information spécifique supplémentaire, mais à transmettre directement des émotions. C'est ce type d'approche que nous avons implémenté dans l'application Touch Through. On constate que les deux approches se distinguent tant par la nature des interactions haptiques proposées que par les processus de conception qu'elles requièrent. Cependant, nous pensons qu'il s'avère pertinent de les combiner. En effet, à partir d'une mise en relation haptique directe, comme celle permise par Touch Through, une communauté d'utilisateurs pourrait construire un espace de signifiants partagés, aboutissant potentiellement à l'émergence d'icônes haptiques. Il ne s'agirait alors pas de signes imposés par les concepteurs parallèlement aux interactions des utilisateurs, mais de signes résultant d'une négociation dans la situation d'interaction. Partant de cette idée, nous avons donc réalisé une expérience exploratoire autour du langage tactile.

La question du signe, de son émergence et de sa négociation est donc celle que nous considérons ici. Nous adoptons à nouveau la posture énonciationniste qui pose la cognition comme un processus incarné, situé et social, tout en considérant la dimension émotionnelle de notre problématique. Eggs définit une émotion comme une façon d'engager son corps dans un scénario

donné, ce qui implique que sa manifestation est toujours liée à des indices corporels (mimique, intonation, gestuelles,...), qui forment un système sémiotique spécifique (in Plantin, Doury, et Traverso, 2000). Certaines approches phénoménologiques nourrissent une perspective compatible avec de nombreux résultats empiriques de la neuroscience cognitive contemporaine : le sens serait le résultat de nos interactions situées avec le monde (Glenberg et Gallese, 2012). Le langage est une part éminente du milieu où nous vivons, il n'est pas dans le cerveau individuel mais dans la société qui nous englobe ; c'est cette dernière qui constitue le véritable organe du langage. Ni interne ni externe, la langue est ainsi un lieu du couplage entre l'individu et son environnement (Rastier, 2006).

Nous souhaitons mettre en place un premier système de signes tactiles que les sujets puissent s'approprier et/ou qu'ils puissent créer eux-mêmes. Nous voulons voir s'il est possible que de premiers exemples de signes puissent se mettre en place. Nous voulons vérifier que, dans le cadre de l'utilisation de nos dispositifs minimalistes (Tactos, Touch Through), il serait possible d'assister à une émergence sémiotique basée sur des signes qui prendraient sens au sein des interactions entre les utilisateurs. Il en va de même dans la communication naturelle, qu'elle soit verbale ou non verbale. Le sens n'est en effet pas circonscrit dans le signe (qu'il soit linguistique ou gestuel), mais émerge en fait d'une perpétuelle (re)négociation au sein de l'interaction (Varela et Maturana, 1994). Ce mécanisme de négociation implique une organisation temporelle très précise de l'interaction, afin de mettre en place une alternance des tours de parole favorisant un échange structuré d'informations.

Nous allons à présent présenter l'expérience préliminaire que nous avons mise en place pour nous saisir de ces questions. Il s'agit d'une expérience exploratoire, visant à tester les capacités d'émergence d'un système de communication de façon plus ou moins spontanée à travers des interactions entre sujets. Nous précisons que nous ne visons pas à faire une théorie sémiotique, étant donné la complexité de ce domaine. Nous voulons cependant étudier les capacités qu'ont les sujets à produire, de façon spontanée à travers leur interaction, des systèmes de signes. C'est ce que nous avons voulu examiner ici au travers de cette expérience. Cette dernière présente de plus un intérêt du point de vue du design, étant donné que l'on va pouvoir aussi étudier l'appropriation du système qu'est Touch Through couplée à un module d'interaction tactile. Le système est bien défini : c'est un ensemble d'objets et de programmes d'interaction. Mais la saisie du système par les utilisateurs est une autre question. Il y a un enjeu d'appropriation sociale, dont nous allons pouvoir amorcer l'étude avec cette expérience.

Cette expérience est préliminaire en ce qu'elle va nous permettre d'explorer nos questions, afin de pouvoir à l'avenir décider de la piste la plus adéquate à suivre pour des études plus poussées. C'est une expérience qui va nous permettre de voir les éléments qui émergent et d'obtenir des pistes d'exploration de nos questionnements.

4.1.2 Population

8 sujets, répartis en 4 binômes, ont participé à cette expérience exploratoire. Les sujets sont des étudiants de l'Université de Technologie de Compiègne âgés de 18 à 28 ans. Ils n'ont pas de connaissance de l'application Touch Through et n'ont pas connaissance des éléments et hypothèses théoriques mobilisés dans l'expérience. Du fait du faible nombre de participants, il n'y a qu'un seul groupe expérimental.

4.1.3 Matériel

Pour cette expérience, nous avons relié les smartphones des participants à un Module d'Interaction Tactile (MIT) ; les utilisateurs vont ainsi pouvoir recevoir des stimulations tactiles en posant leur doigt sur deux cellules brailles. Chaque participant est installé devant un poste de travail qui comprend un smartphone et un Module d'Interaction Tactile (MIT). Le Module d'Interaction Tactile possède deux cellules brailles côte à côte ; elles constituent une matrice tactile de 16 picots qui peuvent être chacun soit abaissés, soit levés. Il est indiqué aux sujets de sentir les picots avec leur main non-dominante et d'utiliser leur main dominante sur l'écran du téléphone.

Le dispositif expérimental comprend un ordinateur et deux smartphones : l'ordinateur constitue le serveur, qui va gérer les connexions-client et enregistrer les données, tandis que les téléphones constituent les clients et sont destinés à chacun des deux sujets de chaque binôme. Les smartphones sont des HTC One S.

L'espace sur lequel agissent les sujets correspond géométriquement à un carré de 3 x 3 cm sur la surface de l'écran des smartphones. Cette zone d'action est quadrillée en 4 x 4 carrés de même taille. À chaque carré est lié l'un des picots du MIT du partenaire ; lorsqu'un doigt d'un sujet touche un (ou plusieurs) carré(s), le(s) picot(s) correspondant(s) se lève(nt) chez son partenaire (et réciproquement). Ainsi, à l'aide de leurs doigts, les sujets modifient leur corps-image, c'est-à-dire la forme qui est retranscrite chez le partenaire par l'activation des picots de sa matrice tactile.

Les sujets sont informés que l'écran est divisé en 4 x 4 carrés qui correspondent aux picots de leur partenaire. Lorsqu'un sujet touche un carré, le picot correspondant se lève donc sous le doigt de son partenaire. Il y a un caractère asymétrique de chaque action : je ne sais pas ce que je fais, c'est mon partenaire qui le reçoit, et de même qu'il ne sait pas ce qu'il fait, c'est moi qui le reçois. Le sujet est averti par une vibration de son smartphone quand il touche une partie de l'écran extérieure à la zone d'action.

4.1.4 Procédure

Les passations se font en binôme. Les sujets sont accueillis dans une salle, et le dispositif expérimental leur est très brièvement présenté, ainsi qu'une consigne globale décrivant l'expérience à laquelle ils vont participer : les sujets sont informés que l'expérience va consister à interagir par le biais du dispositif afin de mettre en place des signaux tactiles. Il est également annoncé que l'expérience est divisée en trois phases : une phase de familiarisation, une phase consacrée à la tâche et une phase de restitution. Après cette consigne, les participants sont installés dans deux salles différentes, de façon à les séparer et à éviter tout indice perceptif autre que les stimulations délivrées par le dispositif expérimental. Chaque sujet est accompagné d'un expérimentateur. Les sujets sont assis devant le smartphone et le MIT qui leur sont respectivement attribués. Lorsque les sujets sont prêts, le premier essai de familiarisation commence. Notons que durant tous les essais, l'ensemble de l'activité des sujets dans l'application est enregistré sous forme numérique.

Dans la première partie de la phase de familiarisation, les sujets découvrent et s'approprient le dispositif. Les sujets sont mis en interaction pendant un essai non-chronométré, afin qu'ils comprennent le fonctionnement de l'interaction et en fassent l'expérience. Ils ont pour consigne d'essayer de s'imiter, c'est-à-dire de synchroniser leurs actions et donc les stimuli qu'ils s'envoient réciproquement. Les expérimentateurs s'assurent que les sujets ont compris le principe de fonctionnement.

La deuxième partie de la familiarisation consiste à élaborer deux signes, servant à exprimer des émotions. On indique aux sujets : « Vous devez réussir à vous mettre d'accord sur un signe pour exprimer ... ». Les sujets doivent convenir d'un signe pour « joie ». Ensuite ils doivent négocier le signe « tristesse ». Pour chacun de ces essais, ils disposent de 2 minutes ; on arrête l'essai soit lorsque le temps est écoulé, soit lorsqu'ils se sont mis d'accord (les sujets le signalent aux expérimentateurs, qui coupent l'essai). Pour vérifier le système mis en place, les sujets doivent refaire « joie », puis refaire « tristesse ». Ces essais de répétition des signes convenus durent une minute.

Dans la phase consacrée à la tâche, les sujets savent qu'ils doivent élaborer en commun des signes associés à des éléments permettant de se donner rendez-vous. Notre stratégie pour la construction des signes a été de demander aux sujets de construire certains mots très généraux, pouvant chacun renvoyer à une famille de signification assez proche, à la fois dans la description des choses, dans la description de la relation et dans la description du langage lui-même. Les notions choisies sont capables à la fois de désigner des notions, et de désigner la relation entre les locuteurs ; elles s'appliquent donc à deux niveaux. Les deux sujets doivent convenir ensemble d'un signe pour 8 notions :

1. « oui »/« je suis d'accord »/« je comprends »,
2. « non »/« je ne sais pas »/« je ne suis pas d'accord »/« je ne comprends pas »,
3. « chez sujet A »,
4. « chez sujet B »,
5. « ailleurs »,
6. « maintenant »
7. « plus tard »
8. « est-ce que tu comprends ? »/« est-ce que tu es d'accord ? »

Certains de ces éléments sont liés à la communication sur la communication (métalangage) et/ou à un scénario de prise de rendez-vous que l'on propose aux sujets. À chaque fois, les sujets disposent d'essais de 2 minutes maximum pour définir leurs signes ; on arrête l'essai soit lorsque le temps est écoulé, soit lorsqu'ils se sont mis d'accord (les sujets le signalent aux expérimentateurs, qui coupent l'essai). Tous les signes sont refaits une seconde fois, mais cette fois-ci les sujets ne disposent que d'une minute.

Dans la dernière phase, les sujets sont tout d'abord invités à restituer les signes mis en place. L'un des deux sujets doit réaliser un signe correspondant à l'un des éléments, désigné par l'expérimentateur. Les éléments étaient dans cet ordre :

- Sujet B joue « chez sujet A »
- Sujet A joue « maintenant »
- Sujet B joue « ailleurs »
- Sujet A joue « est-ce que tu comprends ? »/« est-ce que tu es d'accord ? »
- Sujet B joue « non »/« je ne sais pas »/« je ne suis pas d'accord »/« je ne comprends pas »
- Sujet A joue « chez sujet B »

Le partenaire doit alors reconnaître le signe dont il s'agit. Chacun des sujets doit ainsi alternativement exprimer et tenter de reconnaître 3 signes. Ensuite, 4 scénarios de rendez-vous sont demandés pour la restitution (2 sont joués par le sujet A et devinés par le sujet B, et 2 sont joués par le sujet B et devinés par le sujet A) :

- Sujet A joue : « chez sujet A, maintenant »
- Sujet B joue : « chez sujet A, plus tard »
- Sujet A joue : « ailleurs, je ne sais pas quand »
- Sujet B joue : « chez sujet B, je ne sais pas quand »

Les deux derniers scénarios nécessitent pour les sujets de faire preuve d'imagination, car ils n'ont pas défini de signe pour « je ne sais pas quand » durant les phases précédentes. Chaque essai de restitution dure 3 minutes.

Il s'agit de voir si les sujets arrivent à mettre en place un système de signes basiques et à se l'approprier.

4.1.5 Résultats

Le premier binôme a mis en place des signaux très simples pour « joie » et « tristesse » : respectivement le picot supérieur gauche et le picot inférieur droit. De même, les sujets ont tout d'abord attribué un seul picot aux signes « oui »/« je suis d'accord »/« je comprends » et « non »/« je ne sais pas »/« je ne suis pas d'accord »/« je ne comprends pas », avant d'opter pour des lignes horizontales, puis des paires de picots voisins pour les signes suivants. À chaque fois que plusieurs picots sont impliqués, ils sont utilisés séquentiellement, et non en même temps. En ce qui concerne la phase de restitution, les résultats sont présentés dans le tableau 13.

Tableau 13. Réponses données par le binôme 1.

Sujet	Essai	Réponse attendue	Réponse donnée	Bonne réponse	Mauvaise réponse
A	1	Chez sujet A	Chez sujet A	1	0
B	2	Maintenant	Ailleurs	0	1
A	3	Ailleurs	Plus tard	0	1
B	4	Est-ce que tu comprends ?/Est-ce que tu es d'accord ?	Oui	0	1
A	5	Non/Je ne sais pas/Je ne suis pas d'accord/Je ne comprends pas	Non/Je ne sais pas/Je ne suis pas d'accord/Je ne comprends pas	1	0
B	6	Chez sujet B	Maintenant	0	1
B	7	Chez sujet A, maintenant	Chez sujet B, maintenant	1	1
A	8	Chez sujet A, plus tard	Chez sujet A, plus tard	2	0
B	9	Ailleurs, je ne sais pas quand	Plus tard, d'accord	0	2
A	10	Chez sujet B, je ne sais pas quand	Non pas, chez sujet B	1	1

Le binôme 1 totalise 42,86% de bonnes réponses. Les sujets du premier binôme ont déclaré avoir réussi facilement à se mettre d'accord, et ont tous les deux fait des remarques sur la

difficulté à différencier les picots sous leurs doigts. Cela s’explique en partie par le fait qu’ils aient choisi des signes simples (1 ou 2 picots activés pour la moitié de leurs signes, 4 picots activés pour l’autre moitié). Les sujets ont opté pour des signes abstraits, mais ont souvent essayé de construire les signes par paire, en choisissant par exemple d’inverser l’ordre d’un signe pour en créer un autre.

En ce qui concerne le second binôme, les signes adoptés étaient plus complexes que ceux du binôme 1 et comportaient toujours au minimum 4 picots, là encore mis en jeu séquentiellement. Pour la phase de restitution, les résultats sont présentés dans le tableau 14.

Tableau 14. Réponses données par le binôme 2.

Sujet	Essai	Réponse attendue	Réponse donnée	Bonne réponse	Mauvaise réponse
A	1	Chez sujet A	Je ne comprends pas, maintenant	0	1
B	2	Maintenant	?	0	1
A	3	Ailleurs	Ailleurs	1	0
B	4	Est-ce que tu comprends ?/Est-ce que tu es d’accord ?	D'accord	0	1
A	5	Non/Je ne sais pas/Je ne suis pas d’accord/Je ne comprends pas	Est-ce que tu es d'accord ?	0	1
B	6	Chez sujet B	Chez sujet B	1	0
B	7	Chez sujet A, maintenant	Maintenant, est-ce que tu comprends ?	1	1
A	8	Chez sujet A, plus tard	Ailleurs, maintenant	0	2
B	9	Ailleurs, je ne sais pas quand	Chez sujet B, oui je suis d'accord	0	2
A	10	Chez sujet B, je ne sais pas quand	?, plus tard	0	2

Le binôme 2 obtient un résultat correct dans seulement 21,43% des cas, ce qui en fait le binôme avec le plus mauvais score. Les sujets de ce binôme ont fait remarquer en débriefing

qu'ils auraient voulu avoir accès à la possibilité d'obtenir une confirmation de la part de leur partenaire ; ils n'ont donc pas eu l'idée d'utiliser les signes en tant que méta-langage. C'est ce binôme dont les sujets ont aussi le plus considéré qu'ils n'arrivaient pas toujours à se mettre d'accord (mais même ainsi, ils n'avaient cette sensation que pour un quart des essais). Le faible score obtenu peut donc s'expliquer par une incompréhension mutuelle et une certaine incertitude quant aux signes définis, malgré la phase où les signes devaient être rejoués une première fois. De plus amples analyses à l'aide de l'enregistrement des traces pourraient nous en apprendre plus sur le mauvais score de ce binôme.

Pour leur part, les membres du troisième binôme ont eux aussi mis en place des signes impliquant quatre picots ou plus de manière séquentielle. Ils ont généralement ajouté le signe « oui » suite à la négociation des signes afin de signaler le fait qu'ils se soient bien mis d'accord sur le signe choisi. Cela est donc une indication que ce binôme a fait appel aux signes pour parler de la conversation elle-même. Les items « oui »/« je suis d'accord »/« je comprends », « non »/« je ne sais pas »/« je ne suis pas d'accord »/« je ne comprends pas » et « est-ce que tu comprends ? »/« est-ce que tu es d'accord ? » jouent alors un rôle de métalangage. Les résultats de ce binôme sont présentés dans le tableau 15.

Tableau 15. Réponses données par le binôme 3.

Sujet	Essai	Réponse attendue	Réponse donnée	Bonne réponse	Mauvaise réponse
A	1	Chez sujet A	Chez sujet A	1	0
B	2	Maintenant	Maintenant	1	0
A	3	Ailleurs	Ailleurs	1	0
B	4	Est-ce que tu comprends ?/Est-ce que tu es d'accord ?	Est-ce que tu comprends ?/Est-ce que tu es d'accord ?	1	0
A	5	Non/Je ne sais pas/Je ne suis pas d'accord/Je ne comprends pas	Non/Je ne sais pas/Je ne suis pas d'accord/Je ne comprends pas	1	0
B	6	Chez sujet B	Oui	0	1
B	7	Chez sujet A, maintenant	Maintenant	1	1
A	8	Chez sujet A, plus tard	Ailleurs, plus tard	1	1
B	9	Ailleurs, je ne sais pas quand	Ailleurs	1	1
A	10	Chez sujet B, je ne sais pas quand	Chez sujet B	1	1

Le troisième binôme a répondu correctement dans 64,29% des cas. Il est à noter que le sujet A a noté sur sa feuille de réponse même : « réception + signe + OK/Oui ». Nous avons là à nouveau l'indication que ce binôme a fait appel au métalangage, ce qui a été confirmé au cours de la phase de débriefing avec les sujets.

Enfin, le quatrième binôme a mis en place des signes de 4 picots ou plus, activés de manière séquentielle ou parallèle (dans le cas présent deux picots par deux picots), mais aussi un signe à un seul picot pour « est-ce que tu comprends ? »/« est-ce que tu es d'accord ? ». En ce qui concerne la restitution, les résultats de ce binôme sont présentés dans le tableau 16.

Tableau 16. Réponses données par le binôme 4.

Sujet	Essai	Réponse attendue	Réponse donnée	Bonne réponse	Mauvaise réponse
A	1	Chez sujet A	Chez sujet A	1	0
B	2	Maintenant	Maintenant	1	0
A	3	Ailleurs	Ailleurs	1	0
B	4	Est-ce que tu comprends ?/Est-ce que tu es d'accord ?	Est-ce que tu comprends ?/Est-ce que tu es d'accord ?	1	0
A	5	Non/Je ne sais pas/Je ne suis pas d'accord/Je ne comprends pas	Non/Je ne sais pas/Je ne suis pas d'accord/Je ne comprends pas	1	0
B	6	Chez sujet B	Chez sujet B	1	0
B	7	Chez sujet A, maintenant	Chez sujet A, maintenant	2	0
A	8	Chez sujet A, plus tard	Chez sujet A, plus tard	2	0
B	9	Ailleurs, je ne sais pas quand	Ailleurs, quand ??	1	1
A	10	Chez sujet B, je ne sais pas quand	Chez sujet B, ne pas	1	1

C'est donc le binôme 4 qui a obtenu le plus de réponses correctes avec un taux de 85,71% de bonnes réponses ; les seules mauvaises réponses fournies concernaient la difficile notion de « je ne sais pas quand », qui n'avait pas été définie au préalable. Il est à noter que ce résultat trouve peut-être son explication dans la dynamique des signes établis par ce binôme. En effet, les signes mis en place par le binôme 4 étaient plus fluides et rapides que ceux des autres binômes, dont les sujets avaient tendance à appuyer sur un picot, puis sur un autre en levant le pouce entre les deux, tandis que les sujets du binôme 4 laissaient leur doigt appuyé pour passer d'une case de la matrice sur l'écran à l'autre. Cette hypothèse sera à explorer dans un travail futur. Dans ce binôme, le sujet A a fait la remarque que leur signe pour « oui » était symbolique : il décrivait une boucle suivant le cercle extérieur des picots, que le participant associait « au O de 'oui' ».

Nous avons répertorié dans l'annexe D des exemples de signes mis en place par les utilisateurs, ainsi que des exemples de traces enregistrées au cours des essais des différents

binômes.

4.1.6 Remarques générales tirées de nos observations

Pour définir les signes, les sujets ont majoritairement adopté une stratégie d'imitation décalée : l'un des sujets propose un signe, et l'autre participant tente de le rejouer. Les sujets rejouent le signe jusqu'à être sûrs qu'ils jouent bien tous les deux la même chose. Certains binômes (1 et 2) se sont arrêtés là, tandis que les deux autres (3 et 4) ont choisi de jouer « oui » suite à cet échange, pour confirmer leur accord quant au choix du signe établi.

Les tours de parole entre les sujets étaient parfois définis au point que celui qui savait devoir être le récepteur de l'information posait son smartphone pour se concentrer plus pleinement sur les stimuli délivrés par le Module d'Interaction Tactile. D'autre part, plusieurs sujets ont déploré le manque d'un signe pour dire « peux-tu répéter ? ».

Nous notons qu'aucun binôme n'a réussi à se communiquer la notion de « je ne sais pas quand » qui était présente pour la première fois dans les scénarios de rendez-vous, et ce même si les deux sujets devaient la jouer l'un après l'autre (on pourrait penser que lors de la deuxième occurrence, les sujets avaient plus de chance de réussir, même si leurs rôles avaient été échangés).

Il est de plus intéressant de noter que le nombre de signes proposés (2 pour la familiarisation, 8 pour la tâche) poussait les sujets à prendre des notes pour être sûrs de se rappeler quel signe avait été négocié pour quelle notion. À l'avenir, il faudrait essayer de savoir comment les sujets réussissent la tâche sans avoir le droit de prendre de notes, mais avec en contrepartie une phase de familiarisation et de restitution des signes plus poussées, ou un nombre de signes à définir plus faible.

Nous soulignons qu'en dépit de résultats globaux plutôt moyens, les sujets n'ont pas été découragés par la tâche, qu'ils ont trouvée intéressante (notamment du fait qu'elle propose une communication exclusivement basée sur le toucher) bien que difficile. De plus, les sujets ont fait part d'un assez grand optimisme : lorsque les expérimentateurs leur demandaient s'ils avaient réussi à se mettre d'accord, ou s'ils pensaient avoir réussi à deviner le sens de ce que leur partenaire avait joué, ils répondaient très majoritairement que oui. Cela nous semble une piste intéressante à creuser ; cette impression de se comprendre est-elle liée au sens du toucher ?

4.1.7 Pistes futures

Cette expérience permet de poser les pistes pour des explorations futures. Quel est le rôle de la dynamique dans la reconnaissance et la mémorisation des signes négociés ? Quelles sont

les différentes façons dont un signe émerge et est négocié ? Quel est l'apport d'un langage tactile ? Comment sont utilisés les signes en même temps que l'interaction directe ? Toutes ces questions émergentes seront à prendre en considération par la suite.

4.2 Conclusion générale

Notre travail s'est centré sur le design d'interfaces pour la communication émotionnelle à distance. Nous avons donc adressé la question des émotions et de leur transmission dans une situation d'interaction particulière. Dans un but d'innovation, nous avons ici adopté une posture méthodologique résolument pluridisciplinaire, visant à intégrer les méthodes et les points de vue variés propres aux différents domaines impliqués. Nous avons procédé à l'articulation de trois pôles d'activité hétérogènes (théorie, expérimentation, design), nous retrouvant ainsi à l'intersection de plusieurs pratiques. Une des difficultés majeures consistait à intégrer le design, qui est plus un métier qu'une discipline, aux deux autres pôles (dont l'inscription dans une perspective de recherche est déjà mise en place).

Naviguant entre ces trois pôles, nous avons tout d'abord proposé une recherche autour de la notion de présence à distance (première étape possible pour une communication émotionnelle) et du sens du toucher. Nous avons pris en considération les interactions tactiles distales, rendues possibles grâce à la connexion en réseau de dispositifs (par exemple les dispositifs de suppléance perceptive comme Tactos, ou encore ceux permettant une interaction haptique). Nous avons alors conçu une application smartphone permettant une suppléance perceptive du toucher. Baptisée Touch Through, elle permet un croisement perceptif à distance et constitue l'une des réalisations majeures de ce travail. Touch Through permet à ses utilisateurs de faire l'expérience de deux modes. D'une part, le mode distal ne respecte pas la propriété de coïncidence du toucher et permet de savoir où se trouve le partenaire dans l'espace d'interaction partagé même lorsqu'il n'est pas au même endroit que soi. D'autre part, le mode proximal respecte quant à lui cette propriété : les utilisateurs n'ont un retour sensoriel que lorsqu'ils sont en même temps au même endroit de l'espace d'interaction partagé. Touch Through utilise la vue, l'ouïe et/ou des vibrations pour transmettre l'information d'interaction. Nous plaçant dans une perspective d'innovation, cette interface nouvelle nous a permis d'explorer notre problématique en adoptant une démarche de conception.

Cette proposition d'interface nous a d'autre part servi de support pour mettre en place une expérience du type « test de Turing » afin de valider notre dispositif tout en explorant la dynamique dyadique et l'émergence du sentiment de la présence d'autrui à distance. Nous avons ainsi pu mettre en évidence le rôle de la dynamique d'interaction dans la rencontre

interpersonnelle. Cette étude de psychologie expérimentale a d'autre part servi à mettre l'application entre les mains d'utilisateurs ; les premiers retours ainsi obtenus sont très encourageants et plaident pour un développement futur, passant notamment par l'accès grand public à l'application. Les utilisateurs se sont saisis du dispositif et nous avons pu assister à l'émergence d'un espace de signes partagés, rendus possibles notamment par des jeux de mimétisme (suivi de l'autre sujet, répétition de son geste,...). Un futur développement pour Touch Through consiste en l'élargissement de l'interaction à plus de deux utilisateurs⁵¹.

Dans un second temps, une deuxième interface a été esquissée. Il s'agit de notre proposition développée à partir du concept d'attention conjointe (qui constitue une deuxième étape dans l'interaction à distance, après la perception d'autrui), appliquée au réseau social facebook. Nous avons donc procédé en deux étapes, en posant tout d'abord autrui comme émotionnellement très important dans le monde, puis en exposant que l'intérêt qu'autrui porte aux choses qui m'entourent suscite aussi mon intérêt. Notre proposition d'interface met d'une part en exergue le caractère public ou privé des différentes parties de l'interface, et d'autre part offre une nouvelle interaction autour d'éléments à explorer ensemble. Au contraire de notre première proposition de design, nous avons ici conçu l'interface après avoir mené une expérience étudiant l'attention conjointe dans un monde minimaliste, ouvrant ainsi la situation dyadique vers l'intégration d'un objet partagé. La tâche consistait pour les utilisateurs à faire la distinction entre ce qui est commun, partagé, et ce qu'eux seuls peuvent percevoir dans l'espace. Il reste du travail pour le développement de notre prototype, notamment en ce qui concerne les aspects techniques (implémentation) et la prise en main par des utilisateurs qui participeront à son amélioration.

Nous avons donc mis en place deux expériences systématiques (parties 2.4 et 3.3), mais aussi une expérience préliminaire (partie 4.1), qui a pour vocation d'explorer la faisabilité d'une telle étude expérimentale, à réaliser ultérieurement de façon plus systématique. Nos expériences ont été sources d'inspiration pour nos interfaces, lesquelles peuvent aussi devenir support d'expériences minimalistes qui en retour conduisent à des évolutions de ces mêmes interfaces. Nous avons ainsi mis en place une boucle itérative entre expérimentation et design.

Notre recherche a abouti à la proposition de deux interfaces, basées respectivement sur :

- la suppléance perceptive du toucher, afin de permettre un type de toucher à distance basé sur deux propriétés fonctionnelles du toucher (la coïncidence et la perception active) ;
- la mise en jeu de l'attention conjointe dans le réseau social facebook.

⁵¹ Un travail préliminaire autour de cette question est déjà en cours : Christoffel Kuenen nous a en effet contactés et a mis en place une version prototype d'un Touch Through à plusieurs, consultable à l'adresse : <http://murmurations.kuenen.org/2014/04/01/scalable-group-touch-spots/>

Nous avons activement collaboré avec des utilisateurs (que ce soit pour Touch Through, le prototype facebook ou encore Tactos, à l'échelle de notre groupe de recherche et sur plusieurs années) afin de pouvoir obtenir des données d'usage, recueillir les attentes et remarques des utilisateurs, et améliorer nos propositions. Les utilisateurs ont ainsi participé à guider les étapes de conception et nous ont permis de réaliser des analyses d'usage, soit en situation écologique, soit en situation expérimentale.

Notre approche nous a permis de tirer bénéfice d'une recherche fondamentale pour le processus du design ; elle donne en effet la possibilité de traiter les questionnements liés à la conception sous de nouveaux points de vue et d'enrichir le design. Il est alors possible de sortir de l'analyse stricte des besoins des utilisateurs, en visant plutôt à proposer de nouveaux usages et à faire *émerger* de nouveaux besoins. Pendant de cette démarche, le design s'est révélé riche pour la recherche étant donné qu'il permet d'explorer une thématique (ici la communication émotionnelle à distance) par l'objet conçu et par le processus d'élaboration d'interfaces. Sur le plan des études de psychologie expérimentale, nous avons mis en place des expérimentations en situation contrôlée, qui ne permettent pas d'évaluer l'utilisation des dispositifs en milieu écologique, mais qui présentent l'avantage de pouvoir accéder à la mise en place de processus cognitifs. Nous avons défini les conditions techniques nécessaires à l'interaction d'utilisateurs distants dans des espaces numériques partagés. Nous avons parallèlement soulevé les questions théoriques qui sont posées par de tels espaces, en nous consacrant particulièrement à la dynamique des interactions comme étant constitutive des processus cognitifs sociaux. Nous avons aussi ébauché une recherche sur la nature des mécanismes qui permettent aux utilisateurs de faire sens des rencontres interpersonnelles.

Les expérimentations et observations menées dans notre travail nous ont permis de constater l'inventivité dont les utilisateurs font preuve. Ils réussissent à se projeter dans des univers très pauvres et artificiels, et à y mener à bien des tâches complexes (se suivre, se reconnaître, se synchroniser autour d'objets, déterminer les objets qu'ils partagent, s'accorder sur un ensemble de signes, etc.). Ils arrivent à se saisir et à s'approprier des dispositifs qui donnent accès à des conditions perceptives inhabituelles, non naturelles. Le potentiel créatif des individus les mène à s'engager dans les interactions directes, dans lesquelles ils développent aisément des usages nouveaux (caresse à distance, suivi, mimétisme, etc.). Cette capacité des utilisateurs à inventer de nouvelles façons d'interagir dans les interfaces qu'ils se voient proposer est une source de possibilités infinies pour les designers. Ces derniers devront chercher à développer leurs interfaces en visant à susciter chez les utilisateurs ces possibilités d'innover. Le designer doit alors concevoir un espace de possibles, un domaine d'actions et de sensations relativement libre. Il s'agit en effet de créer des conditions favorables pour que le dispositif soit ressaisi par les utilisateurs. Le designer a aussi pour but de voir son dispositif intégré dans les usages des

futurs utilisateurs. Le design envisagé ici tient compte de l'inventivité des utilisateurs ; avec Touch Through, nous avons typiquement proposé un espace d'interaction libre.

Nous avons articulé une recherche théorique fondamentale portant sur les mécanismes de la cognition sociale à une recherche appliquée, ayant pour but de concevoir des dispositifs enrichissant la communication à distance. Nous avons cherché à définir les conditions techniques de l'interaction tactile sur le réseau, tout en questionnant la nature des processus qui permettent aux utilisateurs de faire sens des rencontres interpersonnelles minimalistes. Nous avons adopté une approche énaïve et interactionniste de la cognition sociale : la coordination des activités perceptives résulte d'un processus actif d'engagement mutuel qui devient le support d'une co-construction de sens. Nous avons alimenté notre recherche fondamentale de questionnements pratiques, ancrés dans la conception, l'usage et l'expérimentation. Et en retour, la théorie nous a permis de guider nos idées de design dans l'élaboration de solutions de design d'interaction pertinentes.

Nous avons ainsi mis en place une navigation circulaire entre les pôles théorie, design et expérimentation. Cette navigation a été permise en partie par notre positionnement minimaliste. L'intérêt principal du minimalisme réside dans sa capacité à forcer le déploiement des activités individuelles et collectives. Les dispositifs mis en jeu (Touch Through et Tactos) étant de plus numériques, ils offrent la possibilité d'enregistrer ces activités, ce qui les rend accessibles à l'analyse. Cette analyse nous a permis d'avancer des éléments objectifs soutenant nos hypothèses. Notons que nous avons majoritairement utilisé des indicateurs dits statiques, qui résument l'interaction ou la séquence d'interaction en une valeur (la moyenne ou l'écart-type). Cela est dans une certaine mesure représentatif de l'interaction, car cela donne accès à une tendance globale par sujet et par essai, mais cela ne permet pas en général de capturer les stratégies individuelles et l'essence même de la dynamique d'interaction. Des outils d'analyse provenant par exemple de l'approche adoptée pour l'étude de systèmes dynamiques permettraient d'affiner nos résultats, tels que les méthodes d'analyse du signal (analyse par temps-fréquence, analyse spectrale et fractale). On pourrait ainsi isoler des patterns d'interaction propres à la coordination interpersonnelle.

Une meilleure compréhension de la coordination interpersonnelle permet d'alimenter les projets du point de vue conception, en plus de l'approche collaborative avec les utilisateurs. Cette dernière constitue un atout indéniable dans le cadre du design. Il est à noter que nous n'avons pu mettre en jeu qu'un nombre relativement faible de participants, ce qui pose la question de la généralisation de nos résultats. La prochaine étape de notre travail consistera donc en la mise à disposition de nos dispositifs à des utilisateurs à plus grande échelle, dans le but notamment de constituer une communauté d'utilisateurs.

Bibliographie

- Adler, A., et Brett, C. 1999. *Social Interest: Adler's Key to the Meaning of Life*. Oneworld Publications.
- Ali-Ammar, A. 2005. « Analyse des explorations haptiques de formes pour la conception d'un dispositif de suppléance perceptive dédié aux personnes aveugles ». Thèse de doctorat, Université de Technologie de Compiègne.
- Ames, M.G., Go, J., Kaye, J., et Spasojevic, M. 2010. « Making love in the network closet: the benefits and work of family videochat ». In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 145- 54. CSCW '10. New York, NY, USA: ACM.
- Archer, B. 1995. « The nature of research ». *Co-Design Journal* 2 (11): 6- 13.
- Argyle, M. 1988. *Bodily communication*. Londres: Methuen.
- Argyle, M., et Cook, M. 1976. *Gaze and mutual gaze*. Cambridge, England: Cambridge University Press.
- Arnold, M.B. 1960. « Emotion and personality. »
- Astington, J.W., et Gopnik, A. 1991. « Theoretical explanations of children's understanding of the mind ». *British Journal of Developmental Psychology* 9 (1): 7- 31.
- Auvray, M. 2004. « Immersion et perception spatiale : l'exemple des dispositifs de substitution sensorielle ». Thèse de, Paris: EHESS.
- Auvray, M., et Fuchs, P. 2007. « Perception, immersion et interactions sensorimotrices en environnement virtuel ». *Intellectica*, n° 45: 23- 35.
- Auvray, M., Lenay, C., O'Regan, J.K., et Lefèvre J. 2005. « Suppléance perceptive, immersion et informations proprioceptives ». *Arobase* 1: 94- 113.
- Auvray, M., Lenay, C., et Stewart, J. 2009. « Perceptual interactions in a minimalist virtual environment ». *New Ideas in Psychology* 27 (1): 32-47.
- Auvray, M., Philipona, D., O'Regan, J.K., et Spence, C. 2007. « The perception of space and form recognition in a simulated environment: The case of minimalist sensory-substitution devices ». *Perception* 36 (12): 1736- 1751.
- Bach-y-Rita, P. 1972. *Brain Mechanisms in Sensory Substitution*. Academic Press Inc.
- . 2002. « Sensory substitution and qualia ». *Vision and mind*, 497- 514.
- Bach-y-Rita, P, Collins, C.C., Saunders, F.A., White, B., et Scadden, L. 1969. « Vision substitution by tactile image projection ». *Nature* 221 (184): 963- 964.
- Bach-y-Rita, P., et Kaczmarek, K.A. 2002. « Tongue placed tactile output device ».
- Bach-y-Rita, P., et Kercel, S.W. 2003. « Sensory substitution and the human-machine interface ». *Trends in cognitive sciences* 7 (12): 541- 46.

- Bailenson, J.N., Yee, N., Brave, S., Merget, D., et Koslow, D. 2007. « Virtual interpersonal touch: expressing and recognizing emotions through haptic devices ». *Human-Computer Interaction* 22 (3): 325-53.
- Baldwin, D.A. 1995. « Understanding the link between joint attention and language. »
- Bales, E., Li, K.A., et Griwsold, W. 2011. « CoupleVIBE: mobile implicit communication to improve awareness for (long-distance) couples ». In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 65-74. CSCW '11. New York, NY, USA: ACM.
- Baron-Cohen, S. 1997. *Mindblindness: An essay on autism and theory of mind*. MIT press.
- Basdogan, C., Ho, C., Srinivasan, M.A. et Slater, M. 2000. « An experimental study on the role of touch in shared virtual environments ». *ACM Trans. Comput.-Hum. Interact.* 7 (4): 443-60.
- Belzung, C. 2007. *Biologie des émotions*. Bruxelles: De Boeck.
- Benali-Khoudja, M., Hafez, M., Alexandre, J.M., et Kheddar, A. 2004. « Tactile interfaces: a state-of-the-art survey ». In *Int. Symposium on Robotics*. Vol. 31.
- Bermúdez, J.L., Marcel, A.J., et Eilan, N. 1998. *The Body and the Self*. MIT Press.
- Bernstein, M.S., Bakshy, E., Burke, M., et Karrer, B. 2013. « Quantifying the invisible audience in social networks ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 21-30. CHI '13. New York, NY, USA: ACM.
- Bigelow, A. 2003. « The development of joint attention in blind infants ». *Development and Psychopathology* 15 (2): 259-75.
- Binder, T., De Michelis, G., Ehn, P., Jacucci, G., Linde, P., et Wagner, I. 2012. « What is the object of design? » In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, 21-30. CHI EA '12. New York, NY, USA: ACM.
- Bisognin, L., et Pesty, S. 2004. « Emotions et systèmes multi-agents: une architecture d'agent émotionnel ». In *Emotions et systèmes multi-agents : une architecture d'agent émotionnel 2004*.
- Bødker, S., Nielsen, C., et Petersen, M.G. 2000. « Creativity, cooperation and interactive design ». In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, 252-61. DIS '00. New York, NY, USA: ACM.
- Bottineau, D. 2010. « Language and Enaction ». *Enaction: Toward a new paradigm for cognitive science*, 267-306.
- Bougnoux, D. 2001. *Introduction aux sciences de la communication*. Paris: La Découverte.
- Boyd, D.M., et Ellison, N.B. 2007. « Social Network Sites: Definition, History, and Scholarship ». *Journal of Computer-Mediated Communication* 13 (1): 210-30.

- Boyer, H. 1982. « La communication épistolaire comme stratégie romanesque ». *Semiotica* 39 (1-2): 21-44.
- Brave, S., et Dahley, A. 1997. « inTouch: A Medium for Haptic Interpersonal Communication ». In , 363-64. ACM Press.
- Brave, S., Ishii, H., et Dahley, A. 1998. « Tangible interfaces for remote collaboration and communication ». In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, 169-78. CSCW '98. New York, NY, USA: ACM.
- Brown, P. 1987. *Politeness: Some universals in language usage*. Vol. 4. Cambridge University Press.
- Brown, P., et Levinson, S.C. 1978. *Universals in language usage: Politeness phenomena*.
- Bruner, J.S. 1983. *Le développement de l'enfant : savoir faire, savoir dire*. Psychologie d'aujourd'hui. Presses universitaires de France.
- Butterworth, G., et Jarrett, N. 1991. « What Minds Have in Common Is Space: Spatial Mechanisms Serving Joint Visual Attention in Infancy ». *British Journal of Developmental Psychology* 9 (1): 55-72.
- Call, J., et Tomasello, M. 2005. « What chimpanzees know about seeing revisited: An explanation of the third kind ». *Joint attention: Communication and other minds*, 45-64.
- Campbell, J. 2005. « Joint attention and common knowledge ». *Joint attention: Communication and other minds*, 287-97.
- Card, S.K., Newell, A., et Moran, T.P. 1983. *The Psychology of Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.
- Cardon, D. 2011. « Eloge des amitiés numériques ». *Le Monde*, novembre 10.
- . 2012. « Le parler privé-public des réseaux sociaux de l'Internet ». In *Médias sociaux : enjeux pour la communication*, 33-45.
- Carpendale, J., et Lewis, C. 2004. « Constructing an understanding of mind: The development of children's social understanding within social interaction ». *Behavioral and Brain Sciences* 27 (01): 79-96.
- Cassell, J., et Thorisson, K.R. 1999. « The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents ». *Applied Artificial Intelligence* 13 (4-5): 519-38.
- Chang, A., Gouldstone, J., Zigelbaum, J., et Ishii, H. 2007. « Simplicity in interaction design ». In *Proceedings of the 1st international conference on Tangible and embedded interaction*, 135-38. TEI '07. New York, NY, USA: ACM.
- Chang, A., et Ishii, H. 2006. « Sensorial interfaces ». In *Proceedings of the 6th conference on Designing Interactive systems*, 50-59. DIS '06. New York, NY, USA: ACM.

- Chang, A., O'Modhain, S., Jacob, R., Gunther, E., et Ishii, H. 2002. « ComTouch: design of a vibrotactile communication device ». In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, 312-20. DIS '02. New York, NY, USA: ACM.
- Chang, A., Resner, B., Koerner, B., Wang, X., et Ishii, H. 2001. « LumiTouch: an emotional communication device ». In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, 313- 14. CHI EA '01. New York, NY, USA: ACM.
- Chapman, M. 1991. « The epistemic triangle: Operative and communicative components of cognitive competence. » In *A preliminary version of this chapter was presented at the Meeting of the Society for Research in Child Development, Kansas City, Apr 1989*.
- Chartier, R. 1994. « Du codex à l'écran : les trajectoires de l'écrit ». *Solaris*, n° 1.
- Chung, H., Jackie Lee, C., et Selker, T. 2006. « Lover's cups: drinking interfaces as new communication channels ». In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, 375- 80. CHI EA '06. New York, NY, USA: ACM.
- Clark, A. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension (Philosophy of Mind)*. Oxford University Press.
- Clark, A., et Chalmers, D. 1998. « The extended mind » 58 (1): 7- 19.
- Collier, G., et DiCarlo, D. 1985. *Emotional expression*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Collins, C.C., et Bach-y-Rita, P. 1973. « Transmission of pictorial information through the skin. » *Advances in biological and medical physics* 14: 285- 315.
- Colombetti, G. 2008. « Enaction, sense-making and emotion ». *Enaction: Toward a New Paradigm for Cognitive Science*, J. Stewart, O. Gapenne, and EA Di Paolo eds, 145- 64.
- Criminisi, A., Shotton, J., Blake, A., et Torr, P. 2003. *Efficient dense stereo and novel-view synthesis for gaze manipulation in one-to-one teleconferencing*. Microsoft Research.
- Csibra, G., et Gergely, G. 1998. « The teleological origins of mentalistic action explanations: A developmental hypothesis ». *Developmental Science* 1 (2): 255- 59.
- Cullen, K.E. 2004. « Sensory signals during active versus passive movement ». *Current opinion in neurobiology* 14 (6): 698-706.
- D'Mello, S., et Calvo, R.A. 2013. « Beyond the basic emotions: what should affective computing compute? » In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2287-94. CHI EA '13. New York, NY, USA: ACM.
- Damasio, A.R. 1999. *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace.
- . 2006. *L'erreur de Descartes - La raison des émotions*. Odile Jacob.

- Darses, F. 2004. « La conception participative : vers une théorie de la conception centrée sur l'établissement d'une intelligibilité mutuelle ». *Le consommateur au cœur de l'innovation : la conception participative*, 25-41.
- Darwin, C. 1872. *The expression of the emotions in man and animals*. Londres: Murray.
- De Fornel, M. 1988. « Contraintes systémiques et contraintes rituelles dans l'interaction visiophonique ». *Réseaux* 6 (29): 33-46.
- . 1992. « « ALORS, TU ME VOIS ? » : Objet technique et cadre interactionnel dans la pratique visiophonique ». *Culture technique*, n° 24: 113-120.
- . 1994. « Le cadre interactionnel de l'échange visiophonique ». *Réseaux* 12 (64): 107-132.
- De Jaegher, H. 2006. « Social interaction rhythm and participatory sense-making. An embodied, interactional approach to social understanding, with implications for autism ». Unpublished DPhil Thesis, University of Sussex.
- De Jaegher, H., et Di Paolo, E. 2007. « Participatory sense-making: An enactive approach to social cognition ». *Phenomenology and the Cognitive Sciences* 6 (4): 485-507.
- De Jaegher, H., Di Paolo, E., et Gallagher, S. 2010. « Can Social Interaction Constitute Social Cognition? » *Trends in Cognitive Sciences* 14 (10): 441-447.
- Deckers, E., Wensveen, S., Ahn, R., et Overbeeke, K. 2011. « Designing for perceptual crossing to improve user involvement ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1929-38. CHI '11. New York, NY, USA: ACM.
- Declerck, G. 2010. « Phénoménologie et psychologie du tangible : éléments pour une théorie de la valeur cognitive et pratique de la résistance ». Thèse de doctorat, Compiègne: Université de Technologie de Compiègne.
- Desavoye, B. 2011. « Quels usages font les jeunes d'Internet ? » *Netalya*.
<http://www.netalya.com/fr/article2.asp?CLE=158>.
- Descartes, R. 1649. *Les Passions de l'âme*. Paris.
- Deschamps, L., Le Bihan, G., Lenay, C., Rovira, K., Stewart, J., et Aubert, D. 2012. « Interpersonal recognition through mediated tactile interaction ». In *2012 IEEE Haptics Symposium (HAPTICS)*, 239-245.
- Deschamps, L., Rovira, K., et Lenay, C. 2008. « The appropriation of prosthetic devices for reading computer screens and network interactions via tactile perception ». *Sciences et Technologies pour le Handicap. n spécial: Technologie Tactile et ses Applications pour les Handicaps* 2 (2): 159-73.
- Deschamps, L. 2013. « Suppléance perceptive et cognition sociale - Etude des interactions tactiles minimalistes ». Thèse de doctorat, Université de Technologie de Compiègne.
- DeVito, J.A. 2008. *Human Communication: The Basic Course*. Pearson College Division.

- DiSalvo, C., Gemperle, F., Forlizzi, J., et Montgomery, E. 2003. « The Hug: an exploration of robotic form for intimate communication ». In *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003*, 403-408.
- Dobson, K., Boyd, D., Ju, W., Donath, J., et Ishii, H. 2001. « Creating visceral personal and social interactions in mediated spaces ». In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, 151-52. CHI EA '01. New York, NY, USA: ACM.
- Dourish, P. 1993. « Culture and Control in a Media Space ». In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW '93*, édité par Giorgio de Michelis, Carla Simone, et Kjeld Schmidt, 125-137. Springer Netherlands.
- Dourish, P., et Bly, S. 1992. « Portholes: supporting awareness in a distributed work group ». In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 541-47.
- Duchamp, R. 1988. *La conception de produits nouveaux*. Paris: Hermès Science.
- Duchenne, G.B. 1862. *Le mécanisme de l'expression faciale humaine*. Paris: Yves Jules Renouard.
- Dunbar, R.I.M. 2010. « The social role of touch in humans and primates: Behavioural function and neurobiological mechanisms ». *Neuroscience & Biobehavioral Reviews* 34 (2): 260-268.
- Dunne, A., et Raby, F. 1994. « Fields and thresholds ». *Proceedings of the Doors of Perception-2*.
- Durlach, N., et Slater, M. 2000. « Presence in Shared Virtual Environments and Virtual Togetherness ». *Presence: Teleoperators and Virtual Environments* 9 (2): 214-217.
- Eilan, N. 2005. « Joint Attention, Communication, and Mind ». *Joint Attention: Communication and Other Minds: Issues in Philosophy and Psychology: Issues in Philosophy and Psychology*, 1.
- Ekman, P. 1992a. « Are there basic emotions ». *Psychological review* 99 (3): 550-53.
- . 1992b. « Facial Expressions of Emotion: New Findings, New Questions ». *Psychological Science* 3 (1): 34 - 38.
- . 1999. « Basic emotions ». *Handbook of cognition and emotion* 98: 45-60.
- Ekman, P., Friesen, W.V., et Ellsworth, P. 1972. *Emotion in the Human Face: Guide-Lines for Research and an Integration of Findings*. Pergamon Press.
- Elfenbein, H.A., et Ambady, N. 2002. « On the universality and cultural specificity of emotion recognition: a meta-analysis. » *Psychological bulletin* 128: 203-35.
- Elias, N. 1939. « La Civilisation des mœurs ». Paris, Calmann-Lévy.

- Fagan, J. 1998. « Thoughts on Using Touch in Psychotherapy ». In *Touch in psychotherapy: Theory, research, and practice*.
- Feltham, F., Vetere, F., et Wensveen, S. 2007. « Designing tangible artefacts for playful interactions and dialogues ». In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, 61-75.
- Fiebich, A., et Gallagher, S. 2012. « Joint attention in joint action ». *Philosophical Psychology*, 1-17.
- Field, T. 2003. *Touch*. MIT Press.
- Finnegan, R. 2002. *Communicating: The multiple modes of human interconnection*. Routledge.
- Fischer, J.A., et Gallant, S.J. 1990. « Effect of touch on hospitalized patients ». In *Advances in touch*, 141-147.
- Fogel, A. 1993. *Developing Through Relationships: Origins of Communication, Self and Culture*. Chicago: University of Chicago Press.
- Fogel, A., Nwokah, E., Dedo, J.Y., Messinger, D., Dickson, K.L., Matusov, E., et Holt, S.A. 1992. « Social Process Theory of Emotion: A Dynamic Systems Approach ». *Social Development* 1 (2): 122-42.
- Franco, F. 2005. « Infant Pointing: Harlequin, Servant of Two Masters ». *Joint attention: Communication and other minds*, 129-64.
- Frank, M.G., et Ekman, P. 1993. « Not all smiles are created equal: The differences between enjoyment and nonenjoyment smiles ». *Humor* 6: 9-9.
- Frank, M.G., Ekman, P., et Friesen, W.V. 1993. « Behavioral markers and recognizability of the smile of enjoyment ». *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*, 217-38.
- Fraser, M., Hindmarsh, J., Best, K., Heath, C., Biegel, G., Greenhalgh, C., et Reeves, S. 2006. « Remote Collaboration Over Video Data: Towards Real-Time E-Social Science ». *Computer Supported Cooperative Work (CSCW)* 15 (4): 257-279.
- Frens, J.W. 2006. « Designing for Rich Interaction: Integrating Form, Interaction, and Function ». Thèse de doctorat, Eindhoven: Eindhoven University of Technology.
- Frith, C.D. 2008. « Social cognition ». *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1499): 2033-39.
- Froese, T. 2012. « From Adaptive Behavior to Human Cognition: A Review of Enaction ». *Adaptive Behavior*, février.
- Fuchs, P. 2007. « Interfaces à action mécanique ». *Techniques de l'ingénieur. Télécoms*, n° TE5904v2.
- Fuchs, T., et De Jaegher, H. 2009. « Enactive intersubjectivity: Participatory sense-making and mutual incorporation ». *Phenomenology and the Cognitive Sciences* 8 (4): 465-86.

- Furukawa, M., Kajimoto, H., et Tachi, S. 2012. « KUSUGURI: a shared tactile interface for bidirectional tickling ». In *Proceedings of the 3rd Augmented Human International Conference*, 9:1-9:8. AH '12. New York, NY, USA: ACM.
- Gallace, A., et Spence, C. 2010. « The science of interpersonal touch: An overview ». *Neuroscience & Biobehavioral Reviews* 34 (2): 246-259.
- Gallagher, S. 2004. « Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind ». *Philosophy, Psychiatry, & Psychology* 11 (3): 199-217.
- . 2008. « Direct perception in the intersubjective context ». *Consciousness and Cognition* 17 (2): 535-543.
- . 2011. « Interactive Coordination in Joint Attention ». A. Seeman (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, Cambridge, MA: MIT Press. <http://ummos.org/gall10ja.pdf>.
- Gapenne, O., Lenay, C., Stewart, J., Bériot, H., et Meidine, D. 2001. « Prosthetic device and 2D form perception: the role of increasing degrees of parallelism ». In *Proceedings of the Conference on Assistive Technology for Vision and Hearing Impairment (CVHI'2001)*, 113-18. Castelvecchio Pascoli, Italie.
- Gapenne, O., Rovira, K., Ali Ammar, A., et Lenay, C. 2003. « Tactos: Special computer interface for the reading and writing of 2D forms in blind people ». *Universal Access in HCI, Inclusive Design in the Information Society* 10: 1270-74.
- Gapenne, O., Rovira, K., Lenay, C., Stewart, J., et Auvray, M. 2005. « Is form perception necessary tied to specific sensory feedback? ». In *Proceedings of the 13th International Conference on Perception and Action (ICPA)*, 16.
- Gaver, B. 2002. « Provocative Awareness ». *Comput. Supported Coop. Work* 11 (3): 475-93.
- Gaver, W.W., Smets, G., et Overbeeke, K. 1995. « A Virtual Window on media space ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 257-64. CHI '95. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co.
- Gibbs, M., Howard, S., Kjeldskov, J., Vetere, F., et Satchell, C. 2006. « « Was it Good for you Darling? » – Intimacy, Sex and Critical Technical Practice ». In *Creative Industries Faculty; Institute for Creative Industries and Innovation*. Montreal.
- Gibbs, M., Vetere, F., Bunyan, M., et Howard, S. 2005. « SynchroMate: a phatic technology for mediating intimacy ». In *Proceedings of the 2005 conference on Designing for User eXperience*. DUX '05. New York, NY, USA: AIGA: American Institute of Graphic Arts.
- Gibson, J.J. 1962. « Observations on active touch ». *Psychological review* 69 (6): 477-91.
- . 1966. *The senses considered as perceptual systems*. Oxford, England: Houghton Mifflin.
- . 1986. *The ecological approach to visual perception*. Lawrence Erlbaum.

- Glenberg, A.M., et Gallese, V. 2012. « Action-Based Language: A Theory of Language Acquisition, Comprehension, and Production ». *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 48 (7): 905-922. doi:10.1016/j.cortex.2011.04.010.
- Goffman, E. 1959. « The presentation of self in everyday life ». *New York: Doubleday Anchor*.
- . 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Free Press of Glencoe.
- . 1974. *Les Rites d'interaction*. Les Editions de Minuit.
- . 1987. « Façons de parler ». *Paris, Editions de Minuit*.
- Goldman, A.I. 1989. « Interpretation Psychologized ». *Mind & Language* 4 (3): 161-85.
- Gomez, J.C. 1991. « Visual behavior as a window for reading the mind of others in primates ». *Natural theories of mind*, 194-207.
- . 2005. « Joint Attention and the Notion of Subject: Insights From Apes, Normal Children, and Children with Autism ». In *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology*. Oxford University Press.
- Gooch, D., et Watts, L. 2012. « It's neat to feel the heat: how can we hold hands at a distance? » In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts*, 1535-40. CHI EA '12. New York, NY, USA: ACM.
- Goody, J. 1979. *La raison graphique*. MINUIT EDITIONS. Les Editions de Minuit.
- Gordon, R.M. 1986. « Folk psychology as simulation ». *Mind & Language* 1 (2): 158-71.
- Gratch, J., et Marsella, S. 2004. « A domain-independent framework for modeling emotion ». *Cognitive Systems Research* 5 (4): 269-306.
- Gumtau, S. 2005. « Tactile semiotics: the meanings of touch explored with low-tech prototypes ». In *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint*, 651-652. doi:10.1109/WHC.2005.124.
- Haans, A., et IJsselsteijn, W.A. 2006. « Mediated social touch: a review of current research and future directions ». *Virtual Real.* 9 (2): 149-59.
- Hanneton, S., Gapenne, O., Genouel, C., Lenay, C., et Marque, C. 1999. « Dynamics of shape recognition through a minimal visuo-tactile sensory substitution interface ». In *Third international conference on cognitive and neural systems*, 26-29. Boston.
- Harlow, J.M.. 1868. « Recovery from the Passage of an Iron Bar through the Head ». *Bulletin of the Massachusetts Medical Society*.
- Hartmut, R. 2010. *Accélération : une critique sociale du temps*. La Découverte, Paris.
- Hashimoto, Y., et Kajimoto, H. 2008. « An emotional tactile interface completing with extremely high temporal bandwidth ». In *SICE Annual Conference, 2008*, 1457-1462.

- Hassenzahl, M., Heidecker, S., Eckoldt, K., Diefenbach, S., et Hillmann, U. 2012. « All You Need is Love: Current Strategies of Mediating Intimate Relationships through Technology ». *ACM Trans. Comput.-Hum. Interact.* 19 (4): 30:1- 30:19.
- Hatwell, Y. 2003. *Touching for knowing: cognitive psychology of haptic manual perception*. John Benjamins Publishing Company.
- Heal, J. 2005. « Joint attention and understanding the mind ». *Joint attention: Communication and other minds*, 34-44.
- Heath, C., et Luff, P. 1991. « Disembodied conduct: communication through video in a multi-media office environment ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 99- 103. CHI '91. New York, NY, USA: ACM.
- . 1992. « Media space and communicative asymmetries: preliminary observations of video-mediated interaction ». *Hum.-Comput. Interact.* 7 (3): 315-46.
- Heath, C., Luff, P., et Sellen, A. 1977. « Reconfiguring media space: Supporting collaborative work ». In *In*, 323-47.
- Heath, C., Svensson, M.S., Hindmarsh, J., Luff, P., et vom Lehn, D. 2002. « Configuring Awareness ». *Comput. Supported Coop. Work* 11 (3): 317-47.
- Heaton, L., Millette, M., et Proulx, S.. 2012. *Médias sociaux : enjeux pour la communication*. PUQ.
- Heidegger, M. 1927. *Etre et temps*.
- Hemmert, F., Gollner, U., Löwe, M., Wohlauf, A., et Joost, G. 2011. « Intimate mobiles: grasping, kissing and whispering as a means of telecommunication in mobile phones ». In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, 21-24. MobileHCI '11. New York, NY, USA: ACM.
- Henley, N.M. 1973. « Power, sex, and nonverbal communication ». *Berkeley Journal of Sociology*, 1-26.
- Herrenschmidt, C. 1999. « Écriture, monnaie, réseaux. Inventions des Anciens, inventions des Modernes ». *Le débat*, n° 106 (octobre).
- Hertenstein, M.J. 2002. « Touch: Its communicative functions in infancy ». *Human Development* 45 (2): 70-94.
- . 2005. « Touch ». *Encyclopedia of human development* 3: 1275-77.
- Hertenstein, M.J., et Campos, J.J. 2001. « Emotion regulation via maternal touch ». *Infancy* 2 (4): 549-66.
- Hertenstein, M.J., Holmes, R., McCullough, M., et Keltner, D. 2009. « The communication of emotion via touch ». *Emotion* 9 (4): 566.
- Hertenstein, M.J., Keltner, D., App, B., Bulleit, B.A., et Jaskolka, A.R. 2006. « Touch communicates distinct emotions ». *Emotion* 6 (3): 528-33.

- Hess, U., et Bourgeois, P. 2010. « You smile–I smile: Emotion expression in social interaction ». *Biological Psychology* 84 (3): 514- 520.
- Heylings, P. N. K. 1973. « The no-touching epidemic – an English disease ». *The British Medical Journal*.
- Hindmarsh, J., et Heath, C. 2000. « Embodied reference: A study of deixis in workplace interaction ». *Journal of Pragmatics* 32 (12): 1855- 1878.
- Hite, S. 1982. *The Hite Report on Male Sexuality*. Ballantine Books.
- Hobson, R.P. 1993. « Through feeling and sight to self and symbol ». *The perceived self: Ecological and interpersonal sources of self-knowledge*, 254- 79.
- Hoggan, E., Stewart, C., Haverinen, L., Jacucci, G., et Lantz, G. 2012. « Pressages: augmenting phone calls with non-verbal messages ». In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, 555- 62. UIST '12. New York, NY, USA: ACM.
- Howard, S., Kjeldskov, J., Skov, M.B., Garnæs, K., et Grünberger, O. 2006. « Negotiating presence-in-absence: contact, content and context ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 909- 12. CHI '06. New York, NY, USA: ACM.
- Howard, S., Vetere, F., Gibbs, M., Kjeldskov, J., Pedell, S., Mecoles, K., Bunyan, M., et Murphy, J. 2004. « Mediating Intimacy: digital kisses and cut and paste hugs ». In *Proceedings of HCI*.
- Hummels, C. 2005. « ‘Descendants’ of a Design Quest for Diversity ». In *Proceedings of 2005 International Conference on Designing Pleasurable Products and Interfaces*, 13- 26. Eindhoven: Technical University of Eindhoven, The Netherlands.
- . 2013. « Designing for social cohesion ». In *Actes du Séminaire Interdisciplinaire Philosophie Technologie Cognition PHITECO 2013*. Compiègne, France.
- Husserl, E. 1900. *Recherches logiques*.
- IJsselsteijn, W.A., de Ridder, H., Freeman, J., et Avons, S.E. 2000. « Presence: concept, determinants, and measurement ». In *Electronic Imaging*, 520- 29.
- Illich, I. 1991. *Du lisible au visible: la naissance du texte : un commentaire du « Didascalicon » de Hugues de Saint-Victor*. Les Editions du Cerf.
- Ingsholt, A. 2000. « Being together, interaction and dialogues in interactions between blind infants and their parents ». In *Proceedings of the 5th European Conference of ICEVI (International Council for Education of People with Visual Impairment), Visions and strategies for the new century*. 9-13 July 2000, Cracow, Poland (<http://www.icevi-europe.org/cracow2000/proceedings/chap01.html#8>).

- Ishii, H. 2008. « Tangible bits: beyond pixels ». In *Proceedings of the 2nd international conference on Tangible and embedded interaction*, xv-xxv. TEI '08. New York, NY, USA: ACM.
- Ishii, H., et Ullmer, B. 1997. « Tangible bits: towards seamless interfaces between people, bits and atoms ». In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, 234-41.
- Izard, C.E. 1977. *Human Emotions*. Plenum press, New-York.
- Jones, S.E., et Yarbrough, A.E. 1985. « A naturalistic study of the meanings of touch ». *Communications Monographs* 52 (1): 19-56.
- Jouët, J. 1993. « Pratiques de communication et figures de la médiation ». *Réseaux* 11 (60): 99-120.
- Jourard, S.M. 1966. « An Exploratory Study of Body-Accessibility ». *British Journal of Social and Clinical Psychology* 5 (3): 221-31.
- Kaitz, M. 1992. « Recognition of familiar individuals by touch ». *Physiology & behavior* 52 (3): 565-67.
- Kaminow, I., Li, T., et Willner, A.E. 2010. *Optical Fiber Telecommunications V. B: Systems and Networks*. Academic Press.
- Kappas, A. 1991. « The illusion of the neutral observer: on the communication of emotion ». *Cahiers de linguistique française* 12: 153-68.
- Kemper, T.D. 1981. « Social constructionist and positivist approaches to the sociology of emotions ». *American Journal of Sociology*, 336-62.
- Kidwell, M., et Zimmerman, D.H. 2007. « Joint attention as action ». *Journal of Pragmatics* 39 (3): 592-611.
- Kirsh, D. 1996. « Adapting the environment instead of oneself ». *Adaptive Behavior* 4: 415-52.
- . 2013. « Embodied cognition and the magical future of interaction design ». *ACM Trans. Comput.-Hum. Interact.* 20 (1): 3:1-3:30.
- Klatzky, R.L., et Lederman, S.J. 2002. « Touch ». *Experimental Psychology* 4: 147-176.
- Klatzky, R.L., Lederman, S.J., et Metzger, V.A. 1985. « Identifying objects by touch: An « expert system » ». *Perception & Psychophysics* 37 (4): 299-302.
- Klatzky, R.L., Loomis, J.M., Lederman, S.J., Wake, H., et Fujita, N. 1993. « Haptic identification of objects and their depictions ». *Perception & Psychophysics* 54 (2): 170-78.
- Kleinginna, P.R., et Kleinginna, A.M. 1981. « A categorized list of emotion definitions, with suggestions for a consensual definition ». *Motivation and emotion* 5 (4): 345-79.
- Knapp, M.L., et Hall, J.A. 1997. *Nonverbal communication in human interaction (4th ed.)*. Fort Worth, TX: Harcourt Brace College.

- Krueger, J. 2011. « Extended cognition and the space of social interaction ». *Consciousness and cognition* 20 (3): 643-57.
- Kuwabara, K., Ohguro, T., Watanabe, T., Itoh, Y., et Maeda, Y. 2002. « Connectedness Oriented Communication: Fostering a Sense of Connectedness to Augment Social Relationships ». In *Proceedings of the 2002 Symposium on Applications and the Internet*, 186-93. SAINT '02. Washington, DC, USA: IEEE Computer Society.
- Kwon, D., et Kim, S. 2008. « Haptic interfaces for mobile devices: a survey of the state of the art ». *Recent Patents on Computer Science* 1 (2): 84-92.
- Landers, A. 1985. « Sex: why women feel short-changed ». *Family circle*, juin.
- Laschke, M., Hassenzahl, M., et Mehnert, K. 2010. « linked.: a relatedness experience for boys ». In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, 839-44. NordiCHI '10. New York, NY, USA: ACM.
- Lenay, C., Canu, S., et Villon, P. 1997. « Technology and perception: the contribution of sensory substitution systems ». In *Second International Conference on Cognitive Technology, 1997. Humanizing the Information Age. Proceedings*, 44-53.
- Lenay, C. 2006a. « Enaction, externalisme et suppléance perceptive ». *Intellectica*, n° 43: 27-52.
- . 2006b. « Réticences et conditions d'adoption des prothèses perceptives ». In *Actes de la conférence Handicap'2006, « Nouvelles technologies au service de l'Homme »*. Paris.
- . 2006c. « La méthode minimaliste : phénoménologie et science objective ». *Intellectica* 1 (43): 125-34.
- . 2010. « « C'est très touchant ». La valeur émotionnelle du contact ». *Intellectica* 1/2 (53-54).
- . 2011. « Suppléance perceptive et perception humaine ». In *Informatique et Sciences Cognitives : influences ou confluences ?* Co-édition Ophrys/MSH.
- Lenay, C., Gapenne, O., Hanneton, S., Marque, C., et Genouëlle, C. 2000. « La substitution sensorielle. Limites et perspectives ». In *Toucher pour connaître. Psychologie cognitive de la perception tactile manuelle*, Paris, PUF, 287-306.
- Lenay, C., et Stewart, J. 2012. « Minimalist approach to perceptual interactions ». *Frontiers in Human Neuroscience* 6: 98.
- Lenay, C., Stewart, J., et Gapenne, O. 2002. « Espace d'action technique et geste perceptif ». *Le geste technique: réflexions méthodologiques et anthropologiques* 2 (14): 215-30.
- Lenay, C., Thouvenin, I., Guénand, A., Gapenne, O., Stewart, J., et Maillet, B. 2007. « Designing the ground for pleasurable experience ». In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, 35-58. DPPI '07. New York, NY, USA: ACM.
- Leroi-Gourhan, A. 1943. *L'homme et la matière*. Paris: Albin Michel.

- . 1964. *Le Geste et la Parole, tome 1 : Technique et Langage*. Albin Michel.
- Li, M., et Jianting, H. 2009. « Ambient environments for emotional physical communication ». In *Proceedings of the 10th International Conference NZ Chapter of the ACM's Special Interest Group on Human-Computer Interaction*, 81-84.
- Liechti, O., et Ichikawa, T. 2000. « A Digital Photography Framework Enabling Affective Awareness in Home Communication ». *Personal Technologies* 4 (1): 6-24.
- Lindley, S.E., Harper, R., et Sellen, A. 2009. « Desiring to be in touch in a changing communications landscape: attitudes of older adults ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1693-1702.
- Lombard, M., et Ditton, T. 1997. « At the heart of it all: The concept of presence ». *Journal of Computer-Mediated Communication* 3 (2): 0-0.
- Loomis, J.M., et Lederman, S.J. 1984. « What utility is there in distinguishing between active and passive touch ». In *Psychonomic Society meeting*.
- . 1986. « Tactual perception ». *Handbook of perception and human performances 2: 2*.
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., et Oyama, S. 2003. « Fractured ecologies: creating environments for collaboration ». *Hum.-Comput. Interact.* 18 (1): 51-84.
- Luff, P., Heath, C., et Svensson, M., S. 2008. « Discriminating Conduct: Deploying Systems to Support Awareness in Organizations ». *International Journal of Human-Computer Interaction* 24 (4): 410-436.
- Maillet, B. 2010. « Design et appropriation d'un dispositif d'interactions haptiques pour l'expression et la perception d'attitudes communicationnelles ». Doctorat de Sciences et Technologies, Compiègne, France: Université de Technologie de Compiègne.
- Maillet, B., Guénand, A., et Lenay, C. 2006. « An experimental work on tactile interaction: how to give to the user the possibility to adopt an engaged or a receptive attitude? » In *Proceedings of Design & Emotion 2006*. Göteborg, Sweden.
- Maillet, B., Guénand, A., Lenay, C., et Chêne, D. 2005. « A study on a user centered design of a tactile mobile phone ». In *Proceedings of Designing Pleasurable Products and Interfaces 2005*, 24-28. Eindhoven, Nederland.
- Maillet, B., Lenay, C., et Guénand, A. 2008. « Designing for interpersonal tactile interaction over distance ». In *Proceedings of the 21st International Symposium on Human Factors in Telecommunications*, 391-98. Kuala Lumpur.
- Malson, L. 2002. *Les enfants sauvages*. Paris: 10-18.
- Markopoulos, P., Romero, N., van Baren, J., IJsselsteijn, W.A., de Ruyter, B., et Farshchian, B. 2004. « Keeping in touch with the family: home and away with the ASTRA awareness

- system ». In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1351-54. CHI EA '04. New York, NY, USA: ACM.
- Marsh, K.L., Richardson, M.J., Baron, R.M., et Schmidt, R.C. 2006. « Contrasting approaches to perceiving and acting with others ». *Ecological Psychology* 18 (1): 1-38.
- Masters, W.H., et Johnson, V.E. 1970. *Human Sexual Inadequacy*. Boston : Little, Brown.
- McGlone, F., et Reilly, D. 2010. « The cutaneous sensory system ». *Neuroscience & Biobehavioral Reviews* 34 (2): 148-159.
- Meijer, P.B.L. 1992. « An experimental system for auditory image representations ». *Biomedical Engineering, IEEE Transactions on* 39 (2): 112-21.
- Merleau-Ponty, M. 1945. *Phénoménologie de la perception*. Paris: Gallimard.
- . 1964. *Le visible et l'invisible*. Paris: Gal.
- Micaëlli, J.P., et Forest, J. 2003. *Artificialisme: introduction à une théorie de la conception*. Lausanne, Suisse: Presses Polytechniques et Universitaires Romandes.
- Mikolajczak, M., Quoidbach, J., Kotsou, I., et Nelis, D. 2009. *Les compétences émotionnelles*. Dunod.
- Minsky, M. 1980. « Telepresence ». *Omni* 2 (9): 45-51.
- Moggridge, B. 2006. *Designing Interactions*. MIT Press.
- Montagu, A. 1986. *Touching: the human significance of the skin*. New York: Harper & Row.
- Moore, C., Angelopoulos, M., et Bennett, P. 1997. « The role of movement in the development of joint visual attention ». *Infant Behavior and Development* 20 (1): 83-92.
- Moran, G., Dumas, J.E., et Symons, D.K. 1992. « Approaches to sequential analysis and the description of contingency in behavioral interaction. » *Behavioral Assessment*.
- Morris, D. 1977. *Manwatching: A Field Guide to Human Behavior*. London: Jonathan Cape.
- Motamedi, N. 2007a. « The aesthetics of touch in interaction design ». In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, 455-60. DPPI '07. New York, NY, USA: ACM.
- . 2007b. « Keep in touch: a tactile-vision intimate interface ». In *Proceedings of the 1st international conference on Tangible and embedded interaction*, 21-22. TEI '07. New York, NY, USA: ACM.
- Mueller, F., Vetere, F., Gibbs, M.R., Kjeldskov, J., Pedell, S., et Howard, S. 2005. « Hug over a distance ». In *CHI'05 extended abstracts on Human factors in computing systems*, 1673-76.
- Mugnaioni, R. 2006. « Les débuts de l'écriture en Mésopotamie ». In *Langues et écritures de la Méditerranée: actes du forum des 9, 10 et 11 mars 2001, Mansion du séminaire, Nice*, 37-48.

- Mundy, P., et Acra, C.F. 2006. « Joint attention, social engagement, and the development of social competence ». *The development of social engagement neurobiological perspectives*, 81-117.
- Mundy, P. 1995. « Joint attention and social-emotional approach behavior in children with autism ». *Development and Psychopathology* 7: 63-63.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Vaughan Van Hecke, A., et Parladé, M.V. 2007. « Individual differences and the development of joint attention in infancy ». *Child development* 78 (3): 938-54.
- Mundy, P., et Newell, L. 2007. « Attention, Joint Attention, and Social Cognition ». *Current directions in psychological science* 16 (5): 269-274.
- Neisser, U. 1988. « Five kinds of self-knowledge ». *Philosophical psychology* 1 (1): 35-59.
- . 1991. « Two perceptually given aspects of the self and their development ». *Developmental Review* 11 (3): 197-209.
- Neustaedter, C., et Greenberg, S. 2012. « Intimacy in long-distance relationships over video chat ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 753-62. CHI '12. New York, NY, USA: ACM.
- Neviarouskaya, A., Prendinger, H., et Ishizuka, M. 2010. « EmoHeart: Conveying Emotions in Second Life Based on Affect Sensing from Text ». *Advances in Human-Computer Interaction* 2010: 1-13.
- Noblet, J. de. 1988. *Design : le geste et le compas*. Paris: Somogy.
- Norman, D.A. 1988. *The Psychology of Everyday Things*. New York: Basic Books.
- . 2008. « THE WAY I SEE IT: Signifiers, not affordances ». *interactions* 15 (6): 18-19.
- O'Regan, J.K., et Noë, A. 2001. « A sensorimotor account of vision and visual consciousness ». *Behavioral and brain sciences* 24 (5): 939-72.
- Oakley, I., ABrewster, S., et Gray, P. 2001. « Communicating with feeling ». In *Haptic Human-Computer Interaction*, 61-68. Springer.
- Oakley, I., et O'Modhrain, S. 2002. « Contact IM: Exploring asynchronous touch over distance ». In *Proceedings of CSCW*, 16-20.
- Oatley, K., et Johnson-Laird, P.N. 1987. « Towards a cognitive theory of emotions ». *Cognition and emotion* 1 (1): 29-50.
- Ochs, E., et Schieffelin, B. 1989. « Language has a heart ». *Text* 9 (1): 7-25.
- Octavia, J.R., van den Hoven, E., et De Mondt, H. 2007. « Overcoming the distance between friends ». In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as we know it - Volume 2*, 79-82. BCS-HCI '07. Swinton, UK, UK: British Computer Society.

- Olson, G.M., et Olson, J.S. 2000. « Distance matters ». *Human-computer interaction* 15 (2): 139-78.
- Overbeeke, K., Djadjadiningrat, T., Hummels, C., et Wensveen, S. 2002. « Beauty in Usability: Forget about Ease of Use! » *Pleasure with products: Beyond usability*, 9-18.
- Pacherie, E. 1997. « Du problème de Molyneux au problème de Bach-y-Rita ». *Perception et intermodalité, approches actuelles du problème de molyneux*, 255-93.
- . 2002. « The role of emotions in the explanation of action ». *European Review of Philosophy* 5: 53-91.
- Pahl, G., et Beitz, W. 1996. *Engineering design: a systematic approach*. Springer-Verlag, London.
- Park, Y.W., Bae, S.H., et Nam, T.J. 2012. « How do couples use CheekTouch over phone calls? » In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 763-66. CHI '12. New York, NY, USA: ACM.
- Park, Y.W., Baek, K.M., et Nam, T.J. 2013. « The roles of touch during phone conversations: long-distance couples' use of POKE in their homes ». In *Proceedings of the 2013 ACM annual conference on Human factors in computing systems*, 1679-88. CHI '13. New York, NY, USA: ACM.
- Park, Y.W., Hwang, S., et Nam, T.J. 2011. « Poke: emotional touch delivery through an inflatable surface over interpersonal mobile communications ». In *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology*, 61-62. UIST '11 Adjunct. New York, NY, USA: ACM.
- Park, Y.W., Lim, C.Y., et Nam, T.J. 2010. « CheekTouch: an affective interaction technique while speaking on the mobile phone ». In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, 3241-46. CHI EA '10. New York, NY, USA: ACM.
- Park, Y.W., et Nam, T.J. 2013. « POKE: a new way of sharing emotional touches during phone conversations ». In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2859-60. CHI EA '13. New York, NY, USA: ACM.
- Pasquero, J. 2008. « Tactile Display for Mobile Interaction ». Thèse de doctorat, Montreal, Canada: McGill University.
- Pelachaud, C. 2011. « Les émotions dans l'interaction homme-machine ». In *Informatique et Sciences Cognitives : influences ou confluences ?*, 47-71. Co-édition Ophrys/MSH.
- Pérez-Pereira, M., et Conti-Ramsden, G. 1999. *Language development and social interaction in blind children*. Hove, UK: Psychology Press.
- [http://books.google.fr/books?hl=fr&lr=&id=9wd1xXuYwmwC&oi=fnd&pg=PP2&dq=Pe%CC%81rez-Pereira,+M.+%26+Conti-Ramsden,+G.,+\(1999\),+Language+Development+and+Social+Interaction+in+Blind+Chi](http://books.google.fr/books?hl=fr&lr=&id=9wd1xXuYwmwC&oi=fnd&pg=PP2&dq=Pe%CC%81rez-Pereira,+M.+%26+Conti-Ramsden,+G.,+(1999),+Language+Development+and+Social+Interaction+in+Blind+Chi)

- ldren,+Hove,+UK:+Psychology+Press&ots=MOQICBofAf&sig=B7rghti4dNKiBI8LriPj
UtVm4RU.
- Pettigrew, J. 2009. « Text Messaging and Connectedness Within Close Interpersonal Relationships ». *Marriage & Family Review* 45 (6-8): 697-716.
- Plantin, C., Doury, M., et Traverso, V. 2000. *Les émotions dans les interactions*. Presses Universitaires de Lyon.
- Platon. 1922. *Phèdre, ou de la Beauté des âmes*. Traduit par Mario Meunier.
- Plutchik, R. 1980. *Emotion: A Psychoevolutionary Synthesis*. New-York: Harper and Row.
- Polanyi, M. 1967. *The tacit dimension*. Peter Smith Gloucester, MA.
- Ponzio, A. 1996. *Sujet et altérité sur Emmanuel Levinas: suivi de deux dialogues avec Emmanuel Levinas*. Editions L'Harmattan.
http://books.google.fr/books?hl=fr&lr=&id=BMoAN2BZLzoC&oi=fnd&pg=PA5&dq=PONZIO+Sujet+et+alte%CC%81rite%CC%81+sur+Emmanuel+Levinas&ots=vdCzgz5Ja u&sig=gjlb0FPju0_dv3DBfC0i-fOmY4E.
- Prattichizzo, D., Chinello, F., Pacchierotti, C., et Minamizawa, K. 2010. « RemoTouch: A system for remote touch experience ». In *2010 IEEE RO-MAN*, 676-679.
- Preisler, G.M. 1991. « Early patterns of interaction between blind infants and their sighted mothers ». *Child: care, health and development* 17 (2): 65-90.
- Premack, D., et Woodruff, G. 1978. « Does the chimpanzee have a theory of mind? » *Behavioral and brain sciences* 1 (04): 515-26.
- Pujol, R.S., et Umemuro, H. 2009. « Productive love: A new approach for designing affective technology ». In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, 2469-78.
- Rabardel, P. 1995. *Les hommes et les technologies: approche cognitive des instruments contemporains*. Paris: Armand Colin.
- Racine, T.P., et Carpendale, J.I.M. 2007. « The Role of Shared Practice in Joint Attention ». *British Journal of Developmental Psychology* 25 (1): 3-25.
- Rallet, A. 2000. « Communication à distance : au-delà des mythes ». *Sciences humaines* N°104 (4): 24-24.
- Rallet, A., et Torre, A. 2007. *La proximité à l'épreuve des technologies de communication*. Editions L'Harmattan.
- Rastier, F. 2006. « De l'origine du langage à l'émergence du milieu sémiotique ». *Marges linguistiques* 11: 297-326.
- Rattray, J., et Zeedyk, M.S. 2005. « Early communication in dyads with visual impairment ». *Infant and Child Development* 14 (3): 287-309.

- Reddy, V. 2011. « A gaze at grips with me ». *Joint Attention: New Developments in Psychology, Philosophy of Mind and Social Neuroscience*, 137- 157.
- . 2003. « On being the object of attention: implications for self–other consciousness ». *Trends in Cognitive Sciences* 7 (9): 397-402.
- Reddy, V., et Morris, P. 2004. « Participants Don't Need Theories Knowing Minds in Engagement ». *Theory & Psychology* 14 (5): 647- 65.
- Reed, C.M., Delhorne, L.A., Durlach, N.I., et Susan D. Fischer. 1990. « A study of the tactual and visual reception of fingerspelling ». *Journal of speech and hearing research* 33 (4): 786.
- Register, L.M., et Henley, T.B. 1992. « The phenomenology of intimacy ». *Journal of Social and Personal Relationships* 9 (4): 467- 81.
- Rettie, R. 2003. « Connectedness, awareness and social presence ».
- Riva, G., Davide, F., et IJsselsteijn, W.A. 2003. « Measuring Presence: Subjective, Behavioral and Physiological Methods ». *Being there: Concepts, effects and measurement of user presence in synthetic environments*, 110- 18.
- Sacks, O. 1986. *A Leg to Stand On*. Picador.
- . 1992. *L'homme qui prenait sa femme pour un chapeau*. Seuil. Seuil.
- Saffer, D. 2006. *Designing for Interaction: Creating Smart Applications and Clever Devices*. Pearson Education.
- Samani, H.A., Parsani, R., Rodriguez, L.T., Saadatian, E., Dissanayake, K.H., et Cheok, A.D. 2012. « Kissenger: design of a kiss transmission device ». In *Proceedings of the Designing Interactive Systems Conference*, 48- 57. DIS '12. New York, NY, USA: ACM.
- Sarriá, E., Gómez, J.C., et Tamarit, J. 1996. « Joint attention and alternative language intervention in autism: Implications of theory for practice ». *Augmentative and alternative communication: European perspectives*, 49- 64.
- Saslis-Lagoudakis, G., Cheverst, K., Dix, A., Fitton, D., et Rouncefield, M. 2006. « Hermes@ Home: supporting awareness and intimacy between distant family members ». In *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, 23- 30.
- Schanberg, S. 1995. « The Genetic Basis for Touch Effects ». *Touch in early development*, 67- 79.
- Scherer, K.R. 1987. « Toward a dynamic theory of emotion: The component process model of affective states ». *Geneva Studies in Emotion and Communication* 1: 1-98.
- . 2000. « Psychological models of emotion ». *The neuropsychology of emotion*, 137- 162.
- . 2005. « What Are Emotions? And How Can They Be Measured? » *Social Science Information* 44 (4): 695- 729.

- Schmeer, J., et Baffi, T. 2011. « Touch trace mirror: asynchronous, collaborative messaging as a concept for creating a relatedness experience ». In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, 303-4. TEI '11. New York, NY, USA: ACM.
- Schmidt, R.C., et Richardson, M.J. 2008. « Dynamics of Interpersonal Coordination », janvier, 281-308.
- Schneider, B. 2005. « Le design comme science et comme recherche ». *La recherche en relation avec l'environnement du design*, Swiss Design Network édition.
- Schutz, A. 2005. « Common-Sense and Scientific Interpretation of Human Action ». *Knowledge Critical Concepts*.
- Searle, A. 1999. *Introducing research and data in psychology: A guide to methods and analysis*. Psychology Press.
- Searle, J.R. 1980. « Minds, brains, and programs ». *Behavioral and brain sciences* 3 (3): 417-57.
- Sebanz, N., Bekkering, H., et Knoblich, G. 2006. « Joint action: bodies and minds moving together ». *Trends in cognitive sciences* 10 (2): 70-76.
- Serino, A., et Haggard, P. 2010. « Touch and the body ». *Neuroscience & Biobehavioral Reviews* 34 (2): 224-236.
- Sheldon, K.M., Elliot, A.J., Kim, Y., et Kasser, T. 2001. « What is satisfying about satisfying events? Testing 10 candidate psychological needs. » *Journal of personality and social psychology* 80 (2): 325.
- Shin, H., Lee, J., Park, J., Kim, Y., Oh, H., et Lee, T. 2007. « A Tactile Emotional Interface for Instant Messenger Chat ». In *Human Interface and the Management of Information. Interacting in Information Environments*, édité par Michael J. Smith et Gavriel Salvendy, 166-175. Lecture Notes in Computer Science 4558. Springer Berlin Heidelberg.
- Short, J., Williams, E., et Christie, B. 1976. *The Social Psychology of Telecommunications*. Wiley.
- Smith, E.R., et Mackie, D.M. 2007. *Social Psychology*. 3rd éd. Psychology Press.
- Smith, J., et MacLean, K. 2007. « Communicating emotion through a haptic link: Design space and methodology ». *International Journal of Human-Computer Studies* 65 (4): 376-387.
- Sodhi, R., Poupyrev, I., Glisson, M., et Israr, A. 2013. « AIREAL: interactive tactile experiences in free air ». *ACM Trans. Graph.* 32 (4): 134:1-134:10.
- Sonneveld, M. 2002. « Aesthetics of tactual experience ». Delft University of Technology.
- Sribunruangrit, N. 2004. « Etude et développement de systèmes de suppléance perceptive tactile pour les personnes aveugles ». Thèse de doctorat, Université de Technologie de Compiègne.

- Sribunruangrit, N., Marque, C., Lenay, C., Gapenne, O., et Vanhoutte, C. 2002. « Braille Box: analysis of the parallelism concept to access graphic information for blind people ». In *Proceedings of the Second Joint Meeting of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society (EMBS-BMES)*, 3:2424- 25. USA, Houston, Texas.
- Stack, D.M. 2008. « The Saliency of Touch and Physical Contact During Infancy: Unraveling Some of the Mysteries of the Somesthetic Sense ». *Blackwell handbook of infant development*, 351- 378.
- Steiner, P. 2010. « Philosophie, Technologie, Cognition. Etat des Lieux Et Perspectives ». *Intellectica* 53: 7- 40.
- Stephan, P.F. 2005. « Le design cognitif – une perspective pour la recherche en design ». *La recherche en relation avec l'environnement du design*, Swiss Design Network édition.
- Stewart, J., Gapenne, O., et Di Paolo, E. 2010. *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press.
- Strong, R., et Gaver, B. 1996. « Feather, scent and shaker: supporting simple intimacy ». In *Proceedings of CSCW*, 96:29- 30.
- Suchman, L.A. 1987. *Plans and situated actions: the problem of human-machine communication*. Cambridge university press.
- Sundström, P., Ståhl, A., et Höök, K. 2005. « A User-Centered Approach to Affective Interaction ». In *Affective Computing and Intelligent Interaction*, édité par Jianhua Tao, Tieniu Tan, et Rosalind W. Picard, 931-938. Lecture Notes in Computer Science 3784. Springer Berlin Heidelberg.
- Takahashi, N., Okazaki, R., Okabe, H., Yoshikawa, H., Aou, K., Yamakawa, S., Yokoyama, M., et Kajimoto, H. 2011. « Sense-Roid: Emotional Haptic Communication with Yourself ». In *Virtual Reality International Conference*.
- Thayer, S. 1982. « Social touching ». *Tactual perception: A sourcebook*, 263.
- . 1986. « History and Strategies of Research on Social Touch ». *Journal of Nonverbal Behavior* 10 (1): 12- 28.
- Tollmar, K., et Persson, J. 2002. « Understanding remote presence ». In *Proceedings of the second Nordic conference on Human-computer interaction*, 41- 50.
- Tomasello, M. 1995. « Joint attention as social cognition ». *C. Moore & P. Dunham (Eds.), Joint Attention: Its Origins and Role in Development*, Hillsdale, NJ: Erlbaum, 103- 30.
- . 1999. *The cultural origins of human cognition*. Cambridge: Harvard University Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., et Moll, H. 2005. « Understanding and sharing intentions: The origins of cultural cognition ». *Behavioral and brain sciences* 28 (5): 675- 90.

- Tronick, E. Z. 1995. « Touch in mother-infant interaction ». In *Touch in early development*, 53- 65.
- Tsakiris, M., et Haggard, P. 2005. « The rubber hand illusion revisited: visuotactile integration and self-attribution ». *Journal of Experimental Psychology-Human Perception and Performance* 31 (1): 80-91.
- Tsujita, H., Tsukada, K., et Siio, I. 2009. « InPhase: a communication system focused on happy coincidences of daily behaviors ». In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, 3401-6.
- Turing, A.M. 1950. « Computing machinery and intelligence ». *Mind, New Series* 59 (236): 433- 60.
- Usoh, M., Catena, E., Arman, S., et Slater, M. 2000. « Using presence questionnaires in reality ». *Presence: Teleoperators & Virtual Environments* 9 (5): 497- 503.
- Van Baren, J., IJsselsteijn, W.A., Markopoulos, P., Romero, N., et de Ruyter, B. 2004. « Measuring affective benefits and costs of awareness systems supporting intimate social networks ». In *CTIT workshop proceedings series*, 2:13- 19.
- Van Overwalle, F. 2009. « Social cognition and the brain: A meta-analysis ». *Human brain mapping* 30 (3): 829- 58.
- Varela, F. J., et Maturana, H.R. 1994. *L'Arbre de la connaissance, racines biologiques de la compréhension humaine*. Ed. Addison-Wesley France, Paris.
- Varela, F.J., Rosch, E. et Thompson, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Vaudable, C. 2012. « Analyse et reconnaissance des émotions lors de conversations de centres d'appels ». Thèse, Ecole doctorale informatique Paris Sud.
- Vaughan Van Hecke, A., et Mundy, P. 2007. « Neural systems and the development of gaze following and related joint attention skills ». *Gaze following: Its development and significance*, 17- 51.
- Vetere, F., Gibbs, M.R., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoless, K., et Bunyan, M. 2005. « Mediating intimacy: designing technologies to support strong-tie relationships ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 471- 80. CHI '05. New York, NY, USA: ACM.
- Vetere, F., Howard, S., et Gibbs, M. 2005. « Phatic technologies: Sustaining sociability through ubiquitous computing ». In *First International Workshop on Social Implications of Ubiquitous Technology. ACM Conference on Human Factors in Computing Systems, CHI*.
- Visser, T., Vastenburger, M.H., et Keyson, D.V. 2011. « Designing to support social connectedness: The case of SnowGlobe ». *International Journal of Design* 5 (3): 129- 42.

- Von Hippel, E. 2009. « Democratizing innovation: the evolving phenomenon of user innovation ». *International Journal of Innovation Science* 1 (1): 29-40.
- Wang, R., et Quek, F. 2010. « Touch & talk: contextualizing remote touch for affective interaction ». In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, 13-20. TEI '10. New York, NY, USA: ACM.
- Wang, R., Quek, F., Tatar, D., Teh, K.S., et Cheok, A. 2012. « Keep in touch: channel, expectation and experience ». In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 139-48. CHI '12. New York, NY, USA: ACM.
- Wang, R., Quek, F., Teh, K.S., Cheok, A., et Lai, S.R. 2010. « Design and evaluation of a wearable remote social touch device ». In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, 45:1-45:4. ICMI-MLMI '10. New York, NY, USA: ACM.
- White, T., et Back, D. 1986. « Telephonic arm wrestling ». In *The Strategic Arts Initiative Symposium*.
- Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., et Yarin, P. 1998. « Ambient displays: Turning architectural space into an interface between people and digital information ». In *Cooperative buildings: Integrating information, organization, and architecture*, 22-32. Springer.
- Witmer, B.G., et Singer, M.J. 1998. « Measuring presence in virtual environments: A presence questionnaire ». *Presence* 7 (3): 225-40.
- Yarosh, S., Cuzzort, S., Müller, H., et Abowd, G.D. 2009. « Developing a media space for remote synchronous parent-child interaction ». In *Proceedings of the 8th International Conference on Interaction Design and Children*, 97-105. IDC '09. New York, NY, USA: ACM.
- Zahorik, P., et Jenison, R.L. 1998. « Presence as being-in-the-world ». *Presence: Teleoperators and virtual environments* 7 (1): 78-89.

ANNEXES

ANNEXE A : Note concernant l'analyse des trajectoires perceptives

Nous avons déjà mentionné que nous pouvons grâce à la méthodologie minimaliste forcer le déploiement des activités perceptives, afin de les rendre accessibles à l'analyse. Nos médiations techniques (Tactos et Touch Through) étant numériques, elles nous ont permis d'enregistrer ce déploiement sous la forme de fichiers « traces », à partir desquelles il est possible de tirer un ensemble d'indicateurs objectifs. Grâce à ces indicateurs, nous pouvons expliquer, ou du moins donner des pistes et des éléments d'explication concernant les réponses et les performances des sujets relativement aux tâches auxquelles ils se sont retrouvés confrontés.

De manière générale, les dispositifs Touch Through et Tactos permettent d'enregistrer la position du ou des sujets dans l'espace numérique (en termes de positions en (x ; y)) en fonction du temps. Pour l'ensemble des expérimentations relatées dans ce travail, nous avons recueilli ces positions à raison d'un enregistrement toutes les 20 ms. Nous obtenons ainsi pour chaque essai d'expérience un fichier trace daté et comportant tous les renseignements de la situation en question (voir par exemple la figure 143).

INFO	VERSION : Tactos1V2 2.0.LYNX.06 du 1er juin 2011						
INFO	=====						
INFO	Passation:	none					
INFO	Scenario:	File:	C:\Users\LaboCred\Desktop\Partage\tactos1V2_LYNX_05_Gabriell				
INFO	Scenario:	Name:	Passation1				
INFO	Scenario:	Session:	session 1				
INFO	Scenario:	Session:	01 08				
INFO	=====						
INFO	Users:	User 0					
INFO	Cursor:	SensorClassical-4-4-4.xml		O1			O2
INFO	Salle:	300_150_All.xml	300		300	150	150
INFO	Body:	..\data\Bodies\Black02.xml					
INFO	Groupe:	0					
TYPE	Tps(ms)	O.X	O.Y	O.Stim	O.Clic	Position clic	Position Objé
DATA	30	83	150	9E+16	0	✓	0
DATA	50	83	150	9E+16	0	✓	0
DATA	70	83	150	9E+16	0	✓	0
DATA	90	83	150	9E+16	0	✓	0
DATA	110	83	150	9E+16	0	✓	0
DATA	130	83	150	9E+16	0	✓	0
DATA	150	83	150	9E+16	0	✓	0

Figure 143. Exemple de fichier trace de Tactos. L'encadré rouge (en haut) comporte les informations générales, comme la version logiciel utilisée ou encore le numéro de passation et d'essai ; l'encadré bleu comporte les informations techniques relativement à l'essai enregistré, comme le type de corps-image ou de corps-percevant ; l'encadré orange comporte les données en tant que telles (position dans l'espace en fonction du temps, stimulations reçues, ou parfois clic des sujets).

À partir des données recueillies (positions en (x ; y) en fonction du temps, stimulations reçues ou encore clic des sujets), il nous est possible de dégager un ensemble d'indicateurs dérivés. Nous pouvons par exemple calculer l'écart temporel entre deux stimulations, ou encore la distance parcourue par un sujet donné à chaque pas de temps. Nous regroupons l'ensemble des données dérivées dans un fichier d'analyse récapitulatif portant sur l'essai considéré de la situation expérimentale étudiée. Ces données dérivées peuvent donc donner lieu à des statistiques descriptives, de façon à obtenir une valeur globale par sujet, par essai et par indicateur donnés (voir par exemple la figure 144).

Indicateurs Globaux	Moyenne	Ecart-type
Nb de changement de direction ou d'arrêts	7,545454545	4,06747524
Nb d'arrêts > 1/2 sec	5	2,60195336
Dont arrêts > 1 sec	65,24242424	40,1493055
Nombre de Stim	54343,0303	13369,5219
Durée Stim	0,719775236	0,17707976
% Durée Stim	3772,034284	4603,20986
Durée Stim Moyenne	26124,84848	18146,2977

Figure 144. Exemple d'un fichier d'analyse (extrait).

Ces valeurs sont ensuite regroupées dans un fichier récapitulatif regroupant les données de tous les sujets pour tous les essais, fichier à partir duquel nous pouvons généraliser les statistiques descriptives en fonction des conditions expérimentales ou des types de réponses données par les sujets (par exemple, la moyenne d'un indicateur donné obtenu par un sujet lorsqu'il a correctement répondu à la tâche) (voir figure 145).

Sujet	Essai	Condition	Bonnes Réponses	Mauvaise réponse	Indicateur 1	Si bonne réponse
1	1	1	1	0	x1	
1	2	1	0	1	x2	
1	3	1	0	1	x3	
1	4	1	1	0	x4	X1
2	1	1	0	1	y1	
2	2	1	1	0	y2	
2	3	1	1	0	y3	
2	4	1	1	0	y4	Y2

Figure 145. Exemple de fichier récapitulatif (extrait).

ANNEXE B : Exemples de trajectoires perceptives (Turing avec Touch Through)

Exemple 1 : Interaction entre deux partenaires (condition Distale)

Cette trajectoire correspond à un essai qui s'est déroulé en face à face : pour chaque sujet, l'interactant était le partenaire (voir figure 146). Cet essai a conduit à une reconnaissance mutuelle, et est caractérisé par un grand nombre de séquences d'interaction. Chaque cercle indique un contact entre les sujets ; plus les cercles sont épais, plus les sujets ont interagi à l'endroit concerné.

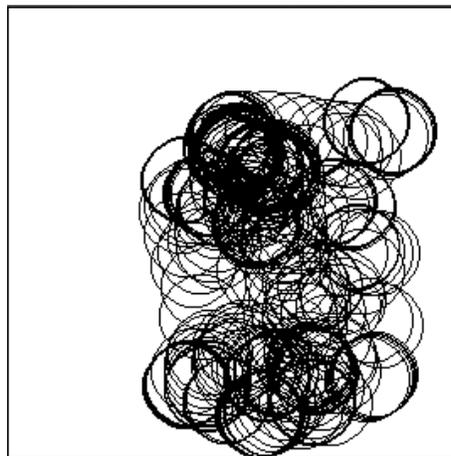


Figure 146. Interaction entre deux partenaires pendant la condition Distale.

Exemple 2 : Interaction entre deux partenaires (condition Proximale Visuelle)

À nouveau, nous présentons une interaction ayant eu lieu en face à face, mais il s'agit cette fois d'un essai en condition Proximale Visuelle, où les sujets se sont à nouveau reconnus mutuellement (voir figure 147). On constate que le nombre d'interactions est moins important et que ces dernières sont bien plus localisées. Cela s'explique par le fait que les sujets ne savent pas à chaque instant où se trouve leur partenaire, contrairement à la condition Distale.

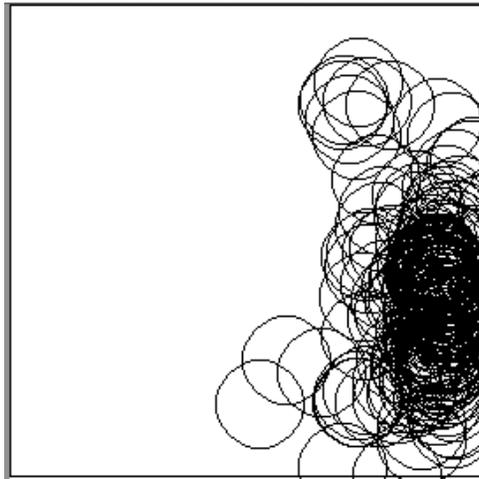


Figure 147. Interaction entre deux partenaires pendant la condition Proximale Visuelle.

Exemple 3 : Interaction entre deux partenaires (condition Proximale Vibratoire)

À nouveau, nous présentons une interaction en face à face où les sujets se sont mutuellement reconnus, mais cette fois-ci en condition Proximale Vibratoire (voir figure 148). On constate à nouveau un nombre de séquences d'interaction plus faible qu'en condition Distale, et des interactions plus localisées.

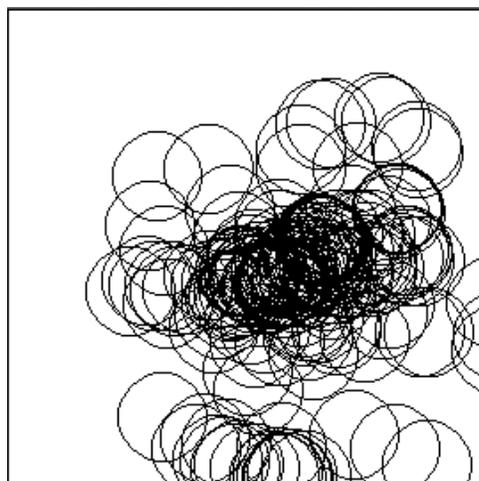


Figure 148. Interaction entre deux partenaires pendant la condition Proximale Vibratoire.

Exemple 4 : Interaction entre un sujet humain et un robot (reconnaissance)

Cette trajectoire correspond à un essai qui s'est déroulé face au robot ; nous montrons ici l'enregistrement de ce qu'un sujet (le rond) a fait face un robot (le carré) (voir figure 149). Cet essai a conduit à une réponse correcte : le sujet a identifié le robot comme étant un robot. Cette interaction est caractérisée par un nombre de séquences d'interaction faible ; le sujet a pu constater que le robot ne recherchait pas son contact.

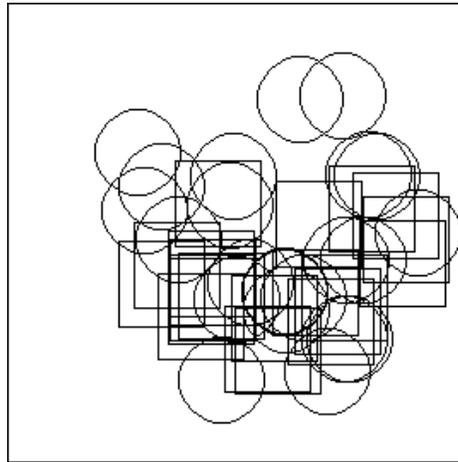


Figure 149. Interaction entre un sujet humain et un robot (reconnaissance : réponse Robot). Le robot est représenté par un carré par souci de distinction, mais bien entendu, les contacts étaient calculés à partir de la forme ronde de tous les corps-images.

Exemple 5 : Interaction entre un sujet humain et un robot (non-reconnaissance)

Cette trajectoire correspond à un essai qui s'est déroulé face au robot en condition distale ; nous montrons ici l'enregistrement de ce qu'un seul sujet (le rond) a fait face un robot (le carré) (voir figure 150). Cet essai a conduit à une réponse incorrecte : le sujet a identifié le robot comme étant son partenaire. Cette interaction est caractérisée par un nombre de séquences d'interaction élevé, ce qui a induit le sujet en erreur ; il a cru que le robot recherchait son contact, et était donc humain.

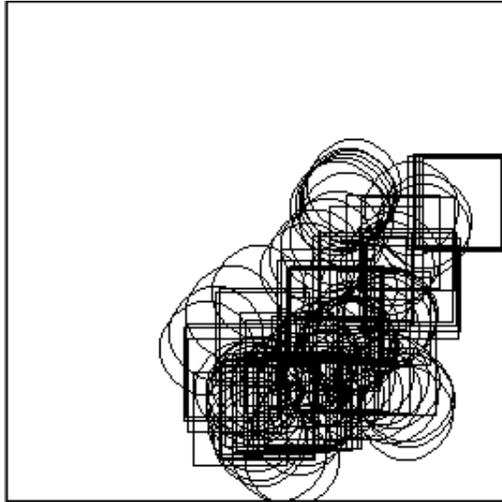


Figure 150. Interaction entre un sujet humain et un robot (non-reconnaissance : réponse Partenaire).

ANNEXE C : Exemples de trajectoires perceptives (Commun/Privé)

Exemple 1 : Clic Mutuel en condition « Mono » (stratégie orientée objets)

La trajectoire perceptive considérée correspond à un essai de la condition « Mono » (voir figure 151). Il est intéressant de noter que les activités perceptives sont d'abord tournées vers la recherche des objets, malgré un premier contact entre les sujets survenu très tôt au début de l'essai. Ainsi, les sujets se rencontrent d'abord successivement autour de leur objet privé respectif, mais l'interaction ne s'y stabilise pas. L'interaction se stabilise en revanche lorsqu'ils se rencontrent autour de l'objet commun ; les sujets coordonnent leurs activités perceptives à proximité de celui-ci. Cet essai a conduit à un clic mutuel sur l'objet commun.

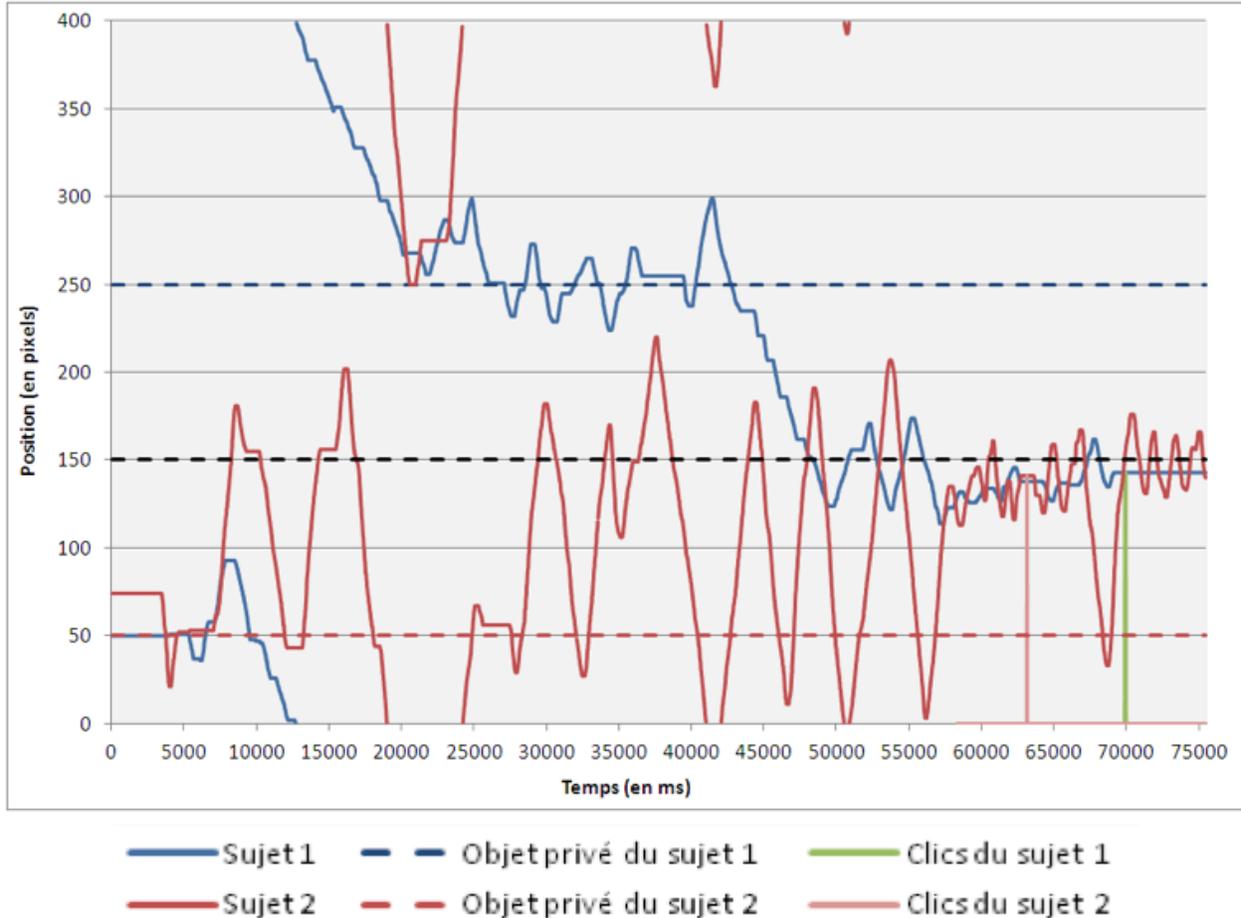


Figure 151. Clic Mutuel en condition « Mono » (stratégie orientée objets).

Exemple 2 : Clic Mutuel en condition « Mono » (stratégie orientée sujet)

La trajectoire perceptive considérée correspond à un essai de la condition « Mono » (voir figure 152). Nous notons que cette fois-ci, les activités perceptives déployées par les deux sujets sont en première intention dirigées l'une vers l'autre. Ainsi, les sujets se suivent dans l'espace d'interaction et explorent ensemble les objets qui y sont présents, même si nous observons ponctuellement que le sujet 2 se désengage de l'interaction pour localiser les objets. La rencontre de l'objet privé du sujet 1 produit une courte séquence d'interaction, mais elle ne se maintient pas dans le temps. Cet essai a conduit à un clic mutuel sur l'objet commun.

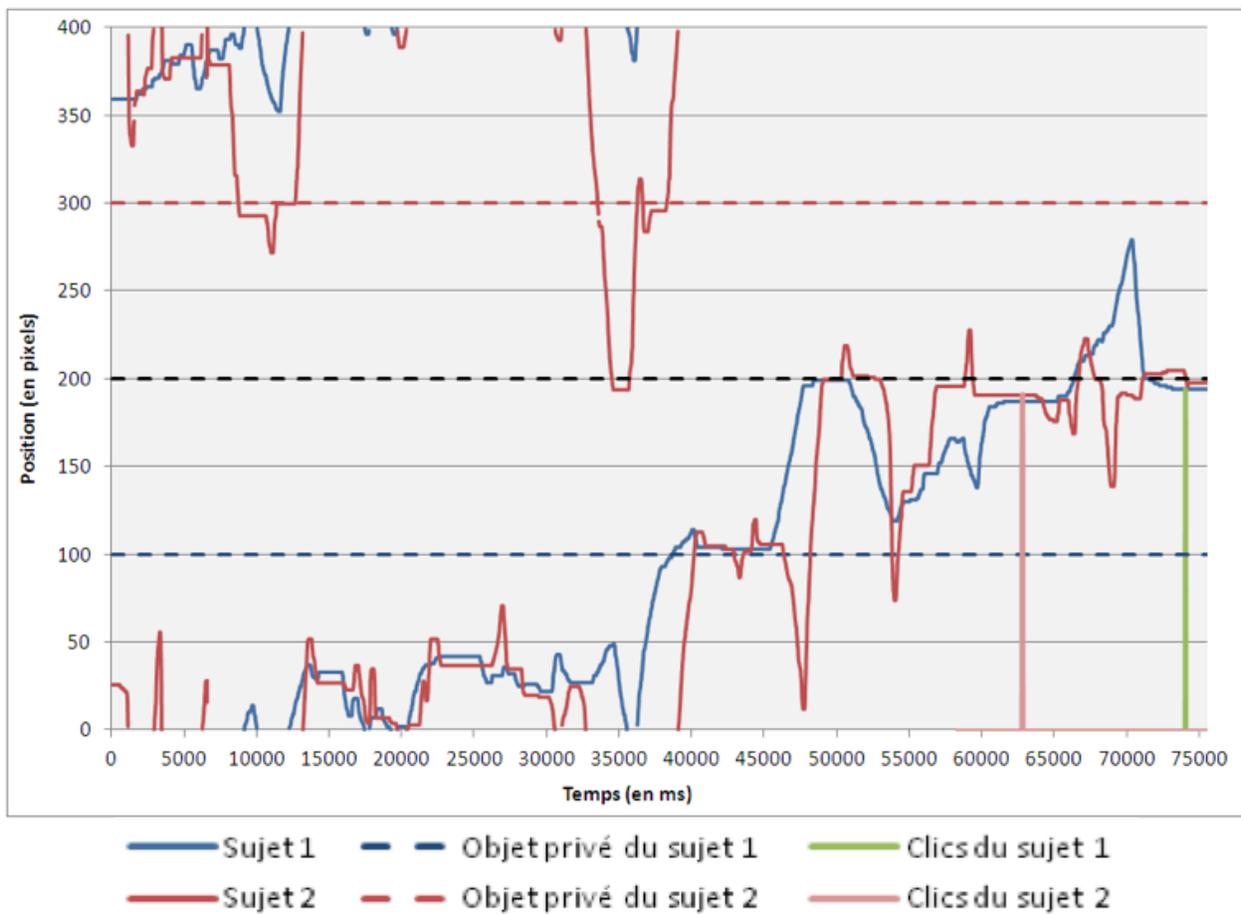


Figure 152. Clic Mutuel en condition « Mono » (stratégie orientée sujet).

Exemple 3 : Clic Mutuel en condition « Différenciée » (stratégie orientée objets)

La trajectoire perceptive considérée correspond à un essai de la condition « Différenciée » (voir figure 153). Les stratégies individuelles sont ici clairement orientées vers les objets : les sujets cherchent les objets en première intention, puis ils s’y arrêtent en attendant l’arrivée de leur partenaire. Ainsi, le déploiement des activités perceptives se caractérise par une succession d’arrêts sur les objets. Les arrêts sur les objets privés ne se maintiennent pas, contrairement à celui qui a lieu sur l’objet commun, qui se maintient et dure pour le reste de l’essai. L’essai présenté ici a conduit à un clic mutuel correct.

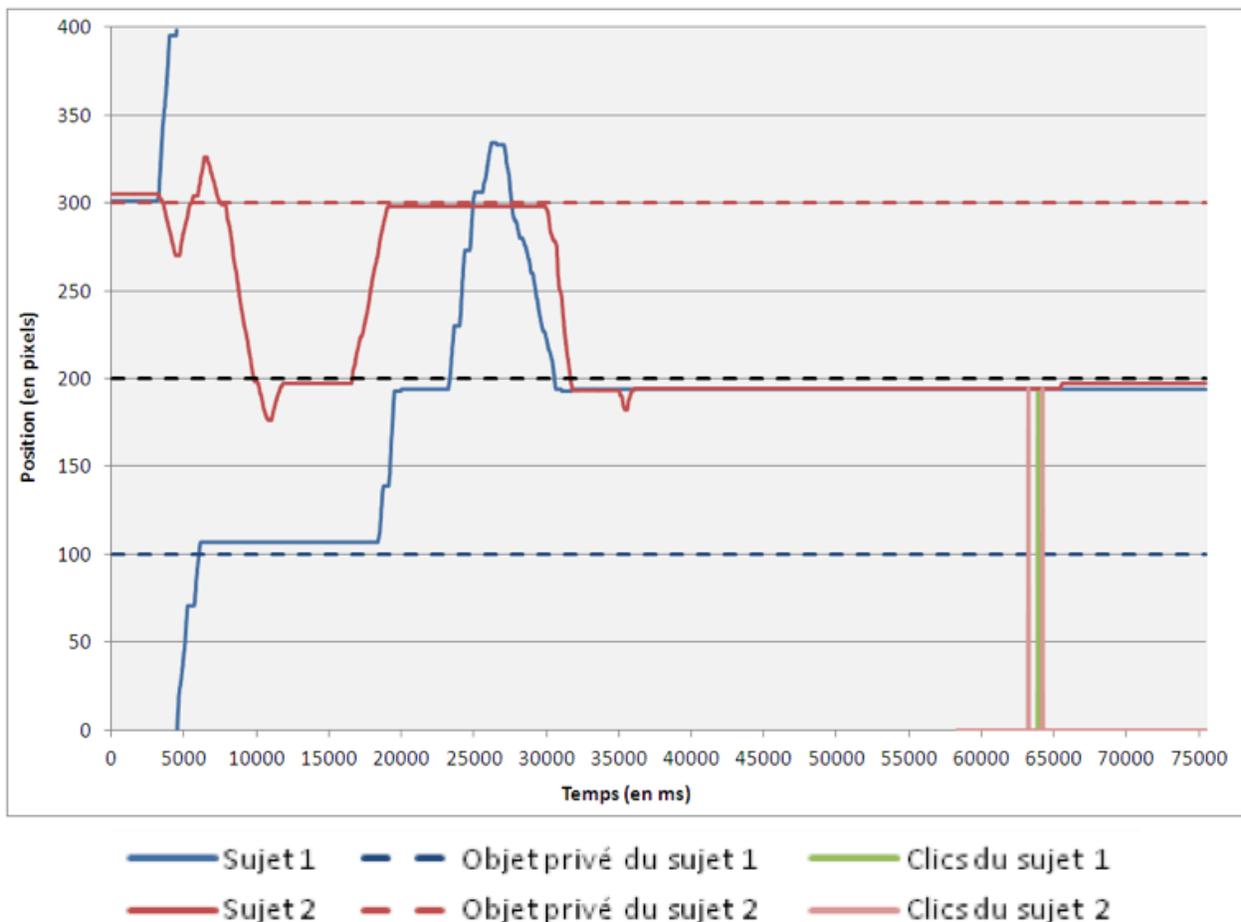


Figure 153. Clic Mutuel en condition « Différenciée » (stratégie orientée objets).

Exemple 4 : Clic Mutuel en condition « Différenciée » (stratégie orientée sujet)

La trajectoire perceptive considérée correspond à un essai de la condition « Différenciée » (voir figure 154). Contrairement à l'exemple 3, les stratégies des sujets sont ici orientées l'une vers l'autre. La première intention des sujets est de chercher à maintenir le contact, de manière à explorer ensemble l'espace d'interaction et les objets présents. Notons que la rencontre de l'objet privé du sujet 1 conduit ce dernier à s'arrêter un peu, alors que le sujet 2 continue ses oscillations. Au contraire, lorsque les deux sujets arrivent sur l'objet commun, l'interaction se stabilise jusqu'à l'arrêt mutuel, ce qui conduit alors à un clic mutuel correct en fin d'essai.

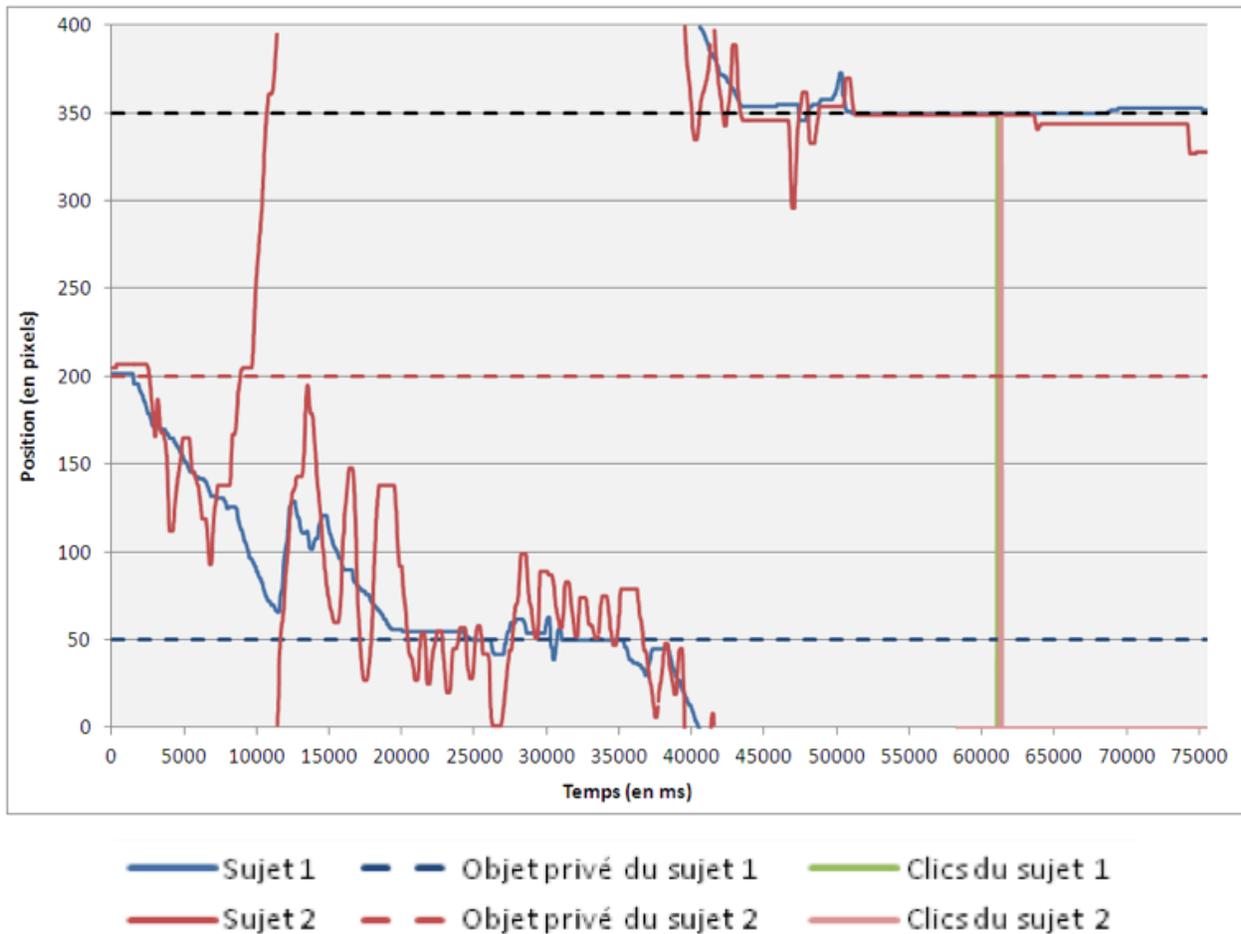


Figure 154. Clic Mutuel en condition « Différenciée » (stratégie orientée sujet).

Exemple 5 : Clic Mutuel en condition « Parallélisme » (stratégie orientée objets)

La trajectoire perceptive considérée correspond à un essai de la condition « Parallélisme » (voir figure 155). Les sujets déploient ici une stratégie plutôt orientée vers les objets. Ils recherchent en effet activement les objets présents dans l'espace d'interaction, en passant de l'un à l'autre. Lorsque les sujets se rencontrent une première fois autour de l'objet commun, ils poursuivent leur exploration. Au milieu de l'essai, lorsqu'ils se rencontrent de nouveau autour de l'objet commun, le sujet 2 signale sa présence à l'autre sujet en effectuant de rapides oscillations avec une faible amplitude. Le sujet 1, de son côté, déploie une activité qui semble plus chaotique, mais maintient tout de même sa présence autour de l'objet et de l'autre participant. L'essai a conduit à un clic mutuel sur l'objet commun.

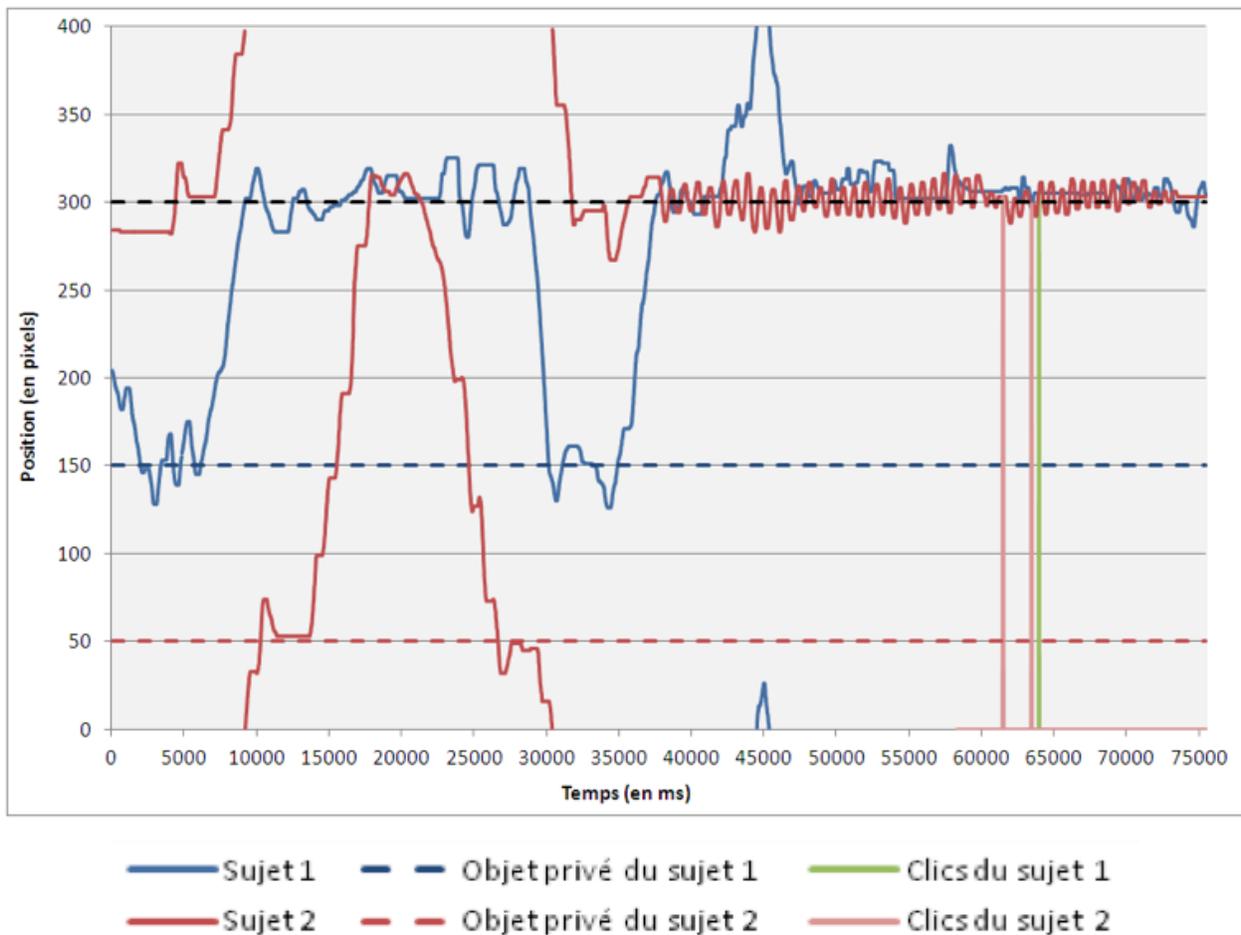


Figure 155. Clic Mutuel en condition « Parallélisme » (stratégie orientée objets).

Exemple 6 : Clic Mutuel en condition « Parallélisme » (stratégie orientée sujet)

La trajectoire perceptive considérée correspond à un essai de la condition « Parallélisme » (voir figure 156). Contrairement à l'exemple précédent, les stratégies des sujets sont ici clairement orientées l'une vers l'autre. Les sujets cherchent donc en première intention à se rencontrer dans l'espace d'interaction, avant de l'explorer ensemble. Lorsque le sujet 2 rencontre son objet privé, il suit d'abord le sujet 1, mais revient ensuite autour de l'objet. Pendant ce temps, le sujet 1 part à la recherche d'un objet : il n'en a pas encore rencontré. Puis il revient au contact du sujet 2, qui était resté sur son objet privé. Le sujet 1 parvient alors à emmener le sujet 2 dans la direction de l'objet qu'il a rencontré, et leur interaction s'organise alors autour de l'objet commun. L'essai a conduit à un clic mutuel correct.

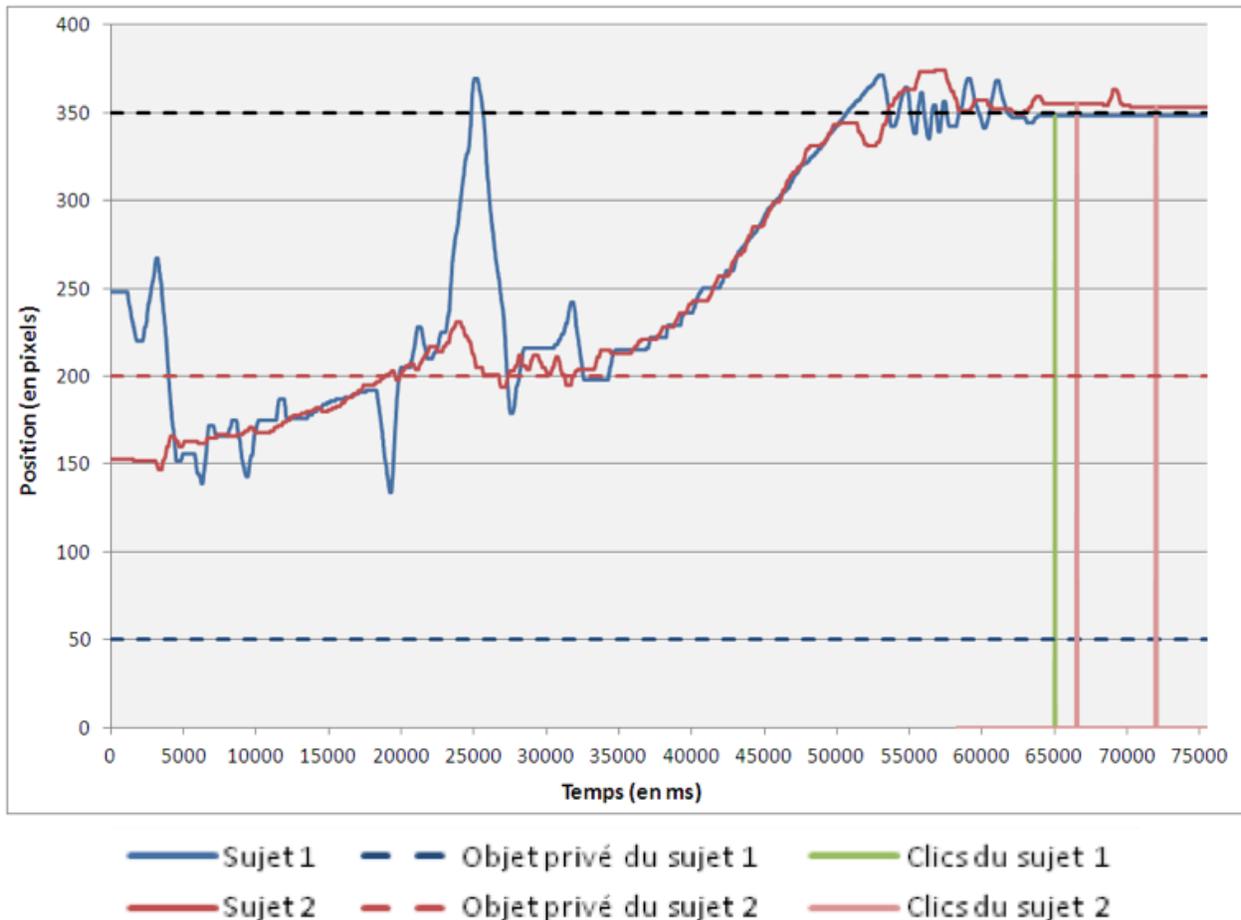


Figure 156. Clic Mutuel en condition « Parallélisme » (stratégie orientée sujet).

ANNEXE D : Exemples de trajectoires perceptives (expérience préliminaire sur un « langage » tactile)

Exemple 1 : binôme 1

La trajectoire perceptive considérée correspond à un essai du binôme 1 (voir figure 157). Il s'agit ici de la phase de restitution des signes. Les sujets rejouent le signe qu'ils ont convenu pour « oui »/« je suis d'accord »/« je comprends » ; à gauche, on voit ce que le sujet A a joué, à droite ce qu'a joué le sujet B. Sont grisées de façon de plus en plus intense en fonction du temps de contact les cases de la matrice 4 x 4 (apparaissant sur l'écran des smartphones et correspondant à la matrice de picots du partenaire). Ici, le sujet A est donc resté un peu plus longtemps à faire le signe (à savoir activer le picot en haut à gauche) que le sujet B.

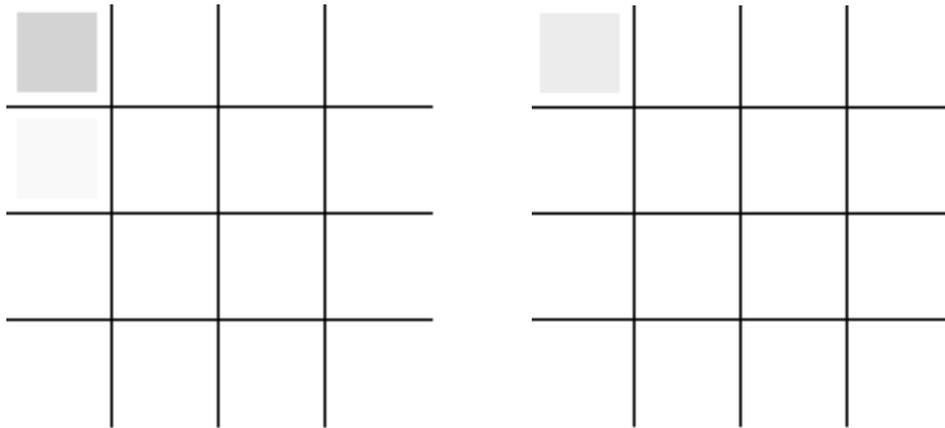


Figure 157. Trajectoire perceptive d'un essai du binôme 1 (sujet A à gauche, sujet B à droite).

Exemple 2 : binôme 2

La trajectoire perceptive considérée correspond à un essai du binôme 2 (voir figure 158). Il est intéressant de noter que les deux sujets n'ont pas investi les mêmes zones de l'espace d'interaction, et il n'est donc pas étonnant qu'ils ne se soient pas compris. L'effet a en effet conduit à une mauvaise réponse. De plus, il s'agissait de l'essai où le sujet B devait jouer « chez sujet B, je ne sais pas quand » pour le faire deviner au sujet A : il fallait donc inventer le signe pour « je ne sais pas quand », ce qui explique la différence de comportements et l'incompréhension entre les sujets.

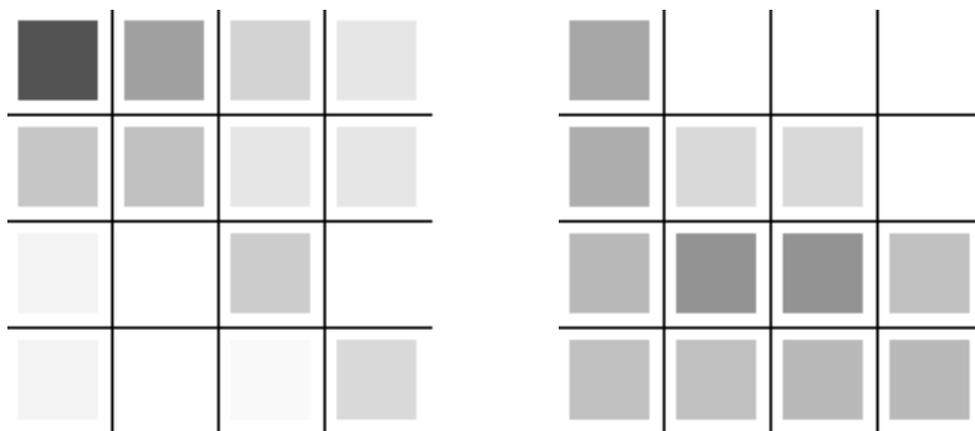


Figure 158. Trajectoire perceptive d'un essai du binôme 2 (sujet A à gauche, sujet B à droite).

Exemple 3 : binôme 3

La trajectoire perceptive considérée correspond à un essai du binôme 3 (voir figure 159). C'est un essai de la phase de restitution ; le sujet B devait jouer « est-ce que tu comprends ? »/« est-ce que tu es d'accord ? » (signe défini par le picot inférieur gauche par ces sujets) pour le faire deviner au sujet A. On notera que le sujet A a fait jouer une zone de la matrice qui correspondait au signe pour « oui »/« je suis d'accord »/« je comprends » (le deuxième picot de la troisième colonne), indiquant ainsi au sujet B qu'il avait bien compris, et répondant effectivement la bonne réponse sur la feuille.

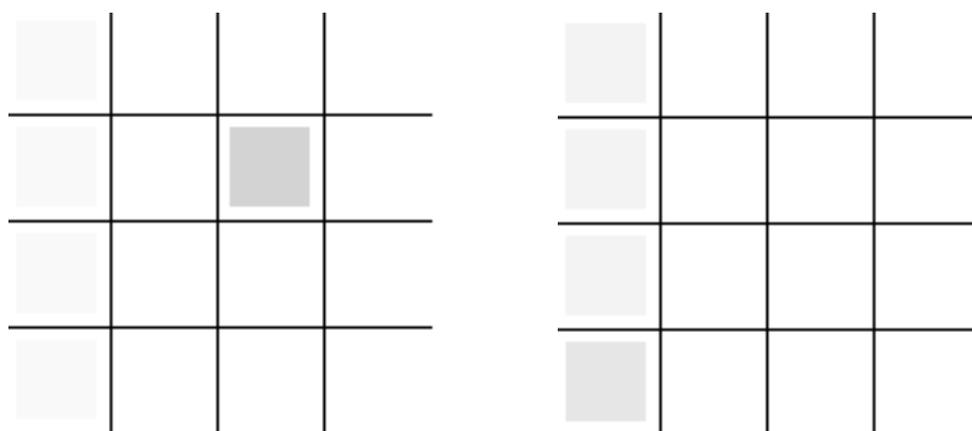


Figure 159. Trajectoire perceptive d'un essai du binôme 3 (sujet A à gauche, sujet B à droite).

Exemple 4 : binôme 4

La trajectoire perceptive considérée correspond à un essai du binôme 4 (voir figure 160). C'est un essai de la phase de restitution ; le sujet B devait jouer « non »/« je ne sais pas »/« je ne suis pas d'accord »/« je ne comprends pas » pour le faire deviner au sujet A. Le sujet B a tout d'abord joué le signe (les deux rangées centrales), puis c'est le sujet A qui l'a joué. Ensuite, le sujet B a confirmé le signe en le rejouant, puis en jouant « oui »/« je suis d'accord »/« je comprends » (le cercle de picot extérieur), signe que le sujet A a rejoué à son tour. Les deux sujets ont donc mis en place une véritable conversation pour assurer qu'ils s'étaient bien compris, et, sans surprise, le sujet A a bien répondu sur sa feuille.

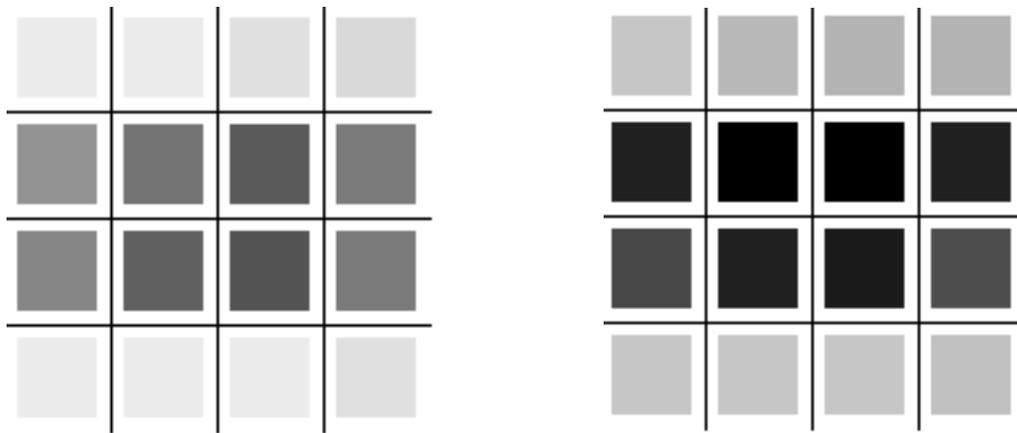


Figure 160. Trajectoire perceptive d'un essai du binôme 4 (sujet A à gauche, sujet B à droite).