



HAL
open science

Étudier une expérience de formation doctorale hybride : dispositif de captation audiovisuelle et montage

Christine Develotte, Justine Lascar

► To cite this version:

Christine Develotte, Justine Lascar. Étudier une expérience de formation doctorale hybride : dispositif de captation audiovisuelle et montage. Visuals Methods International Conference in Social Sciences 2021, Universidad La Laguna, Dec 2021, Ténérife, Espagne. halshs-03948474

HAL Id: halshs-03948474

<https://shs.hal.science/halshs-03948474v1>

Submitted on 20 Jan 2023

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.

Étudier une expérience de formation doctorale hybride : dispositif de captation audiovisuelle et montage

Christine Develotte et Justine Lascar

Dans ce chapitre c'est en tant que responsable du projet de recherche « Présences numériques » (Develotte et al., 2021) et en tant qu'ingénieure (Lascar et Dujour, in Develotte et al. 2021) que nous allons rendre compte de cette expérience scientifique sous ses dimensions visuelles. Le projet s'est déroulé de 2016 à 2021. Son objectif réside en l'analyse d'un séminaire doctoral hybride ayant eu lieu à la fois en présentiel à Lyon et en distanciel (dans d'autres lieux en France mais aussi en Angleterre et en Chine). Ce séminaire est qualifié de « polyartefacté » car il a utilisé différents artefacts tels que des robots téléprésence (Kubi et Beam), une webcam directionnelle et une plate-forme interactive multimodale (Adobe Connect). L'analyse de cette situation concrète de formation à la recherche a conduit à la conceptualisation des interactions liées aux expériences en contexte hybride et les résultats scientifiques viennent d'être publiés en ligne¹. Nous nous intéresserons ici à deux moments spécifiques de notre travail scientifique : tout d'abord, en amont, la mise en place du dispositif que nous avons choisi pour récupérer des données étudiables par les chercheurs, puis en aval, l'apport technologique privilégié pour la valorisation des vidéos nécessaires à la méthodologie d'éthologie visuelle retenue.

Avant d'en venir plus précisément aux choix techniques et matériels qui ont été faits, nous allons donner quelques précisions sur le séminaire doctoral qu'il s'agissait d'étudier en en donnant les caractéristiques concernant son organisation et ses participants puis nous préciserons notre méthodologie d'éthologie réflexive visuelle.

Le séminaire IMPEC hybride synchrone

Le séminaire mensuel du groupe IMPEC² (Interactions Multimodales Par ECran) avait pour but de fournir un accompagnement scientifique aux doctorants encadrés par Christine Develotte en stimulant leurs réflexions par l'apport de conférenciers invités et en leur offrant la possibilité de présenter l'avancement de leur thèse à intervalles réguliers. La grande majorité des doctorants avaient des sujets de recherche liés à la communication numérique, généralement en contexte éducatif. Au fil des années, les anciens doctorants, devenus docteurs voire enseignants-chercheurs ont souhaité continuer à suivre ces séminaires. Comme la distance géographique n'autorisait pas leur présence physique, l'utilisation de la visioconférence a commencé à se mettre en place par le biais d'un ordinateur placé au milieu de la table autour de laquelle se tenait le séminaire. Ce dispositif « bricolé », s'il permettait une présence à distance des participants, était peu confortable. C'est pourquoi, en 2016, nous avons cherché d'une part à améliorer le dispositif et d'autre part à l'associer à la construction d'un corpus propre à étudier différents aspects des interactions par écran qui constituent, depuis l'origine, l'objet des recherches du séminaire.

¹ <http://ateliers.sens-public.org/fabrique-de-l-interaction-parmi-les-ecrans/index.html>

² <https://impec.sciencesconf.org/resource/page/id/4>

De la visioconférence au projet « Présences numériques »

L'idée de départ a donc été de continuer à accueillir des participants à distance mais sous différentes formes, de façon à pouvoir analyser les effets de la plus ou moins grande autonomie des artefacts dans la dynamique du séminaire.

Seize personnes ont, à des titres divers, participé à cette recherche : une étudiante, quatre doctorants, une post-doctorante, une attachée temporaire d'enseignement et de recherche, 6 maîtres de conférences, 3 professeurs (dont un honoraire). Ces participants se trouvaient en France (Aix en Provence, Besançon, Lyon), en Angleterre (Londres) ou encore en Chine (Hangzhou) et chacun possédait une compétence différente des artefacts (allant de l'inexpérience à la maîtrise). Concernant leurs disciplines, si les sciences de l'information et de la communication et les sciences de l'éducation sont représentées, les représentants des sciences du langage sont majoritaires à l'image du laboratoire ICAR au sein duquel se tient le séminaire IMPEC.

Le programme général de la recherche a été établi ainsi :

- 2016-2017 – design de la recherche et recueil des données ;
- 2017-2018 – traitement des données, archivage, choix des axes de recherche, début de discussion concernant l'éditorialisation ;
- 2018-2019 – analyse des données et écriture du livre, réflexions préparatoires à l'édition en ligne ;
- 2019-2020 – éditorialisation en ligne et ouverture de la base Ortolang³ contenant le corpus des données de cette recherche au public scientifique.
- Juin 2021 : publication en ligne

La démarche méthodologique : l'éthologie réflexive visuelle

Prenant appui sur les recherches naturalistes réalisées par Jacques Cosnier au sein de son laboratoire lyonnais (1978, 1987), nous avons repris son approche d'éthologie compréhensive en l'adaptant à notre contexte d'étude. L'éthologie compréhensive consiste pour Cosnier en une analyse descriptive des comportements humains associant également le point de vue des individus observés à partir d'entretiens (Hotier, 2001). Nous avons donc développé une recherche en deux volets : une description objective des comportements des participants au séminaire et un recueil de leur ressenti et de leur subjectivité.

Premier volet :

La description des comportements humains s'effectue par le biais de films, vidéos, ou encore de captures dynamiques d'écran s'il s'agit de capter des interactions par écrans, telles que la conversation en ligne (Cosnier, Develotte 2011). Dans notre recherche et en lien avec les travaux de l'ethnographie visuelle (Pink 2007, Ibnelkaïd 2018), l'observation et l'analyse des pratiques s'appuient sur tous les enregistrements vidéo des données interactionnelles. Ces vidéos recueillies le jour des 5 séances de séminaires choisies ont en effet été traitées ensuite pour l'analyse puis enrichies de façon à pouvoir comporter en elles-mêmes les résultats de recherche. C'est en ce sens que l'on parle d'éthologie visuelle, dans la mesure où le visuel est central dans la façon d'étudier et de rendre compte des comportements des participants.

Deuxième volet :

En vue de prendre en compte le ressenti des participants, on a cherché à recueillir leurs avis au fur et à mesure de l'expérience, à partir d'entretiens d'explicitation (Vermersch, 1994). Des

³ <https://www.ortolang.fr/market/corpora/impec/v1>

entretiens semi-dirigés ont également été proposés pour documenter des aspects précis. Au total dix huit entretiens⁴ ont été conduits, puis transcrits et réunis dans un document de 231 pages. D'autres recueils des perceptions individuelles ont eu lieu par écrit ; tant à la fin de chaque séance, en vue d'une évaluation du ressenti de la présence des un·e·s et des autres, que de manière asynchrone, donc plus réfléchis, en aval des séminaires (quatre questions relatives au ressenti des participant·e·s⁵)

Ces deux types de données ont concerné les participants au séminaire doctoral et ont été recueillies et travaillées par ces mêmes participants au séminaire. C'est en sens que nous parlons d'une recherche éthologique réflexive. Nous définissons ainsi l'éthologie réflexive visuelle comme « 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, Ibelkaïd, 2021)

1. La mise en place du dispositif de recueil de données

La Cellule Corpus Complexes (CCC)

La Cellule Corpus Complexes, structure transversale au laboratoire ICAR dans laquelle évolue Justine Lascar, vise à la mutualisation et au développement des outils et pratiques portant sur la collecte et l'exploitation des corpus pluridisciplinaires et multimodaux. C'est dans ce cadre que Christine Develotte a sollicité la CCC pour collaborer au projet *Présences numériques*. Les questions soulevées par le recueil et le traitement de corpus multimodaux relèvent d'une conception technique mais aussi d'une réflexion méthodologique pour donner accès aux détails pertinents des interactions lors de l'analyse. En effet, les enregistrements ethnographiques à des fins d'études interactionnelles nécessitent de conserver les détails de l'interaction, notamment les éléments suivants : la chronologie des actions, l'écologie de l'interaction, le cadre participatif et les objets qui sont mobilisés par les participants. Pour pouvoir rendre compte de cette complexité, nous avons réalisé des captations audiovisuelles multiscopes des séances du séminaire *Présences numériques*.

Le dispositif de captation

L'emplacement des caméras se détermine en partie en fonction des questions de recherche posées en amont : dans notre cas l'intérêt se centrait en particulier sur les robots et la plateforme Adobe Connect ce qui a conduit à positionner des caméras dans la salle en fonction de ces éléments et de chercher à documenter le point de vue des utilisateurs des robots à leur domicile. Nous avons utilisé entre 3 et 4 caméras pour multiplier les points de vue sur la situation *in situ* à l'aide caméscopes classiques sur trépied mais aussi grâce à des caméras action (Gopro) disséminées dans la salle de séminaire à l'aide de trépied magnétique, et également une camera 360° dôme SP360 de Kodak. Nous avons également enregistré ce qui se passait sur les différents artefacts numériques à l'aide de capture d'écran dynamique : pour Adobe Connect, pour le robot Beam et pour le Kubi. Nous avons également demandé aux chercheuses qui pilotaient ces outils de téléprésence de s'enregistrer chez elles à l'aide d'une camera externe

⁴https://hdl.handle.net/11403/impec/v2/2_corpus_presences_numeriques/presences_numeriques_transcriptions_entretiens

⁵https://repository.ortolang.fr/api/content/impec/v2/2_CORPUS_Presences_Numeriques/Presences_Numeriques_questionnaires_bilan

sur trépied. En plus des pistes vidéos, nous avons récupéré les pistes audio des différent.e.s participant.e.s à l'aide de 4 micros HF. C'est donc entre 8 et 12 pistes selon les séances que nous avons recueillies lors de chaque séminaire. (voir fig.1)

Dispositif de captation

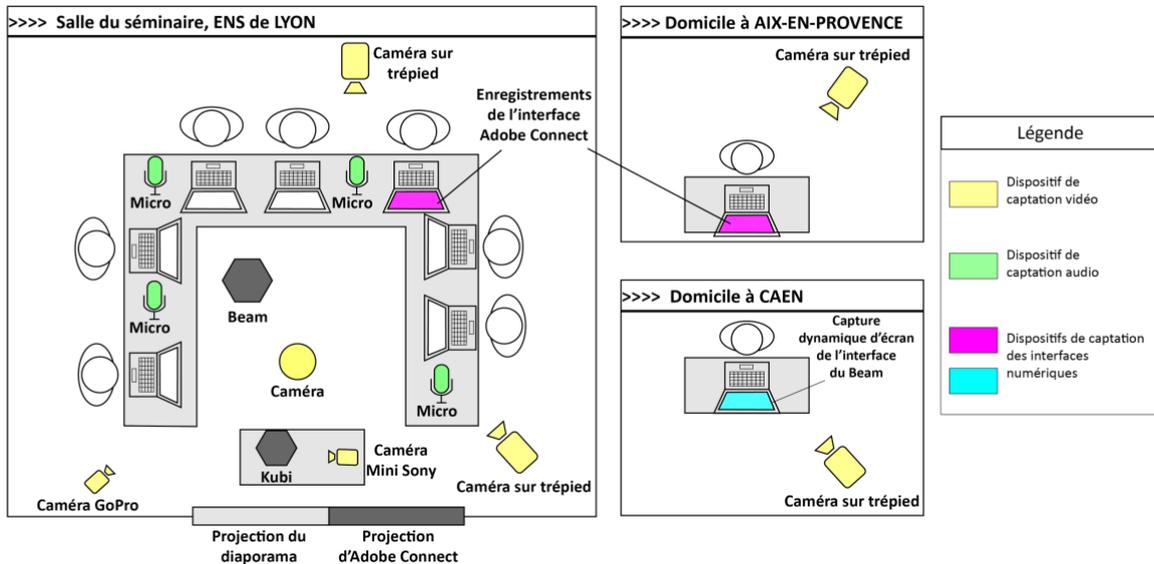


Figure 1 : dispositif de captation audiovisuelle du séminaire

Le travail sur les vidéos

Les 8 à 12 pistes audio et vidéos ont été synchronisées puis exportées de façon indépendante et en plusieurs montages multiscope. Par exemple, nous avons réalisé pour un même séminaire un montage de 3 vues *in situ* et la capture d'écran dynamique d'Adobe Connect (fig.2) et un montage entre 2 vues *in situ* et 2 vues « à distance » (fig. 3). L'agencement des différentes vues a été effectué de concert avec les chercheur.e.s. pour répondre aux exigences de l'analyse. Cette méthode de captation permet de prendre en compte le point de vue des différents participants. Elle permet également de rendre disponibles et donc analysables les paroles mais aussi les gestes, les postures, les regards, les actions des participant.e.s. et d'avoir accès aux détails des interactions à partir de la diversité des vues.



Figure 2 : montage multiscope in situ

Figure 3 : montage multiscope mixte

Une fois les exports et montages terminés, ceux-ci ont été mis à disposition de l'équipe via la plateforme Ortolang (Outils et Ressources pour un Traitement Optimisé de la LANGue) d'Huma-Num (fig.4). Dès le départ du projet, l'équipe avait décidé de mettre le corpus à disposition de la communauté scientifique et avait pensé sa diffusion via cette plateforme. L'accès a été réservé aux membres du projet pendant 2 ans puis a été ouvert au public quelques mois avant la parution de l'ouvrage *Fabrique de l'interaction parmi les écrans*. Cela nous a permis d'archiver et de documenter le corpus dès sa confection, offrant à la fois un espace de stockage des vidéos mais aussi le partage de documents de travail. Les métadonnées (description, auteurs, fichiers...) ont été renseignées au fur et à mesure des ajouts et modifications sur la plateforme Ortolang ce qui a ensuite permis une mise en ligne facilitée du corpus à la fin du projet.

The screenshot shows the Ortolang website interface. At the top, there is a dark blue header with the Ortolang logo and navigation links for 'Informations', 'Langue', and 'Connexion'. A left sidebar contains a menu with 'Accueil', 'Corpus', 'Lexiques', 'Terminologies', 'Outils', 'Projets Intégrés', 'Actualités', 'Informations', and 'Producteurs'. The main content area is titled 'Présences numériques' and features a small image of a meeting. Below the title, it lists the 'Laboratoire producteur' (ICAR - Interactions, corpus, apprentissages et représentations - UMR 5191 (Lyon, FR)) and 'Soutien institutionnel' (ICAR, LPL - Laboratoire parole et langage - UMR 7309 (Aix-en-Provence, FR), and ASLAN - LabEx ASLAN (Lyon, FRANCE)). There are buttons for 'Version 1' and 'Version 2'. A 'Référence à citer' section provides citation information for Christine Develotte and Justine Lascar (2021). A 'Description' section details the research context and objectives. A 'Téléchargement' section includes a Creative Commons license notice. A 'Contacter le producteur' box with an 'Envoyer un mail' button is also present.

Figure 4 Page d'accueil du corpus *Présences numériques* sur Ortolang

2. Valorisation des vidéos montrant le travail d'analyse

L'ouverture à l'édition numérique

La phase de confection et de montage du corpus a eu lieu lors de la tenue du séminaire en 2016-2017. S'en sont suivies deux années de travail collectif et collaboratif autour des trois premiers axes de recherche choisis : corporativité, attention et présence artefactuelle. La rencontre avec les éditeurs des Ateliers de [sens public] a concrétisé le projet d'ouvrage initial en un projet d'édition numérique, en ligne et en édition augmentée. Cette collaboration a ouvert le champs des possibles de l'écriture de l'ouvrage mais également de la forme des matériaux de recherche tels qu'ils pouvaient être diffusés. Dans les articles papier traditionnels, les matériaux audiovisuels sont la plupart du temps réduits en quelques captures d'écran contextualisées. L'édition numérique nous permettait tout à coup de pouvoir non seulement diffuser des extraits vidéos intégrés aux chapitres de l'ouvrage, mais également de repenser ces extraits vidéos non plus comme de simples illustrations mais comme des objets d'analyse à part entière.

Nous avons donc créé une méthodologie se basant sur les outils du montage audiovisuels pour traduire à l'écran les analyses des chercheur.e.s. La vidéo devient ainsi à la fois objet d'étude et média d'analyse.

Le changement de statut de la vidéo

Nous avons fait le choix d'intégrer les capsules vidéo dans les chapitres comme des éléments indissociables de l'analyse, plutôt de créer des vidéos indépendantes. Les capsules vidéo sont interdépendantes de l'analyse du chapitre dans lequel elles apparaissent. En accord avec l'équipe, nous avons décidé de réaliser des capsules courtes (environ 3 minutes) sur le modèle un concept = une capsule.

En ce qui concerne le processus de création, nous avons collaboré étroitement avec les auteur.e.s. Dans un premier temps, nous avons isolé les extraits dans la bibliothèque Final Cut du projet pour avoir accès aux différentes vues et micros disponibles. Une fois ce repérage effectué, nous avons demandé aux auteur.e.s de réaliser un synopsis précis pour chaque capsule, soit en explicitant quels changements de vue ils souhaitaient opérer, soit en explicitant quel était l'effet désiré : par exemple, dans cette capsule au timing 00 :00 :20, je souhaite focaliser sur les mains de Morgane qui attrape le robot Beam pour le reculer.

Suite à cette première phase de synopsis, nous avons réalisé un premier montage avec toutes les vues disponibles. Ensuite, nous avons en présentiel et en distantiel⁶ rencontré les auteur.e.s pour traduire les analyses avec les outils du montage audiovisuel. Dans le même exemple de Morgane qui attrape le robot, nous avons recadré le clip, incrusté la vue de l'écran de contrôle du Beam, encercler les mains de Morgane et ajouter des commentaires écrits. (fig 5)

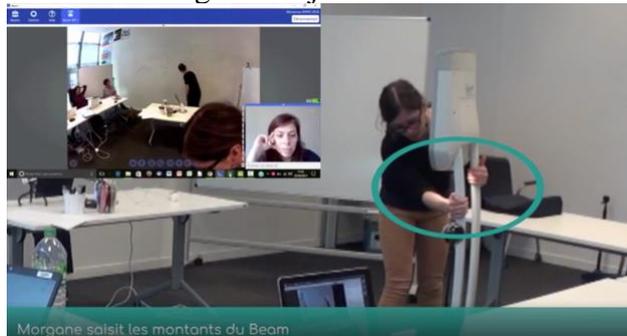


Figure 5 Morgane saisit les bras du Beam

Ainsi le statut de la vidéo change : on passe de la donnée brute à un produit issu de l'analyse de cette donnée. Le but est de montrer le point de vue du ou de la chercheur.e, c'est-à-dire de rendre visible l'analyse par le montage. Nous avons créé des conventions audiovisuelles pour avoir accès à une palette d'outils. Par exemple, nous avons traité les sous-titres sur une bande de fond noir, les commentaires sur fond bleu. Nous avons utilisé plusieurs techniques pour focaliser l'attention du spectateur : le zoom progressif, le recadrage, le flou, l'ajout de forme et de masques. En ce qui concerne l'audio, nous avons joué sur les différentes pistes et leurs volumes pour atténuer, ou à l'inverse, mettre en saillance un élément. Nous avons utilisé la rampe de vitesse pour accélérer ou à l'inverse ralentir un geste, une action, ou encore des transitions en atténuation de couleur pour créer des ellipses. Enfin, le split-screen nous a permis de rendre compte des actions réalisées in situ et à distance de façon synchronisée.

Les nombreux allers-retours et discussions avec les auteur.e.s nous ont permis de créer un langage commun d'écriture audiovisuelle.

⁶ le travail d'édition numérique de l'ouvrage commencé en 2019 s'est poursuivi en 2020, lors du confinement mis en place en raison de la pandémie de la COVID-19. Une partie des capsules a donc été réalisée à distance.

L'hébergement des vidéos

En ce qui concerne le stockage, l'archivage et la diffusion des capsules, nous avons choisi l'hébergement sur la plateforme ATV (Archivage et Transcodage Vidéo) créée et gérée au sein de l'ENS. Les capsules en format .mp4 sont indexées et mises à jour automatiquement dans les formats vidéos pérennes.

Les ingénieures de la Cellule Corpus Complexes ont été intégrées dans tous les processus de la recherche, de la prise de vue à l'écriture de l'ouvrage. Ces 5 années de travail collaboratif ont été riches en réflexions et expériences et ont donné lieu à une nouvelle façon de penser les matériaux audiovisuels, non plus uniquement comme une donnée brute mais aussi comme un média d'analyse à part entière.

Conclusion

Ces premières phases de la recherche *Présences numériques* que nous venons de documenter entrent en complémentarité avec les phases d'analyse⁷ et d'éditorialisation⁸ qui ont suivi et qui figurent dans les actes de ce colloque. Le fait que les vidéos se trouvent au centre de notre travail a en effet aussi entraîné une modification des protocoles habituels d'analyse et de rédaction vers une édition numérique augmentée.

Le contexte historico-technique actuel apparaît propice à l'émergence d'un nouveau positionnement scientifique de la vidéo dans les Sciences Humaines et Sociales : tant le développement de l'open science (Éditions Ateliers-sens public⁹) que l'accès récent à des plateformes de stockage pour de gros volumes de données (Ortolang, Huma-Num) permettent de dessiner des scénarios scientifiques fondés sur la vidéo. S'ouvrent aujourd'hui de nouvelles possibilités de montrer les résultats de la science (et également la science en train de se faire) où la part de l'écriture se réduit au profit de celle des images fixes et dynamiques.

Mots clés : séminaire doctoral polyartefacté ; capsule vidéo ; dispositif de captation audiovisuel ; éthologie réflexive visuelle ; édition numérique augmentée ; vidéos multiscopes.

Références :

Cosnier, Jacques. 1978. « Spécificité de l'attitude éthologique dans l'étude du comportement humain ». *Psychologie Française* 23 (1): 19-26.

Cosnier, Jacques. 1987. « L'éthologie du dialogue ». In *Décrire la conversation*, édité par Jacques Cosnier et Catherine Kerbrat-Orecchioni, 291-315. Linguistique et Sémiologie : travaux du Centre de recherches linguistiques et sémiologiques de l'Université de Lyon II. Lyon: Presses universitaires de Lyon.

⁷ Bouquain, A., Combe, C. « De la captation à la diffusion de données de téléprésence : points de vue des participants à distance ».

⁸ « El Hachani, M., Rémon, J. « Editorialisation scientifique à partir de données vidéo : l'exemple du projet Présences numériques ».

⁹ <https://ateliers.sens-public.org/>

Cosnier, Jacques, et Christine Develotte. 2011. « Éthologie compréhensive de la conversation en visioconférence poste à poste ». In *Décrire la conversation en ligne : le face à face distanciel*, édité par Christine Develotte, Richard Kern, et Marie-Noëlle Lamy. Lyon: ENS Éditions.

Christine Develotte, Morgane Domanchin, Samira Ibelkaïd, 2021, « Cadrage théorique et méthodologique pour l'éthologie réflexive visuelle », Fabrique de l'interaction parmi les écrans : formes de présences en recherche et en formation (édition augmentée), Les Ateliers de [sens public], Montréal, isbn:978-2-924925-13-3, <http://ateliers.sens-public.org/>. version, 15/06/2021 Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Develotte, Ibelkaïd, 2021 « Définition de l'éthologie réflexive visuelle » in Christine Develotte, Morgane Domanchin, Samira Ibelkaïd, 2021, « Cadrage théorique et méthodologique pour l'éthologie réflexive visuelle », Fabrique de l'interaction parmi les écrans : formes de présences en recherche et en formation (édition augmentée), Les Ateliers de [sens public], Montréal, isbn:978-2-924925-13-3, <http://ateliers.sens-public.org/>. version, 15/06/2021 Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Hotier, Hugues. 2001. « Entretien avec Jacques Cosnier ». *Communication et organisation*, n° 19. <https://doi.org/10.4000/communicationorganisation.2537>.

Ibelkaïd Samira. (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, « 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, isbn:978-2-924925-13-3, <http://ateliers.sens-public.org/>. version , 15/06/2021 Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Pink, Sarah. 2007. *Doing visual ethnography: images, media, and representation in research*. 2nd ed. London ; Thousand Oaks, Calif: Sage Publications.

Vermersch, Pierre. 1994. *L'entretien d'explicitation en formation initiale et en formation continue*. Paris: ESF Éditeur.