

Corporéité partagée et agentivité distribuée en interaction par écran

< Samira Ibnelkaïd¹ >

1. Laboratoire ICAR (CNRS, Université Lyon 2, ENS de Lyon)
15, parvis René Descartes, BP7000, 69342 Lyon Cedex 07
samiraibnelkaid@gmail.com

DOI:10.3199/RIN.1.1-n © AFDI 2012

< RESUME >

Les technologies numériques ont engendré une révolution culturelle et cognitive modifiant le rapport de l'homme à son environnement, à son prochain et à lui-même. En effet, les sujets se trouvent, à l'apparition de chaque nouvelle technologie de l'information et de la communication, impliqués dans des environnements technoperceptifs inédits (Vial, 2013). Et l'ère numérique introduit de nouvelles formes de connaissance et de communication qui ne renvoient plus à une pensée linéaire à l'espace et au temps délimités mais constituent un réseau. Il y a alors reconfiguration de la perception et de la représentation du et par le corps dans son écologie réticulaire. Les nouvelles modalités d'agentivité (Butler, 2002) issue de la médiation de la *tekhnê* de plus en plus intelligente, immersive et engagée dans l'interaction sociale introduisent un risque de confusion entre le Soi et l'Autre. Entre le locuteur, l'artefact et l'interlocuteur se pose alors la question de l'attribution du geste interactionnel de son émission à sa perception au cours d'interactions par écran. Sont donc ici analysées des interactions vidéo par écrans – fixes (ordinateurs) et mobiles (robots de téléprésence) – au cours desquelles la multimodalité interactionnelle se trouve complexifiée par l'agentivité distribuée et donne lieu à une transsubjectivité du geste communicatif.

< ABSTRACT >

Digital technologies have led to a cultural and cognitive revolution, one that reshapes humans' relationship to their environment, to others, to themselves. Indeed, at every new information and communication technology emergence, individuals find themselves involved in unprecedented techno-perceptive environments (Vial, 2013). And the digital era introduces new forms of knowledge and communication that no longer consist in a linear perception of a delimited time and space but constitute a network. We are therefore witnessing a reconfiguration of perception and representation of the body and through the

body in its reticular ecology. The new agency (Butler, 2002) modalities, originating from the mediation of the *tekhnê* which is more and more intelligent, immersive and involved in social interactions, introduce a risk of confusion between the Self and the Other. Between the locutor, the artefact and the interlocutor, the attribution of the interactional gesture from its emission to its perception during the digital interaction, becomes problematic. To address this issue, this paper examines data of video-synchronous digital interactions (motionless screens (computers) and movable screens (robots)) within which interactional multimodality is made more complex by the distributional agency and introduces transsubjectivity in the communicative gesture.

< MOTS-CLES >

Interaction numérique ; écran ; agentivité ; corporéité ; multimodalité ; plurisémiotité ; transsubjectivité ; linguistique interactionnelle ; ethnographie visuelle ; phénoménologie.

< KEYWORD >

Digital interaction ; screen ; agency ; body ; multimodality ; plurisemioticity ; transsubjectivity ; interaction analysis ; visual ethnography ; phenomenology.

1. Introduction

Le langage humain « réalise, pour le sujet parlant et pour ceux qui l'écoutent, une certaine structuration de l'expérience, une certaine modulation de l'existence, exactement comme un comportement de mon corps investit pour moi et pour autrui les objets qui m'entourent d'une certaine signification » (Merleau-Ponty, 1945 : 225). La prise d'existence et la structuration de l'expérience se déploient en effet au cours de l'interaction sociale par des ressources symboliques – tant verbales que corporelles – se trouvant à la disposition des interactants qui cherchent à se positionner face à l'altérité. Les participants à l'interaction sociale font ainsi usage de ressources multimodales – formes linguistiques, gestes, regards, mimiques, postures, etc. (Greco, Mondada & Renaud, 2014) – mettant en œuvre leur capacité à se rendre visibles et rendre intelligibles leurs productions et les activités dans lesquelles elles s'inscrivent.

Et l'hypermodernité ouvre de surcroît un nouvel univers de communication qui augmente et modifie la capacité humaine de manipulation symbolique ; par là c'est « l'être même de l'humanité – sa

singularité ontologique – qui est appelé à se reconstruire » (Lévy, 2013 : 16). Les interactions sociales se réalisent, de plus en plus, dans et par les écrans. L'écran, omniprésent dans les sociétés hypermodernes, fait figure d'interface des activités humaines de communication, information, médiation et affecte chez l'Homme sa pensée visuelle et la perception de son corps dans la culture matérielle (Frau-Meigs, 2011). Loin de n'être qu'un simple support, l'écran – objet dynamique, immersif – ouvre de nouvelles formes de liens sociaux et culturels. Il se révèle être un « *vecteur de communication et d'échange qui permet de dialoguer sur le monde et sur les autres, notamment dans l'espace potentiel où le sujet est à la fois relié à l'autre et séparé de lui* » (Ibid. : 124). Une dialectique s'établit entre l'utilisateur créateur d'une nouvelle sémiotique et l'écran générateur de nouvelles significations. Des modalités inédites d'interaction se déploient alors par l'intermédiation de la *tekhnê*¹.

Il apparaît dès lors nécessaire d'interroger la responsabilité de chaque action produite dans la communication par écran. Rappelons en effet qu'au cours de leurs interactions avec autrui, les sujets produisent des actions dont, quelle que soit leur modalité – verbale, para-verbale, non-verbale –, ils se trouvent tenus pour responsables. Cette capacité d'action des sujets sur leur environnement, sur les objets et sur Autrui, ainsi que la perception de cette faculté par le sujet relève de la notion d'agentivité (ou *agency*) (Butler, 2002). L'agentivité garantit à l'interactant qu'il est maître de ses actes, évitant ainsi toute confusion entre Soi et Autrui. Cette garantie apparaît pourtant mise à l'épreuve de la médiation de la *tekhnê* de plus en plus intelligente, immersive et engagée dans l'interaction sociale. Entre le locuteur, l'artefact et l'interlocuteur se pose la question de l'attribution du geste interactionnel, de son émission à sa perception au cours d'interactions numériques. À la multimodalité interactionnelle s'associe la plurisémiotité des échanges (hypertextualité, graphisme, audio, vidéo, action à l'écran, etc.) et émerge la problématique à la fois théorique et

¹ Selon Heidegger, le terme grec *tekhnê* renvoie à un mode du savoir et une production de l'étant ; « *il ne désigne jamais un genre de réalisation pratique, mais le fait d'appréhender, d'éprouver la présence du présent en tant que tel. La tekhnê est une production qui fait venir l'être à découvert, hors de sa réserve, dans sa décloison, c'est-à-dire dans sa vérité* » ([1935] 1962 : 66).

empirique de leur fonctionnement dans le langage et de leur documentation par le chercheur.

C'est pourquoi nous proposons d'étudier des interactions vidéo par écrans – fixes (ordinateurs) et mobiles (robots de téléprésence) – au cours desquelles la multimodalité interactionnelle se trouve complexifiée par l'agentivité distribuée. Notre corpus d'étude consiste en une auto-ethnographie de séminaires de recherche du sous-groupe de Recherche sur les Présences Numériques du groupe IMPEC (Interactions Multimodales Par Écran) du Laboratoire ICAR de Lyon (UMR 5191). Ces séminaires, impliquant des participant.e.s géographiquement distant.e.s, font usage d'artefacts communicationnels multiples (des ordinateurs et tablettes équipés de logiciels visio tels que Skype, Google Hangout et Adobe Connect, ainsi que des robots de téléprésence Beam et Kubi). Nous avons capturé ces interactions au moyen d'un dispositif technique consistant en une capture dynamique d'écrans de participants en ligne et un enregistrement vidéo du contexte spatial (caméra externe). La restitution de ces données, par assemblage des vues, incrustation des transcriptions verbales et enrichissement sémiotique des vidéos, permet d'analyser simultanément les comportements communicatifs des participants sur et hors écran. Nous en présenterons des extraits dans ce qui suit. À partir d'une approche interdisciplinaire – ethnographie visuelle (Ruby, 1996 ; Banks & Morphy, 1997 ; Pink, 2007 ; Dion, 2007), linguistique interactionnelle multimodale (Goffman, 1973 ; Cosnier, 2004 ; Mondada, 2008 ; Kerbrat-Orecchioni, 2010 ; Traverso, 2012) et analyse phénoménologique (Husserl, 1929 ; Merleau-Ponty, 1945 ; Le Breton, 2001 ; Vial, 2013) – nous mettons en exergue la multimodalité et la plurisémiotité de l'agentivité technico-corporelle en interaction par écran. Nous développerons les fondements de cette approche théorique avant de la mettre en pratique sur des séquences interactionnelles issues de notre corpus et dites problématiques en termes de médiation et d'agentivité multimodale. Ces analyses nous inviteront à réviser et prolonger l'appareillage théorique d'étude de l'agentivité corporelle en interaction par écran.

2. Interaction et sujets de l'action

2.1. L'approche interactionniste

La notion d'interaction recouvre des définitions plus ou moins restreintes en fonction de la posture adoptée par les chercheur.e.s à son égard. Goffman, linguiste et sociologue figurant parmi les fondateurs de l'analyse des interactions, explique que « *par interaction on entend à peu près l'influence réciproque que les partenaires exercent sur leurs actions respectives lorsqu'ils sont en présence physique immédiate les uns des autres* » (1973 : 23). Kerbrat-Orecchioni précise, quant à elle, que pour qualifier une situation d'interaction « *il faut et il suffit que l'on ait un groupe de participants modifiable mais sans rupture, qui dans un cadre spatio-temporel modifiable mais sans rupture, parlent d'un objet modifiable mais sans rupture* » (1990 : 216). Dans une acception plus large, Vion affirme que le terme interaction « *intègre toute action conjointe, conflictuelle et/ou coopérative mettant en présence deux ou plus de deux acteurs* » (1992 : 17). Joseph (1998 : 27), dans une orientation sociologique, définit l'interaction comme :

« *un système interactif comportant au moins quatre composantes : un ensemble d'unités qui interagissent les unes avec les autres ; un code ou un ensemble de règles qui structurent aussi bien l'orientation de ces unités que l'interaction elle-même ; un système ou un processus ordonné de l'interaction ; enfin un environnement dans lequel opère le système et avec lequel ont lieu les échanges systématiques* ».

Par ailleurs, que l'interlocuteur soit immédiatement physiquement présent ou non, l'activité de parole implique nécessairement une adaptation à son auditoire correspondant au *recipient design principle*. Ce concept implique que « *tout au long de son travail de production l'émetteur tient compte projectivement de l'interprétation qu'il suppose que l'auditeur va faire de ses propos* » (Kerbrat-Orecchioni, 2005 : 16). En développant cette notion de *recipient design*, Sacks Schegloff et Jefferson (1974) font référence aux multiples ressources, visibles dans le tour de parole d'un locuteur, qui témoignent d'une orientation manifeste vers les co-participants. Ce procédé s'inscrit dans la sélection des unités lexicales et thématiques, dans la manière d'ordonner les séquences, et

également dans les obligations et alternatives retenues pour ouvrir et clore une interaction (Sacks *et al.*, 1974 : 727). L'ensemble de ces procédés interactionnels permet aux participants d'assurer une coordination continue durant l'échange en s'orientant manifestement vers leurs partenaires, projetant une action coordonnée de leur part. Ces procédures locales, « *bien qu'hétérogènes et mobilisées en des niveaux très divers de la structure interactionnelle, participent dans leur ensemble d'un procédé d'organisation générale de l'échange* » (Denouël, 2008 : 107). Le principe de *recipient design* permet aux interactants de structurer leurs ressources linguistiques multimodales de manière à créer un foyer d'attention conversationnel commun, construire et contrôler conjointement le cours de l'interaction, garantir l'intelligibilité des éléments qui leur semblent pertinents et préserver la stabilité du lien interactionnel (*Ibid.*).

Le lien interactionnel apparaît d'autant plus fragile dès lors qu'il ne se fonde pas sur des échanges en face-à-face physique immédiat mais à distance géographique *via* des technologies numériques. Proposant de faire émerger un champ d'« analyse du discours numérique », Paveau introduit la notion de « *technologie discursive* », à savoir « *l'ensemble des processus de mise en discours de la langue dans un environnement technologique* » (2015²). Elle précise qu'il s'agit d'un « *dispositif au sein duquel la production langagière et discursive est intrinsèquement liée à des outils technologiques en ligne ou hors ligne (ordinateurs, téléphones, tablettes, logiciels, applications, sites, blogs, réseaux, plateformes, etc.)*. *La technologie discursive implique une nature composite des productions langagières* » (Paveau, 2015). Cette nature composite tient au fait que les productions des locuteurs sont constituées d'un assemblage entre du langagier et du technique formant un tout hybride. Émergent alors des formes technolangagières : du technomot (mot cliquable) au technogenre de discours (genres de discours natifs du web et relevant du composite discursif) en passant par le technosigne (segment iconique ou verbo-iconique cliquable permettant la diffusion et le partage de technodiscours, l'expression d'affect, la demande d'affiliation,

² Définitions issues du *Dictionnaire d'analyse du discours numérique* en cours de production en ligne par Marie-Anne Paveau depuis 2015.

etc.) (*Ibid.*). La co-construction du langagier et du technique se trouve constitutive des interactions et discours « *natifs du web* » (*Ibid.*).

Une analyse des interactions numériques logocentrée paraît alors inenvisageable. Dès lors, « *décrire la conversation en ligne* » revient à décrire « *la frontière entre nouvelles pratiques et structures normatives, et l'appropriation par les acteurs humains à la fois des outils et des pratiques discursives ou sémiotiques qu'ils induisent* » (Develotte *et al.*, 2011 : 19). Il apparaît nécessaire de renouveler les analyses traditionnelles logocentrées, et d'identifier et adapter les méthodes d'analyse des interactions en intégrant leur multimodalité – voco-posturo-mimo-gestualité – et leur plurisémiotité – notamment le graphisme, l'audio, la vidéo. Des conditions de possibilité de l'analyse des interactions numériques sont ainsi énoncées par Develotte *et al.* (2011). En premier lieu les cadres d'analyse doivent souligner « *l'interrelation entre les composantes du discours et la matérialité de l'environnement* » (notamment les affordances communicatives). En deuxième lieu, les cadres doivent se fonder sur une théorie du discours n'attribuant pas les effets de sens à la seule linguistique mais intégrant également un rôle dans la *semiosis* aux conditions de production, distribution et *design* de l'environnement numérique. Enfin, les cadres doivent, par une théorisation des actes des usagers, rendre compte des « *valeurs et représentations à travers lesquels ils vivent l'espace numérique* » (Develotte *et al.*, 2011 : 25). Il s'agit d'appréhender de manière multimodale et plurisémiotique les « *expériences écraniques* » des usagers tant hors écran que par écran.

2.2. L'action dans l'interaction

Quelle que soit la définition retenue, la notion d'interaction implique littéralement l'action réciproque de ses participants. Aussi, au cours de leurs interactions avec autrui, les sujets produisent-ils des actions multimodales dont ils sont tenus pour responsables. Cette capacité d'action des sujets sur leur environnement, sur les objets et sur Autrui, ainsi que la perception de cette faculté par le sujet relève de la notion d'agentivité (ou *agency*) (Butler, 2002).

L'agentivité (nos actes, nos pensées, nos désirs sont nôtres et nous sommes relativement conscients de les causer et de les contrôler) associée à la résonance (capacité automatique, non consciente qui nous pousse à faire résonner en nous les émotions d'autrui) et à l'empathie (processus actif permettant de comprendre la cause de l'état émotionnel d'autrui et lui témoigner reconnaissance et compréhension de ses émotions) constituent les trois dimensions des interactions qui autorisent les interactants à construire des liens sociaux et qui garantissent la cohésion du groupe, aussi bien que l'autonomie de chacun vis-à-vis des autres (Nadel & Decety, 2006). Les individus sont en effet liés les uns aux autres par résonance et par empathie : par résonance, ils reflètent automatiquement les attitudes et les mimiques des autres, et par empathie ils ressentent ce qu'autrui éprouve, ce qui permet par exemple de lui venir en aide. En revanche, l'agentivité garantit à chacun qu'il est maître de ses actes, évitant ainsi toute confusion entre soi et autrui.

Il apparaît, au niveau cognitif, qu'un ensemble de régions cérébrales est activé à la fois lorsque le sujet exécute une action et lorsqu'il observe cette action produite par autrui. Ce sont les neurones dits « miroirs » qui induisent cette résonance motrice. Des expérimentations révèlent néanmoins que la résonance motrice s'active uniquement lorsque le modèle observé est un être humain et non un robot (*Ibid.*). Ces neurones distinguent en effet les agents biologiques des objets ; ils ne s'activent que lorsque l'agent est un être vivant. Par ailleurs une région cérébrale joue un rôle pivot dans l'expérience de l'agentivité : elle compare les signaux qui sont issus du soi et ceux issus de l'environnement. Par-là, l'individu est à même de distinguer les conséquences d'une action qu'il a déclenchée de celles liées à un événement extérieur. L'individu est donc conduit à reconnaître cognitivement et intuitivement qu'il est l'auteur et le responsable de ses propres actions – l'agent (*Ibid.*).

Pourtant cette distinction se révèle délicate à opérer dès lors qu'une médiation entre en jeu entre la production d'une action et sa perception par autrui. En effet au cours d'une interaction par écran, le locuteur produit physiquement des activités langagières multimodales qui sont en partie retransmises par l'outil à l'interlocuteur. Les éléments apparaissant à l'écran de l'interlocuteur peuvent être définis comme des

« indices » de l'action, au sens de Peirce (1903). Ce sémiologue distingue en effet plusieurs rapports que le signe entretient avec son objet (et qui ne s'excluent pas nécessairement) :

- L'indice : signe qui fait référence à l'objet qu'il dénote. L'indice est réellement affecté par cet objet, il en est le signe immédiat. L'indice est une expression directe de l'objet manifesté. Il a alors nécessairement certaines qualités en commun avec cet objet (empreinte de pas, fumée, action du vent sur une girouette).
- L'icône : signe qui fait référence à l'objet qu'il dénote par les caractères qu'il possède, par une similarité qualitative ou ressemblance. L'icône ressemble à l'objet et est utilisé comme le signe.
- Le symbole : signe qui se réfère à l'objet qu'il dénote par une loi, une association d'idées. Le symbole ne représente pas l'objet. Le lien entre le symbole et son objet tient de la connaissance par l'interprétant de la règle qui le régit.

En se référant à la typologie Peircienne, les productions verbales et posturo-mimo-gestuelles apparaissant à l'écran des interactants semblent alors être de l'ordre de l'indice en ce que ces éléments numériques sont induits par une activité physique de l'utilisateur. La restitution de ces éléments se réalise conjointement par l'outil et l'utilisateur, et est subordonnée aux affordances. Ces dernières « *entrent en jeu au cours d'une activité instrumentée et se définissent comme l'ensemble des possibilités et contraintes de l'environnement, qui donnent aux agents différentes options pour agir* » (Lamy, 2010 : 3). La notion d'affordance ne se comprend donc que comme une relation de réciprocité entre les acteurs et l'environnement. Gibson, dont les recherches sur la perception visuelle de l'animal sont à l'origine du concept d'affordance, souligne que :

« Il est important de noter que les affordances de l'environnement sont objectives, réelles et physiques, contrairement aux valeurs et significations, que l'on suppose fréquemment subjectives, phénoménales et mentales. À vrai dire une affordance n'est ni une propriété objective ni une propriété subjective ; ou si l'on préfère elle est les deux... L'affordance tient à la fois de l'environnement et de l'observateur. » (Gibson, 1979 : 129)

Hutchby, quant à lui, adapte cette notion aux technologies numériques et parle d'« affordances communicatives » du média, *i.e.* « *les multiples possibilités actionnelles que l'artefact s'avère capable d'ouvrir à l'utilisateur* » (2001 : 123). La technologie se comprend alors comme « *un ensemble d'affordances qui se dévoilent dans et par les efforts que déploient les acteurs pour interagir avec l'artefact* » (*Ibid.* : 146). Ainsi l'objet technique est à considérer comme « *un partenaire agissant d'une relation autant habilitante que contraignante* » dans la mesure où il agit comme « *une interface entre le projet d'action qui a été déposé en lui et le sujet actif dans ses usages de ces interfaces techniques* » (Voirol, 2013 : 149). De ce fait, Voirol introduit une nouvelle dimension à la notion d'intersubjectivité en mettant en exergue la manière dont l'utilisateur interagit avec un dispositif qui agit avec lui comme un partenaire de l'interaction. L'utilisateur doit alors faire preuve d'inventivité et de créativité dans son interaction avec le dispositif.

Les sujets et leurs artefacts communicationnels recourent alors tant au langage naturel que non naturel comme outil de médiation et d'intercession en mobilisant l'ensemble des ressources sémiotiques à leur disposition pour agir les uns sur les autres (De Fornel, 2013). Aussi la notion d'agentivité dans le domaine de l'anthropologie nous apporte-t-elle un éclairage précieux sur la distinction de l'attribution de la responsabilité de l'action entre sujet et objet. L'anthropologie opère en effet un décentrement du sujet humain intentionnel au profit d'une multiplicité d'agents, qu'ils soient humains ou non humains (*Ibid.*). À partir notamment d'études de pratiques rituelles au sein d'ethnies amérindiennes et dans une volonté de rejet du point de vue objectiviste, l'anthropologie autorise une double promotion de l'agentivité en ce que « *les êtres vivants sont conçus comme des humains et les objets sont traités avec un statut d'êtres vivants, voire d'êtres humains participant comme tels à la réalité sociale* » (*Ibid.* : 39). Deux orientations entrent ainsi en opposition. La première considère l'agentivité comme stable et s'inscrit dans la perspective ontologique. La seconde définit au contraire l'agentivité comme instable et contextuelle : les animaux, les plantes ou les artefacts ne sont conçus comme des personnes qu'au sein d'activités spécifiques (de nature rituelle, relationnelle, etc.). Pour que ces entités deviennent des agents, il doit être opéré « *un travail de perception et de catégorisation réalisé par des êtres humains au sein d'activités*

quotidiennes ou rituelles » (De Fornel, 2013 : 39). De là, il nous apparaît que cette seconde approche de la notion d'agentivité, comme artefactuelle et contextualisée, interroge la responsabilité des actions exercées en interaction par écran notamment le rôle joué par la *tekhné* dans les actions de communication.

3. Corporéité de l'action

3.1. L'interaction comme corps à corps

La corporéité dans l'interaction sociale peut être appréhendée à partir d'une approche phénoménologique. En effet, l'origine étymologique du terme « phénomène » se trouve dans le verbe grec *φαινεσθαι* : apparaître, se montrer. Or comme le rappelle Vannotti et Gennart (2014), *apparaître* est un mode privilégié de la rencontre et se réalise par la corporéité des sujets qui se rencontrent. L'apparition préfigure l'entrée en contact du sujet et de l'objet ou de l'autre, de leur décisive « *prise de connaissance* » (Heidegger, 1927). Comme l'explique Merleau-Ponty dans sa *Phénoménologie de la perception*, le corps ne peut être considéré comme un objet au monde mais comme moyen de notre communication avec lui. Il se définit comme « *horizon latent de notre expérience, présent sans cesse avant toute pensée déterminante* » (Merleau-Ponty, 1945 : 109). Ainsi, le contour de mon corps constitue une frontière que les relations d'espace ordinaires ne peuvent franchir (Merleau-Ponty, 1945 : 114). L'appréhension par le sujet du corps et de la position de chacun de ses membres passe par un « *schéma corporel* » (*Ibid.*). Ce dernier ne renvoie pas au simple résumé des expériences corporelles, mais bien plus à une « *prise de conscience globale de ma posture dans le monde intersensoriel, une "forme"* » (*Ibid.* : 116). Néanmoins, au-delà d'être une forme le schéma corporel est une dynamique en ce sens que le corps apparaît au sujet comme posture en vue de l'accomplissement d'une certaine tâche actuelle ou possible. Il convient alors de considérer la spatialité du corps différemment de celle des objets extérieurs. Les objets tiennent d'une « *spatialité de position* », tandis que le corps tient d'une « *spatialité de situation* » (*Ibid.*).

Dès lors, l'application d'un déictique du type « ici » à la position du corps ne désigne pas, comme pour un objet extérieur, une position déterminée par rapport à d'autres positions mais l'ancrage du corps actif à un objet, la situation du corps face à ses tâches. Il apparaît que le corps est « *polarisé par ses tâches, il existe vers elles, il se ramasse lui-même pour atteindre son but, et le schéma corporel est finalement une manière d'exprimer que mon corps est au monde* » (Merleau-Ponty, 1945 : 117). Le corps tend à incorporer les actions dans lesquelles il s'engage au point que ces actions participent à la structure du corps. Un accord s'établit entre ce que vise le sujet et ce qui est donné – entre l'intention et l'effectuation. Le corps se révèle alors l'ancrage du sujet au monde. Il en résulte que le corps n'est « *ni dans l'espace ni dans le temps ; il habite l'espace et le temps.* » (Ibid. : 162). Selon Merleau-Ponty, « *en tant que j'ai un corps [...] je suis à l'espace et au temps, mon corps s'applique à eux et les embrasse* » (1945 : 164).

Reste que les études phénoménologiques ont jusqu'alors centré leur intérêt sur la relation entre le Soi et l'Autre et entre Soi et l'objet, délaissant de ce fait la médiation technique des relations interindividuelles. Or, comme le souligne Vial « *tout phénomène est en soi phénoménotechnique. Il y a une technicité transcendantale de l'apparaître, c'est-à-dire une dimension technique a priori dans toute manifestation phénoménale ou "phanie"* » (2013 : 152). En effet, pour qu'un Autre ou un objet apparaisse à un Soi et inversement, une médiation technique est nécessaire quelle qu'en soit la forme. La phénoménotechnique générale impulsée par Bachelard nous rappelle que l'« *ontophanie* » – manifestation de l'être – nécessite une technique tant pour se réaliser que pour être observée.

Et la technique se révèle elle-même porteuse de phénoménalité autant qu'elle permet d'engendrer la phénoménalité, à savoir « *la possibilité d'apparaître ou d'apparaître comme réel* » (Vial, 2013 : 16). L'idée derrière le concept de phénoménotechnique est que les techniques ne consistent pas uniquement en des outils ; elles se révèlent être des structures de la perception. En effet, les systèmes techniques sont qualifiés par Vial de systèmes technoperceptifs en ce qu'ils « *structurent au plan phénoménologique notre expérience du monde*

possible en créant un Umwelt³ perceptif dans lequel baignent toutes nos perceptions » (Ibid. : 31). De ce fait, les perceptions du sujet relèvent de structures techno-transcendantes qui elles-mêmes dépendent de la technique de l'époque (imprimerie, téléphone, internet, etc.) (Ibid.).

3.2. Le corps à l'ère numérique

Avec l'avènement des nouvelles technologies, le corps est perçu par certains comme « *un indigne vestige archéologique amené à disparaître* » (Le Breton, 2001 : 20). Si la problématique de la relation entre le corps et les technologies fascine et inquiète tant, c'est qu'elle renvoie au mythe d'un esprit séparé du corps, d'un être artificiel que le savant pourrait créer, d'une communication parfaite sans malentendu (Flichy, 2009 : 11). Cette crise de sensibilité du corps implique une tension entre deux modes différents de se rapporter au corps ; en l'occurrence la tension entre une vision biomécanique héritée de la modernité et une vision virtuelle du corps issue de la postmodernité (Casilli, 2012 : 6). Or, la crainte de la disparition du corps « *englouti par un écran d'ordinateur est moins un risque réel qu'une réaction paradoxale à son hypertrophie imaginaire, à son omniprésence* », et ce, car notre société exalte le corps en référent ultime (Casilli, 2009 : 3). L'ontophanie numérique affectant globalement la présence phénoménologique des choses mêmes, il ne s'agit pas d'une disparition du corps mais de nouvelles formes d'apparitions corporelles (Vial, 2013 : 239). Les interactions numériques révèlent des traces corporelles monodimensionnelles (pseudonymes, émoticônes, etc.), bidimensionnelles (photos, avatars 2D, profils, etc.) et tridimensionnelles (avatars 3D, personnages virtuels, etc.) qui sont autant de repères cognitifs permettant de dessiner les caractéristiques physiques, les sensibilités et le comportement des interlocuteurs en ligne (Casilli, 2012).

Le corps du sujet en ligne se rend présent par une mise en scène, une forme de monstration naturaliste ou idéalisée, lui permettant d'interagir avec les autres sujets en ligne (Casilli, 2012 : 16). Il s'agit de « customiser », adapter les images, projeter des aspirations ; en

³ La notion d'*umwelt* désigne l'environnement sensoriel propre à une espèce ou un individu.

somme, faire preuve d'un « souci de soi » par la négociation d'une « *présence corporelle assistée par ordinateur* » (*Ibid.* : 20). Casilli fait le lien avec les travaux de Michel Foucault en évoquant une « *technologie du soi* », dans la mesure où « *à travers la projection des traces corporelles se met en place un procédé réalisant dans le corps même le travail de réflexion sur soi, de déchiffrement des désirs et des aspirations personnelles* » (*Ibid.*).

Par ailleurs, en plus de la manifestation numérique *du* corps, une manipulation de l'outil numérique *par* le corps est indispensable. D'une part les outils technologiques ne fonctionnent pas *ex nihilo* – leur utilisation requiert un sujet corporel, d'autre part le corps associé aux interfaces numériques peut permettre une activité sensori-motrice dans un univers artificiel. Aussi, contraire à l'idée d'une disparition postmoderne du corps et d'un dualisme cartésien âme-corps, Frias soutient qu'avec l'avènement du numérique, la corporéité de l'individu est certes transfigurée mais demeure « *un vecteur symbolique central dans les échanges scripturaires sur les tchats comme dans les usages cognitifs de l'ordinateur* » (Frias, 2004 : 2). En effet, si le numérique concerne les « immatériaux » – images et textes virtuels – la réalisation de ceux-ci suppose de passer par le digital, à savoir la matérialité et le toucher : les doigts de la main et la tactilité. C'est pourquoi l'interaction numérique met en présence trois entités : la technique, l'intellectuel et le sensitif ; l'artefact, le conceptuel et le corporel (*Ibid.* : 6). Et l'écran devenu tactile, visuel et sonore en est l'illustration. De là, Frias peut affirmer que « *la porosité de ces interfaces fait s'interpénétrer la chair du sujet et le corps de l'objet en une inextricable hybridation* » (2004 : 6).

En outre, la nécessaire corporéisation numérique est rendu possible par des opérations symboliques : graphies, codes, affects. Comme l'explique Sauvageot « *les jargons, les détournements de caractère, l'alphabet "smiley" tirent le texte vers le geste pour instaurer le contact, créer la proximité, abolir la distance* » (1996 : 216). Ces marqueurs symboliques font office de substituts et de prolongements de « l'individu-substrat » et rendent possible, par-delà et au travers de l'écran, une coprésence à distance à la fois sociale et symbolique (Frias, 2004 : 10). C'est ce que Casilli nomme un « *régime de métaphores corporelles* » (2009 : 2). L'écran est touché par le corps qui lui-même est

inspiré par l'écran et la corporéité postmoderne se vit dans cette dialectique entre la technique et les sens. Le numérique « *impose une nouvelle écoute du corps, il engage à une recherche originale de sensations et de formes de l'apparence* » (Casilli, 2009 : 2). Le corps est finalement au centre exact de la société numérique (*Ibid.*).

4. Ethnographie visuelle des interactions

4.1. La manifestation de symboles visibles

Afin d'appréhender les modalités corporelles d'action au cours des interactions par écran, il nous apparaît nécessaire de recueillir des données empiriques documentant les nouvelles pratiques sociotechniques. Une approche ethnographique visuelle nous permet de rendre compte de l'usage de ressources multimodales et plurisémiotiques par les sujets au cours de leurs interactions. La description ethnographique, à savoir « *l'observation directe des comportements sociaux particuliers à partir d'une relation humaine, la familiarité avec des groupes que l'on cherche à connaître en partageant leur existence* » (Laplantine, 2005), s'est vue progressivement enrichie d'une nouvelle méthodologie fondée sur une approche sensorielle et plus spécifiquement visuelle constituant une nouvelle branche de la discipline. L'ethnographie visuelle trouve son origine dans l'idée selon laquelle la culture et les pratiques sociales se manifestent au travers de *symboles visibles* incarnés dans les gestes, cérémonies, rituels et artefacts situés dans des environnements autant naturels que construits (Ruby, 2000 : 1345). Il est alors considéré que dès lors que les pratiques sociales se rendent visibles, le-a chercheur.e doit être en mesure d'employer des technologies (audio)visuelles afin de les recueillir et en constituer des données pouvant être exploitées, analysées, diffusées (*Ibid.*). S'opère ainsi un glissement d'une ethnographie fondée sur les énoncés verbaux à une ethnographie fondée sur les images et séquences vidéo (MacDougall, 1997 : 292). L'image fixe ou animée constitue un « *élément intrinsèque et non extrinsèque du processus de recherche* » en ethnographie visuelle (Dion, 2007 : 62). L'ethnographie visuelle relève d'une méthodologie heuristique cherchant à « *graphier* » (étudier et

représenter) « *l'ethnos* » (culturalités, pratiques et relations sociales) par des données et supports (audio)visuels.

L'approche visuelle s'est trouvée impulsée par les travaux de Bateson et Mead sur une ethnographie au sein d'un village de Bali en 1942. Au cours de cette recherche de deux ans, Mead interrogeait les participants et prenait des notes de terrain tandis que Bateson photographiait et filmait. L'ouvrage issu de cette enquête ethnographique visuelle « *offre non seulement une vision originale de l'apprentissage de la culture mais il constitue aussi un renouvellement des méthodes de terrain* » (Winkin, 1981). La méthodologie développée par Mead et Bateson ne consiste pas en un recours à des supports visuels « *comme une simple preuve mais comme un véritable matériau de recherche à part entière* » (Dion, 2007 : 64). Aussi Mead a-t-il défini l'anthropologie visuelle comme « *l'étude de l'homme dans ce qu'il donne seulement à voir et qu'on appréhende par des outils d'investigation non verbaux* » (1974). Le support visuel, image fixe ou animée, se révèle être à la fois un outil et un objet de recherche (Dion, 2007).

4.2. Une technologie de négociation des relations sociales

L'ethnographie visuelle ne peut être une copie ou un substitut à l'ethnographie verbale mais doit développer une méthodologie et des objectifs alternatifs bénéficiant à l'anthropologie dans sa globalité (MacDougall, 1997 : 292). En portant l'attention sur des données (audio)visuelles, l'ethnographie visuelle propose de nouvelles modalités d'appréhension des individus, relations sociales, cultures matérielles et de la connaissance ethnographique elle-même (Pink, 2007 : 22). La méthodologie de recherche repose alors sur trois activités principales (Banks & Morphy, 1997) : constituer des données (audio)visuelles (analyser les pratiques sociales en produisant des images), examiner les données (audio)visuelles préexistantes (analyser les images fournissant des connaissances sur la société), collaborer avec les acteurs sociaux dans la production des données (audio)visuelles. La finalité de l'analyse ne consiste pas à traduire des images en mots mais à explorer la relation entre les données (audio)visuelles entre elles et en relation avec les autres formes de connaissance. L'étude de l'agentivité multimodale en interaction par écran, dans une démarche ethnographique visuelle, se

fonde ainsi sur l'exploitation des données audio-visuelles numériques à la fois telles qu'elles se présentent en amont à l'écran des interactants et telles que nous les restituons en aval au moyen des enregistrements et montages des captations écraniques et spatio-corporelles. La production des données audio-visuelles d'interactions transmédiatiques ne se réalise en outre qu'avec la collaboration des participant.e.s à la constitution du corpus de données. Par là il nous est possible de voir ce que les sujets voient, d'entendre ce qu'ils entendent, au cours de leurs interactions par écran et ainsi appréhender les actions entreprises et leurs effets sur la mise en présence de chacun.

Au sein des méthodes visuelles, l'enregistrement vidéo représente, plus qu'un outil de recueil de données, une technologie participant à la négociation des relations sociales et un média par lequel la connaissance ethnographique est produite (Pink, 2007 : 173). De surcroît, les nouvelles technologies numériques et les nouvelles interfaces et réseaux socionumériques introduisent progressivement des études ethnographiques portant sur les pratiques communicationnelles digitales quotidiennes des individus et communautés (*Ibid.* : 197). Émerge alors, au-delà de l'ethnographie visuelle, une ethnographie numérique se voulant délinéarisée, multimodale et plurisémiotique (*Ibid.*).

5. Méthodologie de recherche

5.1. Positionnement scientifique

Notre démarche d'analyse des modalités technico-corporelles d'action en interaction par écran se fonde sur une méthodologie interdisciplinaire au croisement entre ethnographie visuelle (Banks & Morphy, 1997 ; Ruby, 2000 ; Pink, 2007 ; Dion, 2007, etc.), linguistique interactionnelle multimodale (Goffman, 1973 ; Cosnier, 2004 ; Mondada, 2008 ; Kerbrat-Orecchioni, 2010 ; Traverso, 2012, etc.) et analyse phénoménologique (Husserl, 1929 ; Merleau-Ponty, 1945 ; Le Breton, 2001 ; Vial, 2013, etc.). Il s'agit d'analyser des séquences d'interactions sociales à partir de l'expérience des sujets, de leur perception et action corporelles hors et par écran. Nous cherchons à étudier les modalités

d'interaction physico-numérique par une attention portée aux ressources langagières multimodales liées à la corporéité (verbal, geste, mimique, regards, postures) et plurisémiotiques liées aux médias (images, graphismes, vidéos, liens, techno-discursivité) convoquées par les interactants. À cet effet notre analyse audiovisuelle des interactions se fonde sur des enregistrements qui constituent à la fois le support et l'objet d'une analyse intrinsèque (transcription incrustée et enrichissement sémiotique⁴). Il s'agit d'étudier les comportements technico-corporels sur et hors écran des participants à l'interaction sociale.

5.2. Terrain et corpus de recherche

Notre terrain d'étude consiste en une auto-ethnographie d'interactions professionnelles faisant usage simultané ou alterné des robots de téléprésence Beam et Kubi (cf. image 1a) ainsi que d'ordinateurs et tablettes connectés à des logiciels de visio (Skype, Adobe Connect ou Google Hangout) afin de mettre en présence au cours de séminaires réguliers des personnes géographiquement distantes. Cette recherche s'intègre au sein des travaux émergents du Groupe de Recherche sur les Présences Numériques se constituant actuellement au sein du Laboratoire ICAR de Lyon (UMR 5191) et dirigé par Christine Develotte à partir du séminaire mensuel de recherche IMPEC (Interactions Multimodales Par ECran). Ce groupe de recherche repose sur la collaboration de treize chercheur.e.s (quatre doctorant.e.s, trois jeunes docteur.e.s, six enseignants-chercheur.e.s) issu.e.s de divers laboratoires de recherche. Notre axe de recherche au sein de ce groupe repose principalement sur l'étude des métamorphoses de la corporéité dans les interactions par écran et se réalise en collaboration avec Dorothée Furnon (doctorante en Sciences de l'Éducation à l'École Centrale de Lyon).

⁴ L'enrichissement sémiotique consiste à dénoter les activités technico-corporelles des participant.e.s par l'incrustation de signes sur le document audio-visuel. Il participe de la démarche ethnographique visuelle traitant l'image comme élément intrinsèque du processus de recherche.

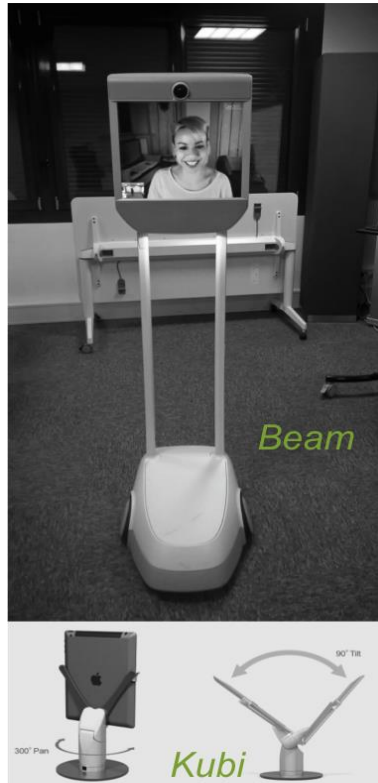


Image 1a : Représentation des robots Beam et Kubi*Image 1b : Terrain de recherche (salle de séminaire ENS-Ifé Lyon)*

Les interactions des participant.e.s à ce séminaire étant hybrides (*in situ* et *ex situ*) et notre étude se voulant multimodale et plurisémiotique, il nous est nécessaire de mettre en place un dispositif d'enregistrement permettant de capturer des données vidéo sur et hors écran. Au cours de cette auto-ethnographie, nous cherchons à recueillir avec nos collègues, l'ensemble des données interactionnelles entre les participant.e.s au séminaire *in situ* et *ex situ*. Il s'agit donc d'enregistrer les écrans des participants pilotant les robots de téléprésence et ceux utilisant des logiciels de visio ainsi que quelques écrans en salle, au moyen de captures dynamiques d'écran *via* le logiciel QuickTime Player sur iOS. Sont par ailleurs filmés au moyen de caméras numériques 360° et de microphones épars la salle de séminaire et le corps des participants pilotant le robot dans leur environnement immédiat. Nous restituons les données recueillies après avoir réalisé un montage avec assemblage des vues, incrustation des transcriptions verbales et

enrichissement sémiotique des vidéos *via* un logiciel de montage vidéo (Final Cut Pro X sur iMac) nous permettant d'analyser simultanément les comportements communicatifs des participants sur et hors écran.

En parallèle de ces interactions, nous cherchons à recueillir des données sur le vécu subjectif des participants à ces interactions transmédiatiques. À cet effet nous menons, toujours en collaboration avec Dorothée Furnon, des entretiens d'explicitation individuels auprès des participants après chaque séance – un participant de chaque média (Beam, Kubi, visio, salle). L'entretien d'explicitation élaboré par Vermersch (1994) consiste en une forme particulière d'entretien qui « s'intéresse au vécu de l'action, et plus précisément aux informations de type procédural, dans le but de reconstituer la structure de l'action » (Martinez, 1997 : 2). Les entretiens d'explicitation en cours de récolte sont également filmés afin d'appréhender la dimension multimodale de la parole sur l'expérience subjective.

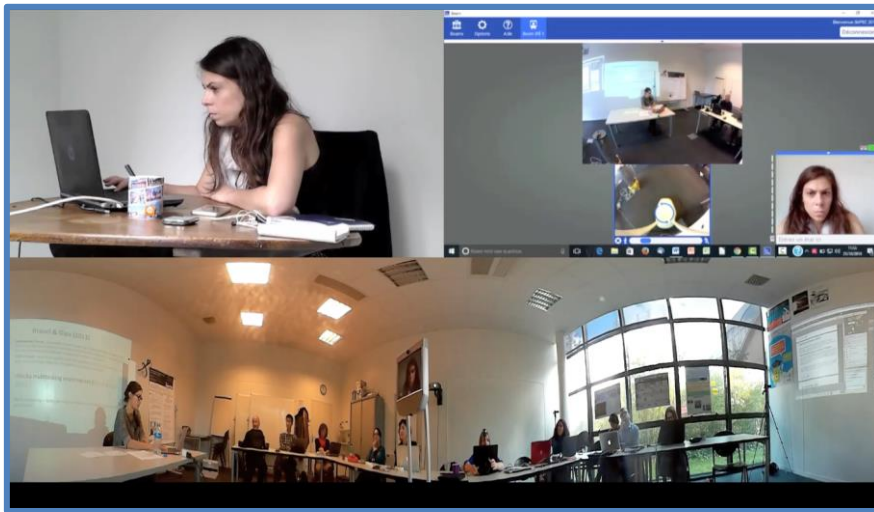


Image 2 : Exemple de montage multiscopes (focus sur l'utilisateur du robot Beam)⁵

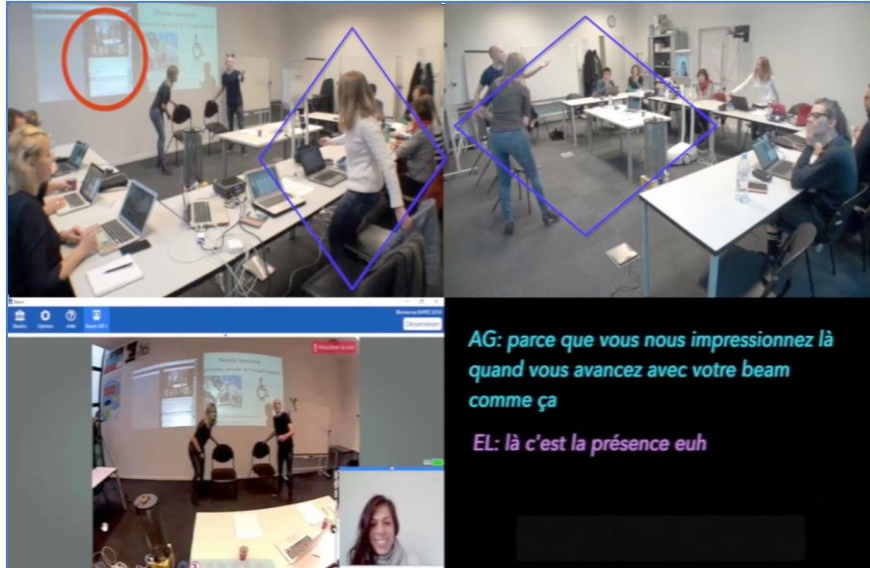


Image 3 : Incrustation des transcriptions et enrichissement sémiotique⁶

6. Analyse empirique des modalités technico-corporelles d'action

Nous proposons de porter notre attention sur une séquence problématique survenant au cours de l'une des interactions du groupe de participants. Il s'agit de l'ouverture de conférence des invités au séminaire; ouverture au cours de laquelle les participants géographiquement distants de la salle de réunion, ne parviennent pas à entendre de façon satisfaisante les conférenciers. Le défaut de qualité sonore, pour être résolu, nécessite à la fois d'être signalé aux autres participants et de faire l'objet d'une remédiation impliquant d'en identifier l'origine.

⁵ Image supérieure gauche : participante à domicile ; Image supérieure droite : capture d'écran de cette participante pilotant le robot Beam ; Image inférieure : vue à 360° de la salle de séminaire.

⁶ Image supérieure gauche : salle de séminaire vue du fond ; Image supérieure droite : salle de séminaire vue de l'avant ; Image inférieure gauche : capture d'écran pilotant le robot Beam ; Cadre inférieur droit : transcriptions verbales.

Nous identifions plusieurs étapes dans le processus de médiation en situation problématique. En premier lieu survient la *démédiation* au cours de laquelle l'incident surgit, la communication n'est plus assurée, le média ne remplit plus son rôle et manque à ses fonctions. En deuxième lieu, se mettent en place des tentatives de *remédiation*; l'incident est pris en charge, la médiation est en cours de réparation et le média peut à cet effet être relancé, remplacé ou complété. Enfin, il y a *immédiation* dès lors que l'incident est résolu, le média assure ses fonctions, la communication est transparente et donne l'illusion d'une communication immédiate. Des actions sont donc menées par les interactants afin de prendre en charge la médiation technique à partir de ressources multimodales et plurisémiotiques.

6.1. Analyse d'une séquence interactionnelle problématique

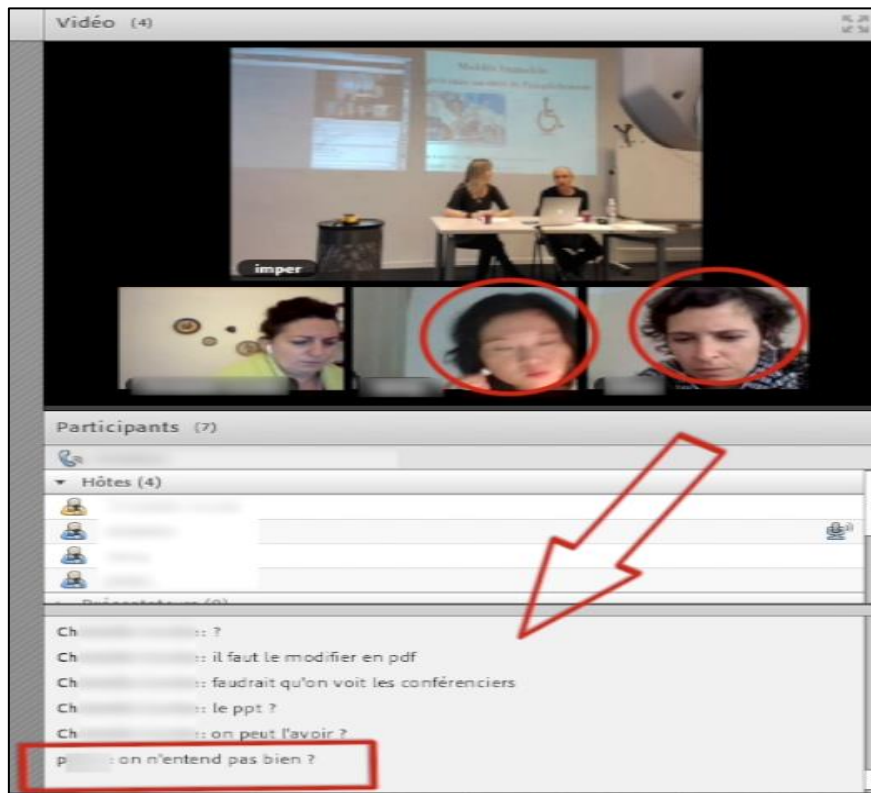
6.1.1. Séquence 1 : Surgissement et signalement de la démediation



Image 4 : Surgissement de la démediation

Au sein de la salle de réunion à Lyon, se trouvent les conférenciers invités qui sont introduits par l'animatrice du séminaire. Dès lors l'un des conférenciers, Axel, prend la parole pour se présenter. Aucune hésitation n'apparaît en salle où chacun des participants peut aisément porter son attention sur la conférence ainsi entamée. En revanche, l'observation des comportements physiques des participantes *ex situ*

nous indique qu'un incident de communication est en cours (Im5). En effet, les expressions faciales des interactantes *via* Adobe Connect⁷ se modifient, les sourcils se froncent, les visages se crispent et se rapprochent de l'écran. L'inconfort interactionnel est ainsi physiquement marqué et l'incident de médiation est rapidement verbalisé sur le tchat écrit de la plateforme Adobe. Le premier message est produit par Prisca : « *On n'entend pas bien ?* ». En parallèle, Christelle fait usage d'un autre média pour signaler l'incident en envoyant un sms (« *On entend mal*») à Morgane, participante en salle en charge de l'ordinateur connecté à Adobe et vidéo-projetant l'image des participantes *ex situ*.



⁷ De gauche à droite sur l'illustration : Christelle, Liping, Prisca.

Image 5 : Signalement de la démediation (tchat)



Image 6 : Signalement de la démediation (sms)

Le signalement de la démediation est donc opéré verbalement mais à l'écrit par tchat et par sms (Im6). Il appartient alors aux autres participants de percevoir le message émis sur la plateforme. C'est alors en effet une participante en salle qui perçoit sur la vidéo-projection au mur le message émis par Prisca (Im7). Nous observons Joséphine changer de posture ; elle se penche et plisse les yeux pour lire le message apparaissant au mur. Elle prend alors la décision de relayer oralement le signalement de démediation « *Ils vous entendent pas là-bas pardon* » en pointant du doigt la vidéo-projection. Son tour de parole étant émis en chevauchement avec la présentation du conférencier, Joséphine émet une excuse et met fin à sa prise de parole, laissant les autres participants prendre en charge l'incident de médiation.

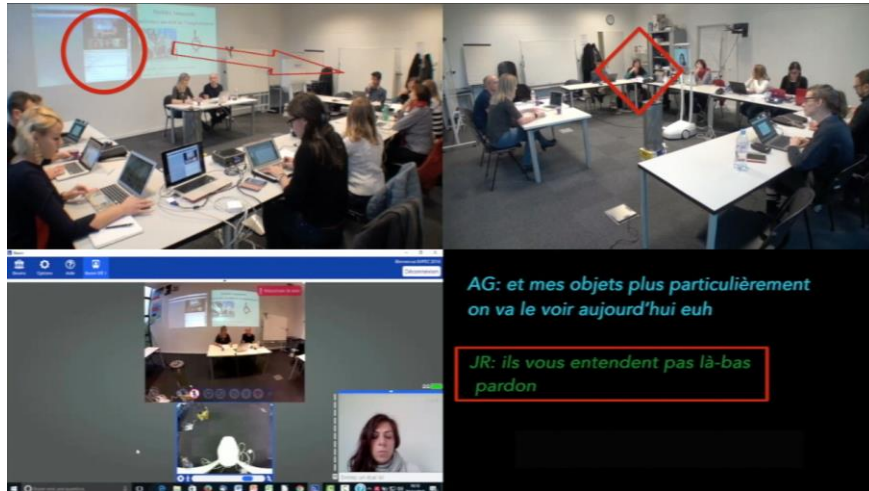


Image 7 : Signalement de médiation relayé

6.1.2. Séquence 2 : Les tentatives de remédiation

La médiation ayant fait l'objet d'un signalement verbal écrit puis oral, l'ensemble des participants à l'interaction se trouve désormais informés. Aussi plusieurs stratégies de remédiation se mettent-elles en œuvre.

1^{ère} tentative de remédiation : modalité orale

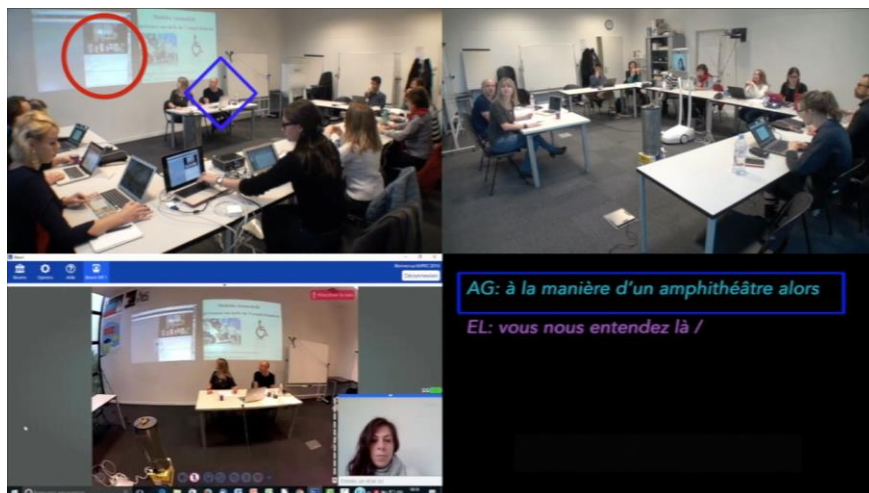
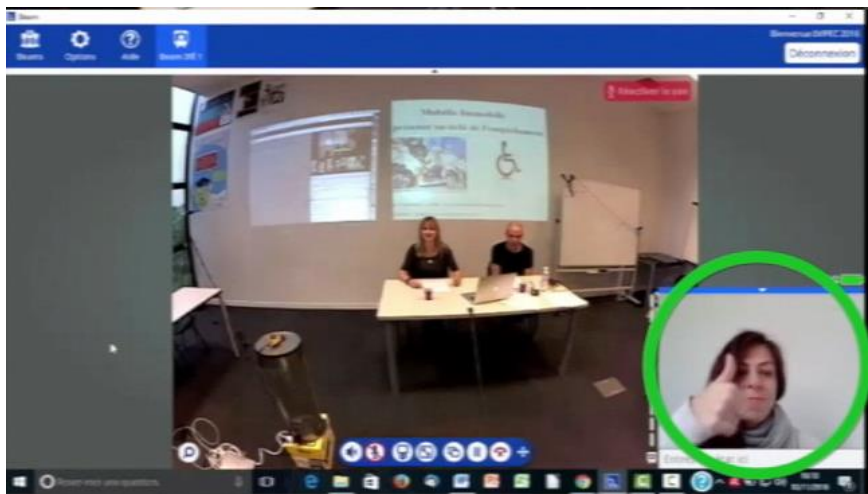


Image 8 : Tentative de remédiation (orale)

En premier lieu, le conférencier locuteur au moment de la démediation se propose de remédier oralement à la situation « *Donc il va falloir que nous parlions euh plus euh à la manière d'un amphithéâtre alors* » en portant la voix plus haut (Im8). Cette première tentative est suivie d'une vérification de remédiation de la part de la seconde conférencière, Evelyne, qui se tourne vers la vidéo-projection et interroge les participantes *via* Adobe « *Vous nous entendez là/* ». N'ayant pas de retour oral ni écrit de leur part, Evelyne conclut « *Eh ben visiblement non* ». Après ce retour négatif sur la communication Adobe, Evelyne se propose d'examiner l'état de la médiation par le robot Beam en questionnant Amélie, son utilisatrice, « *Et euh sur Beam euh c'est bon/* ». Étant positionnée face à son interlocutrice, *via* le robot Beam, Amélie émet une réponse non verbale ; elle lève le pouce en le cadrant à l'écran, confirmant ainsi le bon état de sa médiation (Im9). Le robot Beam fait alors l'objet de commentaires mélioratifs quant à la qualité de fonctionnement de ce type de dispositif.

*Image 9 : Vérification de la médiation robot***2^{ème} tentative de remédiation : modalité technique**

La problématique de la démediation de la communication *via* Adobe n'étant toujours pas résolue, une autre participante émet une nouvelle proposition. Caroline suggère à Morgane de modifier la position de

l'ordinateur connecté à Adobe en salle pour mieux capter le son émis par les conférenciers « *P't être tourne euh l'ordi pour que le micro soit vers eux* » (Im10). Morgane exécute alors le déplacement de l'artefact. Il apparaît ici que l'action est pensée et verbalisée par Caroline mais effectuée par le geste de Morgane. En outre c'est l'artefact qui captera alors différemment le son de la salle retransmis sur la plateforme Adobe. L'action se réalise dans une continuité du sujet à autrui et à l'artefact. L'agentivité se distribue multimodalement (de l'émission verbale de l'une au geste de l'autre et à la captation de l'artefact). Cependant cette tentative de remédiation n'aboutit pas au résultat escompté dans la mesure où il y a confusion dans la source audio captant l'interaction. En effet le micro diffusant sur Adobe n'est pas celui intégré à l'ordinateur mais celui d'une webcam mobile branchée à ce dernier.

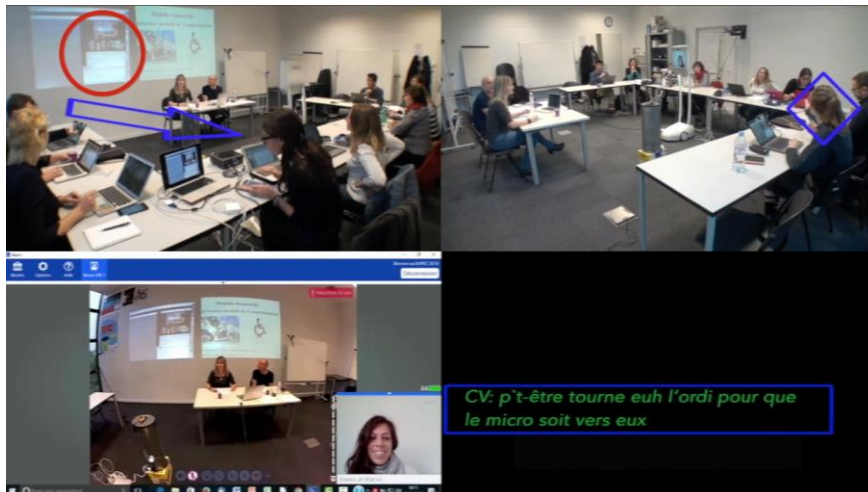


Image 10 : Tentative de remédiation (technique)

3^{ème} tentative de remédiation : modalité spatio-corporelle

Les deux premières tentatives de remédiation, orale et technique, n'ayant abouti, Axel soumet une nouvelle modalité d'action : la restructuration de l'organisation spatiale. Il propose en effet de se rapprocher des participants au séminaire tant pour modifier le rapport au dispositif technique que la nature de l'interaction (Im11). Il opère par là un changement de cadre. Axel explique alors « *Sinon on peut s'*

rapprocher un peu hein on peut essayer d'être un peu plus euh présent de tout l'arsenal euh techno euh pour être moins peut être conférenciers e::t plus en table ronde peut être hein/ on peut essayer ça/». Les deux conférenciers se lèvent alors et saisissent leur table pour la rapprocher des autres participants et du dispositif technique. En modifiant le cadre spatio-corporel de l'interaction, Axel et Evelyn modifient également leur rapport physique à Amélie et son robot de téléprésence. Ils se trouvent excessivement proche du robot, ce qui induit des commentaires de la part d'Evelyne « *Bon du coup on va se rapprocher de Beam alors du coup euh pour vous ça va vraiment être près (.) on va faire du corps à corps avec vous* ». Le lien est fait ici dans la relation corporelle des interactants depuis la disposition technique et spatiale des artefacts au ressenti physique des sujets. Le rapport à la présence et au corps se révèle intrinsèquement liée à l'organisation spatiale de la *tekhnê* qui influe sur les perceptions et affects malgré la distance géographique.

Cette troisième action de remédiation fait l'objet d'une vérification auprès des participantes *via* Adobe qui cette fois valident positivement. Christelle émet sur le tchat le message « *Oui c'est mieux* » avec un visage souriant. S'en suit un message identique de Prisca « *Oui c'est mieux* » et de Liping « *Mieux, maintenant* ». Après cette confirmation verbale écrite et avant de reprendre, Axel recommande à Evelyne de parler plus fort, cette dernière ajoute « *On va essayer de parler distinctement* ».

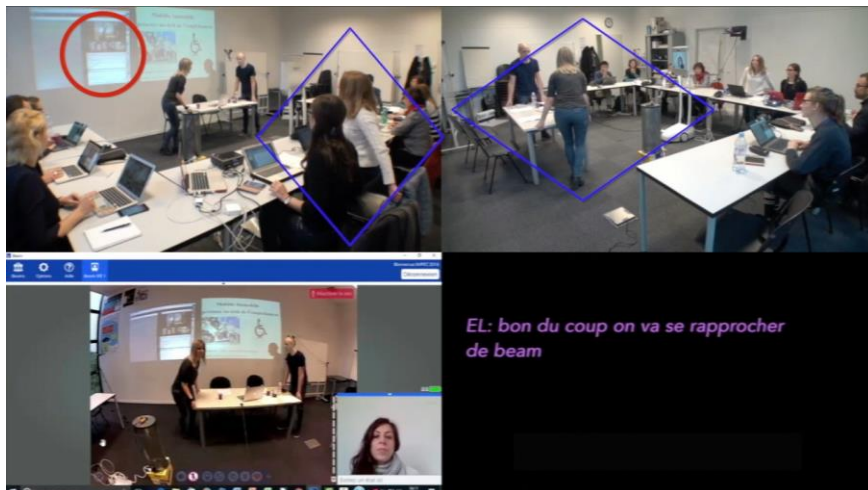


Image 11 : Tentative de remédiation (spatio-corporelle)

4^{ème} remédiation complémentaire : modalité spatio-technique

Malgré la ratification de la remédiation spatio-corporelle opérée par les conférenciers, une remédiation complémentaire est parallèlement opérée. Cette dernière a été initiée par Dorothée dès que les conférenciers se sont levés. En écho à leur geste, Dorothée s'est levée de sa chaise pour participer à la remédiation. Sa proposition est complémentaire au déplacement de la table des conférenciers en ce qu'il s'agit de rapprocher la webcam de ces derniers en la déplaçant de sa table à la leur. L'action entreprise par Dorothée fait l'objet d'une agentivité distribuée et d'une corporéité partagée (Im12). En effet le geste est initié par Dorothée qui se lève pour déplacer le micro et verbalise cette intention à Samira qui se lève à son tour pour prendre le relais, étant plus proche de l'accès à la table des conférenciers. De surcroît le déplacement de la webcam se réalise à plusieurs mains. Samira récupère la webcam et lance un regard à Morgane indiquant uniquement par le regard que le câble reliant la webcam à l'ordinateur est trop court. Aussi Morgane rapproche-t-elle l'ordinateur du bord de la table et Dorothée étend-elle le câble. Suite à quoi le dépôt de la webcam sur la table des conférenciers passe des mains de Samira à celle d'Evelyne qui l'y installe. Ainsi l'action de déplacement de l'artefact ne peut être attribuée à un sujet mais à une pluralité de sujets dont le schéma corporel s'est étendu par intersubjectivité multimodale.

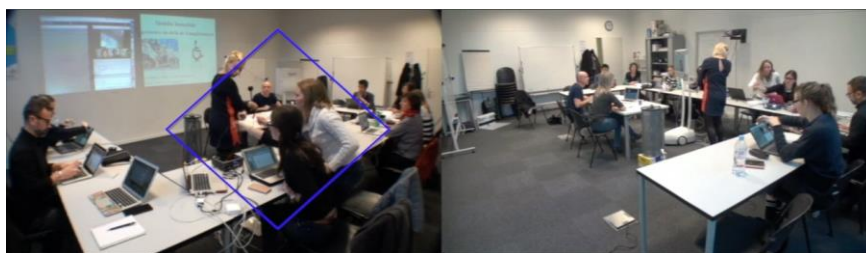


Image 12 : Tentative de remédiation (spatio-technique)

6.1.3. Séquence 3 : L'immédiation et sa ratification

Suite à cette dernière action de remédiation, Samira et Evelyne tourne leur regard vers le tchat Adobe Connect vidéo-projeté sur le mur afin de vérifier l'état de la médiation (Im13). La ratification de la

remédiation est opérée par Christelle qui émet le message « *Super comme ça* » (Im14).

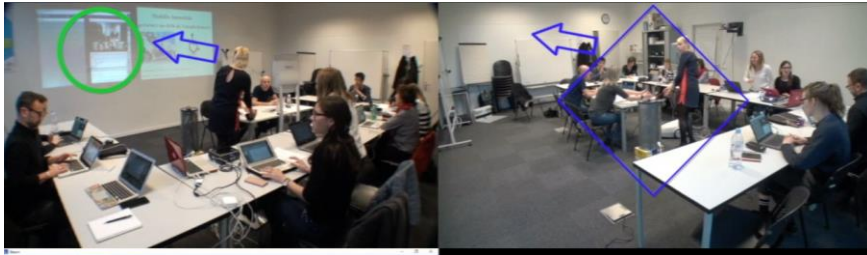


Image 13 : Ratification de l'immédiation (salle)

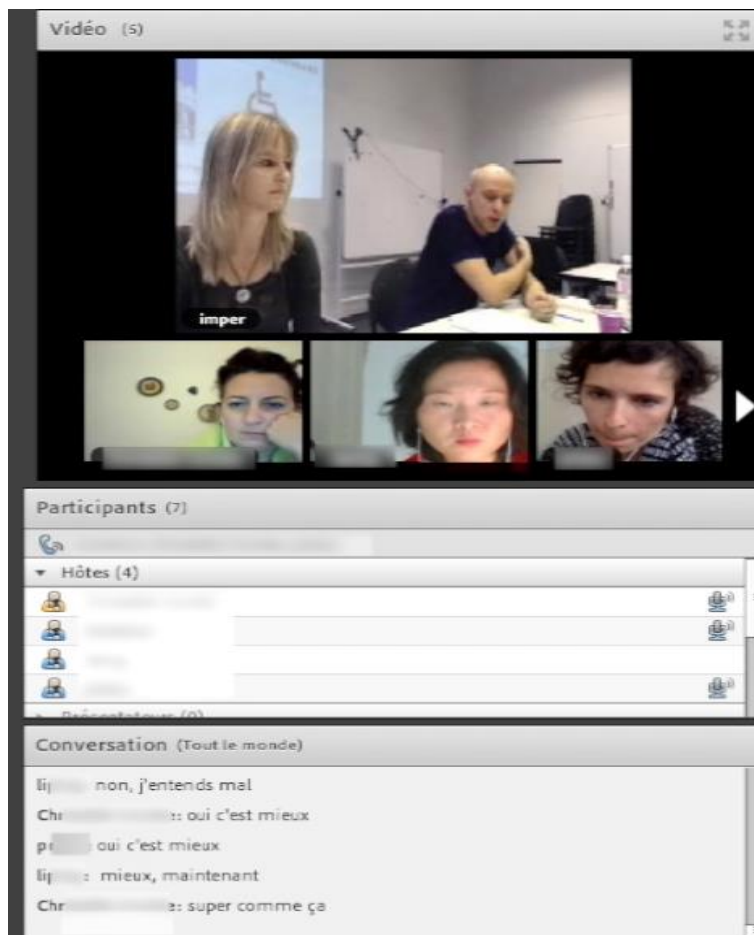


Image 14 : Ratification de l'immédiation (tchat)

Dès lors, l'activité principale de conférence, ou table ronde telle que requalifiée par Axel, reprend son cours. Les participant.e.s ne se préoccupent plus de la médiation technique. L'artefact ne se rend plus visible, il n'est plus au cœur de l'interaction et se rend au contraire transparent. Il y a immédiation en ce sens qu'il y a illusion d'un accès immédiat à l'interaction.

Ainsi, l'incident de médiation, dès son signalement, est pris en charge par les interactants qui mettent en œuvre des actions à partir de ressources multimodales et plurisémiotiques (actions à l'écran, déplacement des artefacts, restructuration spatiale, gestualité, etc.). Ces activités technico-corporelles de remédiation révèlent une agentivité distribuée aux réseaux de sujets et d'artefacts impliqués dans l'interaction. Chaque action ne peut être définitivement attribuée à un seul sujet mais se réalise dans un flux intersubjectif au travers d'une corporéité partagée. Le geste est initié par l'un et poursuivi par l'autre et par là les corps font corps.

Ces métamorphoses de la corporéité et de l'agentivité affectant les sujets, il nous semble nécessaire de porter notre attention sur leur vécu subjectif. Aussi proposons-nous ici de poursuivre cette analyse par des extraits d'entretiens de participants. Cette recherche constituant une auto-ethnographie, les participants réalisent les entretiens entre eux. Le premier extrait ici concerne un entretien semi-directif mené par Amélie auprès de Jean-François. Le second extrait est issu de l'entretien d'explicitation mené par Dorothée et Samira auprès de Joséphine.

6.2. Analyse d'extraits d'entretiens individuels

6.2.1. Entretien 1 « frustration »

Interrogé par Amélie au sujet des frustrations qu'il aurait pu ressentir au cours du séminaire (« *Quels éléments ont été générateurs de frustration/* »), Jean-François aborde la connexion avec les participants *ex situ*. Il indique « *lors du 2^{ème} séminaire j'ai été un peu frustré de pas avoir suffisamment de connexion avec les gens à distance* ». Notons que les participants en salle ont la possibilité de se connecter à la plateforme

Adobe Connect depuis leur ordinateur propre leur évitant de n'être que spectateur de la vidéo-projection au mur de la connexion Adobe depuis l'ordinateur de Morgane. En se connectant lui-même à Adobe chaque participant.e en salle peut interagir directement avec les participantes *ex situ* par modalité écrite ou orale. Jean-François comptait donc sur cette connexion pour se rendre présent et interagir avec les participantes en ligne. Pour autant, le déroulement du séminaire a contrarié ses attentes.

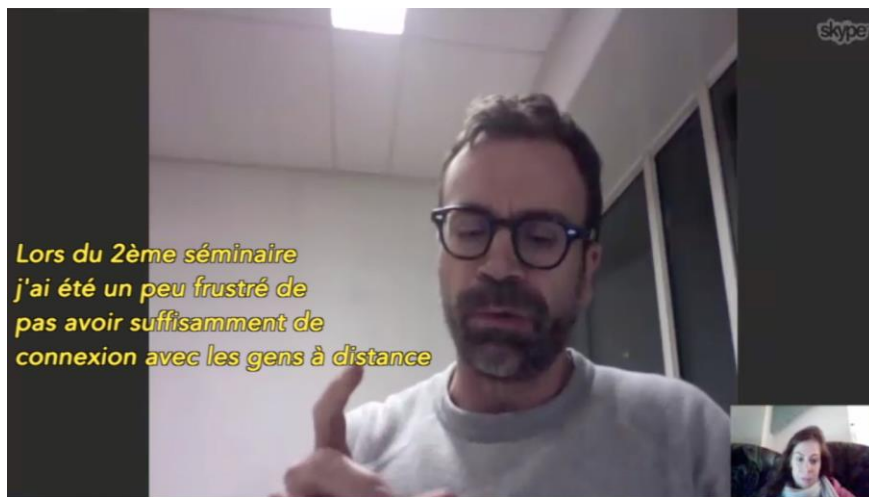


Image 15 : Entretien « frustration »

Jean-François propose alors une auto-analyse de ce défaut de connexion avec les participantes *ex situ* « *Alors moi j'identifie la chose comme ça (.) C'est que je connaissais pas les intervenants extérieurs (.) Je connaissais pas leur travail je connaissais rien (.) Et donc ça m'a demandé plus d'écoute d'attention et ça a monopolisé un moment mon attention* ». Ainsi ce participant en salle met en parallèle l'attention portée à un groupe de participants, les conférenciers, et celle portée à une autre communauté d'interactant.e.s, les participantes en ligne. Ces deux attentions ont été mises en concurrence au cours du séminaire. Et dans la mesure où la méconnaissance de l'une de ces communautés implique de nombreuses lacunes à combler, l'attention a été principalement dirigée vers celle-ci. Le déficit d'attention porte de surcroît autant sur

les interactantes *via* Adobe que sur l'usagère du robot Beam ; « *Le robot c'est à dire toi en fait je l'ai totalement occulté* ».

Cette insuffisance attentionnelle à l'égard des participantes *ex situ* induit une frustration du fait du hiatus existant entre intention et réalisation. Jean-François précise « *C'est comme si je m'étais senti redevable de leur prêter attention* ». Aussi lors d'un précédent séminaire s'attachait-il à se connecter immédiatement à la plateforme Adobe et à y entretenir des interactions notamment par tchat réduisant par là la distance entre participant.e.s en salle et participant.e.s hors salle. Dès lors Jean-François distingue son intention d'action de son action concrète conduisant à la frustration « *Voilà ça c'est la frustration (.) Ne pas avoir la possibilité de prêter attention suffisamment aux gens à distance alors que j'en avais l'intention et que je trouve ça important* ». Notons que Jean-François n'a pas pris part à la séquence interactionnelle problématique analysée plus haut.

Il apparaît alors que les actions de remédiation entreprises au cours de l'interaction par écran se trouvent intrinsèquement liées à l'attention. En effet pour entreprendre des actions de préservation de la communication entre les multiples sujets et artefacts impliqués dans l'interaction, il est nécessaire de percevoir les incidents de médiation et les productions tant verbales (orales et écrites) que non verbales (expressions faciales, postures, etc.) émises par chacun.e. L'action n'est possible que par la perception qui elle-même dépend de l'attention.

6.2.2. Entretien 2 « libération »

A contrario, Joséphine interrogée par Dorothee et Samira au sujet de son choix de positionnement spatio-corporelle à son arrivée dans la salle de séminaire (« *Est-ce qu'on peut revenir sur ce moment-là où tu te demandes est-ce que je me mets à cet endroit là [la même place qu'au séminaire précédent, devant l'entrée (ndla)] comment ça se passe en toi comment tu prends ta décision/*»), met en exergue la simplicité et la liberté. Joséphine distingue en effet plusieurs aires dans la salle, dont notamment un pôle technique plus au fond « *Je crois que c'est plus que, c'est compliqué d'aller vers le fond de la salle parce que y'a tous les câbles euh* » qu'elle compare à l'entrée de la salle « *J'ai l'impression que voilà c'est plus simple de s'asseoir juste en entrant que d'aller vers le fond là-bas* ».

eah ». Deux pôles sont mis en opposition : un pôle technique contraignant et un pôle d'entrée et sortie sans contrainte.



Image 16 : Entretien « libération »

De surcroît Joséphine met en lien sa position spatiale avec ses modalités d'activités. Elle précise ainsi « *C'est juste pour me simplifier la vie en plus moi je travaillais sur papier ce jour là (.) J'étais sur papier donc j'étais dans la simplicité* ». Le choix de position du corps dans l'espace se trouve donc induit par l'organisation artefactuelle technologique et les modalités d'activité interactionnelles et professionnelles. Le choix de se positionner près de la porte implique un détachement vis-à-vis de la *tekhne* et une liberté de mouvement « *Les dispositifs techniques moi je me les représentais dans l'autre coin tandis que là près de la porte on est un peu plus libéré* ». Les capacités d'action sur la médiation sont alors particulièrement réduites dans la mesure où l'accès aux artefacts communicationnels est restreint au profit d'une plus grande « *liberté* ».

6.2.3. Perception de la salle et positionnement des corps

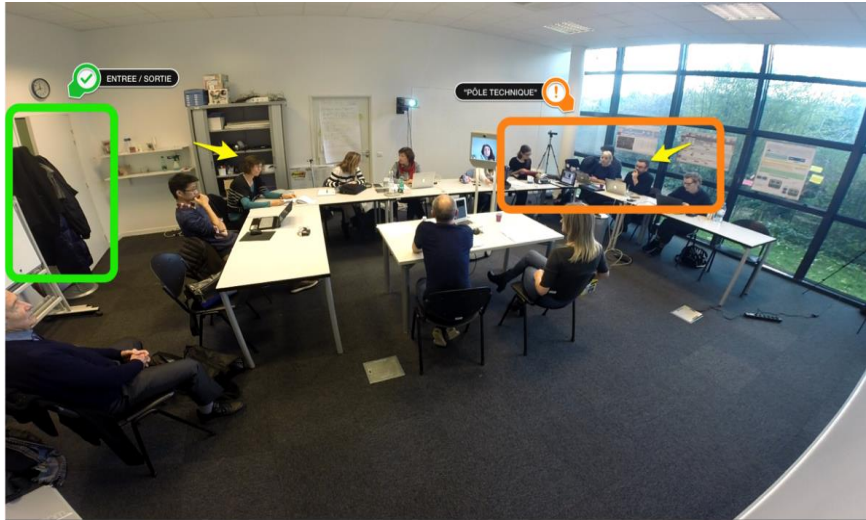


Image 17 : Perception spatio-technique de la salle

Nous observons alors, à partir de la représentation subjective de Joséphine, que la salle comporte deux pôles au sein desquels ces deux interactants ont choisi de se positionner relativement à leurs intentions (Im17). Joséphine à la recherche de simplicité et de liberté interactionnelle prend ses distances vis-à-vis de la *tekhnê* et s'installe près de la porte d'entrée/sortie. À l'inverse, Jean-François en quête de connexion, équipé de son ordinateur, et avec l'intention de porter attention aux participants en ligne, se place au cœur du pôle technique.

Reste que le hiatus entre intention et réalisation conduit ces interactants à aller à l'encontre de leur prévision et agir en opposition aux attentes. Jean-François ne produit aucune action liée à l'incident de médiation et aux tentatives de remédiation. À l'inverse, Joséphine est à l'origine même de la remédiation de par son attention portée à la vidéo-projection de la connexion Adobe et ce qui se jouait sur le tchat. Aussi l'agentivité se révèle-t-elle être contextuelle et située.

7. Conclusion

La présence des sujets en interaction par écran traverse les espaces physiques et les médias communicationnels par les actions opérées multimodalement et plurisémiotiquement par ces sujets, leur corporéité et leurs artefacts. Il s'agit d'exploiter les ressources technico-corporelles à leur disposition afin de préserver la communication malgré les multiples cadres spatio-temporels. Ces ressources technico-corporelles sont mises en œuvre autant pour signaler les incidents de médiation – *démédiation* – que pour tenter de les résoudre – *remédiation* – et parvenir à un état de transparence du média – *immédiation*. À cet effet les sujets entre eux et avec les artefacts se coordonnent et coopèrent tant explicitement qu'implicitement. La gestualité est partagée. Il y a extension du schéma corporel par les artefacts et les autres sujets impliqués dans l'interaction et dans la mise en présence de chacun. Il s'agit d'introduire non seulement l'intersubjectivité dans le schéma corporel qui s'étend à autrui mais au-delà, une forme de transsubjectivité en ce qu'il dépasse et traverse la technique et la distance. Les corps font corps pour interagir dans un réseau technico-corporel complexe.

Le geste transsubjectif et son action s'inscrivent dans une chaîne d'agentivité distribuée. Chacun des interactants a la possibilité d'apporter sa contribution à la préservation de la communication par le moindre geste signifiant. Les affordances communicationnelles et le positionnement des corps et des artefacts dans l'espace impliquent une nécessaire coopération des sujets qui ne peuvent prendre en charge individuellement la médiation physico-numérique complexe et réticulaire. La multimodalité de l'interaction se présente sous la forme d'un flux traversant sujets et artefacts. Le geste se constitue au sein d'une chaîne, il est initié par les uns, poursuivi et achevé par d'autres, qu'ils soient sujets ou *tekhne*. L'agentivité distribuée garantit un champ d'action élargi.

Il y a alors extension actionnelle et perceptive. Il apparaît en effet que l'action est dépendante de la perception et de l'attention. Il est indispensable de percevoir et être perçu pour agir autant qu'il est nécessaire d'agir pour percevoir et être perçu au cours de ces

interactions par écran. Cette coordination de la perception et de l'action intrinsèquement liées, constitue une même opération de « *percepaction* » (Roquet, 2002), l'idée d'une sortie de soi pour percevoir et être perçu. Il s'agit par ailleurs pour Godard (1994), dans son analyse du mouvement, de « *considérer la perception comme un geste* », dans le sens d'un mouvement, une action à portée signifiante (1994 : 68).

De surcroît l'agentivité apparaît formellement liée à la structuration spatio-corporelle de l'interaction. L'organisation spatiale du réseau d'artefacts et de sujets impliqués dans l'interaction favorise ou contraint la capacité d'action des sujets, autant que la nature des activités en cours. Peut alors être mise en œuvre une configuration ou une reconfiguration multimodale et plurisémiotique collaborative et transsubjective – traversant sujet et *tekhnê* – du réseau. La spatialité de l'interaction par écran se révèle nécessairement dynamique et incorporée. Peut alors émerger un hiatus entre intention et réalisation agentive en ce que le contexte interactionnel configure les capacités d'action.

Aussi l'analyse des interactions multimodales par écran ne peut-elle se satisfaire d'une analyse logocentrée et peut au contraire tirer bénéfice d'une approche ethnographique visuelle et interdisciplinaire permettant de mettre en lumière sa co-construction langagière (linguistique interactionnelle), phénoménale (phénoménologie de la manifestation de soi) et incarnée (par la corporéité et les artefacts). La présence par écran émerge de l'expression plurisémiotique, multimodale et sensorielle des sujets et implique la chair et ses extensions. Les sujets sont donc des êtres sensibles qui se co-construisent dans l'interaction sociale en s'équipant de technologies leur permettant de dépasser la distance physique et se manifester dans des configurations spatio-corporelles multiples et réticulaires. Les sujets se rendent ainsi présents les uns aux autres par agentivité technique, incorporée, spatialisée et transsubjective. Il s'agit pour eux de se rendre perceptibles dans un corps à corps hybride qui mêle sujets et objets, étend le schéma corporel et déroule l'agentivité technico-corporelle.

Aussi les technologies numériques ne sonnent-elles pas le glas de la corporéité mais en souligne au contraire la nature fondamentalement

expressive et extensive, dès lors que les sujets collaborent dans leur mise en visibilité mutuelle et intègrent ces technologies externes comme outil et non substitut de présence.

Bibliographie

- Banks Marcus, Morphy Howard (1997). *Rethinking visual anthropology*. London, Yale University Press.
- Bernard Michel (1993). « Sens et fiction, ou les effets étranges de trois chiasmes sensoriels ». *Nouvelles de Danse*, n° 17, pp. 56-64.
- Butler Judith (2002). *La Vie psychique du pouvoir. L'Assujettissement en théories*. Paris, Leo Scheer.
- Casilli Antonio (2009). « Culture numérique : l'adieu au corps n'a jamais eu lieu ». *Esprit*, n° 353, pp. 151-153.
- Casilli Antonio (2012). « Être présent en ligne ». *Idées économiques et sociales*, n° 169, pp. 16-29.
- Cosnier Jacques (2004). « Le corps et l'interaction (empathie et analyseur corporel) ». *Texte de communication Société française de Psychologie*, Paris, 8-9 octobre 2004.
- De Fornel Michel (2013). « Pour une approche contextuelle et dynamique de l'agentivité ». *Ateliers d'anthropologie*, n° 39, vol. 2, pp. 1-8.
- Denouël Julie (2008). *Les interactions médiatisées en messagerie instantanée, organisation située des ressources sociotechniques pour une coprésence à distance*. Thèse en sciences du langage, université Paul Valéry Montpellier III.
- Develotte Christine, Kern Richard, Lamy Marie-Noëlle (2011). *Décrire la conversation en ligne, le face à face distanciel*. Lyon, ENS Éditions.
- Dion Delphine (2007). « Les apports de l'anthropologie visuelle à l'étude des comportements de consommation ». *Recherche et Applications en Marketing*, vol. 22 n° 1, pp. 61-78.
- Flichy Patrice (2009). « Le corps dans l'espace numérique ». *Esprit*, n° 353, pp. 163-174.
- Frau-Meigs Divina (2011). *Penser la société de l'écran*. Paris, Presses Sorbonne Nouvelle.
- Frias Anibal (2004). « Esthétique ordinaire et chats : ordinateur, corporéité et expression codifiée des affects ». *Techniques & Culture*, n° 42, pp. 1-22.
- Godard Hubert (1994). « Le geste manquant ». *Revue internationale de Psychanalyse*, n° 5, pp. 63-75.

- Goffman Erwing (1973). *La mise en scène de la vie quotidienne, 1. La présentation de soi*. Paris, Les éditions de minuit.
- Gibson James J. (1979). *The ecological approach to visual perception*. Hillsdale, Lawrence Erlbaum.
- Greco Lucas, Mondada, Lorenza, Renaud, Patrick (2014). *Identités en interaction*, Limoges, Editions Lambert-Lucas.
- Heidegger Martin [1935] (1962). « L'origine de l'œuvre d'art ». *Chemins qui ne mènent nulle part*, trad. Wolfgang Brokmeier. Paris, Gallimard, pp.13-98.
- Heidegger Martin [1927] (1972, 1985). *Être et Temps*. Paris, Gallimard.
- Heidegger Martin [1927] 1985). *Les Problèmes fondamentaux de la phénoménologie*. Paris, Gallimard.
- Husserl Edmund (1929). *Méditations cartésiennes*. Paris, Vrin.
- Hutchby Ian (2001). « Technologies, Texts and Affordances ». *Sociology*, vol. 35 (2), pp. 441-456.
- Joseph Isaac (1998). *Erving Goffman et la microsociologie*. Paris, Presses universitaires de France.
- Kerbrat-Orecchioni Catherine (2010). « Pour une analyse multimodale des interactions orales. L'expression des émotions dans les débats politiques télévisuels ». *Cadernos de Letras da UFF*, n° 40, pp. 17-45.
- Kerbrat-Orecchioni Catherine (2005). *Le discours en interaction*. Paris, Armand Colin.
- Kerbrat-Orecchioni Catherine (1990). *Les interactions verbales, Tome 1*. Paris, Armand Colin.
- Lamy Marie-Noëlle (2010). « Apprentissage des langues médié par ordinateur : discours critique sur l'outil ». *Le français dans le monde*, n° 48, pp. 135-149.
- Le Breton David (2001). « La délivrance du corps. Internet ou le monde sans mal ». *Revue des Sciences Sociales*, n° 28, *nouveaux mondes ?*, pp. 20-26
- Levy Pierre. « Préface ». In Vial Stéphane (2013). *L'être et l'écran, comment le numérique change la perception*. Paris, Presses universitaires de France.
- Macdougall David (1997). « The visual in anthropology ». Banks Marcus & Morphy Howard (1997). *Rethinking visual anthropology*. London, Yale University Press.
- Martinez Claudine (1997). « L'entretien d'explicitation comme instrument de recueil de données ». *Expliciter*, n° 21, pp. 2-7.
- Merleau-Ponty Maurice (1964). *L'œil et l'esprit*. Paris, Gallimard.
- Merleau-Ponty Maurice (1964). *Le visible et l'invisible*. Paris, Gallimard.

- Merleau-Ponty Maurice (1945). *Phénoménologie de la perception*. Paris, Gallimard.
- Mondada Lorenza (2008). « Using video for a sequential and multimodal analysis of social interaction : Videotaping institutional telephone calls ». *Forum: Qualitative Social Research*, vol. 9 (3), pp. 1-35.
- Nadel Jacqueline, Decety Jean. (2006). « Résonance et agentivité ». *Cerveau et Psycho*, n° 13, pp. 50-53.
- Paveau Marie-Anne (2015-). *Dictionnaire d'Analyse du Discours Numérique*. Production en cours en ligne : <http://technodiscours.hypotheses.org/245>.
- Peirce Charles S. [1903] (1960). « Elements of logic ». *Collected Papers*. Cambridge, Harvard University Press.
- Pink Sarah (2007). *Doing visual ethnography*. Londres, Sage Publications.
- Roquet Christine (2002). *La scène amoureuse en danse. Codes, modes et normes de l'intercorporéité dans le duo chorégraphique*. Thèse de Doctorat en Danse, Université Paris 8.
- Ruby Jay (2000). *Picturing culture : exploration of film and anthropology*. Chicago, University of Chicago Press.
- Sacks Harvey, Schegloff Emanuel, Jefferson Gail (1974). « A Simplest Systematics for the Organisation of Turn-Taking for Conversation ». *Language*, n° 50, pp. 696-735.
- Sauvageot Anne (1996). « Art, technologie et recomposition du sensible ». In Borillo Mario & Sauvageot Anne. (dir.), *Les cinq sens de la création. Art, Technologie, Sensorialité*. Seyssel, Champ Vallon, pp. 211-218.
- Traverso Véronique (2012). « Analyses de l'interaction et linguistique : état actuel des recherches en français », *Langue française*, n° 175, pp. 53-73.
- Vermersch Pierre (1992). *L'entretien d'explicitation en formation continue et initiale*. Paris, ESF.
- Vial Stéphane (2013). *L'être et l'écran*, Paris, Presses universitaires de France.
- Vion Robert (1992). *La communication verbale : analyse des interactions*. Paris, Hachette.
- Voirol Olivier (2013). « La lutte pour l'interobjectivation. Remarques sur l'objet et la reconnaissance ». In Ferrarese, E. (dir.), *Qu'est-ce que lutter pour la reconnaissance ?*, Lormont, Bord de l'Eau.