

Documenter le spectacle vivant. Apports théoriques et méthodologiques de l'analyse vidéo des interactions

Laurent Camus

*Université de Bâle, Département de Linguistique et Littérature
Université de Lausanne, Institut des Sciences Sociales*

Mylène Tanferri

Université de Lausanne, Institut des Sciences Sociales

Sarah Waeber

Université de Lausanne, Institut des Sciences Sociales

Dominique Vinck

Université de Lausanne, Institut des Sciences Sociales

Résumé

L'article offre une réflexion méthodologique et théorique sur les apports de l'analyse vidéo des interactions pour l'étude des pratiques de documentation du spectacle vivant. Contre une appréhension ontologique qui présupposerait l'existence a priori d'une incommensurabilité entre l'expérience scénique live et sa documentation par l'image, il propose d'appréhender empiriquement la documentation filmique du live en étudiant les pratiques locales de mise en image depuis les sites de production du spectacle. On envisage ainsi la manière dont un programme de recherche inspiré de l'analyse conversationnelle et de l'ethnométhodologie, mobilisant la vidéo à titre de ressource empirique, permet de rendre compte de

l'expérience scénique et de la temporalité vécue du direct par l'analyse des problèmes pratiques que les technicien.ne.s rencontrent lorsqu'ils la documentent.

Mots-clés : spectacle vivant ; documentation ; vidéo ; analyse des interactions.

This article offers a methodological and theoretical reflection on the contributions of video analysis of interactions to the study of documentation practices of live performance. In contrast to an ontological approach that would presuppose the existence of an a priori incommensurability between the live stage experience and its documentation by images, it proposes an empirical approach to the audiovisual documentation of live performance by scrutinizing local practices of filming from the perspective of the control room. The aim is to examine how a research program inspired

by conversation analysis and ethnomethodology, using video as an empirical resource, can provide an account of the stage experience and the temporality of the live performance by analyzing the practical problems technicians encounter when documenting it.

Keywords: live performance; documentation; video; interaction analysis.

La documentation du spectacle vivant soulève des enjeux fondamentaux relatifs à l'enregistrement et à la préservation du moment présent. La démarche poursuivie dans cet article vise à déplacer la question de la documentation de la performance scénique d'un problème ontologique – tenant aux statuts respectifs, présumés incommensurables, de la performance *live* et de son compte rendu audiovisuel – à un problème pratique, par l'étude empirique de l'activité située présidant à sa production audiovisuelle en temps réel. À travers ce déplacement, il s'agit de concevoir la documentation non pas comme un résultat déjà constitué mais comme processus émergent accompli par les technicien-ne-s, suivant les trajectoires de l'œuvre en train d'être jouée.

Pour ce faire, l'article propose d'adopter un cadre théorique et mé-

thodologique permettant d'étudier les pratiques de documentation en les appréhendant dans leur temporalité vécue, et en rendant compte des savoir-faire des technicien-ne-s (réalisateur-trice-s, monteur-se-s, cadreur-se-s, etc.) dans la perception, la captation et le montage de la performance scénique. Le propos développé ici s'inscrit dans une démarche, inspirée de l'ethnométhodologie (Garfinkel, 1967) et de l'analyse conversationnelle (Sacks, 1992 ; Sacks, Schegloff & Jefferson, 1974), se caractérisant par l'usage de dispositifs d'enregistrement audiovisuel pour saisir et analyser des activités sociales dans leur temporalité fine. Cette démarche permet de s'attacher au rôle des diverses « ressources multimodales » (Mondada, 2014) ou « sémiotiques » (Goodwin, 2000) – c'est-à-dire verbales, visuelles, matérielles et corporelles – à l'œuvre dans l'organisation et la coordination de

l'action. L'article constitue en premier lieu une réflexion théorique et méthodologique sur les usages de la vidéo pour l'étude des pratiques de documentation du spectacle vivant et des savoir-faire qui leur sont propres.

Dans une première partie de cet article, nous discuterons du supposé écart entre l'événement vécu par le spectateur dans l'ici et maintenant du *live* et sa représentation audiovisuelle. Nous proposerons une voie alternative à ce dualisme à travers l'étude empirique des pratiques situées de documentation.

Dans une deuxième partie, nous montrerons les possibilités offertes par l'approche conversationnaliste et multimodale pour étudier ce type de pratique. Cela nous amènera d'abord à montrer les spécificités de ce courant de recherche vis-à-vis d'autres usages de la vidéo en sciences sociales et à témoigner du cheminement conceptuel ayant conduit à l'étude minutieuse des pratiques de production audiovisuelle dans différents contextes et, partant, des pratiques de documentation du spectacle vivant.

Dans une troisième partie empirique, nous analyserons un extrait issu du corpus que nous avons constitué lors de la mise en image d'un grand spectacle vivant, la Fête des Vignerons de Vevey, filmé en multi-caméras et en temps réel. Ce sera l'occasion de mettre en avant l'importance de la prise en compte du contexte pour l'analyse des pratiques de documentation.

Le problème de la documentation du spectacle vivant

Le spectacle vivant, et de façon générale toute performance accomplie devant un public, pose des problèmes spécifiques pour sa documentation. Ces difficultés tiennent aux singularités de ce type de performance scénique, lesquelles peuvent être résumées en trois points : son caractère incarné (i), son caractère indexical (ii) et son caractère émergent (iii).

(i) La performance scénique est incarnée. Par des mouvements corporels, des postures, des chants, des prises de parole, des vocalisations, des regards adressés, la performance scénique est incarnée physiquement. De façon évidente, ces actions incorporées ont une épaisseur temporelle et s'enchaînent séquentiellement (telle action suit telle autre). La performance réside dans cet accomplissement ordonné.

(ii) La performance scénique est indexicale. La performance scénique se déroule dans un cadre situé, l'ici et le maintenant de la représentation scénique, qui lui confère son sens, notamment lié à la présence d'un public particulier. Ce qui singularise la performance en tant qu'œuvre, c'est ce moment vécu dans l'être ensemble de l'espace théâtral, qui fait de la représentation un moment unique, caractérisé par la simultanéité de son accomplissement et de sa perception collective.

(iii) La performance scénique est émergente. Bien qu'elle soit généralement scriptée et que, du point de vue des acteur-ric-e-s sur scène, l'action qui se déroule est (en partie) prévisible, cela n'est pas le cas pour les spectateur-ric-e-s qui découvrent le spectacle qui émerge.

Compte tenu de ces spécificités, l'enregistrement audiovisuel apparaît comme la modalité privilégiée pour documenter le spectacle vivant. Il paraît donner la possibilité à la fois de capturer les mouvements et trajectoires des corps en action, de préserver les propriétés acoustiques de la scène, de donner à voir des actions situées en rendant compte de leur séquentialité, et de suivre le déroulement d'une scène qui émerge dans sa continuité.

Pourtant, malgré ces caractéristiques avantageuses, l'enregistrement audiovisuel est régulièrement contesté dans sa capacité à pouvoir documenter pleinement le spectacle vivant. Cette tendance, héritée de Benjamin (2013 [1935]), veut qu'une part de la performance serait altérée à travers sa médiation technique par l'image. Le moment vécu, le *hic et nunc* de la performance scénique (ce à quoi nous avons fait référence sous le terme d'indexicalité) serait nécessairement perdu à travers les opérations de médiation. La captation n'est pas pure transparence : les effets de cadrage (par la restriction du champ) et de montage (par la mise en succession des plans) modifient l'action scénique et son intelligibilité, et produisent un hiatus

irrémediable entre la réalité vécue de la scène et la réalité montrée à l'écran. Ces arguments classiques sont régulièrement convoqués pour critiquer les modes de représentation des scènes *live*, en particulier pour le sport¹.

L'idée d'une incommensurabilité entre la documentation de l'œuvre et l'œuvre se base sur une conception ontologique de la performance scénique (Phelan, 1993). Cette approche tend à réifier la performance *live* en faisant de l'absence de médiation technique l'une de ses propriétés essentielles. Ainsi l'immédiateté temporelle du *live* est conçue comme une « *im-médiateté* » (autrement dit l'absence de médiation) qui lui serait constitutive. Cette conception a été critiquée en montrant notamment qu'elle était le fait d'une confusion à propos de la caractéristique d'immédiateté propre au *live*. Celle-ci n'est pas un donné mais une construction socio-historique, le résultat des pratiques par lesquelles les artistes, le public et les médiateurs produisent un événement conçu comme *live* (Auslander, 2008 [1999] ; Vinck & Tanferri, 2020 ; Vinck *et al.*, 2021).

L'approche que nous défendons ici refuse également de s'en tenir à l'argument de l'incommensurabilité de la performance et de sa documentation par l'image. Elle propose de déplacer la question de sa documentation d'un problème ontologique – constitutif de l'établissement *a priori* d'une dualité

¹ Pour une discussion de ces points, voir Camus (2021a).

de nature entre l'œuvre et sa médiation – à un problème pratique – en rendant compte des manières concrètes dont la documentation de l'œuvre est accomplie. Une telle conception, qu'on qualifiera de praxéologique, invite à traiter la performance scénique dans son caractère unitaire, en montrant comment les pratiques de médiation (et par conséquent les techniques) participent de la constitution de l'événement en temps réel.

De ce point de vue, le propos est redevable d'une conception ethnométhodologique du social (Garfinkel, 1967) et s'inscrit dans le sillage de travaux ayant relevé la pertinence de cette approche pour l'étude des pratiques de conservation et de documentation (Kreplak, 2017 ; 2018). La notion ethnométhodologique d'*accountability* permet notamment d'envisager la documentation dans sa dimension élémentaire, comme une activité ordinaire dans laquelle les participant-e-s s'interrogent sur le statut de l'objet matériel, le qualifient, le catégorisent, identifient des problèmes, etc., et, ce faisant, transforment l'œuvre en projetant de nouvelles trajectoires. En ce sens, la documentation est conçue comme une activité performative ; elle n'est pas un simple compte rendu transparent de l'œuvre mais participe de son existence en tant qu'objet et de sa réception pour un public. Cette conception amène également, sous une modalité praxéologique, à prendre en considération la dimension culturelle et morale des pratiques de documentation. L'analyse des pratiques fines permet ainsi

d'appréhender la manière dont une communauté professionnelle perçoit et qualifie une œuvre selon les standards, les attentes et les valeurs propres à la profession. Elle permet enfin de resituer la documentation dans son environnement matériel et technologique d'émergence. L'ethnométhodologie propose d'accéder au niveau de détail pertinent pour les participant-e-s dans la conduite de leur activité. Au prix d'analyses techniques et minutieuses, cette approche permet de rendre raison de leurs compétences au travail, de leurs savoir-faire incarnés et, par là même, de leur vision de l'œuvre.

Une telle entreprise, appliquée à la performance scénique, impose néanmoins un certain nombre d'aménagements méthodologiques afin que soit rendue possible l'appréhension des trois dimensions évoquées plus haut (son caractère incarné, indexical et émergent).

Il s'agit notamment de montrer comment les participant-e-s qui documentent la performance scénique tiennent compte de, et s'orientent vers, ces singularités du *live*, en examinant les modalités par lesquelles ils perçoivent et rendent compte par l'image des actions incorporées (i), dans la temporalité de leur accomplissement (ii) et de son caractère émergent et situé (iii).

Cela suppose de mettre en œuvre une méthodologie susceptible de rendre compte des impératifs pratiques dans lesquels les participant-e-s sont engagés. Dans les pages qui suivent, nous nous

attachons à un courant en particulier qui a problématisé ces questions – l’analyse conversationnelle d’inspiration ethnométhodologique – et retraçons le cheminement emprunté (de l’analyse des conversations à l’analyse des pratiques) pour parvenir à capturer les pratiques de documentation audiovisuelle.

Enregistrer et analyser des interactions sociales : de l’analyse de conversation à l’analyse des pratiques de mise en image

L’idée que l’analyse conversationnelle soit une approche à même de servir de fondement à l’étude empirique des pratiques de documentation ne va évidemment pas de soi. Dans cette partie, nous allons montrer le cheminement théorique et méthodologique ayant conduit ce mouvement de l’analyse des échanges conversationnels *stricto sensu* à celle des différentes activités multimodales, mobilisant le corps et la perception dans la communication. Ce mouvement, parallèle à l’évolution des techniques d’enregistrement, se fonde sur l’adoption de la vidéo pour étudier les interactions sociales. Nous verrons ensuite comment cette approche basée sur l’usage de la vidéo a servi de fondement à l’émergence d’une réflexion empirique sur les pratiques de production audiovisuelle et les possibilités qu’elle ouvre pour l’étude des pratiques de documentation.

L’enregistrement audio et l’analyse des échanges conversationnels

Constitutivement liée à l’ethnométhodologie, l’analyse conversationnelle est initialement pratiquée par des chercheur-euse-s, pour la plupart inscrit-e-s dans le domaine de la sociologie interactionniste, ayant pour visée de rendre compte de la dimension élémentaire du social par l’étude des phénomènes d’ajustements mutuels tacites. La conversation représente ainsi un site favorable pour l’investigation de phénomènes organisationnels émergents, dont l’accès est facilité à partir des années 1960-1970 par la démocratisation des possibilités d’enregistrement. Sur la base de grands corpus audio (en particulier d’interactions téléphoniques), les chercheur-euse-s en analyse conversationnelle mettent ainsi en lumière le caractère ordonné et contingent de l’organisation de la conversation, à travers notamment l’importance accordée à la notion de séquentialité. L’enchaînement des tours de parole (Sacks, Schegloff & Jefferson, 1974) obéit à une dimension normative vers laquelle les participant-e-s s’orientent sans y prêter attention : les tours de parole sont à la fois ajustés au contexte (et en particulier au tour de parole qui précède) et contribuent à le reconfigurer en projetant une suite d’actions (de réponses) socialement normées (*context sensitive/context renewing*).

La conversation apparaît ainsi comme le résultat de ces micro-ajuste-

ments tacites. L'analyse conversationnelle (qui donne son nom à ce courant) n'est pas une méthodologie analytique en tant que telle, mais une activité menée par les participant-e-s eux-mêmes : au cours de leurs échanges, celles-ci et ceux-ci exhibent par leurs actions des formes ordinaires et tacites de compréhension de la situation. Autrement dit, ils sont les premiers analystes. Le rôle du-de la chercheur-euse est de rendre compte de ce phénomène de compréhension et de ses implications pour l'organisation de l'activité. L'analyse conversationnelle partage avec l'ethnométhodologie au moins deux de ses idées directrices : d'une part, la prise en considération de la dimension émergente de l'action et de son organisation et, d'autre part, le refus d'une position de surplomb du-de la chercheur-euse et la mise en lumière des compétences analytiques tacites dont font preuve les participant-e-s.

Pour ce faire, les travaux en analyse conversationnelle reposent sur la constitution de corpus d'interactions authentiques et naturelles documentées par l'usage de dispositifs d'enregistrement. Ceux-ci permettent d'accéder et de ré-accéder *ad libitum* aux actions enregistrées, à leur enchaînement et à leurs détails. L'analyse conversationnelle revendique une vision naturaliste des enregistrements : il s'agit de capturer l'activité dans laquelle sont engagés les participant-e-s sans interférer et en la laissant se déployer telle qu'elle se déroulerait naturellement sans la présence du-de la chercheur-euse.

Un point important ici, conforme à l'esprit de l'ethnométhodologie, tient au fait que ces enregistrements ne sont pas des documents mobilisés à l'appui d'une thèse déjà établie. Ce ne sont pas des ressources pour l'argumentation mais le thème des analyses. La démonstration se fonde sur la monstration (le pointage) d'actions et de formats rendus visibles (ou audibles) sur les enregistrements. La constitution de collections de cas, regroupés par formats d'actions spécifiques, permet d'accréditer l'idée que la démarche analytique ne repose pas de façon accidentelle sur la structure d'un cas isolé, mais qu'elle est normativement inscrite dans le type de situations investiguées. L'un des principes singuliers de l'analyse conversationnelle tient également dans la manière dont les compétences des participant-e-s sont mises en exergue, à travers notamment la « *next turn proof procedure* » (« *procédure d'établissement de la preuve par le tour de parole qui suit* ») (Sacks, Schegloff & Jefferson, 1974, pp. 728-729). Celle-ci consiste à faire reposer la démonstration sur la mise en évidence empirique de l'orientation des participant-e-s vers les phénomènes mis en avant par l'analyste, en montrant comment ceux-ci rendent manifeste leur compréhension du tour de parole qui précède à travers leur action et contribuent ainsi à en faire émerger le sens.

Un tel parti pris analytique n'est rendu possible que par l'adoption de modalités conventionnelles de visualisation scripturale des interactions verbales. La

transcription est conçue comme une première étape du travail analytique avec la pleine conscience que les choix de transcription ne sont pas neutres pour l'analyse des séquences. Ces choix scripturaux participent à la mise en visibilité de phénomènes, aux dépens d'autres, résultant ainsi d'opérations de sélection (Duranti, 2006) en lien avec des choix théoriques (cf. Ochs, 1979). Ces conventions, développées dès les années 1960, permettent de rendre compte des échanges en préservant leur dimension séquentielle (un locuteur parle après l'autre) et leur temporalité, à travers notamment l'indication de la simultanéité des prises de parole, les pauses et les silences. L'étude du langage en interaction et sa transcription incitent à l'abandon d'une perspective centrée sur le discours pour rendre possible l'analyse des micro-ajustements séquentiels structurants.

L'enregistrement vidéo et l'analyse des interactions dans leur écologie complexe

L'histoire des techniques d'enregistrement utilisées en vue d'étudier l'action humaine est relativement ancienne. Dès la fin du XIX^e siècle, la chronophotographie a été créée et utilisée en vue d'analyser la gestualité humaine et animale (Marey, 1899). Le courant qu'on a appelé en français l'anthropologie de la communication (Winkin, 2001 [1996]) a joué un rôle majeur dans la réflexion et l'expérimentation des méthodes audiovisuelles en sciences sociales, par le biais de cher-

cheurs tels que Bateson, Birdwhistell ou Schefflen. Ce courant adopte une conception naturaliste de l'enregistrement audiovisuel amenant à capturer les scènes en veillant à ne pas en altérer le déroulement (absence de scénarisation, minimisation des interactions avec les participant-e-s). Dès les années 1930, Bateson a expérimenté l'usage d'enregistrements filmiques lors du terrain qu'il a mené avec Mead à Bali (Bateson & Mead, 1942). Les analyses reposent essentiellement sur la gestualité et mettent souvent de côté les aspects langagiers (lesquels peuvent être traités séparément, *via* des enregistrements audio), Bateson documentant pour l'essentiel des actions incarnées (et notamment des danses) à partir de films muets (cf. Erikson, 2011). La technique utilisée, relativement lourde et dissociant image filmique et enregistrement sonore, ne permet pas de documenter finement des interactions dans leurs dimensions à la fois verbale et gestuelle.

Ce mouvement de recherche se caractérise notamment par la prise en compte du contexte pour l'analyse de l'interaction et la mise en question du modèle cybernétique de la communication. Notamment mise en œuvre dans le célèbre projet « The Natural History of an Interview »², la démarche a connu des héritages divers en sciences sociales, à

2 McQuown N. (ed.) (1971). « The natural history of an interview » (Series 15, Nos. 95-98), Chicago, Microfilm Collection of Manuscripts on Cultural Anthropology, Microfilm Department, Joseph Regenstein Library, University of Chicago. Pour un développement, voir Leeds-Hurwitz (1987).

travers différents courants étudiant la gestualité et la corporéité en contexte : l'analyse conversationnelle est l'un de ceux qui a le plus intégré la vidéo à ses problématiques de recherche et tiré les bénéfices de son usage³. Si l'analyse conversationnelle et l'ethnométhodologie sont souvent présentées dans les termes d'une rupture vis-à-vis des traditions établies, la réalité de l'émergence et de l'institutionnalisation des usages de la vidéo pour analyser les interactions est pourtant plus complexe. Elle révèle, si ce n'est une filiation directe, des liens interpersonnels et théoriques avec les différentes traditions interactionnistes : appartenance commune à des réseaux d'interconnaissance, engagement dans des discussions collectives avec des chercheurs issus de traditions diverses intéressées aux pratiques communicationnelles en contexte (Bateson, Labov, Gumperz, Goffman, Hymes, Schefflen, Birdwhistell) et souci commun d'appréhender l'action humaine, conçue en tant que phénomène global incluant le corps, dans son écologie⁴.

Dès la fin des années 1970, Charles et Marjorie Goodwin, qui ont eux-mêmes été en lien avec les différents membres de ces courants (Sacks & Schegloff, 2002,

³ Le propos ici ne prétend pas faire l'énumération de l'ensemble de ces courants, mentionnons tout de même l'importance des *gesture studies* (Kendon, 2004 ; McNeill, 1985) et sa filiation revendiquée avec l'anthropologie de la communication (Kendon, 2004 ; 2007).

⁴ Pour une discussion approfondie des enjeux biographiques et théoriques autour de l'histoire des usages de la vidéo en analyse de l'interaction, voir Mondada (2021).

pp. 133-134), ont été précurseurs dans l'adoption des techniques audiovisuelles pour étudier l'interaction dans une acception conversationnaliste. Leur travail participe d'une utilisation novatrice de la vidéo consistant à étendre à la gestualité, aux postures corporelles et aux regards les enjeux de compréhension mutuelle et les ajustements réciproques inhérents aux conversations ordinaires. Goodwin (1979) met en avant le rôle actif de l'auditeur-ice dans la conversation et dans la production du sens, et montre que la phrase est un phénomène accompli interactionnellement. Il donne ainsi naissance à un champ de recherche qui sera plus tard regroupé sous l'étiquette d'analyse de l'interaction multimodale (cf. Deppermann, 2013 ; Mondada, 2016 ; 2017).

L'adoption de la vidéo par l'analyse conversationnelle se fait cependant au prix de réaménagements méthodologiques et de nouvelles directions théoriques qui la distinguent des entreprises précédentes et, en particulier, de l'œuvre pionnière de Birdwhistell (1952 ; 1970). Un point central tient au statut de la gestualité. Celle-ci n'est pas étudiée pour elle-même, mais en tant qu'elle est insérée dans des interactions au cours desquelles les participant-e-s lui donnent sens. Les actions non-verbales sont ainsi transcrites conjointement aux actions verbales (les tours de parole) et insérées à l'intérieur des tours de parole ou des silences par le recours à des repères temporels.

En utilisant la vidéo, l'analyse conversationnelle s'est ouverte à des

phénomènes plus larges que la conversation en face-à-face proprement dite, pour étudier différentes pratiques sociales reposant sur des formes subtiles de coordination et d'ajustement, en coprésence et à distance, où le corps ou des ressources visuelles sont mobilisés. L'usage de dispositifs audiovisuels (et non plus seulement audio), permet de décentrer l'analyse du *logos* en l'étendant aux actions incorporées (Goodwin, 2000) et, de façon plus générale, à tout phénomène visuellement accessible pour les participant-e-s. Le propos s'élargit également de l'étude des activités langagières où le corps est mobilisé comme ressource (Mondada, 2007a ; Broth & Mondada, 2013), à celle de l'accomplissement conjoint d'activités corporelles où le langage joue un rôle dans la synchronisation des corps en mouvement tel que le sport (Evans & Lindwall, 2020) ou la danse (Keevalik, 2015).

Un tel décentrement offre également la possibilité d'étudier la manière dont les objets (Brassac *et al.*, 2008 ; Goodwin, 2010 ; Nevile *et al.*, 2014, Mondada, 2019) ou les technologies (Heath & Luff, 2000 ; Mondada, 2007b ; Licoppe, 2009) sont mobilisés en interaction, ou de s'intéresser à d'autres modes d'appréhension sensible, tels que le toucher (Cekaite & Mondada, 2020, *cf.* en particulier Kreplak & Mondémé (2014) pour le contexte muséal) ou l'ensemble des cinq sens (Mondada *et al.*, à paraître). Si la question n'est pas de se prononcer sur le statut ontologique des objets et de la matérialité, l'analyse vidéo des interactions permet d'étudier empiriquement les modalités

pratiques par lesquelles l'environnement matériel participe de la conduite collective de l'action, que ce soit au titre de ressource ou de partenaire (Goodwin, 2006).

Quelle que soit la diversité des positions respectives de ces auteurs vis-à-vis de l'analyse conversationnelle et de l'ethnométhodologie originelles, ces travaux partagent un certain nombre de vues qui peuvent être caractérisées par les points suivants.

Il y a d'abord l'adoption d'une perspective émique, c'est-à-dire la prise en compte de l'appréhension du monde et de sa compréhension par les participant-e-s en situation, plutôt que le recours à des explications exogènes sur leurs actions. Une telle perspective amène à attacher une importance particulière aux *accounts* verbaux, scripturaux ou visuels qui sont produits par les participant-e-s en situation dans leur tâche de documentation de la réalité vécue.

La prise en compte de la perspective des participant-e-s amène à mettre en valeur les compétences pratiques dont ils-elles font preuve dans la conduite de leurs activités concertées. Il s'agit de rendre compte des techniques, opérations et méthodes par lesquelles les participant-e-s perçoivent, décrivent et reconfigurent le monde, en se reposant sur des savoir-faire incorporés, en particulier pour faire face à des situations imprévues. L'un des points importants ici réside dans le refus d'une conception mentaliste de l'action, et la défense d'une connaissance

par corps et d'un sens pratique observable en situation, et descriptible au moyen de la vidéo.

Un troisième point réside dans la mise en exergue du caractère émergent et contingent de l'action. En adoptant le point de vue de la situation vécue par les participant-e-s, les événements auxquels ils se confrontent et dont ils rendent compte apparaissent comme une suite d'actions contingentes et ordonnées de façon endogène. Cette conception se caractérise par le refus du déterminisme. Les dispositifs et formats interactionnels contraignants pour l'action supposent des formes de compréhension et micro-ajustements : les instructions (Amerine & Bilmes, 1988) ou les plans d'action (Suchman, 2007 [1987]) sont mobilisées comme des ressources pour penser, agir et interpréter les situations, et les participant-e-s accomplissent des efforts pour s'y conformer ou rendre visibles leurs actions comme étant conformes au plan.

Le dernier point tient à l'importance de l'indexicalité, autrement dit des facteurs contextuels et de la singularité de la situation dans l'accomplissement de l'action. L'enjeu analytique est donc d'accéder à l'expérience vécue par les participant-e-s dans l'ici et le maintenant de la situation, et les prises offertes par elle pour agir.

Dans la partie qui suit, nous allons voir comment ces différents points peuvent servir de pivot pour l'étude des

pratiques de documentation de la performance vivante par la vidéo.

L'enregistrement vidéo et l'analyse des pratiques de documentation du spectacle vivant

Documenter la performance vivante par la vidéo est à la fois une activité scientifique pour l'analyste intéressé par l'action et son déroulement, et une activité professionnelle ordinaire pour différents participant-e-s du monde social engagé-e-s dans des tâches de mise en image de la réalité. De ce point de vue, l'usage de la vidéo apparaît à la fois comme une ressource analytique pour étudier les interactions sociales, et à la fois, comme le thème des analyses, la focale étant placée sur les manières concrètes dont les participant-e-s perçoivent, mobilisent et produisent des images filmiques.

L'analyse multimodale propose une approche originale pour étudier ce type d'activités. En plaçant des caméras dans les lieux de la production audiovisuelle de la performance, on accède ainsi au travail, ordinairement invisible (Camus, 2021b), de coordination de l'action filmique. Les activités de perception de la scène à travers les écrans y apparaissent concomitantes aux prises de vues et aux interactions fines qui les rendent possibles.

La méthode, en ce qu'elle vise à étudier le détail de l'action et à restituer la perspective des participant-e-s sur l'action, permet d'accéder empirique-

ment aux modes ordinaires d'enquête que les participant-e-s mènent sur l'œuvre lorsqu'ils la perçoivent et s'interrogent sur son statut (que se passe-t-il sur scène ? quelle image sélectionner ? quelle caméra choisir ? comment respecter l'intégrité d'une scène ?). Une fois replacées dans le contexte de la régie, ces questions n'apparaissent pas simplement comme des questions spéculatives ou métaphysiques liées aux statuts respectifs de l'œuvre et de sa documentation, mais comme une question pratique, généralement non formulée et rendue observable à travers des actions routinières.

Étudier les pratiques de documentation par la vidéo implique nécessairement de questionner la dimension réflexive de ces pratiques. À travers le filmage et le montage, les participant-e-s mettent en avant des formes de compréhension ordinaire des scènes qu'ils filment ou montent. Les mouvements de caméra, de focale ou les changements de plan coordonnés aux actions filmées rendent visible une certaine vision du monde et de l'œuvre. Cette réflexivité des pratiques filmiques, évoquées pour différentes pratiques en lien avec la vidéo (Lomax & Casey, 1998 ; Mondada, 2009 ; Camus, 2021a), n'est pas à concevoir comme le fait d'une opération mentale mais plutôt comme un ajustement routinier à ce qui arrive, ajustement basé sur des manières de faire et de voir incorporées et professionnellement normées (Goodwin, 1994). C'est particulièrement le cas pour la réalisation en direct, où la temporalité de la

mise en images est synchronisée avec celle de la scène filmée : l'urgence qui caractérise la captation en direct suppose des formes de routines organisationnelles du travail permettant de cadrer l'événement qui émerge (Camus, 2017a) et de s'ajuster à sa temporalité (Camus, 2017b).

À travers ces activités filmiques, on retrouve donc un thème classique de l'analyse conversationnelle et de l'ethnométhodologie – évoqué *supra* –, celui des formes endogènes de compréhension de l'action et de leur dimension séquentiellement ordonnée. Elles ne concernent pas uniquement des interactions ayant lieu dans un même cadre participatif mais également, et de façon plus essentielle, des actions se déroulant à distance (en l'occurrence sur scène) et auxquelles les participant-e-s donnent sens par leurs actions. Ces actions sont à la fois des actions verbales (descriptions, évaluations, instructions, etc.) et non-verbales (mouvements de caméras, montage). Ces pratiques complexes sont donc ajustées à une scène qui est perçue et filmée en temps réel : elle donne à voir des comédien-ne-s, chanteur-euse-s, danseur-euse-s, etc. en train d'accomplir des actions à la fois vocales (échange de répliques, chant, etc.) et incorporées (postures du corps, gestes, regards, mouvements, etc.). L'un des points essentiels de ces actions réside dans la séquentialité de ces actions : l'enchaînement des plans doit rendre possible la saisie du déroulement de l'action scénique et, donc, l'accès à son intelligibilité propre.

Le thème de la réflexivité filmique des actions permet ainsi d'aborder à nouveaux frais les activités de documentation de la performance vivante. Ainsi conçues, les activités de filmage ne sont ni des pratiques transparentes, faisant accéder à un événement sans médiation, ni des pratiques déformantes, travestissant l'événement⁵. En replaçant l'analyse dans le contexte situé dans lequel sont opérés le cadrage et le montage en temps réel, on a ainsi accès à l'expérience vécue des participant-e-s dans la temporalité de son accomplissement. La documentation n'apparaît pas comme une activité extérieure à l'événement mais comme un cours d'action ajusté réflexivement à ses trajectoires, rendant manifestes des formes de compréhension de son action et donc une certaine vision de la performance.

Partie analytique

Données et terrain

Afin d'appuyer empiriquement notre propos, nous allons désormais mener l'analyse d'une séquence issue d'un corpus que nous avons constitué, relatif aux activités de documentation et de préservation d'un spectacle donné lors d'une fête traditionnelle suisse, la Fête des Vignerons de Vevey.

La Fête des Vignerons est une fête patronale traditionnelle se déroulant à chaque génération (soit tous les 20 à 25 ans) et s'organisant autour d'un grand spectacle vivant. Au fil des éditions, la fête est devenue un événement d'envergure, la dernière édition (celle que nous avons observée) attirant chaque jour durant un mois, un public suisse et international dans une enceinte de 20000 places construite pour l'occasion sur la place de la ville. Depuis 1977, l'une des représentations est retransmise sur la chaîne télévisée nationale suisse (RTS). La dernière édition présente en outre la particularité d'avoir été filmée et montée chaque soir en temps réel, préservant ainsi la trace de chacune des représentations. Les images collectées servent de rushes pour le montage d'un DVD disponible pour le grand public.

Dans cette partie, nous entendons rapporter les activités de documentation du spectacle vivant à leur dimension élémentaire, c'est-à-dire aux pratiques de captation et de montage en temps réel. Pour saisir la documentation de la performance vivante, nous proposons donc de l'envisager non pas comme un résultat déjà constitué, mais comme un processus émergent, en donnant accès aux pratiques d'ordinaire invisibles par lesquelles les images sont produites. Cette perspective s'écarte donc d'une traditionnelle analyse sémiologique des images pour proposer une analyse des interactions multimodales participant de la production des images vidéo de la performance, c'est-à-dire en rendant

⁵ Cf. Camus (2021a) pour une discussion de cette question à propos de la réalisation télévisée du football.

compte des différentes « *ressources sémiotiques* » (« *semiotic resources* »), qui peuvent être verbales, corporelles et visuelles) Goodwin, 2000, p. 1490) dont les participant-e-s disposent en situation.

La mise en image de la fête est produite et réalisée par la RTS, chaîne qui met à disposition des moyens logistiques et humains pour la captation. Les représentations sont filmées par treize caméras. Le réalisateur sélectionne les plans en direct et coordonne la mise en image depuis un car-régie situé en dehors de l'arène. Cet espace constitue le « *centre de coordination* » (« *center of coordination* ») (Suchman, 1996) de l'ensemble des activités de captation. Là, sont rendus visibles les retours de caméras de chacun des cadres-euses sur un mur d'écrans. Ces images, qui constituent l'unique moyen d'accès visuel à l'événement pour les membres de la régie, sont mobilisées pour coordonner le cadrage et sélectionner les plans. Le réalisateur et la scripte communiquent avec les cadres-euses situé-e-s dans l'espace de la représentation par le biais de casques et micro, afin de donner des instructions de cadrage et de positionnement aux techniciens. La réalisation apparaît ainsi comme le fait d'une coordination à distance complexe médiée par les technologies visuelles et audio (cf. Broth, 2009).

À l'occasion de l'enquête que nous avons menée sur les coulisses de la production de cette fête (Vinck, 2019), nous avons suivi l'activité de production sur

différents sites par le recours à ethnographie⁶ et par l'usage de la vidéo à des fins analytiques. L'analyse portera en particulier ici sur un extrait issu des données récoltées en régie image. Deux caméras ont été placées en régie lors de la réalisation de la performance, une face au réalisateur et à la scripte et une dans leur dos, donnant à voir l'environnement visuel leur permettant de coordonner la captation. Nous avons en outre récupéré les images produites, de façon à pouvoir les insérer dans les transcriptions qui vont suivre. Le corpus « régie RTS » comprend l'enregistrement de cinq représentations (soit environ 20 heures de données).

Les transcriptions sont effectuées selon les conventions traditionnelles de Jefferson (2004) pour le verbal et Mondada (2018) pour le multimodal. Ces transcriptions permettent notamment de rendre compte des actions verbales (en caractères gras) et non-verbales (en caractères minces), en respectant leur temporalité (durée des silences et succession temporelle). Les locuteur-ric-e-s (ou participant-e-s accomplissant une action non-verbale) sont indiqué-e-s dans la colonne de gauche, aligné-e-s à leur tour de parole ou leur action. Les actions non-verbales sont indiquées sur des lignes non-numérotées et sont indexées à la ligne de parole du dessus

6 Voir notamment Tanferri, Waeber & Vinck (à paraître) pour la production du DVD, Vinck, Waeber & Tanferri (2021) pour la production du son et Waeber, Tanferri & Vinck (2021) pour une approche longitudinale sur la narration audiovisuelle de la fête.

(en gras et numérotées) par l'usage de balises temporelles précisant le moment exact de début et de fin d'un phénomène. Des images sont également insérées par le symbole # pour aider la compréhension. La ligne verbale rend compte de toutes productions vocales des participant-e-s, sans les modifier. Elle est également utilisée comme un flux temporel, faisant mention précise des silences (entre parenthèses, secondes et dixièmes de seconde). Si elle peut paraître technique pour le non-initié, une telle manière de transcrire permet d'indexer les événements non-verbaux à la temporalité vécue des participant-e-s, d'envisager la durée des phénomènes et de rendre compte de l'enchaînement séquentiel des actions.

La documentation comme processus émergent, analyse d'un cas

L'analyse qui va suivre présente une visée didactique et entend témoigner des apports possibles de l'analyse multimodale des interactions, et donc des usages *de la vidéo* pour étudier les pratiques de documentation *par la vidéo*.

L'extrait est relativement court, une vingtaine de secondes, et présente une scène du spectacle lors de laquelle plusieurs actions simultanées se produisent, comme c'est souvent le cas lors d'une performance vivante. Il s'agit d'une situation de spectacle ordinaire, choisie d'une part pour sa relative simplicité (au regard

de l'ensemble du corpus), et d'autre part car elle permet de mettre en avant les dimensions gestuelles, matérielles et visuelles des actions accomplies sur scène et visibles depuis la régie ainsi que la manière dont celles-ci sont ressaisies par les membres de la régie.

La scène s'organise de la façon suivante : pendant le chant d'un soliste, une danseuse suspendue en l'air jouant dans le spectacle le rôle d'une libellule, survole l'espace scénique. Sur scène, des comédiens projettent en l'air de gros ballons gonflables transparents (des gouttes de sève du point de vue de la scénographie) qui montent et descendent lentement. L'analyse qui va suivre adopte le point de vue des participant-e-s de la régie qui perçoivent l'action depuis un mur d'écrans et travaillent à filmer et monter la scène en temps réel. Le réalisateur (REA) donne des instructions aux cadreurs (ca) en s'adressant à eux par l'usage de leur numéro de caméra. Il leur indique qui passe à l'antenne, quelle sera la prochaine caméra à être sélectionnée et quel type de phénomène il souhaite capter.

Au début de la séquence, différentes images sont disponibles pour le réalisateur (REA) dans ses tâches de montage. Le soliste est filmé en plan serré (ca3), les comédiens (comX) poussant les ballons sont disponibles sur différents plans et notamment celui de la caméra 2 (ca2) et de la caméra 9 (ca9) (fig. #1). Enfin, au moment où débute cet extrait, le plan sélectionné est un plan de *spidercam*, en plongée et en mouvement, donnant à voir

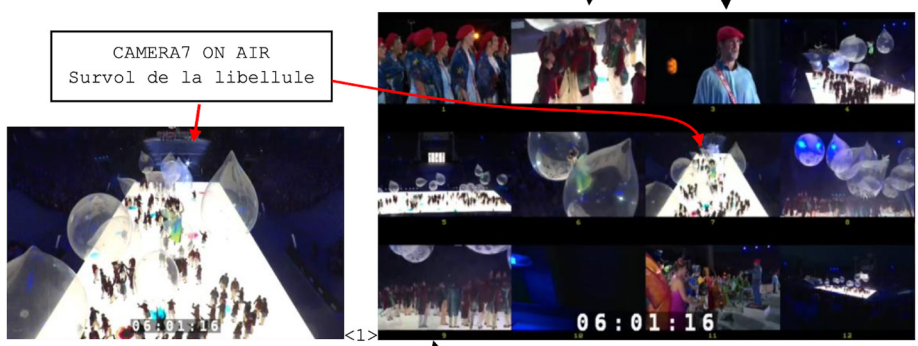
extrait_FEVI_live_200719_012229

1 REA #et la deux
ca7 >plan en plongée de la libellule suspendue au-dessus de la scène-->
fig #fig.1

CAMERA2 : acteurs poussent les ballons


CAMERA3 : plan du soliste

CAMERA7 ON AIR
Survol de la libellule



CAMERA9 : acteurs poussent les ballons

2 (1.0)%+#(0.3) %# (0.5) #
com1 %.....%pousse le ballon en l'air
ca9 + plan acteur1 poussant le ballon et maintien focale sur actX-->
fig #fig.2 #fig.3 #fig.4



CAMERA9


```

3 REA neuf c'est bien avec le mouvement des gestes
4      •(0.5)
comX  •...-->
5 REA •+#neuf (.) top*#
comX  ->•poussent le ballon
ca9   ->+suit la trajectoire du ballon qui monte-->
réa   (7*9)
fig   #fig.5      #fig.6

```



5

CAMERA9



6

CAMERA9 ON AIR

```

6      (1.0)#(0.8)
fig   #fig.7

```



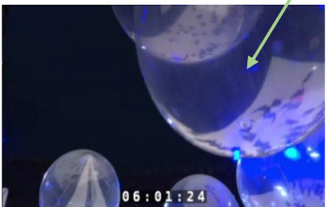
7

CAMERA9 ON AIR

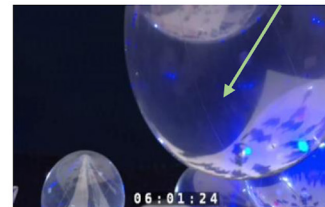
```

7 REA resdescends (.) rede+#scends la #neuf sur eux
ca9 ->+suit la trajectoire d'un autre ballon qui descend-->
fig  #fig.8      #fig.9

```

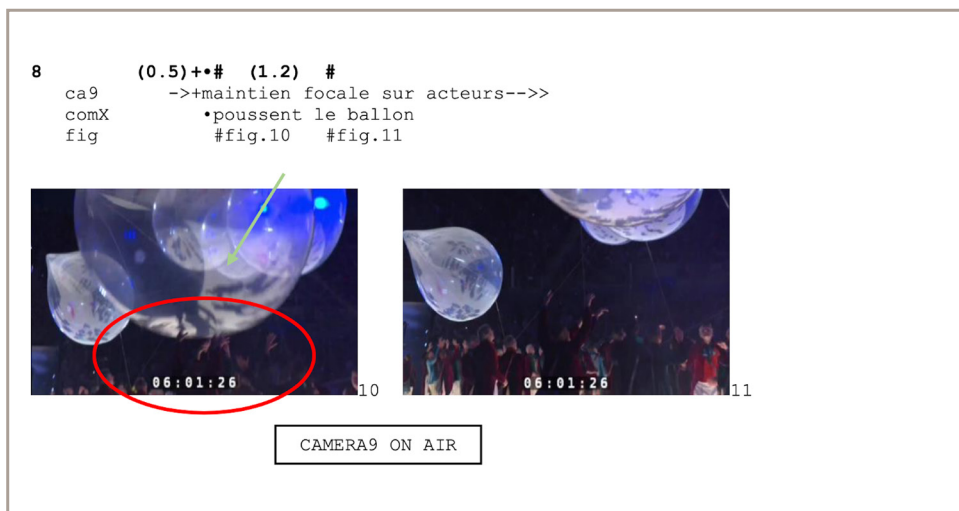


8



9

CAMERA9 ON AIR



la libellule (une danseuse) suspendue au-dessus de la scène.

En ligne 1, le réalisateur (REA) annonce le plan qu'il projette de prendre, à savoir celui de la caméra 2 (ca2). Le tour de parole permet au cadreur ca2 de se préparer au passage à l'antenne, ce tour pouvant être compris comme une instruction pour stabiliser son plan sur le phénomène qu'il est en train de filmer. La caméra 2 donne à voir les comédiens au sol, mais pas les ballons gonflables qu'ils accompagnent par leurs gestes.

Les gestes des comédiens (comX) sont en revanche bien visibles et intelligibles sur le plan de la caméra 9, plus large, qui montre les ballons descendre (fig. #2), être poussés par les acteur-ice-s (fig. 3), puis remonter (fig. 4, ligne 2). En ligne 3, le réalisateur s'oriente vers ce plan en faisant une évaluation positive : « *neuf c'est bien avec le mou-*

vement des gestes ». À travers cet évaluatif, le réalisateur rend manifeste la pertinence de cette scène et de cette perspective pour la captation. En ligne 5, il sélectionne ce plan par une annonce verbale et commute vers la caméra 9. De façon intéressante, le réalisateur réoriente donc l'activité de filmage par la saisie d'une opportunité visuelle offerte par un cadreur (ca9) : ce type d'action est tout à fait caractéristique de la documentation en temps réel et donne à voir celle-ci comme un phénomène émergent et contingent, laquelle prend appui sur les ressources locales offertes par les situations singulières.

Il est également intéressant de voir le rôle joué par les interactions langagières dans l'activité de filmage. Le tour de parole du réalisateur est à la fois orienté rétrospectivement vers l'action qui précède (à savoir le plan des comédiens dont il fournit une évaluation posi-

tive) et a des implications pour la suite de la mise en image (la sélection du plan). Si les évaluatifs sont ordinairement orientés vers un phénomène passé, les instructions sont orientées vers des phénomènes à venir. C'est ce à quoi la suite de la séquence permet de donner accès.

Alors qu'il passe à l'antenne, le plan du cadreur 9 (ca9) rend visibles les gestes des comédiens poussant le ballon vers le ciel, puis accompagne la trajectoire du ballon qui s'envole lentement (fig. 5-6-7, lignes 5-6). Le réalisateur formule une nouvelle instruction à ca9 en ligne 7 l'invitant à « *redescendre* » sur les comédiens. Le cadreur, toujours à l'antenne, profite de l'entrée au premier plan d'un autre ballon qui retombe au sol, pour l'utiliser comme objet focal afin de redescendre délicatement vers les comédiens (comX) (fig. 8-9, ligne 7). Cela lui permet de finalement produire un plan focalisé sur les comédiens sans rupture dans le cadrage.

Discussion

À partir de l'observation de ce court extrait, nous pouvons tirer une série de conséquences pour l'analyse des pratiques de documentation du spectacle vivant. En adoptant la perspective des participant·e·s engagé·e·s dans l'activité concrète de documentation (en l'occurrence réalisateur·rice et cadreur·se·s), celle-ci apparaît comme un phénomène

contingent et émergent dans le cours de la performance. Les images produites sont utilisées comme des ressources visuelles pour percevoir la scène et réorienter des choix qui ont été préalablement opérés. Les images qui restent, autrement dit celles qui sont effectivement sélectionnées par le réalisateur, sont le fruit d'un ajustement continu aux contingences de la scène et de l'activité des cadresurs. Il existe donc une historicité locale de la documentation à laquelle l'analyse vidéo des interactions peut donner accès : elle montre comment des scènes fugaces, capturées de façon incidente, peuvent être élevées au rang de trace matérielle et mémorielle de l'œuvre. Ce processus contingent peut être rapproché d'autres pratiques de réalisation en temps réel, par lesquelles des phénomènes éphémères sont transformés en des faits objectifs par l'usage de la vidéo, en particulier dans le sport (*cf.* Camus, 2021c).

Une telle conception amène à refuser l'idée d'une incommensurabilité de la performance scénique et de sa représentation par l'image. En effet, la dimension indexicale de *ce* spectacle singulier et de *cette* scène particulière est traitée par l'analyse qui met précisément en avant les micro-ajustements des participant·e·s aux actions scéniques qu'ils-elles mettent en image. Cet ajustement repose sur des savoir-faire perceptifs et une division du travail visuel permettant de voir et de rendre visible les phénomènes qui émergent et, ainsi d'adapter constamment l'organisation de la captation aux trajectoires de l'ac-

tion filmée. Accéder à ces savoir-faire pratiques et à cette organisation émergente n'est possible qu'à la condition de les replacer dans les situations dans lesquels ils prennent vie, aussi bien d'un point de vue spatial (en resituant la documentation en régie) que temporel (en montrant comment les participant-e-s se coordonnent moment par moment à la conduite de l'action sur scène).

Un tel accès à ces savoir-faire est permis par l'usage des dispositifs d'enregistrement audiovisuel mais également par des pratiques conventionnellement établies de transcription des phénomènes. La transcription, telle qu'elle est envisagée ici, repose sur des opérations de sélection adaptées à l'orientation des participant-e-s vis-à-vis des activités dans lesquelles ils sont engagés. Lorsque, comme c'est le cas ici, la performance scénique repose sur les actions d'une centaine de comédien-ne-s et figurant-e-s, mêlant danse, chants, chœurs, musique et gestualité, il serait illusoire pour le-la chercheur-euse de prétendre représenter scripturalement l'ensemble des actions constitutives d'une scène, en transcrivant la totalité des gestes accomplis par les comédien-ne-s ou la totalité des mouvements de caméras. L'intérêt de la transcription réside donc dans la sélection des phénomènes rendus pertinents par les participant-e-s eux-mêmes. Ceux-celles-ci, engagé-e-s dans des tâches ordinaires de documentation, s'orientent vers la diversité des phénomènes et, par leurs actions verbales et non-verbales (cadrage, montage), éta-

blissent la pertinence de certaines perspectives sur l'événement. À travers leur activité, ils rendent donc manifeste une certaine vision de l'événement en train de se faire. On retrouve là l'argument conversationnaliste évoqué plus haut consistant à prendre pour thème d'analyse les modes de compréhension et d'appréhension situés de l'action par les participant-e-s.

Cet article s'est proposé d'examiner les apports de l'analyse vidéo des interactions pour l'étude des pratiques de documentation du spectacle vivant. Nous avons passé en revue le mouvement théorique et technologique autorisant l'analyse empirique des pratiques de documentation, en accordant une attention particulière à la manière dont l'événement est perçu, compris et rendu visible par les différents participant-e-s travaillant à sa production audiovisuelle.

Dans la première partie de cette contribution, nous avons évoqué trois dimensions de la performance scénique – la gestualité, l'indexicalité et l'émergence – rendant problématiques sa documentation mais également son traitement analytique en sciences sociales. Nous avons tenté de montrer comment l'analyse vidéo des interactions pouvait tenir compte de ces spécificités. La documentation de la performance scénique y est comprise comme une activité située et dynamique, à travers la manière dont les participant-e-s de la régie perçoivent et font sens des conduites corporelles qui émergent.

Contre une vision dualiste faisant de la documentation et de la performance scénique deux réalités substantiellement distinctes, la performance scénique peut ainsi être comprise de façon holiste, en considérant, conformément à une conception performative de sa documentation (Auslander, 2006), que les activités de documentation participent précisément de l'accomplissement de la performance. Celles-ci sont ajustées moment par moment, par des tours de parole, des mouvements de caméras ou par le montage, à la performance scénique en train de se déployer. La mise en image est réalisée par des participant-e-s (réalisateur-riche, script-e, cadreur-euse-s, etc.) qui constituent le premier public de la représentation. Leurs actions rendent manifestes non seulement des savoir-faire professionnels incorporés, mais également une forme de compréhension ordinaire de l'événement dans sa singularité.

Conventions de transcription

Les conventions utilisées sont celles de Jefferson (2004) pour le verbal, de Mondada (2018) pour le multimodal et de Broth (2014) pour l'activité spécifique de production audiovisuelle. Les conventions multimodales sont accessibles sous leur forme détaillée à l'adresse suivante :

<https://www.lorenzamondada.net/multimodal-transcription>

Références

- Amerine R. & Bilmes J. (1988). « Following instructions ». *Human Studies*, 11(2/3), pp. 327-339.
- Auslander P. (2006). « The performativity of performance documentation ». *PAJ : A Journal of Performance and Art*, 28(3), pp. 1-10.
- Auslander P. (2008 [1999]). *Liveness: Performance in a mediatized culture*. Abingdon et New York : Routledge.
- Bateson G. & Mead M. (1942). *Balinese character: A photographic analysis*. New York : New York Academy of Sciences.
- Benjamin W. (2013 [1935]). *L'œuvre d'art à l'époque de sa reproductibilité technique*. Paris : Payot.
- Birdwhistell R.L. (1952). *Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture*. Washington DC : Department of State, Foreign Service Institute.
- Birdwhistell R.L. (1970). *Kinesics and Context*. Philadelphia : University of Pennsylvania Press.
- Brassac C., Fixmer P., Mondada L. & Vinck D. (2008). « Interweaving objects, gestures and talk in context ». *Mind, Culture and Activity*, 15(3), pp. 208-223.
- Broth M. (2009). « Seeing through screens, hearing through speakers: Managing distant studio space in television control room interaction ». *Journal of Pragmatics*, 41(10), pp. 1998-2016.
- Broth M. (2014). « Pans, Tilts, and Zooms Conventional Camera Gestures in TV Production ». In M. Broth, E. Laurier & L. Mondada (eds.), *Studies of video practices: Video at work*. New York : Routledge, pp. 63-96.
- Broth M. & Mondada L. (2013).

- « Walking away: The embodied achievement of activity closings in mobile interaction ». *Journal of Pragmatics*, 47(1), pp. 41-58.
- Camus L. (2017a). « La faute et son ralenti. Le cadrage temporel et visuel de l'action normée ». *Temporalités*, 2017/1, 25 [URL : <http://journals.openedition.org/temporalites/3635>].
- Camus L. (2017b). « Attendre le signal. Une analyse des pratiques de synchronisation de l'événement avec sa retransmission télévisée en direct ». *Réseaux*, 202-203, pp. 341-372.
- Camus L. (2021a). « Le problème de la fidélité au réel dans la réalisation télévisée du football ». *Télévision*, 12, pp. 103-120.
- Camus L. (2021b). « Le technicien invisible de l'événement médiatique ». *Questions de communication*, 39, pp. 209-238.
- Camus L. (2021c). « La production télévisuelle de l'objectivité. La révélation du hors-jeu en football ». *Réseaux*, 230, pp. 141-170.
- Cekaite A. & Mondada L. (eds.) (2020). *Touch in Social Interaction: Touch, Language, and Body*. Londres : Routledge.
- Deppermann A. (ed.) (2013). « Conversation Analytic Studies of Multimodal Interaction ». Special issue of *Journal of Pragmatics*, 46(1).
- Duranti A. (2006). « Transcripts, like shadows on a wall ». *Mind, Culture and Activity*, 13(4), pp. 301-310.
- Drew P. & Heritage J. (eds.) (1992). *Talk at work. Interaction in institutional settings*. Cambridge : Cambridge University Press.
- Erickson F. (2011). « Uses of video in social research: a brief history », *International journal of social research methodology*, 14(3), pp. 179-189.
- Evans B. & Lindwall O. (2020). « Show Them or Involve Them? Two Organizations of Embodied Instruction ». *Research on Language and Social Interaction*, 53(2), pp. 223-246.
- Garfinkel H. (1967). *Studies in ethnomethodology*. Englewood Cliffs : Prentice-Hall.
- Goodwin C. (1979). « The interactive construction of a sentence in natural conversation ». In G. Psathas (ed.), *Everyday language: Studies in ethnomethodology*. Irvington : John Wiley & Sons, pp. 97-121.
- Goodwin C. (1994). « Professional Vision ». *American Anthropologist*, 96(3), pp. 606-633.
- Goodwin C. (2000). « Action and embodiment within situated human interaction ». *Journal of Pragmatics*, 32(10), pp. 1489-1522.
- Goodwin C. (2003). « Pointing as situated practice ». In S. Kita (ed.), *Pointing: Where Language, Culture and Cognition Meet*. Hillsdale NJ : Erlbaum, pp. 217-241.
- Goodwin C. & Goodwin M.H. (1996). « Seeing as a situated activity: Formulating planes ». In D. Middleton & Y. Engestrom (eds.), *Cognition and Communication at Work*. Cambridge : Cambridge University Press.
- Goodwin C. (2006). « Human sociality as mutual orientation in a rich interactive environment ». In N. Enfield et S. Levinson (eds.), *Roots of Human Sociality*. Londres : Berg Press, p. 96125.
- Goodwin C. (2010). « Things and their embodied environments ». In L. Malafouris & C. Renfrew (eds.), *The Cognitive Life of Things*. Cambridge : McDonald Institute, pp. 103-120.
- Heath C. & Luff P. (2000). *Technology in action*. Cambridge : Cambridge University Press.
- Jefferson G. (2004). « Glossary of transcript symbols with an introduction ». In G.H.

- Lerner (ed.), *Conversation analysis. Studies from the first generation*. Amsterdam : John Benjamins Publishing Company, pp. 13-31.
- Keevallik L. (2015). « Coordinating the temporalities of talk and dance ». In A. Depermann & S. Günthner (eds.), *Temporality in Interaction*. Amsterdam/Philadelphia : John Benjamins Publishing Company, pp. 309-336.
- Kendon A. (2004). *Gesture: Visible action as utterance*. Cambridge : Cambridge University Press.
- Kendon A. (2007). « On the origins of modern gesture studies ». In S. Duncan, J. Cassell & E. Levy (eds.), *Gesture and the dynamic dimensions of language*. Amsterdam : John Benjamins Publishing Company, pp. 13-28.
- Kreplak Y. (2017). « Voir une œuvre en action. Une approche praxéologique de l'étude des œuvres ». *Cahiers du CAP*, 4, pp. 189-213.
- Kreplak Y. (2018). « On thick records and complex artworks: A study of record-keeping practices at the museum ». *Human Studies*, 41(4), pp. 697-717.
- Kreplak Y. & Mondémé C. (2014). « Artworks as touchable objects: Guiding perception in a museum tour for blind people ». In M. Neville, P. Haddington, T. Heinemann, M. Rauniomaa (eds), *Interacting with Objects: Language, Materiality, and Social Activity*. Amsterdam/Philadelphia : John Benjamins, pp. 295-318.
- Leeds-Hurwitz W. (1987). « The social history of The Natural History of an Interview: a multi-disciplinary investigation of social communication ». *Research on Language and Social Interaction*, 20, pp. 1-51.
- Licoppe C. (2009). « Recognizing mutual 'proximity' at a distance: Weaving together mobility, sociality and technology ». *Journal of Pragmatics*, 41(10), pp. 1924-1937.
- Lomax H. & Casey N. (1998). « Recording social life: Reflexivity and video methodology ». *Sociological research online*, 3(2), pp. 121-146.
- Marey E. (1899). *La chronophotographie*. Paris : Gauthier-Villars.
- McNeill D. (1985). « So you think gestures are nonverbal? » *Psychological review*, 92(3), pp. 350-371
- Mondada L. (2007a). « Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers ». *Discourse studies*, 9(2), pp. 194-225.
- Mondada L. (2007b). « Imbrications de la technologie et de l'ordre interactionnel. L'organisation de vérifications et d'identifications de problèmes pendant la visioconférence ». *Réseaux*, 2007/5, 144, pp. 141-182.
- Mondada L. (2009). « Video recording practices and the reflexive constitution of the interactional order: some systematic uses of the split-screen technique ». *Human Studies*, 32(1), pp. 67-99.
- Mondada L. (2014). « The local constitution of multimodal resources for social interaction ». *Journal of Pragmatics*, 65, pp. 137-156.
- Mondada L. (2016). « Challenges of multimodality: Language and the body in social interaction ». *Journal of sociolinguistics*, 20(3), pp. 336-366.
- Mondada L. (2017). « Le défi de la multimodalité en interaction ». *Revue française de linguistique appliquée*, 22(2), pp. 71-87.
- Mondada L. (2018). « Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality ». *Research on Language and Social Interaction*, 51(1), pp. 85-106.
- Mondada L. (2019). « Contemporary issues in conversation analysis: Embodiment and

materiality, multimodality and multisensoriality in social interaction ». *Journal of Pragmatics*, 145, pp. 47-62.

Mondada L. (2021). « The Natural History of an Interview and the emergence of video-based multimodal studies of social interaction ». In J. McElvenny & A. Ploder (eds), *Holisms of communication: the early history of audio-visual sequence analysis*. Berlin : Language Science Press.

Mondada L., Bouaouina S.A., Camus L., Gauthier G., Svensson H. & Tekin B. (2021). « The local and filmed accountability of sensorial practices: The intersubjectivity of touch as an interactional achievement ». *Social Interaction. Video-Based Studies of Human Sociality*, 4(3) [URL : <https://tidsskrift.dk/socialinteraction/article/view/128160>].

Nevile M., Haddington P., Heinemann T., & Rauniomaa M. (eds.) (2014). *Interacting with objects: Language, materiality, and social activity*. Amsterdam : John Benjamins Publishing Company.

Ochs E. (1979). « Transcription as theory ». In E. Ochs & B. Schieffelin (eds.), *Developmental pragmatics*. New York : Academic, pp. 43-72.

Phelan P. (1993). *Unmarked: The Politics of Performance*. London/New York : Routledge.

Rawls A.W. (2008). « Harold Garfinkel, ethnomethodology and workplace studies ». *Organization studies*, 29(5), pp. 701-732.

Sacks H. (1992). *Lectures on Conversation*. Oxford : Basil Blackwell.

Sacks H., Schegloff E.A. & Jefferson G. (1974). « A Simplest Systematics for the Organization of Turn-Taking for Conversation ». *Language*, 50(4), pp. 696-735.

Sacks H. & Schegloff E.A. (2002). « Home position ». *Gesture*, 2(2), pp. 133-146.

Suchman L. (1997). « Centers of coordination: A case and some themes ». In L.B. Resnick, R. Säljö, C. Pontecorvo & B. Burge (eds.), *Discourse, tools and reasoning*. Berlin : Springer, pp. 41-62.

Suchman L. (2007 [1987]). *Human-machine reconfigurations: Plans and situated actions*. Cambridge : Cambridge University Press.

Tanferri M., Waeber S. & Vinck D. (à paraître). « “Peut-être que je peux retomber sur ses pattes ?” Agentivités multiples dans la constitution d’un DVD de spectacle à partir de sa captation en direct ». *SociologieS*.

Vinck D. (2019). *Les métiers de l’ombre de la Fête des vigneronns*. Lausanne : Antipodes.

Vinck D. & Tanferri M. (2020). « Taking the metaphor of theatre seriously: from staging a performance toward staging design and innovation ». In C. Clausen, D. Vinck, S. Petersen, J. Dorland (eds.). *Staging Collaborative Design and Innovation: An Action-Oriented Participatory Approach*. Cheltenham : Edward Elgar, pp. 232-251.

Vinck D., Waeber S. & Tanferri M. (2021). « Produire un son “naturel” : imbrication et différenciation de l’analogique et du numérique pour créer l’univers sonore d’un spectacle de grande envergure ». *Tsantsa*, 26, pp. 51-70.

Waeber S., Tanferri M. & Vinck D. (2021). « Comparer l’invisibilisation des techniques dans le travail de narration audio-visuelle d’un grand spectacle : un siècle de production filmique sur la Fête des Vigneronns ». *Ethnographiques.org*, 41 [URL : https://www.ethnographiques.org/2021/Waeber_Tanferri_Vinck].

Winkin Y. (2001 [1996]). *Anthropologie de la communication : de la théorie au terrain*. Paris : Le Seuil.