



HAL
open science

De la captation à la diffusion de données de téléprésence : points de vue des participants à distance

Amélie Bouquain, Christelle Combe

► To cite this version:

Amélie Bouquain, Christelle Combe. De la captation à la diffusion de données de téléprésence : points de vue des participants à distance. Méthodes visuelles en sciences humaines et sociales : de la captation à la diffusion des résultats. Données, méthodes, résultats, valorisation., Dec 2021, La Laguna, Espagne. hal-03481156

HAL Id: hal-03481156

<https://hal.science/hal-03481156>

Submitted on 15 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

De la captation à la diffusion de données de téléprésence : points de vue des participants à distance

Amélie BOUQUAIN, Rectorat de la région académique Normandie, amelie.bouquain@gmail.com, +33698124026

Christelle COMBE, LPL UMR 7309, christelle.combe@univ-amu.fr, +33 695583036

Introduction

Cet article s'inscrit dans le projet de recherche « Présences numériques » (Develotte et al., 2021) qui a étudié sur une année universitaire un séminaire de recherche doctoral hybride polyartéfacté, c'est-à-dire un séminaire dont une partie des participants se trouvaient en présence et l'autre à distance, la communication se faisant notamment via trois dispositifs distincts de téléprésence : une plateforme interactive multimodale (Adobe Connect) et deux robots de téléprésence (Kubi, et Beam). Il s'ancre dans le champ de recherche de la Communication Multimodale Interactive (*interactive multimodal communication*) et celui plus émergent de la Communication Médiée par robot (*robot-mediated communication*) (Herring, 2015, 2016). Notre démarche méthodologique relève principalement de l'ethnographie visuelle numérique (Ibnelkaïd, 2018) et de l'éthologie réflexive visuelle (Develotte, Domanchin et Ibnelkaïd, à paraître 2021), qui se veut être « une approche transdisciplinaire plaçant la multimodalité, le vécu subjectif et la sensorialité au cœur de l'analyse des interactions, de l'observation à la diffusion scientifiques » (Develotte et Ibnelkaïd, 2021). Dans cet article, nous nous intéressons plus spécifiquement aux données des dispositifs de téléprésence, de leur captation à leur diffusion au sein d'une édition numérique augmentée, en nous focalisant sur les participants à distance et les différents artefacts auxquels ils ont été confrontés. C'est donc en tant que chercheuses mais aussi en tant que participantes à distance que nous allons rendre compte de cette expérience. Après avoir rapidement rappelé les trois artefacts de téléprésence, nous présenterons le recueil des données, effectué du point de vue des participants à distance, de ces trois dispositifs particuliers. Nous nous attarderons d'une part sur les spécificités du recueil propres à chacun des dispositifs (enregistrement vidéo interne, recueil textuel, capture d'écran dynamique, captation par caméra externe). Nous montrerons ensuite comment le montage multiscopes élaboré à partir de l'ensemble des différentes données recueillies a permis une analyse multimodale rendant compte des différents points de vue. Nous présenterons enfin un nouvel objet de diffusion scientifique : la « capsule vidéo » pour l'édition numérique augmentée (Lascar et Dujour, 2021) et son évolution. Nous verrons comment à partir des données recueillies, nous appuyant sur le langage cinématographique, et après différentes expérimentations, nous sommes parvenues à retranscrire l'expérience subjective des dispositifs de téléprésence et son analyse multimodale.

Trois dispositifs distincts de téléprésence

Comme nous l'avons rapidement mentionné précédemment, trois dispositifs distincts de téléprésence ont été utilisés au cours de ce séminaire doctoral polyartéfacté :

- ❖ une plateforme interactive multimodale Adobe Connect intégrant principalement des espaces de visioconférence, de conversation écrite, de partage de document et de prise de notes ;
- ❖ un robot de téléprésence Kubi, à savoir une tablette Ipad sur un socle articulé ;
- ❖ et un robot de téléprésence Beam, à savoir un robot à taille humaine mobile sur deux roues.

La plateforme interactive multimodale Adobe Connect

Adobe Connect est une plateforme interactive multimodale qui intègre un certain nombre de modules dont une visioconférence, une conversation écrite, un partage de document et une prise de notes parmi ceux qui ont été principalement utilisés pendant le séminaire. Elle offre également la possibilité d'effectuer des enregistrements qui sont stockés sur la plateforme et qui peuvent être ensuite téléchargés au format vidéo. La conversation écrite peut aussi être envoyée par courriel et ainsi sauvegardée en fichier texte.

Le robot de téléprésence Kubi

Le robot de téléprésence Kubi se compose d'une tablette Ipad, posée sur un socle articulé et connectée au logiciel Skype. L'utilisateur a une interface qui lui permet de positionner les autres interlocuteurs pour mieux pouvoir ensuite s'orienter sur son socle au cours de la visioconférence.

Le robot de téléprésence Beam

Le robot de téléprésence Beam est un robot à taille humaine mobile sur deux roues. Il est dirigé par l'utilisateur à partir d'une interface accessible sur un écran d'ordinateur. L'interface lui permet de le piloter en prenant en compte l'espace et les objets présents dans le lieu où il se trouve.

Pour chacun de ces dispositifs, le recueil des données s'est fait de manière adaptée en tenant compte d'une part des besoins pour les futures analyses mais aussi des possibilités qu'offraient les artefacts et enfin en fonction de ce que les participants à distance identifiaient au fur et à mesure de leur usage des artefacts.

Typologie des données recueillies du point de vue de la distance

Enregistrement vidéo interne

Le recueil des données de la plateforme Adobe consiste en un enregistrement vidéo interne à la plateforme de chaque séminaire et une sauvegarde écrite du *chat*. L'enregistrement interne permet ainsi de rendre compte précisément du vécu des participants à distance. Les participants à distance sont connectés à la salle Adobe Connect, sur le lieu du séminaire (salle du Lipen), certains participants en présentiel s'y connectent également. Toutefois c'est en visionnant effectivement l'enregistrement que l'ensemble des participants ont pu vivre l'expérience du point des participants à distance et que certains ajustements ont pu être effectués.

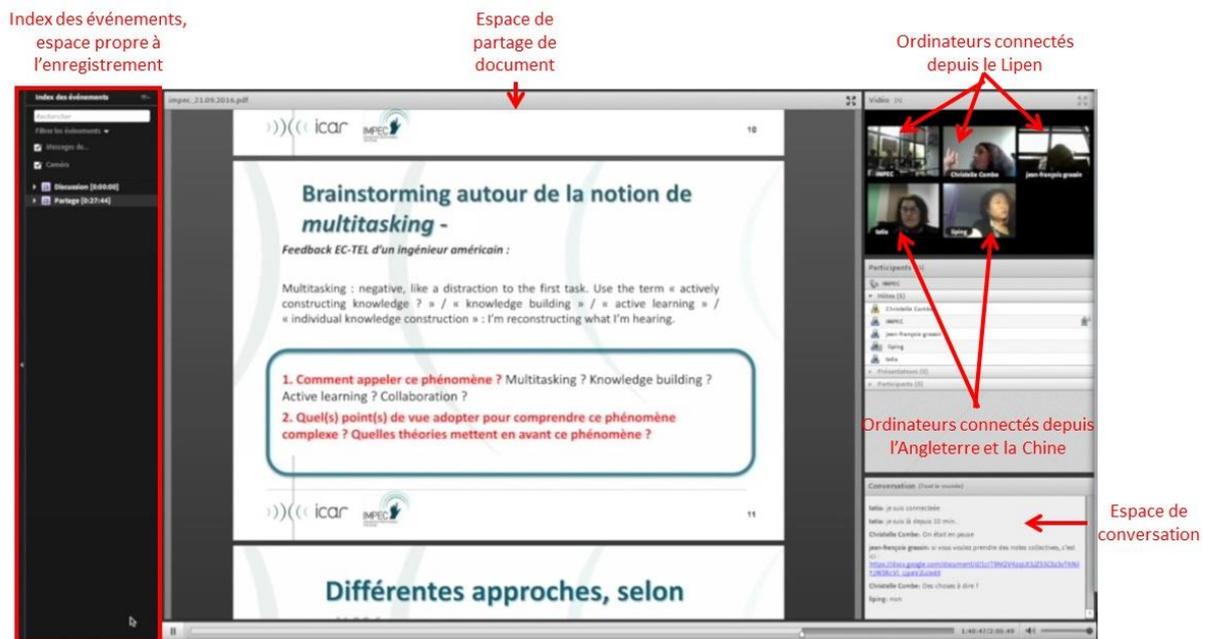


Figure 1 : Capture d'écran d'un enregistrement vidéo sur Adobe Connect

Recueil textuel

Le *chat* dans Adobe Connect peut être sauvegardé indépendamment au format textuel. Le recueil textuel de la conversation permet une recherche par mot clé sans avoir à visionner l'intégralité de l'enregistrement vidéo.

tatia:je suis là depuis 10 min...

Christelle Combe:On était en pause

jean-françois grassin:si vous voulez prendre des notes collectives, c'est ici : https://docs.google.com/document/d/1cIT8NQVKzqUt3JZS5CSz3vTKNiIYJW3RcVl_IJpaV2U/ed
it

Christelle Combe:Des choses à dire ?

liping:non

Figure 2 : Recueil textuel d'un extrait de la conversation dans Adobe Connect

Capture d'écran dynamique

Le recueil des données des robots de téléprésence Kubi et Beam consiste en des captures d'écran dynamiques de l'ordinateur de l'utilisateur comme nous pouvons le voir dans la figure ci-dessous.

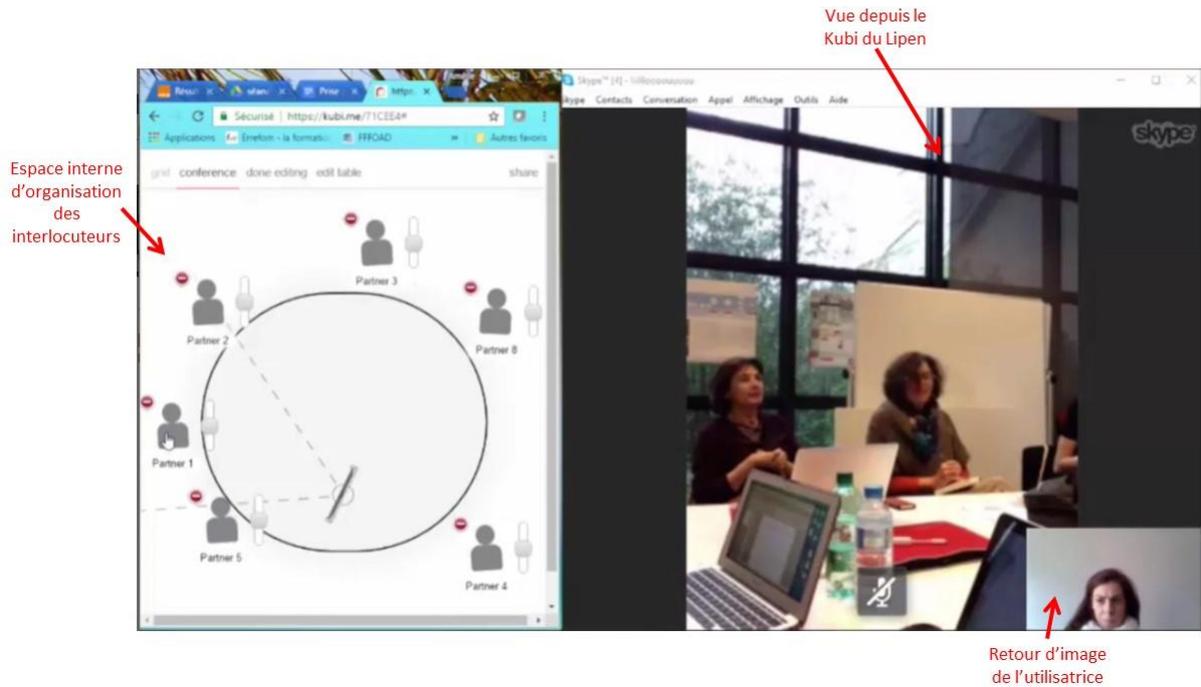


Figure 3 : Capture d'écran dynamique de l'ordinateur de l'utilisatrice du Kubi

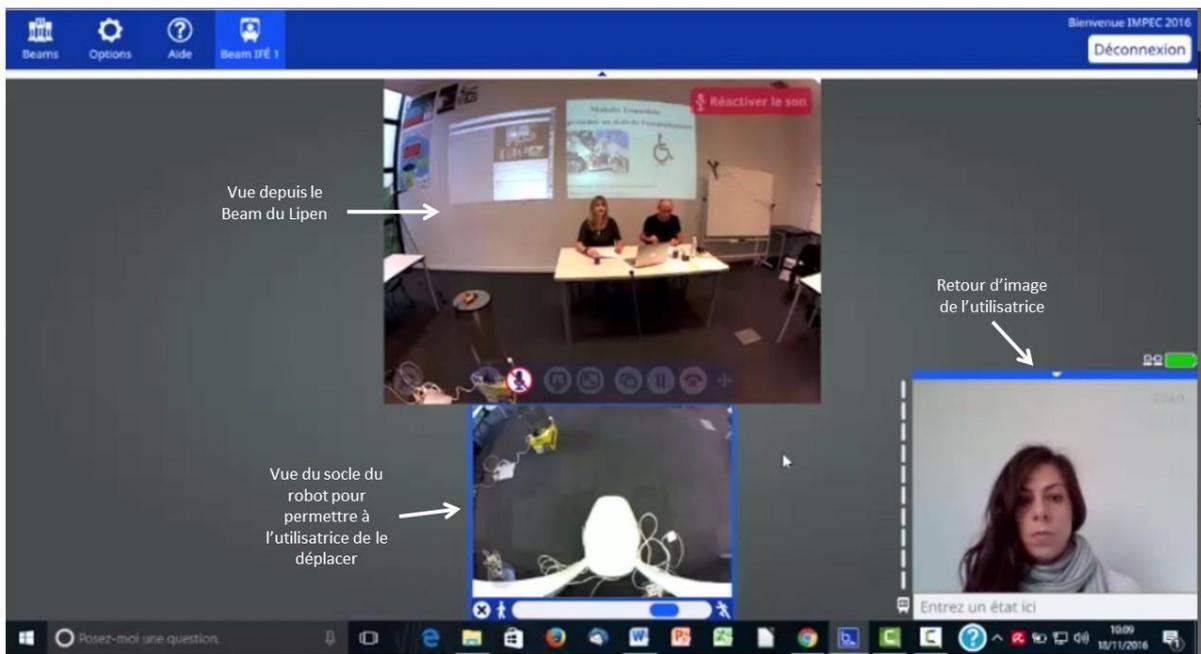


Figure 4 : Capture d'écran dynamique de l'interface du robot Beam

La capture d'écran dynamique de l'ordinateur de l'utilisatrice à distance permet ainsi de suivre l'activité selon son point de vue.

Captation par caméra externe

Un autre type d'enregistrement a été réalisé par les participantes à distance utilisant un robot de téléprésence (Kubi ou Beam) à partir d'une caméra sur pied : à savoir un enregistrement vidéo externe d'elles-mêmes dans leur environnement personnel et notamment de l'écran de son ordinateur. On observe dans les deux exemples ci-dessous que les deux utilisatrices ont opté pour un positionnement différent.

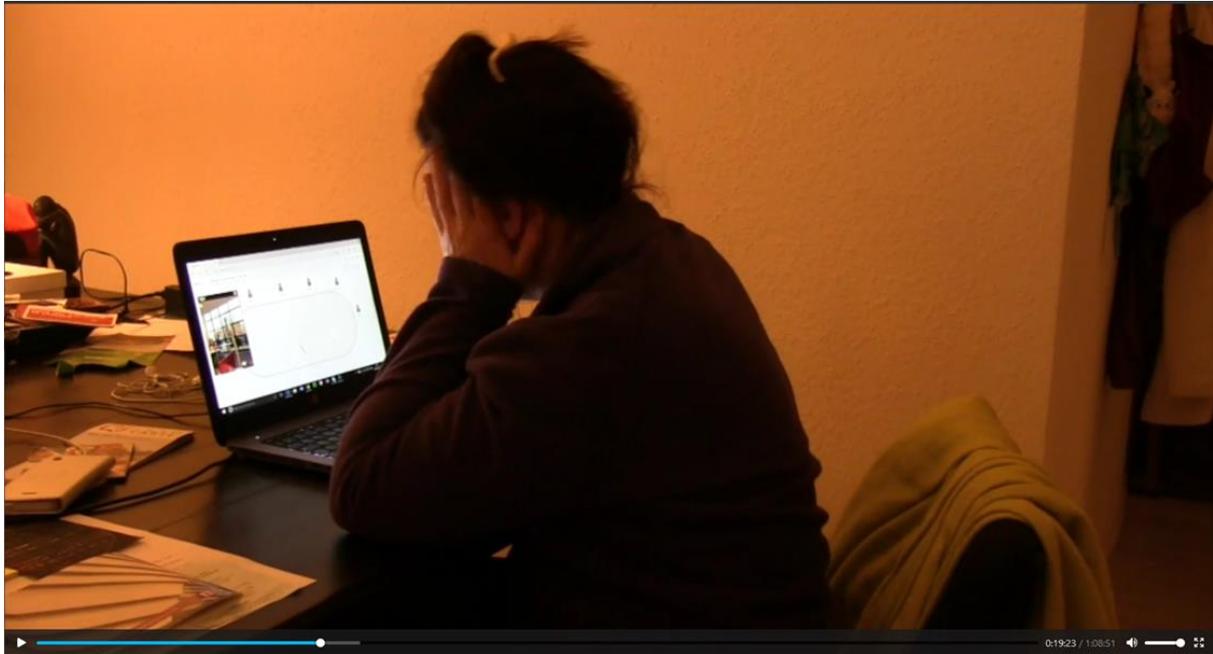


Figure 5 : Enregistrement vidéo externe de l'utilisatrice du Kubi chez elle

La première a positionné la caméra légèrement derrière elle. Ce recueil permet non seulement d'observer l'écran de l'ordinateur de l'utilisatrice - on remarque notamment l'espace d'organisation des interlocuteurs et la vue depuis le Kubi de la salle du séminaire - mais aussi l'utilisatrice face à son écran dans son environnement personnel.

La seconde a positionné la caméra face à elle.



Figure 6 : Enregistrement vidéo externe de l'utilisatrice du robot Beam chez elle

L'objectif de ce recueil de données à travers ce type d'enregistrement est d'observer le comportement de la participante à distance dans le cadre du suivi du séminaire à distance dans son propre environnement : pilotage de l'artefact, attention, participation, implication, interactions avec les autres, réactions en lien avec un événement, impacts de l'environnement personnel, etc.

Ce type d'enregistrement est à mettre en parallèle des enregistrements réalisés de l'environnement où se déroulait le séminaire de recherche avec les participants en présentiel.

Captations vidéo par caméra externe de la salle du séminaire

Les enregistrements vidéo réalisés sur le lieu du séminaire permettent quant à eux d'observer la présence des participants à distance à travers les artefacts qu'ils utilisent et leurs interactions avec les autres participants.

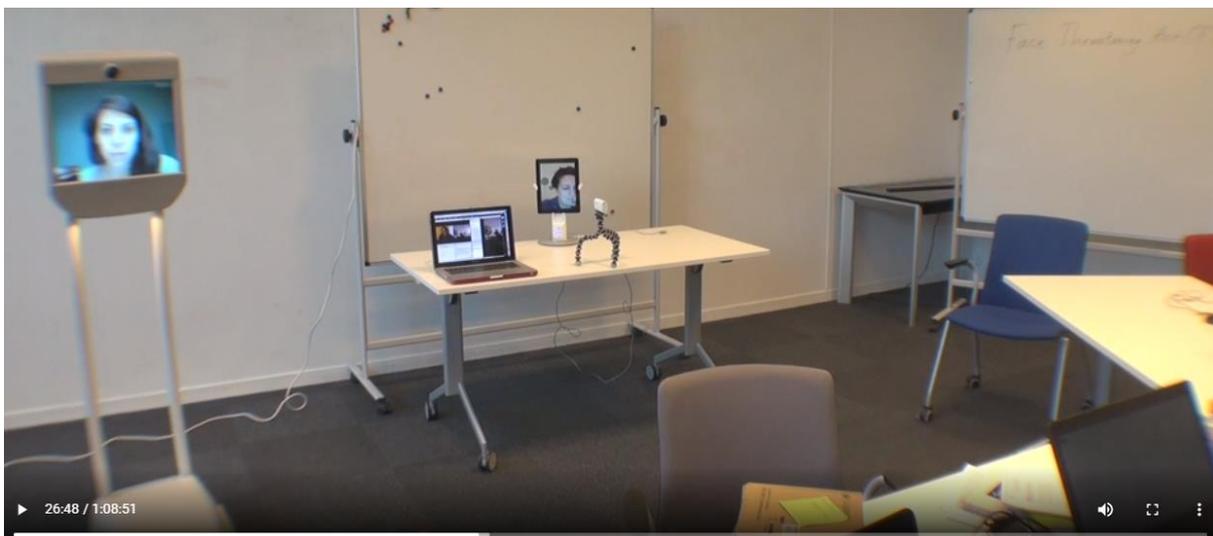


Figure 7 : Enregistrement vidéo de la salle du séminaire, robots Beam et Kubi

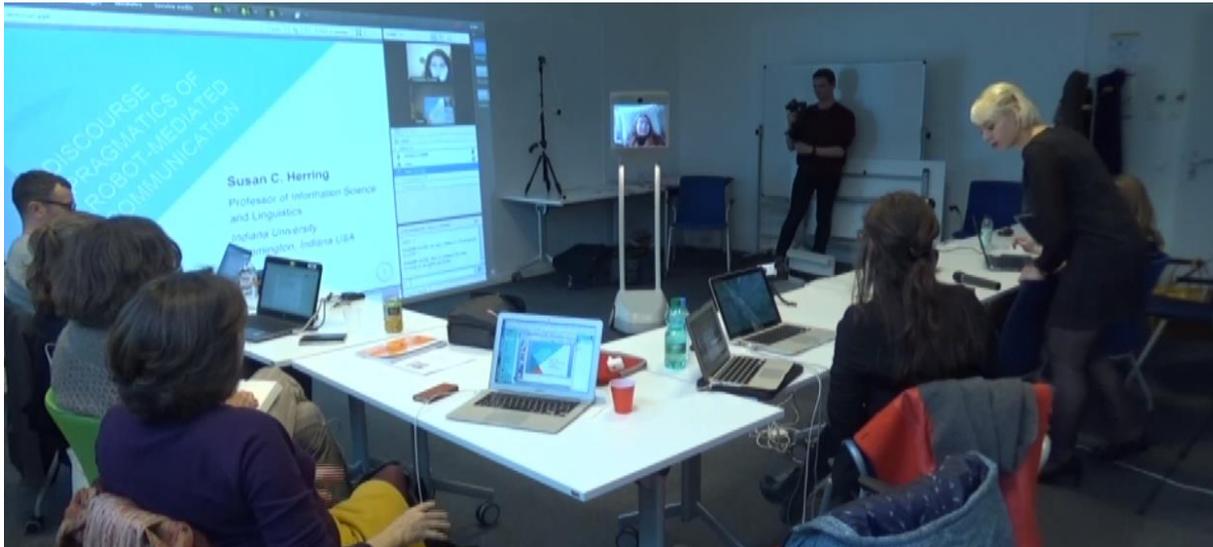


Figure 8 : Enregistrement vidéo de la salle du séminaire, robot Beam et Adobe Connect

Ils permettent également dans le cadre d'un croisement des données entre les différents enregistrements des participants distance (utilisateurs de différents artefacts) : enregistrement interne (Adobe Connect), captures d'écran dynamiques et enregistrements extérieurs.

Le montage multiscopes

Ensuite, à partir des différentes données vidéos recueillies, un montage multiscopes a été effectué par la cellule Corpus Complexe du Laboratoire ICAR (voir "Étudier une expérience de formation doctorale hybride : dispositif de captation audiovisuelle et montage" Christine Develotte et Justine Lascar).

Le montage multiscopes est un montage vidéo qui consiste à assembler plusieurs captations vidéo en les synchronisant permettant ainsi de constituer une vidéo avec différentes vues à partir des données vidéo recueillies. Ce type de montage permet ainsi d'observer et analyser des données à partir d'enregistrements réalisés à une même instant T dans différents lieux, environnements et points de vue.

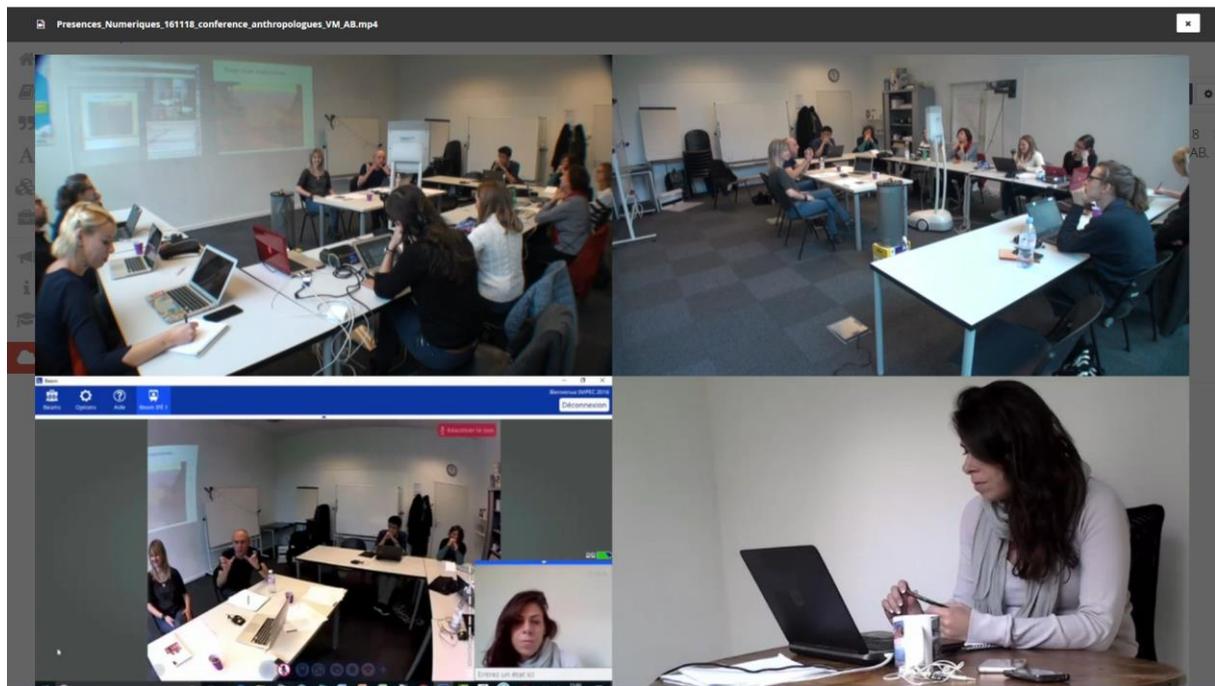


Figure 9 : Montage multiscope des enregistrements vidéo avec un focus sur le robot Beam

Ainsi, pour un artefact donné, les différentes captations vidéo ont permis d'observer :

- l'artefact dans l'environnement présentiel du séminaire, ses mouvements ou déplacements dans le cadre de sa présence à distance,
- la participation au séminaire par l'utilisatrice de l'artefact à travers la capture dynamique de son écran d'ordinateur,
- l'utilisatrice de l'artefact à distance dans son environnement personnel participant au séminaire.

Le montage multiscope des différents enregistrements a permis de mettre en relation et d'analyser de manière multimodale les données recueillies depuis différents points de vue en fonction des objectifs de recherche parcourus.

Retranscrire le point de vue des participants à distance

Par ailleurs, dans un second temps, cette réflexion sur les points de vue ainsi que la perspective d'une édition numérique augmentée (Lascar et Dujour, 2021) nous a conduit à élaborer un nouvel objet de diffusion scientifique : la « capsule vidéo ». Nous allons donc présenter cet objet et son évolution et voir comment à partir des données recueillies, nous appuyant sur le langage cinématographique, et après différentes expérimentations, nous sommes parvenues à retranscrire l'expérience subjective des dispositifs de téléprésence et son analyse multimodale.

La phase de prototypage

Pour la réalisation des capsules vidéo de l'édition numérique augmentée de l'ouvrage *Fabrique de l'interaction parmi les écrans* (Develotte et al., 2021), une phase de prototypage de capsules vidéo a

permis de faire émerger des idées et d'alimenter la réflexion du groupe afin de tendre vers un produit abouti permettant d'illustrer les analyses réalisées à partir du corpus.

Les premiers prototypes réalisés permettent d'identifier les différents procédés et leur pertinence afin de mettre en évidence les éléments de l'analyse (zoom, encadrage, flèches, etc.). Ces prototypes réalisés à partir des montages multiscopes ont également facilité le choix des vues les plus pertinentes pour le montage des capsules vidéo.

Le prototypage est une première étape qui permet rapidement d'envisager des pistes de montage, de visualiser un rendu final, de rechercher des améliorations et d'effectuer des choix éditoriaux. Le prototypage est donc une étape essentielle dans la réalisation de capsules vidéo de données de recherche.

Montage des capsules vidéo

Nous prendrons l'exemple d'une capsule vidéo en particulier réalisée pour expliquer et illustrer un exemple d'actes menaçants pour la face et les procédés réparateurs qui ont suivi (Bouquain, Codreanu et Combe, 2021).

Points de vue

Dans un premier temps, il nous a fallu choisir les extraits que nous allions retenir pour le montage à partir des données vidéos du corpus en croisant notamment les points de vue.

Ainsi dans la capture ci-dessous, on voit que l'on a opté, pour démarrer cette capsule, par un plan qui montre le point de vue de la salle du séminaire. L'audio dans ce plan est excellent et correspond à ce que les participants *in situ* entendent.



Figure 10 : Point de vue depuis la salle du séminaire

Le plan suivant montre l'enregistrement vidéo interne de la plateforme Adobe Connect, le son est inaudible et dévoile bien le point de vue des participants d'Adobe connect.

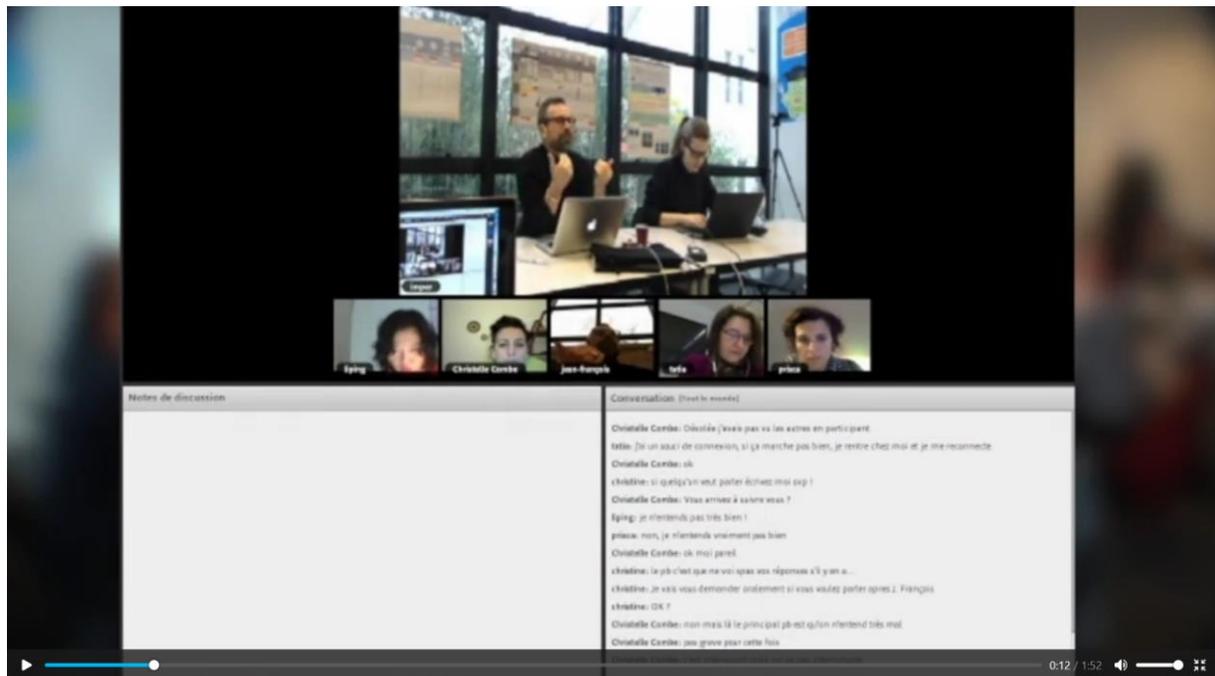


Figure 11 : Point de vue depuis la plateforme Adobe Connect

Mises en exergue et explicitations

Les mises en exergue dans une capsule vidéo peuvent être effectuées par différents procédés : encadrement, encerclement, flèche pointant vers le point d'attention, etc. Quant aux annotations, dans notre cas, notre choix s'est porté sur un système de bandeau dans le bas de la vidéo, elles permettent d'explicitier à l'aide de texte ce que l'on souhaite mettre en évidence avec les différents plans choisis et les mises en exergue.

Le plan suivant met en exergue la partie *chat* de la plateforme et notamment (encadrés en vert) les propos qui montrent que les participants n'entendent pas.

christine: si quelqu'un veut parler écrivez moi svp !

Christelle Combe: Vous arrivez à suivre vous ?

liping: je n'entends pas très bien !

prisca: non, je n'entends vraiment pas bien

Christelle Combe: ok moi pareil

christine: le pb c'est que ne voi spas vos réponses s'il y en a....

christine: Je vais vous demander oralement si vous voulez parler apres J. François

christine: OK ?

Christelle Combe: non mais là le principal pb est qu'on n'entend très mal

Christelle Combe: pas grave pour cette fois

Christelle Combe: c'est intéressant mais on va pas interrompre

Figure 12 : Mise en exergue des propos illustrant l'analyse dans le *chat*

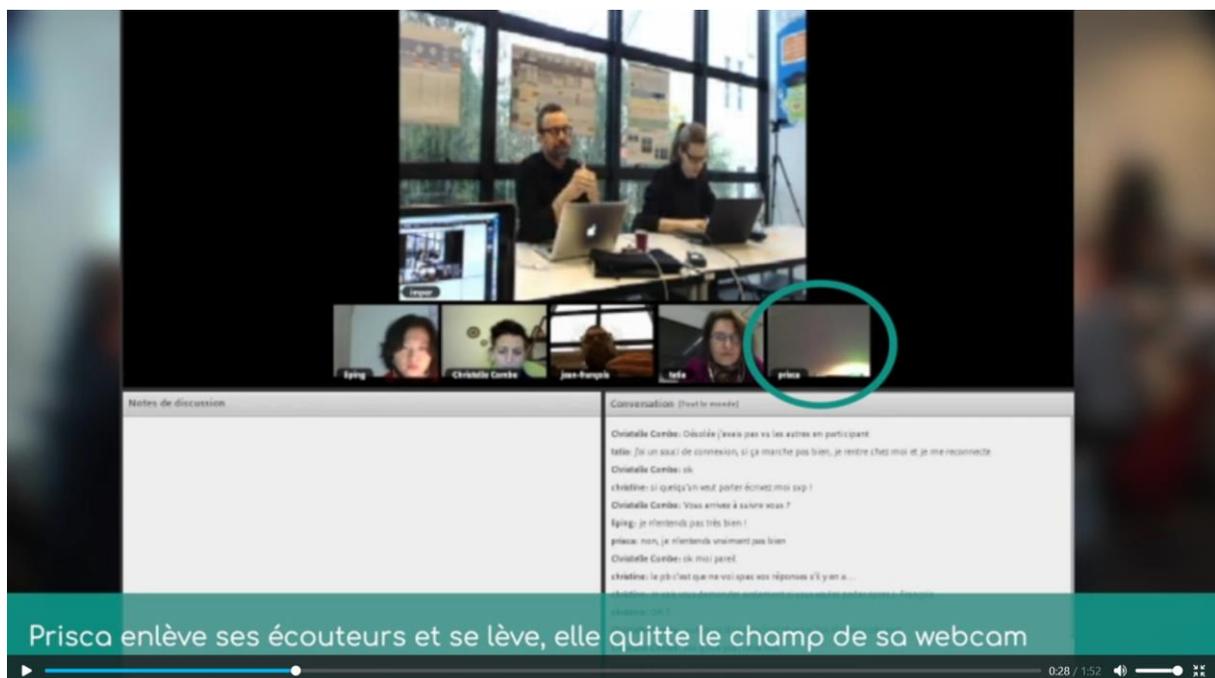


Figure 13 : Mise en exergue et explicitation de l'élément mis en exergue

La vignette du participant à distance qui quitte l'écran de son ordinateur est encerclée pour mettre en évidence son désengagement. Cette mise en exergue par l'encerclement est renforcée par une annotation qui explicite l'action du participant dans le bandeau prévu à cet effet de la capsule vidéo.

Accéléré et sous-titrage

Par ailleurs, nous avons utilisé le procédé cinématographique de l'accélééré pour passer rapidement sur certains contenus verbaux qui n'étaient pas utiles à la démonstration, tout en explicitant l'important (cf. ci-dessus).

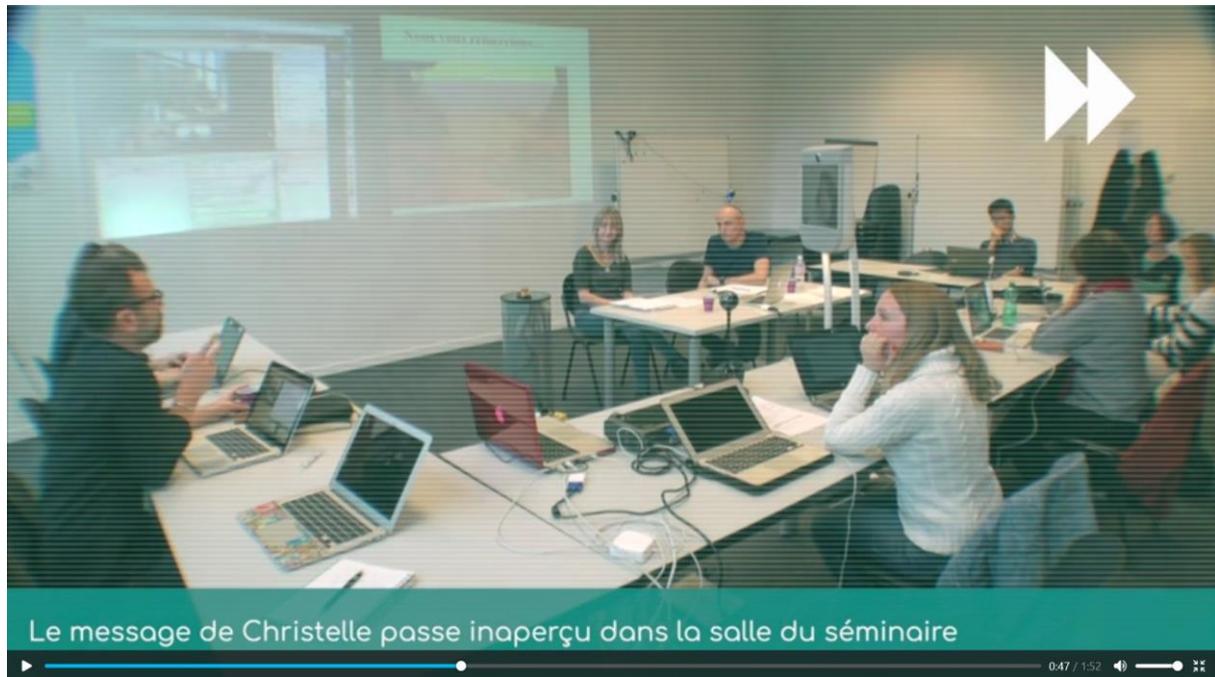


Figure 14 : Accélééré et explicitation

Dans l'exemple ci-dessous, on peut voir que les paroles de Christelle qui s'exprime depuis Adobe Connect sont sous-titrées bien que parfaitement audibles dans la vidéo, cela permet ainsi de mettre en évidence l'acte menaçant pour la face puisque ce sont ces paroles et leur résonance dans la salle du séminaire qui vont être à l'origine de cet acte menaçant.

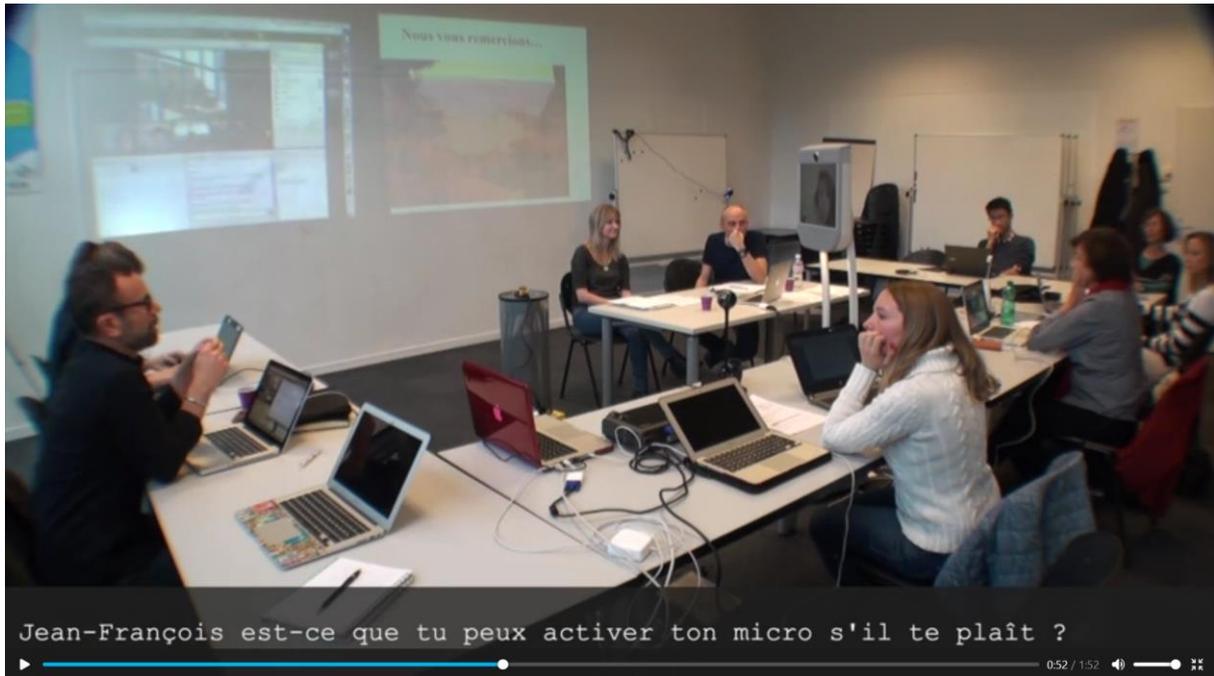


Figure 15 : Sous-titrage des paroles de Christelle dans Adobe Connect

Enfin, dans ce dernier exemple, on voit comment on met en évidence les deux points de vue, celui depuis la salle du séminaire accompagné d'une vignette du point de vue de la plateforme Adobe Connect avec l'adoucisseur (procédé de réparation) extrait du chat en exergue.

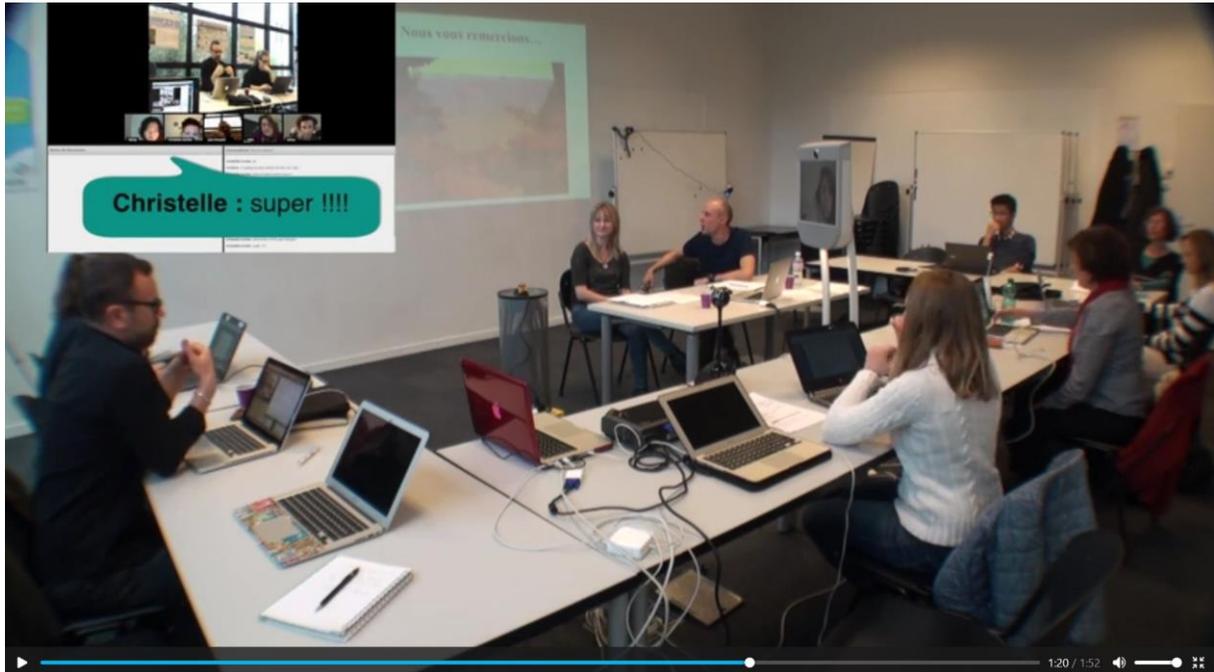


Figure 16 : Double point de vue et mise en exergue du procédé de réparation

Conclusion

Nous avons donc retracé les différentes étapes de traitement de données vidéo de recherche qui ont permis d'aboutir à la conception de capsules vidéo illustrant une analyse scientifique en nous intéressant plus spécifiquement aux données des dispositifs de téléprésence et en nous focalisant sur les participants à distance et les différents artefacts auxquels ils ont été confrontés. Tout d'abord, nous avons décrit la phase de recueil qui a permis d'obtenir des données brutes que nous avons classées selon une certaine typologie de données vidéo : enregistrement vidéo interne, capture d'écran dynamique, captation par caméra externe. Ensuite, nous avons montré comment le croisement de ces données vidéo et la réalisation de montages multisopes a facilité l'analyse des données synchronisées à partir des différents points de vue, notamment celui des participants à distance qui utilisaient un artefact pour suivre le séminaire. Enfin, nous avons montré comment, à partir des données recueillies, la sélection de plans, de vues, mais aussi les choix éditoriaux tels que les annotations, les explicitations, le sous-titrage, les différentes formes de mise en exergue, donnant lieu à une phase de prototypage, ont permis d'aboutir à la conception de capsules vidéo qui illustrent concrètement l'analyse réalisée à partir de données vidéo de recherche.

Références

Amélie Bouquain, Tatiana Codreanu, Christelle Combe (2021). « Ménager les faces par écran : vers de nouvelles règles de politesse ». *Fabrique de l'interaction parmi les écrans* (édition augmentée), Les Ateliers de [sens public], Montréal, <http://ateliers.sens-public.org/>.

Christine Develotte, Amélie Bouquain, Tatiana Codreanu, Christelle Combe, Morgane Domanchin, Mabrouka El Hachani, Dorothée Furnon, Jean-François Grassin, Samira Ibnelkaïd, Justine Lascar, Joséphine Rémon, Caroline Vincent, (2021). *Fabrique de l'interaction parmi les écrans* (édition augmentée), Les Ateliers de [sens public], Montréal, <http://ateliers.sens-public.org/>.

Christine Develotte, Samira Ibnelkaïd, , « Glossaire », in *Fabrique de l'interaction parmi les écrans : formes de présences en recherche et en formation* (édition augmentée), Les Ateliers de [sens public], Montréal, <http://ateliers.sens-public.org/>.

Christine Develotte, Morgane Domanchin, Samira Ibnelkaïd (2021). « Cadrage théorique et méthodologique pour l'éthologie réflexive visuelle », in *Fabrique de l'interaction parmi les écrans : formes de présences en recherche et en formation* (édition augmentée), Les Ateliers de [sens public], Montréal, <http://ateliers.sens-public.org/>.

Susan C. Herring (2016). « Robot-mediated communication ». In R. A. Scott, M. Buchmann, & S. M. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource* (pp. 1–16). Hoboken, NJ: John Wiley & Sons. Prepublication version: <http://ella.slis.indiana.edu/~herring/rmc.pdf>

Susan C. Herring (2015). « New frontiers in interactive multimodal communication ». In A. Georgapoulou & T. Spilloti (Eds.), *The Routledge handbook of language and digital communication* (pp. 398–402). London, England: Routledge.

Samira Ibnelkaïd. (2018) « De l'intérêt d'une ethnographie visuelle de la présence transmédiatique » In Ibanez-Bueno, J. & Chabert, G. & Lamboux-Durand, A. & Wanono, N. (Coords.) *Les méthodes visuelles appliquées à la communication numérique*, Cuadernos Artesanos de Comunicación. <http://www.cuadernosartesanos.org/2017/cac136.pdf>

Justine Lascar, Oriane Dujour (2021). « Enjeux techniques et défis méthodologiques de l'ingénierie de terrain au service de la recherche », Fabrique de l'interaction parmi les écrans : formes de présences en recherche et en formation (édition augmentée), Les Ateliers de [sens public], Montréal, <http://ateliers.sens-public.org/>.