



<https://bulletin.vals-asla.ch> · e-ISSN: 2813-7973 · p-ISSN: 1023-2044

De Stefani, E. (2007). Regards sur la langue. Les données vidéo dans la recherche linguistique. Introduction. *Bulletin suisse de linguistique appliquée*, 85, 1-9.

<https://doi.org/10.26034/ne.vals-asla.2007.0824>

Cet article est publié sous la licence *Creative Commons Attribution 4.0 International* (CC BY):

<https://creativecommons.org/licenses/by/4.0>



© Elwys De Stefani, 2007

Regards sur la langue. Les données vidéo dans la recherche linguistique

Introduction

Elwys DE STEFANI

Depuis quelques années, on observe dans le domaine linguistique un intérêt croissant pour les phénomènes non verbaux de la communication humaine et pour la manière dont ceux-ci s'articulent aux activités langagières des acteurs sociaux. Les progrès technologiques récents et la démocratisation des caméscopes ont sans doute contribué à l'engouement des chercheurs pour les différents modes de communication que sont la gestualité, les regards, la mimique, le mouvement dans l'espace et la manipulation des objets. La possibilité d'accéder aux aspects non verbaux à travers les enregistrements vidéo se reflète non seulement dans la multiplication des études linguistiques consacrées à ces thématiques, mais également dans la remise en question de certaines idées et notions ancrées dans la tradition épistémologique de la linguistique. Ainsi, la focalisation sur la *parole* et sa constitution en objet de recherche ont mené inévitablement à un certain logocentrisme¹ qui consiste à attribuer à la production verbale un rôle prédominant dans la communication. Le travail sur les données vidéo a permis de définir un nouveau champ de recherche qui se caractérise par une approche holistique de la communication humaine et que l'on identifie par le terme de *multimodalité*. La prise en compte des pratiques multimodales dans l'interaction oblige les chercheurs notamment à reconsidérer certains concepts largement utilisés en linguistique (comme les notions de *contexte* et de *référence*). Ce numéro spécial interroge dès lors les retombées de l'analyse de données vidéo sur la recherche linguistique.

Si la possibilité de récolter facilement des corpus vidéo a sans doute accéléré la recherche sur les activités non verbales, l'importance de ces pratiques avait suscité de l'intérêt bien avant l'émergence des dispositifs d'enregistrement. Ainsi, dans son traité de rhétorique *Institutio oratoria* (1^{er} siècle ap. J.-C.) Quintilien avance une description détaillée des gestes – parmi lesquels il compte aussi les mouvements du corps, la posture, la mimique et le regard

¹ Nous utilisons la notion de *logocentrisme* pour indiquer la tendance de nombreux chercheurs à se limiter à l'analyse de la communication verbale aux dépens des autres modalités communicatives. C'est une utilisation différente de celle que l'on retrouve habituellement en philosophie, où le concept de *logocentrisme* est employé pour indiquer la propension de la recherche philosophique à privilégier l'oralité à l'écrit (cf. Derrida, 1968).

(Kendon, 2004: 17-19). La gestualité a intéressé également de nombreux esprits pendant la Renaissance (cf. Kendon, 2004: 20-28) et dans les époques successives: à titre d'exemple, nous rappelons le traité *L'Arte de' Cenni* (Venise, 1616) de Giovanni Bonifacio, dans lequel l'auteur décrit minutieusement les mouvements que les êtres humains peuvent faire avec leur corps². C'est à un autre italien, Andrea de Jorio, que nous devons une étude approfondie de la gestualité des Napolitains. Son ouvrage *La mimica degli antichi investigata nel gestire italiano* (1832)³ contient de nombreuses réflexions qui n'ont pas perdu leur validité à la lumière de la recherche actuelle, comme par exemple l'insistance sur l'analyse contextualisée des gestes qui tient compte des phénomènes verbaux et non verbaux co-occurents.

En linguistique, il faudra attendre les années cinquante du XX^e siècle avant que le comportement non verbal des êtres humains suscite l'intérêt des chercheurs: le linguiste George L. Trager se réfère à la gestualité en utilisant le terme *paralangage* (Trager, 1958) alors que quelques années plus tard l'anthropologue Edward T. Hall introduit la notion de *proxémique* pour indiquer la relation que les individus entretiennent avec l'espace (Hall, 1966). Dans un ouvrage écrit à quatre mains, les auteurs soulignent que l'étude de la communication ne doit pas se limiter à l'analyse du langage verbal, mais qu'elle doit tenir compte aussi des altérations vocales, des mouvements physiques ainsi que des manipulations d'objets (Hall & Trager, 1953). L'intérêt pour les gestes s'est développé parallèlement en psychologie, comme le montre, par exemple, l'article amplement cité d'Ekman & Friesen (1969), dans lequel les auteurs proposent une classification des modalités communicatives non verbales.

A cette époque, les chercheurs ne recouraient que rarement aux enregistrements vidéo, même si la technologie nécessaire était déjà disponible. Les pionniers de la documentation filmée – les anthropologues et les ethnographes – avaient commencé vers la fin du XIX^e siècle à recueillir des corpus visuels. Le premier film ethnographique est tourné aux îles de Torres Straits par Albert C. Haddon en 1898 (Mondada, 2006: 316, n. 2). Gregory Bateson et Margaret Mead commencent à se servir de séquences filmées à Bali et en Nouvelle-Guinée à partir des années trente, contribuant ainsi considérablement au développement de l'anthropologie visuelle⁴. L'un

² Le titre complet de l'ouvrage – dont l'auteur est issu d'une noble famille de Rovigo (Vénétie) – est *L'Arte de' Cenni con la quale formandosi favella visibile, si tratta della muta eloquenza che non è altro che un facondo silenzio*.

³ L'ouvrage – qui a été traduit en anglais par Adam Kendon (de Jorio, 2000) – a suscité récemment l'intérêt de nombreux chercheurs (cf. Gadeau, 1986; Kendon, 1995).

⁴ Quelques années auparavant, en 1922, le cinéaste Robert Flaherty avait réalisé un film de 79 minutes sur une famille Inuit.

des anthropologues les plus influents, Ray L. Birdwhistell, a ensuite également utilisé des données vidéo pour analyser la *kinésique* des êtres humains (cf. Birdwhistell, 1952)⁵. Mais les disciplines anthropologiques et ethnographiques n'étaient pas les seules à utiliser des séquences filmées à des fins analytiques: ainsi, les psychologues sociaux Kurt Lewin et Arnold Gesell réalisent une série de films dans les années vingt et trente et proposent des réflexions méthodologiques pour l'approche qu'ils baptisent *cinemanalysis* (cf. Knoblauch, Schnettler & Raab, 2006: 17-18). Plus tard, grâce aux travaux d'Albert E. Scheflen sur la posture des patients pendant les séances psychothérapeutiques, l'analyse vidéo fait son entrée dans la psychiatrie (Scheflen, 1964)⁶.

Dans le domaine des sciences du langage, l'analyse de données vidéo a surtout été appliquée à l'étude de la conversation et des mécanismes permettant de mener à bien la rencontre sociale. Ce sont notamment des chercheurs influencés par l'ethnométhodologie et par l'analyse conversationnelle qui récoltent des données à l'aide d'enregistrements vidéo. Aux Etats-Unis, Marjorie et Charles Goodwin réalisent et analysent à partir des années soixante-dix des enregistrements d'interactions recueillies dans leurs habitats naturels. Les premiers résultats mènent Charles Goodwin à souligner l'importance du regard dans l'échange interactionnel en général et dans l'organisation des tours de parole en particulier (Goodwin, 1980). Parallèlement, Christian Heath commence à recueillir des données vidéo en Angleterre en vue de sa thèse de doctorat sur les consultations médicales (Heath, 1986). C'est un premier pas vers la définition d'un nouveau domaine de recherche qui se penche sur les contextes professionnels et qui intéresse tout particulièrement les chercheurs s'inscrivant dans la linguistique appliquée: les nombreuses recherches issues des *studies of work* (Drew & Heritage, 1992) ou des *workplace studies* (Luff, Hindmarsh & Heath, 2000) thématisent en effet la manière dont les activités non verbales s'articulent aux pratiques verbales des participants.

Les articles contenus dans ce numéro sont tous héritiers, à différents degrés, des traditions scientifiques que nous avons décrites ci-dessus. Les auteurs partagent un intérêt commun pour les aspects multimodaux de la communication humaine, telle qu'elle se manifeste tant dans les domaines de prédilection de la linguistique appliquée – comme les situations professionnelles (Filliettaz, Galatolo & Traverso) et l'enseignement en classe

⁵ D'autres anthropologues se sont basés sur des séquences filmées. Ainsi, David Efron – élève de Franz Boas – a mené une étude comparative de l'utilisation des gestes parmi les Juifs et les Italiens du sud immigrés à Manhattan (Efron, 1941).

⁶ Dans le domaine de la psychologie, il faut également citer les travaux récents de McNeill (1992).

(Pitsch, Pepin & Steinbach, Veronesi) –, que dans les interactions dyadiques (De Stefani & Mondada, Strebel) et dans la communication télévisuelle (Bovet, Malbois). Leurs analyses se fondent sur des enregistrements d'interactions naturelles (c.-à-d. non expérimentales) réalisés à l'aide d'un caméscope ou alors sur la mise en image des émissions de télévision. Si le travail sur les corpus visuels constitue le point commun des articles, les lecteurs constateront une certaine variation dans la présentation des données. L'insertion dans le texte des séquences analysées – sous forme de transcriptions ou d'une série d'images – pose, en effet, des problèmes pratiques (chaque transcription, chaque image allonge l'article) et méthodologiques. Quels sont les phénomènes non verbaux dont il faut tenir compte dans la transcription? Comment les noter sans compromettre la lisibilité du texte? Dans quels cas la reproduction d'une image fixe est-elle utile? Est-il nécessaire d'anonymiser les images? Les auteurs ont répondu de manière différente à ces questions qui émergent inévitablement dans les processus de récolte, d'analyse et de publication de données vidéo. Le traitement particulier que chaque auteur réserve aux données émerge à travers la manière dont celles-ci sont présentées et intégrées dans les articles. Si nous avons renoncé à uniformiser les normes de transcription et de présentation, c'est parce que nous sommes convaincus que le traitement des données s'articule aux choix méthodologiques et analytiques de chaque auteur et que, pour cette raison, toute intervention éditoriale aurait altéré l'intégrité scientifique des contributions.

Les textes ont été répartis en quatre sections en raison des différents terrains étudiés. La première partie traite des interactions en milieu professionnel et réunit deux études qui montrent comment la manipulation des objets joue un rôle capital dans l'interaction des participants (apprentis et professionnels).

Laurent Filliettaz s'intéresse à l'analyse des aspects sensoriels qui sont constitutifs des processus d'apprentissage. Basant son analyse sur un corpus d'interactions se déroulant entre des apprentis automaticiens et leur moniteur, l'auteur étudie le processus d'enseignement et d'apprentissage d'une pratique professionnelle précise, à savoir le trempage de l'acier. L'auteur montre comment la recherche basée sur des données vidéo permet d'étudier les processus d'apprentissage des connaissances et des pratiques professionnelles en tenant compte des expériences multimodales et multisensorielles qui se déploient de manière séquentiellement structurée, tant au niveau temporel que dans la spatialité (les participants investissant différents espaces au fur et à mesure que l'interaction avance). Ainsi, l'article montre qu'une analyse basée uniquement sur les productions verbales des participants n'arriverait pas à saisir la manière dont le transfert des connaissances est organisé par le moniteur et par les apprentis. Il est dès lors important de tenir compte d'autres modalités communicatives – comme

l'orientation du regard, la gestualité, la manipulation des objets – et, tout particulièrement, des expériences sensorielles – visuelles, tactiles, olfactives, gustatives – auxquelles les participants recourent visiblement.

S'intéressant au déroulement des activités dans un restaurant, *Renata Galatolo* et *Véronique Traverso* analysent la manière dont les différents acteurs sociaux (re)constituent et utilisent les bons de commande qu'une serveuse rédige lorsqu'elle prend une commande. Les auteures montrent comment, loin d'être uniquement un objet physique qui sert de support pour l'écriture, ce document constitue un artefact cognitif que les différents acteurs utilisent de manières distinctes. La structuration verticale des bons de commande – indiquant une séquentialité précise qui s'articule en *entrée*, *premier plat (pâtes)*, *plat principal*, *dessert* – rend visible l'ordre de la consommation de la part du client, mais elle ne revêt pas la même pertinence dans la cuisine, où les cuisinières préparent simultanément, p.ex., les différents premiers plats qui figurent sur les bons de commande que la serveuse leur a transmis. L'orientation diversifiée sur les bons de commande est rendue visible, notamment, dans l'interaction verbale qui se déploie entre les participants lorsqu'ils se réfèrent aux commandes en cours. De manière intéressante, les acteurs se réfèrent systématiquement aux bons de commande en verbalisant une information précise, à savoir le nom du premier plat qui y est inscrit. Les auteures montrent que cette manière de se référer aux bons de commande n'est pas aléatoire, mais qu'elle est liée étroitement au déploiement de l'activité: les participants accordent aux premiers plats (c.-à-d. aux pâtes) effectivement une saillance particulière, du fait que la préparation de ceux-ci demande une organisation temporelle très précise. L'article montre dès lors comment les pratiques de référencement s'articulent aux contingences locales et actionnelles de l'interaction.

La deuxième partie se penche sur l'enseignement en classe, avec trois articles qui s'intéressent respectivement à l'utilisation du tableau noir, au rôle que jouent les émotions dans les processus d'apprentissage et aux déplacements spatiaux de l'enseignant pendant le cours.

Le recours aux données vidéo permet à *Karola Pitsch* d'étudier le rôle fonctionnel que les objets – plus spécifiquement le tableau noir – jouent dans l'interaction en tant que ressources communicatives. Son analyse détaillée montre que les processus d'apprentissage et de constitution du savoir ne se font pas uniquement à travers des pratiques verbales, mais que les objets matériels y jouent un rôle central. En s'appuyant sur le concept d'*objet intermédiaire* comme il a été défini dans la sociologie du travail de tradition française, l'auteure montre que les inscriptions réalisées par l'enseignant sur le tableau noir sont utilisées par celui-ci ainsi que par les élèves comme des points de repère qui contribuent à l'ancrage et au développement de la communication. De manière intéressante, l'auteure montre qu'une même

inscription peut acquérir des sens différents durant l'interaction: l'enseignant peut en effet s'y référer – p.ex. par des gestes de pointage – à différents moments séquentiels et en leur attribuant des significations divergentes. On voit, en d'autres termes, que le sens n'est pas *inscrit* dans les mots qui figurent sur le tableau, mais qu'il est constitué séquentiellement et interactionnellement.

Nicolas Pepin et *Fee Steinbach* analysent une pratique d'enseignement répandue en classes de langue, à savoir le travail en groupes. Ils examinent l'interaction qui se déploie entre trois élèves d'une école bâloise qui ont reçu la consigne de préparer un entretien en français. Les auteurs s'intéressent notamment aux problèmes de langue que les participantes manifestent et qu'elles résolvent de manière collaborative. C'est durant ces phases de résolution conjointe d'un problème que le processus d'apprentissage devient visible. La stabilisation de formes linguistiques correctes se fait aussi en recourant à la manifestation d'un engagement émotionnel qui intéresse tout particulièrement les auteurs. L'expression d'un engagement émotionnel intervient de façon récurrente lors du passage du cadre scolaire (que les élèves rendent reconnaissable à travers leur engagement dans la résolution d'une tâche qui leur a été imposée) au cadre relationnel au sein duquel se déploie la conversation spontanée (en suisse allemand) des jeunes filles étudiées. Dès lors, les auteurs montrent que l'apprentissage de formes linguistiques ne se fait pas seulement à travers la langue, mais qu'interviennent également des comportements multimodaux (gestes, mouvements du corps) qui contribuent à rendre reconnaissable et à traiter un problème linguistique.

Daniela Veronesi s'attache à un aspect de la communication en classe qui avait été négligé jusqu'ici par la recherche, à savoir l'analyse des déplacements des enseignants pendant les cours. Elle montre que le comportement proxémique des enseignants n'est pas aléatoire, mais qu'il s'articule à leurs activités verbales, tant monologiques (leçons, exposés) que dialogiques (interactions avec les étudiants). En particulier, l'auteure observe que dans les phases monologiques le déplacement spatial coïncide souvent avec le début ou la fin d'une activité communicative (p.ex. l'annonce d'un nouveau sujet, l'initiation d'une synthèse du discours précédent, etc.). En revanche, dans un contexte de discussion avec les élèves, les enseignants se déplacent fréquemment à des moments saillants de l'interaction, p.ex. lorsqu'ils posent une question ou – pour utiliser des termes plus conversationnels – lorsqu'ils initient une paire adjacente qui projette la participation active des élèves. L'étude des déplacements spatiaux par les acteurs sociaux fera également l'objet de l'article suivant qui inaugure la troisième section, dédiée aux interactions entre deux personnes.

Elwys De Stefani et *Lorenza Mondada* étudient les phases de réorientation de couples en mouvement sur la base de deux corpus différents: le premier documente les activités des couples qui font leurs courses dans un supermarché, le deuxième rend compte des interactions qui se déroulent dans l'habitacle d'une voiture. Dans ces situations, l'analyse de la réorientation spatiale des participants doit tenir compte du paysage qui évolue constamment avec le déplacement des acteurs sociaux. Les auteurs montrent que la recherche collaborative d'un itinéraire s'articule étroitement à la constitution du couple qui devient publiquement visible à travers les phases d'orientation commune des membres qui le composent. L'article propose dès lors une description située et praxéologique des pratiques de référence spatiale. Ce faisant, il met en évidence le caractère dynamique des descriptions d'itinéraires étudiées: ainsi, les éléments déictiques qui émergent dans l'interaction acquièrent une plasticité considérable qui rend compte du caractère éminemment indexical de la conversation et qui remet en question la description statique que certains courants de la linguistique réservent aux déictiques.

L'article d'*Ignaz Strebel* montre comment l'utilisation de données vidéo permet de décrire sous un nouvel angle une manière répandue de récolter des données – c.-à-d. les interviews basées sur un questionnaire. En adoptant une approche inspirée de l'ethnométhodologie et de la sociologie des sciences, l'auteur montre que le transfert des données empiriques de l'endroit "naturel" de leur production au laboratoire scientifique dans lequel elles sont analysées est fondé sur une adaptation, une négociation continue des catégories pertinentes pour la tâche que les participants (intervieweur et interviewé) sont en train d'accomplir. Le questionnaire, en tant que document qui contribue à structurer l'interview et qui sert à la fois de gabarit accueillant les réponses des interviewés, est un artefact important – son existence permet aux passants, p.ex., de catégoriser l'individu qui le tient entre ses mains comme étant un "intervieweur". Après avoir brièvement analysé le passage entre la phase d'ouverture d'une interview (l'abordage des passants) et la formulation de la première question, l'auteur consacre une partie importante de son article à la problématique qui émane du caractère général et sous-spécifié des questions figurant dans le questionnaire. Il montre par exemple que des expressions déictiques comme *ici* peuvent se révéler problématiques – même quand les participants ne sont pas en mouvement.

La dernière section accueille deux articles qui analysent les émissions télévisées d'un point de vue ethnométhodologique, en questionnant les rôles que ces données peuvent jouer pour la description et l'analyse des pratiques langagières.

Alain Bovet étudie la manière dont un débat télévisé est réalisé et diffusé en direct. S'intéressant à l'articulation entre l'activité interactionnelle sur le plateau

et la mise en image effectuée par l'équipe de réalisation, l'auteur montre comment les techniciens de l'image utilisent les tours de parole des participants (et notamment leur *projectability*) en tant que ressources pour organiser la mise en image. De ce point de vue, la mise en image opérée par l'équipe de réalisation rend visible la manière dont les techniciens analysent l'activité qui se déroule sur le plateau. Par la mise en image, ils rendent également visibles les changements du cadre de participation qui émergent au cours de l'émission, ainsi que les identités relatives des participants: c'est surtout le rôle de l'animatrice du débat que l'auteur analyse de ce point de vue, en soulignant qu'elle participe au débat non seulement en tant que modératrice (impartiale), mais aussi en tant qu'actrice sociale qui peut être tenue responsable pour ses propos et qui, par conséquent, peut être louée, critiquée etc. En somme, l'article montre que la mise en image d'un débat télévisé n'est pas simplement une représentation "objective" de ce qui se passe sur le plateau, mais que c'est à travers elle que l'équipe de réalisation rend intelligible l'interaction qui s'y déroule, par le biais d'une analyse immédiate et continue.

Le dernier article, signé *Fabienne Malbois*, analyse l'émergence d'une identité de genre dans le reportage d'un journal télévisé français. L'auteure montre notamment quelles sont les ressources langagières et de mise en image que les journalistes et les équipes de réalisation mobilisent pour rendre pertinente l'appartenance à une catégorie sexuelle des personnes présentées dans le reportage. Ce faisant, elle observe que ni l'utilisation de formes verbales féminines/masculines, ni les indicateurs physiques perceptibles sont suffisants pour affirmer la pertinence des catégories sexuelles dans l'échange verbal (et, plus généralement, dans l'activité sociale). Ainsi, l'appartenance à une catégorie de genre est construite dans et rendue pertinente par le discours, tant dans l'interaction spontanée que dans le discours planifié et unidirectionnel des reportages télévisés. Comme le rappelle l'auteure, la pertinence de la catégorie sexuelle varie au sein d'un même discours et peut s'articuler à d'autres catégories rendues pertinentes, comme p.ex. l'appartenance à une catégorie de métier: plus précisément, l'auteure étudie la manière dont le terme d'*agricultrice* dévoile deux facettes identitaires, l'une professionnelle, l'autre sexuelle.

Les contributions réunies dans ce numéro spécial témoignent des nombreux champs d'application qui se prêtent à une analyse fondée sur l'enregistrement de données vidéo. Ainsi, une activité de recherche remarquable est en train de se développer dans un terrain traditionnel de la linguistique appliquée, l'enseignement en classe. Dans une perspective méthodologique, l'interaction en classe n'est qu'une situation naturelle parmi d'autres: dès lors, les analyses menées sur les terrains professionnels ou publiques qui complètent ce

numéro spécial contribuent à représenter la variété des thèmes que la linguistique appliquée peut aujourd'hui aborder.

D'un point de vue épistémologique, la prise en compte des aspects multimodaux de la communication permet de repenser certaines catégories linguistiques, sans diminuer pour autant l'importance que la parole revêt dans l'interaction, mais, au contraire, en l'enrichissant par une analyse contextuelle et praxéologique.

BIBLIOGRAPHIE

- Birdwhistell, R. L. (1952): Introduction to kinesics. An annotation system for the analysis of body motion and gesture. Louisville (University of Louisville).
- de Jorio, A. (2000): Gesture in Naples and gesture in classical antiquity. A translation of *La mimica degli antichi investigata nel gestire napoletano* (1832). Bloomington (Indiana University Press).
- Derrida, J. (1968): La différance. In: Foucault, M., Barthes, R. & Derrida, J. (éds.): *Théorie d'ensemble*. Paris (Editions du Seuil). 41-66.
- Drew, P. & Heritage, J. (eds.) (1992): *Talk at work. Interaction in institutional settings*. Cambridge (Cambridge University Press).
- Efron, D. (1941): *Gesture and environment*. New York (King's Crown Press).
- Ekman, P. & Friesen, W. V. (1969): The repertoire of nonverbal behavior. Categories, origins, usage and coding. In: *Semiotica*, 1, 49-98.
- Gadeau, P. (1986): Regards sur les classiques. Andrea de Jorio. In: *Geste et Image*, 6/7, 215-225.
- Goodwin, C. (1980): Restarts, pauses and the achievement of a state of mutual gaze at turn-beginning. In: *Sociological Inquiry*, 50, 272-302.
- Hall, E. T. (1966): *The hidden dimension*. Garden City NY (Doubleday).
- Hall, E. T. & Trager, G. L. (1953): *The analysis of culture*. Washington DC (American Council of Learned Societies).
- Heath, C. (1986): *Body movement and speech in medical interaction*. Cambridge (Cambridge University Press).
- Kendon, A. (1995): Andrea de Jorio. The first ethnographer of gesture. In: *Visual Anthropology*, 7, 375-394.
- Kendon, A. (2004): *Gesture. Visible action as utterance*. Cambridge (Cambridge University Press).
- Knoblauch, H., Schnettler, B. & Raab, J. (2006): Video-Analysis. Methodological aspects of interpretive audiovisual analysis in social research. In: Knoblauch, H., Schnettler, J., Raab, J. & Soeffner, H.-G. (eds.). *Video analysis. Methodology and methods*. Frankfurt am Main (Peter Lang), 9-26.
- Luff, P., Hindmarsh, J. & Heath, C. (eds.) (2000): *Workplace studies. Recovering work practice and informing system design*. Cambridge (Cambridge University Press).
- McNeill, D. (1992): *Hand in mind*. Chicago (University of Chicago Press).
- Mondada, L. (2006): La pertinenza del dettaglio. Registrazione e trascrizione di dati video per la linguistica interazionale. In: Bürki, Y. & De Stefani, E. (a cura di). *Trascrivere la lingua. Dalla filologia all'analisi conversazionale*. Bern (Peter Lang), 313-344.
- Schefflen, A. E. (1964): The significance of posture in communication systems. In: *Psychiatry*, 27, 316-331.
- Trager, G. L. (1958): Paralanguage. A first approximation. In: *Studies in Linguistics*, 13, 1-12.