

CONFÉRENCE • ASSOCIATION SUISSE DE LINGUISTIQUE APPLIQUÉE  
CONGRESSO • ASSOCIAZIONE SVIZZERA DI LINGUISTICA APPLICATA  
CONFERENCE • SWISS ASSOCIATION FOR APPLIED LINGUISTICS  
TAGUNG • VEREINIGUNG FÜR ANGEWANDTE LINGUISTIK IN DER SCHWEIZ  
❄️ UNIVERSITÄT BASEL 6.-8. JUNI 2018

AVI▶E●

TURN

IN LINGUISTICS?  
METHODOLOGIE • ANALISI • APPLICATIONS

ABSTRACTS



# TABLE OF CONTENTS

<b>Foreword</b>	<b>  4</b>
<b>Committees</b>	<b>  5</b>
<b>Conference Programme</b>	<b>  6</b>
<b>Keynote Lectures</b>	<b>  12</b>
<b>Individual Papers</b>	<b>  23</b>
<b>Word cloud</b>	<b>  82</b>
<b>Sponsors</b>	<b>  83</b>



# FOREWORD



## A VIDEO TURN IN LINGUISTICS?

University of Basel

06-08/06/2018

<https://vals-asla2018.philhist.unibas.ch>



We are very proud of presenting a rich program, dense with international contributions displaying the vitality of the field of video studies and the diversity of topics it covers.

The conference is motivated by the evidence of an increasing use of video data and methodologies in the human and social sciences in general, and linguistics in particular – and at the same time by the observation that video production and video analysis still need to be fully reflected both conceptually and analytically.

The use of video is spreading in all domains of the human and social sciences, and particularly in linguistics and applied linguistics. This « video turn » has several methodological, theoretical, and applied consequences. Methodologically, video allows the researcher to collect data of increasing complexity, in order to document not only the gestural and embodied dimension of spoken language but also the visual dimension of written language, as well as the visual and embodied aspects of technologically mediated communication. Crucially, the use of video raises theoretical issues, inviting to re-think language, action, cognition, culture and social order, by recognizing the fundamental importance of embodiment and more generally of multimodality. The applied consequences are important too, since video offers new ways of communicating, sharing, and circulating scientific results, in particular with people filmed and targeted by video research.

Despite a real boom experienced by video these last decades, a lot of theoretical and analytical issues, methodological and technological problems, and questions about the usability and impact of video data remain. Often the potentialities of video remain underestimated and under-exploited. This conference aims at showcasing and reflecting about the diversity of uses of video in linguistics and neighboring disciplines.

In particular, it gathers together scholars contributing to the following topics:

- a) multimodal analysis through the use of video recordings
- b) methodological issues concerning the use of video data
- c) applied issues concerning the use of videos for communicating with the public

We are very happy to welcome you in Basel and very excited about discussing these issues during the conference.

Lorenza Mondada

Chair of the conference

On behalf of the organizing committee

## ORGANIZING COMMITTEE

### Chair

Lorenza Mondada | Universität Basel

### Members

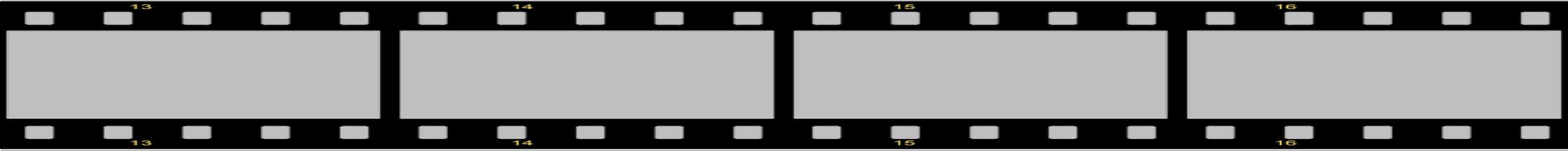
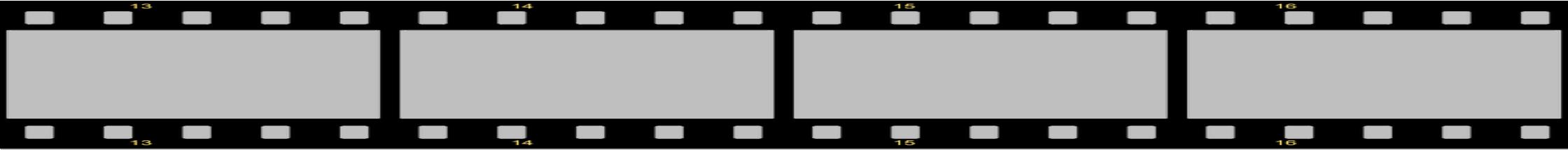
Eva Askari | Universität Basel  
Sofian A. Bouaouina | Universität Basel  
Anne-Sylvie Horlacher | Universität Basel  
Sara Keel | Universität Basel  
Sara Merlino | Universität Basel  
David Monteiro | Universität Basel  
Hanna M. Svensson | Universität Basel  
Burak S. Tekin | Universität Basel

## SCIENTIFIC COMMITTEE

Alain Bovet | HE Arc, Neuchâtel  
Rebekka Bratschi | Bundeskanzlei Zentrale Sprachdienste, Bern  
Mathias Broth | Linköpings universitet  
Barry Brown | Stockholms universitet  
Marcel Burger | Université de Lausanne  
Elwys De Stefani | Universiteit Leuven  
Daniel Elmiger | Université de Genève  
Giolo Fele | Università di Trento  
Laurent Fillietaz | Université de Genève  
Esther González-Martínez | Université de Fribourg & PH Bejune, Bienne  
Marianne Gullberg | Lunds universitet  
Heiko Hausendorf | Universität Zürich  
Judith Holler | MPI, Nijmegen  
Alain Kamber | Université de Neuchâtel  
Ulla Kleinberger | ZHAW, Winthertur  
Hubert Knoblauch | TU Berlin  
Eric Laurier | University of Edinburgh  
Martin Luginbühl | Universität Basel  
Johanna Miecznikowski Fünfschilling | Università della Svizzera Italiana, Lugano  
Karola Pitsch | Universität Duisburg-Essen  
Reinhold Schmitt | IdS Mannheim  
Anja Stukenbrock | Université de Lausanne  
Marta Zampa | ZHAW, Winthertur

**6 JUNE | WEDNESDAY MORNING**

08:00 - 09:15	Registration		
09:15 - 09:30	Opening		
09:30 - 10:30	<p><b>Keynote Lecture   Hörsaal -101</b>  <b>Arnulf Deppermann   Inst. Deutsche Sprache, Mannheim</b>  <i>Why does linguistics need video-analysis and video-corpora?</i></p>		
10:30 - 11:00	Coffee Break		
11:00 - 12:30	<p><b>Seminarraum -201</b></p> <p><b>Adrian Yip<sup>1,2</sup> &amp; Veronika Schoeb<sup>1</sup></b>  <sup>1</sup>Hong Kong Polytechnic U., <sup>2</sup>Queen Mary U.  <i>"Remember what exercises we did?": Responses to instructions during exercise training in physiotherapy</i></p>	<p><b>Seminarraum +207</b></p> <p><b>Biagio Ursi</b>                  Lab ICAR Lyon  <i>Le refus suite à une offre pendant des repas entre ami(e)s : une étude séquentielle et multimodale</i></p>	<p><b>Hörsaal -101</b></p> <p><b>Lorenza Mondada</b>                  U. Basel  <i>On filming and transcribing tasting: Public visibility and intersubjectivity</i></p>
	<p><b>Sara Merlino</b>                  U. Basel  <i>Multimodal cueing practices for the recovery of language in aphasia speech therapy</i></p>	<p><b>Ioana-Maria Stoenica</b>                  U. Neuchâtel  <i>A multimodal analysis of the use of relative clauses in French talk-in-interaction</i></p>	<p><b>Thomas Hillman<sup>1</sup>, Oskar Lindwall<sup>1</sup> &amp; Barry Brown<sup>2</sup></b>  <sup>1</sup>U. Gothenburg, <sup>2</sup>Stockholm U.  <i>Choosing and using instructional videos</i></p>
	<p><b>Karola Pitsch</b>                  U. Duisburg-Essen  <i>Address, reference and reciprocity: Participation in training of mass-casualty incidents</i></p>	<p><b>Evelyne Berger, Simona Pekarek Doehler &amp; Klara Skogmyr Marian</b>                  U. Neuchâtel  <i>From multi-word expression to interaction-organizational resource: A multimodal analysis of 'comment on dit' (how do you say) in second language interaction over time</i></p>	<p><b>David Monteiro</b>                  U. Basel  <i>Adjusting mobile camera operation and participation to interactional and material contingencies in doing CA research</i></p>
12:30 - 14:00	Lunch		



6 JUNE | WEDNESDAY AFTERNOON

	Seminarraum -201	Seminarraum +207	Hörsaal -101
14:00 - 16:00	<p><b>Polina Mesinioti<sup>1</sup>, Jo Angouri<sup>1</sup>, Dimitrios Siassakos<sup>2</sup>, Stephen O'Brien<sup>2</sup> &amp; Tim Draycott<sup>2</sup></b>  <sup>1</sup>U. Warwick, <sup>2</sup>North Bristol NHS Trust  <i>'Okay blood pressure now': Doing leadership and managing directives in obstetric emergencies</i></p>	<p><b>Sara Keel</b>            U. Basel  <i>Articulating touch and instruction in parents' and young children's accomplishment of everyday tasks</i></p>	<p><b>Petteri Laihonen &amp; Tamás Péter Szabó</b>            U. Jyväskylä  <i>Advantages and challenges of the digital turn in linguistic landscape projects on educational settings</i></p>
	<p><b>Jessica La &amp; Ann Weatherall</b>            Victoria U. Wellington  <i>Videod displays of pain in medical consultations</i></p>	<p><b>Anne-Sylvie Horlacher</b>            U. Basel  <i>Becoming the third party: Apprentices embodying a summons in hair salons</i></p>	<p><b>Tom Koole &amp; Paulien Harms</b>            U. Groningen  <i>Note-taking as recipient action</i></p>
	<p><b>Ana Cristina Ostermann &amp; Minéia Frezza</b>            UNISINOS  <i>Zooming into what matters: The double bind affordance of fetal ultrasound technology</i></p>	<p><b>Monika Messner</b>            U. Innsbruck  <i>Der Einsatz multimodaler Ressourcen in der Orchesterprobe</i></p>	<p><b>Sofian A. Bouaouina, Lorenza Mondada, Hanna M. Svensson &amp; Burak S. Tekin</b>            U. Basel  <i>Measuring distance in pétanque games: Participants' practices and ways of filming</i></p>
	<p><b>Brian Due</b>            U. Copenhagen  <i>Interacting with a telepresence robot: Mobility, space and embodied practices</i></p>	<p><b>Johanna Kehr</b>            TU Dresden  <i>Multimodale Analyse von Instruktionsgesprächen im Instrumentalunterricht</i></p>	<p><b>Laurent Camus</b>            U. Basel  <i>Filmic reflexivity in action: The interactional production of objectivity</i></p>
	16:00 - 16:30	Coffee Break	
16:30 - 18:00	VALS-ASLA GENERAL ASSEMBLY		
18:30	Welcoming Reception		

A  
V  
I  
D  
E  
O  
T  
U  
R  
N  
I  
N  
L  
I  
N  
G  
U  
I  
S  
T  
I  
C  
S  
?

V  
A  
L  
S  
A  
S  
L  
A  
2  
0  
1  
8

**7 JUNE | THURSDAY MORNING**

08:30 - 09:30	Registration		
09:30 - 10:30	<b>Keynote Lecture   Hörsaal -101</b> Christian Licoppe   Télécom ParisTech <i>Video-mediated communication (VMC): The organization of visibility, talk-in-interaction and participation in fragmented ecologies</i>		
10:30 - 11:00	Coffee Break		
11:00 - 12:30	<b>Seminarraum -201</b>  <b>Jessica Douglah</b> Stockhom U. <i>The interactional role of the mirror in show dance classes</i>	<b>Seminarraum +207</b>  <b>Safinaz Büyükgüzel &amp; Johannes Wagner</b> U. Southern Denmark <i>Interactional accomplishment of calculations in scrum meetings</i>	<b>Hörsaal -101</b>  <b>Wolfgang Kesselheim, Christina Brandenberger &amp; Christoph Hottiger</b> U. Zurich <i>What is the phenomenon? The multimodal achievement of joint discoveries in science centres</i>
	<b>Monica Simone &amp; Renata Galatolo</b> U. Bologna <i>Instructional practices in indoor climbing with blind and partially-sighted athletes</i>	<b>Mick Smith</b> UCLA <i>Navigating physical &amp; epistemic landscapes: How mobility in the field structures practitioners' understanding of geologic phenomena</i>	<b>Kristian Mortensen &amp; Johannes Wagner</b> U. Southern Denmark <i>Inspecting objects in social interaction</i>
	<b>Francesca Astrid Salvadori</b> U. Milano <i>Translating feelings into actions: Communication and collaboration between motorcycle race drivers and their crew chief</i>		<b>Florence Oloff</b> U. Oulu & U. Basel <i>Pointing at displays: Using smartphones as social objects</i>
12:30 - 14:00	Lunch		

V  
I  
D  
E  
O  
  
T  
U  
R  
N  
  
I  
N  
  
L  
I  
N  
G  
U  
I  
S  
T  
I  
C  
S  
  
?

V  
A  
L  
S  
  
A  
S  
L  
A  
  
2  
0  
1  
8

**7 JUNE | THURSDAY AFTERNOON**

	<b>Seminarraum -201</b>	<b>Seminarraum +207</b>	<b>Hörsaal -101</b>
<b>14:00 - 16:00</b>	<p><b>Philippe Sormani</b> U. Lausanne <i>Playing and Displaying "Move 37": the "Story of 'Oh'," a Multimodal Achievement in Situ</i></p>	<p><b>Mandira Halder</b> U. Genève <i>La vidéo comme un outil d'auto-confrontation à l'habitus des enseignant-e-s : le cas des enseignant-e-s généralistes dans le Canton de Vaud</i></p>	<p><b>Kenan Hochuli, Wolfgang Kesselheim, Christoph Hottiger &amp; Christina Brandenberger</b> U. Zurich <i>Eye-tracking and Conversation Analysis: A theoretical discussion of the approach to gaze 'from within' interactants</i></p>
	<p><b>Burak S. Tekin</b> U. Basel <i>On the incongruences in human machine interactions: Mutual intelligibility, responsibility and agency</i></p>	<p><b>Christine Develotte, Heike Baldauf-Quilliatre &amp; Justine Lascar</b> Lab ICAR Lyon <i>Sensibiliser les pouvoirs publics au rôle du médiateur numérique via la vidéo in situ</i></p>	<p><b>Geert Brône<sup>1</sup>, Clarissa Weiß<sup>2</sup> &amp; Elisabeth Zima<sup>2</sup></b> <sup>1</sup>U. Leuven, <sup>2</sup>U. Freiburg <i>On the role of eye gaze in competition for talk. Simultaneous starts in triadic interactions</i></p>
	<p><b>Isabell Neise &amp; Axel Schmidt</b> Inst. Deutsche Sprache, Mannheim <i>When spectators talk back: Practices of audience involvement in broadcasted video games</i></p>	<p><b>Emilie Jouin-Chardon<sup>1</sup>, Justine Lascar<sup>1</sup>, Vanessa Piccoli<sup>1</sup>, Anna Claudia Ticca<sup>2</sup>, Véronique Traverso<sup>3</sup> &amp; Biagio Ursi<sup>1</sup></b> <sup>1</sup>Lab ICAR Lyon, <sup>2</sup>Lab LIDILEM Grenoble, <sup>3</sup>Inst. Proche-Orient Beyrouth <i>Promoting research and raising awareness on social interaction and societal issues through video data</i></p>	<p><b>Clarissa Weiß</b> U. Freiburg <i>Das Blickverhalten des nicht-blickselektierten Sprechers während nicht-blickselektierten Turnübernahmen: Multimodale Interaktionsanalysen triadischer Gesprächssituationen mit Hilfe von Eye-Tracking</i></p>
	<p><b>Heike Baldauf-Quilliatre &amp; Isabel Colon De Carvajal</b> Lab ICAR Lyon <i>Instructing and giving advices: The spectator as coach in video game interactions</i></p>		<p><b>Elisabeth Zima<sup>1</sup> &amp; Geert Brône<sup>2</sup></b> <sup>1</sup>U. Freiburg, <sup>2</sup>U. Leuven <i>On the role of gaze for eliciting and addressing listener feedback in jointly construed narratives</i></p>
	<b>16:00 - 16:30</b>	<b>Coffee Break</b>	
<b>16:30 - 17:30</b>	<p><b>Keynote Lecture   Hörsaal -101</b> <b>Bina Elisabeth Mohn   U. Siegen &amp; Z. Kamera-Ethnographie</b> <i>Video als Material und Mittel gemeinsamen Hinschauens beim Forschen und darüber hinaus</i></p>		
<b>19:00</b>	<b>Conference Dinner</b>		

**8 JUNE | FRIDAY MORNING**

08:30 - 09:30	Registration		
09:30 - 10:30	<b>Keynote Lecture   Hörsaal -101</b> Federico Rossano   UCSD <i>On video in experimental approaches to babies and primates</i>		
10:30 - 11:00	Coffee Break		
11:00 - 12:30	<b>Seminarraum -201</b>  <b>Zoe DuPree Fine</b> U. South Florida <i>Composing terror: Analyzing video interviews of "women of ISIS"</i>	<b>Seminarraum +207</b>  <b>Steven Schoonjans</b> U. Innsbruck & KU Leuven <i>Verbale und nonverbale Abtönung im Vergleich</i>	<b>Hörsaal -101</b>  <b>Hanna Magdalena Svensson</b> U. Basel <i>Visible problems of hearing: Correcting how to speak in public</i>
	<b>Jacob Geuder</b> U. Basel <i>Translating positions into perspectives: A comparison of activist and journalist videos of protests in Rio de Janeiro</i>	<b>Robert Reinecke</b> U. Basel <i>Kreativität in Interaktion – Eine multimodale Interaktionsanalyse der «Creative Economy»</i>	<b>Elwys De Stefani &amp; Isabelle Heyerick</b> KU Leuven <i>Recipients' gaze in guided tours with sign language interpreters: On the temporal dimension of spatial reference</i>
	<b>Mathias Broth &amp; Mattias Arvola</b> Linköping U. <i>Online text editing in pursuit of readers in news production for the web</i>	<b>Gökhan Özkayın</b> U. Koblenz-Landau <i>Kohäsion und Kohärenz durch Verflechtung von Sprache und Gestik: Eine empirische Untersuchung mündlicher Erzählungen deutschsprachiger Schülerinnen und Schüler</i>	<b>Antonia Steger</b> U. Zurich <i>Flüchtige urbane Begegnungen – Eine Videanalyse interaktiver Minimalformen im öffentlichen Raum</i>
12:30 - 14:00	Lunch		

V  
I  
D  
E  
O  
  
T  
U  
R  
N  
  
I  
N  
  
L  
I  
N  
G  
U  
I  
S  
T  
I  
C  
S  
  
?

V  
A  
L  
S  
  
A  
S  
L  
A  
  
2  
0  
1  
8

8 JUNE | FRIDAY AFTERNOON

	Seminarraum -201	Seminarraum +207	Hörsaal -101
14:00 - 15:30	<p><b>Chloé Mondeme<sup>1</sup> &amp; Laurent Camus<sup>2</sup></b>  <sup>1</sup>CNRS Lyon, <sup>2</sup>U. Basel  <i>Capturing emotion with video: Relocating aggressive behaviors in their ecological environments</i></p>	<p><b>Thomas C. Messerli &amp; Steve Duman</b>            U. Basel  <i>Linguistic and paralinguistic aspects of interaction with non-robotic Voice-User Interfaces</i></p>	<p><b>Frode Guribye<sup>1</sup> &amp; Oskar Lindwall<sup>2</sup></b>  <sup>1</sup>U. Bergen, <sup>2</sup>U. Gothenburg  <i>Sequencing and unpacking powerpoint slides in university lectures</i></p>
	<p><b>Vanessa Piccoli<sup>1</sup>, Anna Claudia Ticca<sup>2</sup> &amp; Véronique Traverso<sup>3</sup></b>  <sup>1</sup>Lab ICAR Lyon, <sup>2</sup>Lab LIDILEM Grenoble, <sup>3</sup>Inst. Proche-Orient Beyrouth  <i>Rendre les émotions du patient dans la traduction orale du récit : le cas des entretiens psychothérapeutiques avec demandeurs d'asile en France</i></p>	<p><b>Miriam A. Locher &amp; Thomas C. Messerli</b>            U. Basel  <i>Participatory watching of telecinematic fiction in an international drama viewing platform</i></p>	<p><b>Monica Clua</b>            UIC Barcelona  <i>More than words can say: Embodied multimodality in an English-medium university lecture</i></p>
	<p><b>Yumei Gan<sup>1</sup>, Christian Greiffenhagen<sup>1</sup> &amp; Christian Licoppe<sup>2</sup></b>  <sup>1</sup>The Chinese U. Hong Kong, <sup>2</sup>Télécom ParisTech  <i>Orchestrating openings: The emotional importance of the first five seconds in video calls between migrant parents and their 'left behind' children in China</i></p>	<p><b>Elliott M. Hoey<sup>1</sup>, David DeLiema<sup>2</sup>, Rachel Chen<sup>2,3</sup> &amp; Virginia J. Flood<sup>2</sup></b>  <sup>1</sup>U. Basel, <sup>2</sup>U. California, Berkeley, <sup>3</sup>San Francisco State U.  <i>Children`s use of imitation of locomotor play</i></p>	<p><b>Martin Luginbühl</b>            U. Basel  <i>The camera as contextualization resource in school children`s peer talk</i></p>
15:30 - 16:00	<b>Coffee Break</b>		
16:00 - 17:00	<b>Keynote Lecture   Hörsaal -101</b> <b>John Hindmarsh   King`s College London</b> <i>Work, interaction and spatial organisation: On naturalistic, video-based studies of technology-in-use</i>		
17:00 - 17:30	<b>Closing Ceremony</b>		
17:30	<b>End of Conference</b>		

A  
V  
I  
D  
E  
O  
  
T  
U  
R  
N  
  
I  
N  
  
L  
I  
N  
G  
U  
I  
S  
T  
I  
C  
S  
  
?

V  
A  
L  
S  
  
A  
S  
L  
A  
  
2  
0  
1  
8

# KEYNOTE LECTURES



## **ARNULF DEPPERMAN**

(IDS Mannheim, Germany)

WHY DOES LINGUISTICS NEED  
VIDEO-ANALYSIS AND VIDEO-CORPORA?

WARUM BRAUCHT DIE LINGUISTIK  
VIDEOANALYSE UND VIDEOKORPORA

## **JON HINDMARSH**

(King's College London, UK)

WORK, INTERACTION & SPATIAL ORGANIZATION:  
ON NATURALISTIC, VIDEO-BASED STUDIES OF  
TECHNOLOGY-IN-USE

TRAVAIL, INTERACTION & ORGANISATION SPATIALE:  
ÉTUDES VIDÉO NATURALISTES DES TECHNOLOGIES  
EN ACTION

## **CHRISTIAN LICOPPE**

(Télécom ParisTech, France)

VIDEO-MEDIATED  
COMMUNICATION (VMC)

LES INTERACTIONS À DISTANCE  
PAR VISIO-CONFÉRENCE

## **BINA ELISABETH MOHN**

(Universität Siegen & Zentrum für Kamera-Ethnographie, Germany)

VIDEO AS A RESOURCE FOR SHARED  
LOOKING IN RESEARCH AND BEYOND

VIDEO ALS MATERIAL UND MITTEL  
GEMEINSAMEN HINSCHAUENS BEIM FORSCHEN  
UND DARÜBER HINAUS

## **FEDERICO ROSSANO**

(UCSD, USA)

ON VIDEO IN EXPERIMENTAL APPROACHES  
TO BABIES AND PRIMATES

STUDI VIDEO SPERIMENTALI  
DI BAMBINI E PRIMATI

## Why does linguistics need video-analysis and video-corpora?

**ARNULF DEPPERMANN**

Institut für Deutsche Sprache Mannheim, Germany

Spoken language is auditory and oral, written language is visual and static. Therefore it is obvious that linguistics needs audio-recordings and texts as data – but why video-recordings? In my talk, I will be dealing with spoken language, more precisely, with talk-in-interaction. I will be arguing that its analysis requires to understand talk as an integral part of bodily multimodal practices of interaction (see Goodwin et al. 2011, Goodwin 2017).

After a short introduction into basic properties of multimodal interaction, I will sketch methodological requirements on video-recordings (Mondada 2012) and multimodal transcription (Mondada 2018). The constitutive role of other bodily resources for the deployment of talk will be demonstrated by looking closer at three sorts of classic linguistic phenomena:

- Semantics: linguistic practices become disambiguated by their coordination with other resources (e.g. gaze and head movements);
- Syntax: complex syntactic structures emerge in response to recipient reactions which occur simultaneously with the production of a turn at talk or immediately after it (cf. Goodwin 1981);
- Lexis: the use of minimal referential forms is based on joint attention and bodily co-orientation of the participants (Deppermann i.Dr., Stukenbrock 2015).

These examples show that video-analysis is not only imperative for the study of action in social interaction among co-present participants; it is equally indispensable for advancing our insights into classic topics of linguistic research.

Studies in interactional linguistics work on the basis of collections of instances of a linguistic practice (Deppermann 2001, Mondada 2005, 2011). For all aspects of the study of collections, the use of large video-corpora is helpful or even indispensable. It informs the identification

- of the range of variants of (the formal realization of) the practice,
- of relations of coordination and co-occurrence between the different bodily resources which are used for the practice,
- of the (pragmatic, interactional, rhetoric etc.) uses of the practice,
- of contextual factors which affect the occurrence and the realization of the practice and
- for an assessment of the generalizability of findings.

Large video-corpora with transcript-image-sound-aligned data and rich meta-data are, e.g., FOLK (<dgd.ids-mannheim.de>) for German and CLAPI (<clapi.ish-lyon.cnrs.fr>) for French. They allow for searches for linguistic items and their co-occurrences, restricted to certain interaction types, turn-constructive environments or groups of speakers. The talk will discuss and show how such corpora can be used in the context of a research project.

## Warum braucht die Linguistik Videoanalyse und Videokorpora?

Gesprochene Sprache ist auditiv-mündlich, geschriebene visuell und statisch. So ist es evident, dass die Linguistik Tonaufnahmen und Texte als Datengrundlage benötigt – warum aber Videoaufnahmen? Der Vortrag befasst sich mit der gesprochenen Sprache. Er argumentiert, dass ein Verständnis ihrer Besonderheiten erfordert, gesprochene Sprache als integralen Bestandteil leiblicher, multimodaler Interaktionspraktiken zu behandeln (vgl. Goodwin et al. 2011, Goodwin 2017).

Nach einer kurzen Einführung in grundlegende Eigenschaften multimodaler Interaktion werden die methodischen Anforderungen an gegenstandsangemessene Videoaufnahmen (Mondada 2012) und multimodale Transkriptionen (Mondada 2018) skizziert. An drei Phänomenbereichen wird exemplarisch gezeigt, wie das Zusammenspiel sprachlicher Praktiken mit anderen leiblichen Ressourcen der Kommunikation anhand von Videoanalysen zu untersuchen ist:

- Semantik: Disambiguierung sprachlicher Praktiken durch ihre Koordination mit anderen Ressourcen,
- Syntax: Die Entstehung komplexer syntaktischer Strukturen aufgrund von nonverbalen Rezipientenreaktionen simultan zur bzw. auf die Produktion sprachlicher Äußerungen (Goodwin 1981),
- Lexik: Die Verwendungen minimaler Referenzformen, die sich auf die multimodale Ko-Orientierung der Beteiligten stützt (Deppermann i.Dr., Stukenbrock 2015).

Die Beispiele zeigen, dass Videoanalyse nicht nur essentiell für die Untersuchung des Handelns in der sozialen Interaktion ist, sondern auch für Erkenntnisse im Bereich klassischer strukturlinguistischer Fragestellungen unabdingbar ist.

Die interaktionslinguistische Analyse geht kollektionsbasiert vor (Deppermann 2001, Mondada 2005, 2011). Für alle Aspekte der Untersuchung von Fällen einer Praktik ist die Benutzung großer Videokorpora hilfreich: Für die Identifikation

- des Variantenspektrums der Formen der Praktik,
- der Kookkurrenz- und Koordinationsverhältnisse zwischen einzelnen Ressourcen, die für eine Praktik eingesetzt werden,
- des Verwendungsspektrums der Praktik (ihrer möglichen Funktionen),
- der relevanten Bedingungsfaktoren im Kontext, die für ihr Auftreten und ihre Realisierung maßgeblich sind sowie für
- die Abschätzung der Generalisierbarkeit der Untersuchungsergebnisse.

Erforderlich dazu sind Text-Bild-Ton-alignierter Videokorpora mit reichhaltigen Metadaten. Vorbilder für die Konstruktion solcher Korpora sind für das Deutsche das Korpus FOLK des Instituts für Deutsche Sprache (Mannheim, <dgd.ids-mannheim.de>) und für das Französische CLAPI (Université Lyon 2, Laboratoire ICAR, Groupe ICOR, <http://clapi.ish-lyon.cnrs.fr/>).

Diese Korpora bieten Möglichkeiten zur strukturierten Suche, ausgehend von bestimmten linguistischen Formen, in Kookkurrenz mit anderen sprachlichen Formen, in bestimmten Gattungen oder Turnkontexten und bei bestimmten Sprechergruppen. Große, wissenschaftsöffentliche Korpora können ein eigenes Korpus ergänzen und zur Testung der Reichweite der am eigenen Korpus gewonnenen Resultate genutzt werden. Oder sie stellen selbst die Grundlage der Untersuchung dar. Sie können explorativ, zur Identifikation von Musterhaftigkeiten, dokumentarisch bzw. korrelativ, zur Feststellung von Verteilungen, oder testend, zur Kontrolle von Bedingungsfaktoren genutzt werden.

# Work, interaction & spatial organisation: On naturalistic, video-based studies of technology-in-use

**JON HINDMARSH**

King's College London, UK

In contemporary organisations, the introduction of new technologies into the workplace is by no means uncommon. What makes these new technologies particularly interesting for studies of social interaction, is that they often lead to the adjustment or reorganisation of everyday work practices. Indeed, using Beane & Orlikowski's (2015) terms, practices are 'entangled' with the distinctive characteristics of tools in use.

To explore and understand forms of 'entanglement', we can consider a range of interactional resources that participants demonstrably use in coordinating their actions with others. In these cases, we need to pay attention to talk, but also gaze, gesture, tool use, position and more. This treatment of social interaction as fundamentally embodied and embedded is key. Indeed, in many situations, talk would make little sense to participants out of visible and material context, just as, at times, visible or material conduct would make little sense to participants without talk.

There are, however, challenges to understanding the interactional resources available to participants within complex organisational and technological domains; challenges that impact on the collection and analysis of relevant and appropriate data. In this presentation, I will consider some of these challenges and the ways that they can be resolved using approaches from ethnomethodology and conversation analysis and in naturalistic, video-based studies of work and interaction.

Furthermore, in order to explicate more fully the issues that arise in coordinating work with and around new technologies, the presentation will explore the value of a praxeological approach to understanding workspace (following Lucy Suchman and Lorenza Mondada). Indeed, it is routinely argued that proximity is a mechanism that supports coordination in organisations. However, I will explore an intriguing case that undermines this seemingly straightforward assertion; indeed, the case has as many similarities with studies of mediated communication as it does with studies of co-present interaction. Again, to explore these issues, we need to consider the range of interactional resources available to participants in particular sequences of conduct.

I will build the presentation around findings from a video-based field study of the implementation of the *da Vinci* robot in colorectal surgery. Using a series of video data extracts, I will explore the ways in which members of interdisciplinary teams in operating theatres coordinate actions moment-by-moment, and in the course of routine surgical procedures. This will involve the discussion of both seemingly seamless interactions, as well as more problematic episodes.

## Travail, interaction & organisation spatiale: études vidéo naturalistes des technologies en action

Dans les organisations contemporaines, l'introduction de nouvelles technologies dans les lieux de travail n'est pas inhabituelle. Ce qui rend ces nouvelles technologies particulièrement intéressantes pour l'étude de l'interaction sociale, c'est qu'elles conduisent souvent à l'ajustement ou à la réorganisation des pratiques de travail quotidiennes. En effet, d'après la terminologie de Beane et Orlikowski (2015), les pratiques sont enchevêtrées avec les caractéristiques distinctives des outils utilisés.

Pour explorer et comprendre ces formes d'enchevêtrement, nous pouvons prendre en considération une gamme de ressources interactionnelles que les participants utilisent manifestement pour coordonner leurs actions avec celles des autres. Dans ces cas, nous devons prêter attention à la parole, mais aussi au regard, aux gestes, à l'utilisation d'outils, au positionnement du corps et plus encore. Ce traitement de l'interaction sociale comme fondamentalement incorporée et imbriquée est crucial. En effet, dans de nombreuses situations, le parler n'a que peu de sens pour les participants s'il est extrait du contexte visible et matériel, tout comme, parfois, la conduite visible ou matérielle n'a pas de sens pour les participants sans la parole.

Cependant, c'est un défi de comprendre les ressources interactionnelles mises à la disposition des participants dans des domaines organisationnels et technologiques complexes; ces défis ont une incidence sur la collecte et l'analyse de données qui soient pertinentes et appropriées. Dans cette présentation, je vais examiner certains parmi ces défis et les façons dont ils peuvent être résolus en utilisant des approches issues de l'ethnométhodologie et de l'analyse conversationnelle et dans des études naturalistes du travail et de l'interaction basées sur la vidéo.

En outre, afin d'expliquer plus en détail les problèmes qui surgissent dans la coordination du travail avec et autour de nouvelles technologies, la présentation explorera la valeur d'une approche praxéologique pour comprendre l'espace de travail (en suivant Lucy Suchman et Lorenza Mondada). En effet, il est régulièrement soutenu que la proximité est un mécanisme qui soutient la coordination dans les organisations. Cependant, je vais explorer un cas intrigant qui questionne cette affirmation apparemment simple ; en effet, le cas présente autant de similitudes avec les études de communication médiatisée qu'avec les études d'interaction en co-présente. Encore une fois, pour explorer ces problèmes, nous devons considérer l'éventail des ressources interactionnelles disponibles aux participants dans des séquences de conduite particulières.

Je vais construire la présentation autour des résultats d'une étude de terrain basée sur la vidéo de la mise en œuvre du robot da Vinci en chirurgie colorectale. À l'aide d'une série d'extraits de données vidéo, j'explorerai la façon dont les membres d'équipes pluridisciplinaires dans les salles d'opération coordonnent les actions moment par moment, et au cours des procédures chirurgicales de routine. Cela impliquera la discussion des deux interactions apparemment transparentes, ainsi que d'épisodes plus problématiques.

## Video-mediated communication (VMC):

### The organization of visibility, talk-in-interaction and participation in fragmented ecologies

CHRISTIAN LICOPPE

Télécom ParisTech, France

VMC settings have a particular interest for researchers interested in video-based analyses of social interaction and multimodal analysis. They provide situations in which the participants collaborate to produce and maintain relevant visual ecologies with respect to the ongoing interaction. This communication analyzes their observable behavior in interpersonal (Skype) and institutional (courtroom) contexts and provide an analysis of the systematic organization of visibility and talk-in-interaction which can be observed in VMC settings. Such an organization can be shown to stem from an orientation towards the Sacks-like maxim “put the face of the current ratified speaker/addressee on screen”, which is relevant both to interpersonal and to institutional VMC. The production of relevant video frames is thus tightly grafted to the sequential organization of talk-in-interaction, the purpose of which is precisely to provide resources for the orderly production of turns-at-talk, and therefore the determination of the current speaker. A consequence of this is to make the ‘talking heads’ configuration the default expectation for dyadic VMC conversation. Another is to make participants accountable for shooting/showing something else, and to provide procedures to recognize and repair such visual “trouble”.

The case of multi-participant settings, which I will discuss here in the case of courtrooms with video links, introduces several interesting twists. First, the person producing the images in a sequentially relevant way may not be the current speaker/hearer any more. Second, there are many different ways for her to shoot the current speaker among others. I will show how participants display an orientation towards the fact that the parties made visible in such video shots may be especially relevant to the current talk-in-interaction, because of their being made visible here and now. Changing the image so that some participants appear or disappear is then done and treated as having important implications in terms of sequence organization and participation frameworks. I will analyze the systematic organization of camera motions in this setting, and show how they constitute interactional moves in their own right, i.e. ‘camera actions’, characteristic of ‘video communication contextures’. To provide an example, I will focus on the production of wide shots, as a situated and timed camera-mediated accomplishment in the course of the hearing, and show : a) how such an accomplishment is sensitive to its sequential environment; b) how it is accountable as a way to mark the particular relevance of a group of participants with respect to the ongoing talk, and therefore oriented-to as a resource to visually highlight collective forms of speakership or reciprocity; c) how such camera motions are sequentially-relevant and sequentially implicative with respect to the ongoing video interaction; d) how such camera motions, and more generally the video communication ecologies which enable them make visible a particular member’s interactional competence, that of being able to recognize the relevance of subtle changes in participation frames, routinely, unreflexively and on the fly. In VMC settings (as elsewhere, but with distinctive concerns) what talk does cannot be understood without attending to the camera work of the participants, and the visual ecologies they craft moment by moment, and vice and versa. The ways speaking and looking are assembled moment by moment as collaborative, practical and situated accomplishments in VMC, reveal underlying expectations of congruence between visibility and talk which may remain unquestioned and hidden in face to face interaction.

## Les interactions à distance par visio-conférence : L'organisation de la visibilité et de la parole en interaction dans des écologies fragmentées

Les configurations impliquant la communication vidéo ont un intérêt particulier pour les chercheurs intéressés à l'analyse de l'interaction sociale basée sur l'enregistrement vidéo. Elles fournissent des situations dans lesquelles les participants collaborent à produire et entretenir des écologies visuelles adaptées de manière endogène à leur interaction. Cette communication analyse ce travail de construction du cadre visuel dans des contextes interpersonnels (Skype) et institutionnels (Tribunal), et cherche à montrer le caractère systématique de l'organisation de la visibilité et de la parole en interaction observable dans ces contextes de vidéo-communication. On peut démontrer que cette organisation s'oriente par une maxime similaire à celles de Sacks, « mettre le visage du locuteur en cours à l'écran », qui est pertinente aussi bien dans les situations interpersonnelles que les situations institutionnelles. La production d'images vidéo pertinentes est donc étroitement articulée à l'organisation séquentielle de la conversation, dont la visée est précisément de fournir des ressources pour la production ordonnée de tours de parole et de transferts de locuteurs. Une conséquence de cette organisation est de faire de la configuration en « têtes parlantes » la configuration attendue par défaut dans les vidéo-conversations dyadiques. Une autre est de rendre les participants « accountable » par rapport au fait de montrer quelque chose d'autre, et de fournir des procédures pour reconnaître et réparer de tels « troubles » visuels.

Le cas des configurations multi-participants, que je discuterai dans le cas d'interventions à distance au tribunal, introduit plusieurs développements intéressants. Tout d'abord la personne qui produit les images de manière séquentiellement pertinente peut ne plus être le locuteur en cours. Il y a ensuite en général diverses manières pour elle de cadrer le locuteur en cours tout en l'ayant à l'écran. Je montrerai comment les participants manifestent une orientation vers le fait que ceux qui sont rendus visibles puissent avoir une pertinence particulière par rapport à l'interaction en cours, du fait de cette visibilité, à ce moment. Tout changement de plan qui fait apparaître ou disparaître des participants est de ce fait produit et traité comme ayant des implications importantes quant à l'organisation des séquences et des formats de participation. J'analyserai ainsi l'organisation systématique des mouvements de caméra dans ce genre de configurations, et montrerai en quoi ils constituent des « actions de caméra », parties intégrantes des « contextures multimodales » qui caractérisent les interactions à distance en vidéo. A titre d'exemple, je me focaliserai sur la production de plans larges comme accomplissement situé lors d'audiences judiciaires par visio-conférence pour montrer: a) comment cet accomplissement est sensible à l'environnement séquentiel; b) comment il est traité comme conférant une forme particulière de pertinence participative par rapport à l'interaction en cours pour les participants ainsi rendus visibles, et susceptible de constituer une ressource pour souligner ou accomplir des formats collectifs de locution ou de réception; c) comment ces mouvements de caméra ont des implications séquentielles vis-à-vis de la conversation institutionnelle en cours; d) comment ce travail de caméra rend visible une compétence interactionnelle et multimodale particulière, celle de produire et reconnaître continûment des changements du cadre de participation en s'orientant sans y réfléchir vers l'écologie visuelle. Dans ces environnements visuellement fragmentés, on ne peut pas comprendre ce que "fait" la parole-en-interaction sans prêter attention au travail de caméra des membres pour produire moment par moment un cadre visuel convenable, et vice-versa. Les manières de parler et de voir sont assemblées moment par moment, et la manière dont le travail d'articulation entre les deux est rendu visible dans les interactions à distance en vidéo en tant qu'accomplissement situé et collaboratif, révèle des attentes normatives plus générales concernant la congruence de la visibilité et de la parole en interaction, qui peuvent être plus facilement tenues pour acquises et donc moins visibles dans le cas des interactions en co-présence.

## Video as a resource for shared looking in research and beyond

**BINA ELISABETH MOHN**

Universität Siegen & Zentrum für Kamera-Ethnographie, Deutschland

How can ethnographers show something visually? Using examples of ‘camera ethnographic’ work (Mohn 2013), the presentation will trace a medial shift from an emphasis on speech, discourse, and textual formats to performative, pictorial, and filmic formats. Wordless practices, socio-material constellations, figures and choreographies of practices are the focus of an ethnography that shows rather than describes, retaining its visual form even at the stage of publication. This makes reception a performative event: the audience become researchers themselves as they explore the visual material presented. The following aspects will be raised and discussed:

**Visibility and showing:** What can be achieved by ethnographic processes of making visible in research and beyond? To what extent are these processes a prerequisite for exhibition and at the same time constitutive of ‘learning to see something *as* something’? Drawing on sociological laboratory studies of the 1980s and 1990s, I argue that the visibility and observability of phenomena and practices is not given *a priori*, but has to be made possible by (medial) research practices.

**Filming and editing:** How do specific camera perspectives and techniques facilitate the investigation of objects of interest? How can editing and montage be undertaken as analytical ethnographic practices of discovery and making visible? How can “thick showing”—analogous to “thick description” (Geertz) or “*übersichtliche Darstellung*” (Wittgenstein)—develop from that? In contrast to “video analysis”, “visual analysis” evolves within and between filmic images, by focusing and concentrating, organizing and arranging, comparing and contrasting.

**Situated methodology:** An understanding of practices as situated solutions to concrete problems can also be applied to research practices. Fieldwork and laboratory phases, phases of reflection, phases of publication, reception, and even application alternate with each other within a circular process. This constitutes a “situated methodology”, within which the practices of an ethnography that shows are integrated.

**Result and event:** Reception is considered an indispensable part of the research process: as a co-construction of results within social events of text, image, or filmic communication (see Hausendorf et al. 2017; Mersch 2002, Pantenburg 2006). This performative understanding of reception renders notions of stable inventories of knowledge or products of research that ‘speak for themselves’ obsolete. How can applied research benefit from such a conceptualization of reception, and how can the audience be interactively involved? When visiting media installations, using interactive archives, or watching observational research films, users can observe, compare, distinguish, and relate; thereby discovering the multiplicity and potential of social phenomena and practices. Ethnographic “*Blicklaboratorien*” (laboratories of looking) constitute one possible performative format in which ethnographic knowledge can be shared.

## Video als Material und Mittel gemeinsamen Hinschauens beim Forschen und darüber hinaus

Wie kommen Ethnograph/innen dazu, etwas bildhaft zeigen zu können? Die Präsentation führt an Beispielen aus der „Kamera-Ethnographie“ (Mohn 2013) mediale Akzentverschiebungen vom Gesprochenen, Diskursiven und Textförmigen zum Performativen, Bildhaften und Filmischen vor: Stumme Praktiken, sozio-materielle Konstellationen, Figuren und Choreographien der Praxis geraten auf diese Weise in den Blick einer zeigenden statt beschreibenden Ethnographie, die bis in die Publikationsformate hinein vorrangig bildhaft bleibt und dabei einen besonderen Rezeptionsmodus erlaubt: „forschende Rezeption“. Angesprochen werden diese Punkte:

**Sichtbarkeit und Zeigen:** Welche Rolle können ethnographische Verbildlichungsprozesse im Forschungsprozess und darüber hinaus spielen? Inwiefern ist Bildgebung beim Forschen eine Voraussetzung späteren Zeigens und zugleich konstitutiv für die Prozesse des Sichtbarmachens und Sehens von etwas „als etwas“? Mit Bezug auf die soziologischen Laborstudien der 1980er/90er Jahre wird davon ausgegangen, dass Sichtbarkeit und Beobachtbarkeit von Phänomenen und Praktiken nicht einfach gegeben sind, sondern durch (Medien-)Praktiken des Forschens erst sukzessive hervorgebracht werden.

**Filmen und Schneiden:** Wie tragen gewählte Kameraeinstellungen und Kameraführung zur Verfertigung epistemischer Objekte bei? Wie können Schnitt und Montage als analytische Praktiken ethnographischer Erfahrungsbildung und Entdeckung gerahmt und gehandhabt werden? Wie entstehen daraus „dichtes Zeigen“ (analog „dichter Beschreibung“ Geertz) oder „übersichtliche Darstellungen“ (nach Wittgenstein)? Anders als in der „Video-Analyse“ entsteht eine „Visuelle Analytik“ nicht an sondern in und zwischen filmischen Bildern, u.a. durch Fokussieren und Verdichten, Sortieren, Arrangieren und Montieren, Vergleichen und Kontrastieren.

**Situierte Methodologie:** Die Grundannahme, dass Praktiken stets im Kontext von Situationen und als Lösungen konkreter Handlungsprobleme sinnvoll zu verstehen sind, lässt sich auf die Situationen des Forschens und ihre Praktiken beziehen. Feld- und Laborphasen, Phasen der Reflexion, Publikation, Rezeption bis hin zu Anwendungsphasen wechseln einander in zirkulären Prozessen ab. Dies begründet eine „situierte Methodologie“, in die auch die Praktiken einer zeigenden Ethnographie eingebunden sind.

**Ergebnis und Ereignis:** Rezeption wird als unverzichtbarer Bestandteil von Forschungsprozessen aufgefasst, als eine Ko-Konstruktion von Ergebnissen im sozialen Ereignis der Text-, Bild- oder Filmkommunikation (vgl. Hausendorf u.a. 2017, Mersch 2002, Pantenburg 2006). Dies löst den Werkbegriff ebenso auf wie die Vorstellung feststehender Wissensbestände. Wie kann dies im Zusammenhang angewandter Forschung genutzt und das Publikum interaktiv eingebunden werden? Beim Besuch von Medieninstallationen, der Nutzung interaktiver Archive oder beim Sichten von auf DVDs angeordneten beobachtenden Forschungsfilmern können Nutzer/innen ihrerseits beobachten und benennen, vergleichen, unterscheiden und verknüpfen und dabei etwas über die Vielfalt und Möglichkeit sozialer Phänomene und Praktiken in Erfahrung bringen: ethnographische „Blicklaboratorien“ als ein mögliches Ereignisformat ethnographischen Wissens.

## On video in experimental approaches to babies and primates

FEDERICO ROSSANO

UCSD, USA

Both ethologists and developmental psychologists have long been interested in studying the communicative repertoire of the individuals they were observing. The goal was to establish reliable maps of a species repertoire or of a human child repertoire at a certain age. Yet videos continued to be used parsimoniously, partly because of costs and partly because of the painstaking effort of coding micro-behavior from low quality video data. Parent's diaries for children's data and ethograms for animal data became standard research methods for investigating children and non-human primates communicative repertoire. In the 80s and 90s the use of video cameras increased significantly, together with an increase in experimental research in both domains and a general concern for the reliability of behavioral coding and the use of video data is pretty much common practice today. Note though that video data is investigated at a level of granularity much more coarse than what is common in studies on multimodal interaction relying on conversation analysis.

In this talk I intend to highlight the importance of video data in two main areas: 1) experimental situations and 2) longitudinal observational studies on the development of communicative abilities in young primates. With respect to the first domain, I will highlight the problem of current practices that consider data only the coded data that has been inserted in an Excel sheet. To this end, edited videos of the actual behaviors during experimental studies are often only presented during talks as exemplifications of the findings as established via statistical means. Moreover, the practice of calculating interrater reliability for the coded behaviors invites a simplification of the behavioral analysis that often gets reduced to behaviors easy to identify for any untrained eye.

With respect to the second domain, I will show how current observational research on animal communication relies on coding methods (e.g., ethograms) that hinder detecting how behaviors change through time. Presenting longitudinal data from mother-infant bonobo dyads, I present an analysis of how signal formation and behavioral change can be detected. I will show the importance of beginning our analysis not by identifying gestures but rather by first identifying the social activities participants are striving to achieve (e.g. carries, grooming, playing) and then looking retrospectively at the behavioral practices that were used to successfully partake in them. This analytical switch is only possible if a more rigorous methodological standard is set that entails extensive and frequent video collection of data.

## Studi video sperimentali di bambini e primati

Lo studio del repertorio comunicativo degli individui ha interessato da tempo sia etologi che psicologi dello sviluppo. Lo scopo è da sempre stabilire una cartografia affidabile del repertorio di una specie o del bambino ad una certa età. Malgrado il suo potenziale per rispondere a questo scopo, l'uso del video è rimasto limitato, in parte a causa dei costi e in parte a causa dell'immenso sforzo rappresentato dalla codifica di micro-comportamenti a partire da dati video di bassa qualità. E' per questo che i diari tenuti dai genitori e gli etogrammi per la descrizione di animali sono diventati metodi di ricerca standardizzati per lo studio dei repertori comunicativi di bambini e primati non umani. Negli anni 80 e 90, l'uso di telecamere e dati video è aumentato sensibilmente, parallelamente ad un aumento della ricerca sperimentale in ambedue i campi, così pure un interesse generale per l'affidabilità della codifica di comportamenti. Oggi l'uso di dati video è diventato una pratica corrente – anche se è importante notare che i dati video sono studiati ad un livello di dettaglio molto più grossolano che quanto è d'uso in studi sull'interazione multimodale in analisi della conversazione.

In questo intervento intendo evidenziare l'importanza dei dati video in due ambiti: 1) situazioni sperimentali e 2) studi longitudinali sullo sviluppo delle capacità comunicative di giovani primati. Per quanto riguarda il primo ambito, evidenzierò il problema dei metodi attuali che considerano in quanto dati solo i dati codificati, etichettati e inseriti in un foglio Excel. A tal fine, i video dei comportamenti reali osservati in studi sperimentali sono spesso presentati ai congressi come esemplificazioni dei risultati stabiliti attraverso mezzi statistici. Inoltre, la pratica di calcolare l'affidabilità della codifica dei comportamenti invita a semplificare l'analisi comportamentale che spesso si riduce a comportamenti facili da identificare per qualsiasi occhio non allenato.

Rispetto al secondo ambito, mostrerò come l'attuale ricerca sulla comunicazione animale si basi su metodi di codifica (ad esempio, gli etogrammi) che impediscono di rilevare come i comportamenti cambino nel tempo. Presentando dati longitudinali da coppie madre-bambino di bonobo, presenterò un'analisi di come sia possibile identificare la formazione del segnale e il cambiamento comportamentale. Mostrerò l'importanza di iniziare la nostra analisi non identificando i gesti ma piuttosto identificando in primo luogo le attività sociali che i partecipanti stanno cercando di realizzare (per esempio, portare, spulciare, giocare) e poi guardando retrospettivamente alle pratiche comportamentali che sono state usate per prendervi parte. Questo passaggio analitico è possibile solo se viene impostato uno standard metodologico più rigoroso che implichi un'ampia e frequente raccolta di dati.

## INDIVIDUAL PAPERS



HEIKE BALDAUF-QUILLIATRE ISABEL COLON DE CARVAJAL EVELYNE BERGER SIMONA PEKAREK DOEHLER  
KLARA SKOGMYR MARIAN SOFIAN A. BOUAOUINA LORENZA MONDADA HANNA M. SVENSSON  
BURAK S. TEKIN GEERT BRÔNE CLARISSA WEIß ELISABETH ZIMA MATHIAS BROTH MATTIAS ARVOLA  
SAFINAZ BÜYÜKGÜZEL JOHANNES WAGNER LAURENT CAMUS MONICA CLUA ELWYS DE STEFANI  
ISABELLE HEYERICK CHRISTINE DEVELOTTE HEIKE BALDAUF-QUILLIATRE JUSTINE LASCAR JESSICA  
DOUGLAH BRIAN DUE ZOE DUPREE FINE YUMEI GAN CHRISTIAN GREIFFENHAGEN CHRISTIAN LICOPPE  
JACOB GEUDER FRODE GURIBYE OSKAR LINDWALL MANDIRA HALDER THOMAS HILLMAN OSKAR  
LINDWALL BARRY BROWN KENAN HOCHULI WOLFGANG KESSELHEIM CHRISTOPH HOTTIGER CHRISTINA  
BRANDENBERGER ELLIOTT M. HOEY DAVID DELIEMA RACHEL CHEN VIRGINIA J. FLOOD ANNE-SYLVIE  
HORLACHER ÉMILIE JOUIN-CHARDON JUSTINE LASCAR VANESSA PICCOLI ANNA CLAUDIA TICCA  
VÉRONIQUE TRAVERSO BIAGIO URSI SARA KEEL JOHANNA KEHR WOLFGANG KESSELHEIM CHRISTINA  
BRANDENBERGER CHRISTOPH HOTTIGER TOM KOOLE PAULIEN HARMS JESSICA LA ANN  
WEATHERALL PETTERI LAIHONEN TAMÁS PÉTER SZABÓ MIRIAM A. LOCHER THOMAS C. MESSERLI  
MARTIN LUGINBÜHL SARA MERLINO POLINA MESINIOTI JO ANGOURI DIMITRIOS SIASSAKOS STEPHEN  
O'BRIEN TIM DRAYCOTT THOMAS C. MESSERLI STEVE DUMAN MONIKA MESSNER LORENZA  
MONDADA CHLOÉ MONDÉMÉ LAURENT CAMUS DAVID MONTEIRO KRISTIAN MORTENSEN JOHANNES  
WAGNER ISABELL NEISE AXEL SCHMIDT FLORENCE OLOFF ANA CRISTINA OSTERMANN MINÉIA  
FREZZA GÖKHAN ÖZKAYIN VANESSA PICCOLI ANNA CLAUDIA TICCA VÉRONIQUE TRAVERSO KAROLA  
PITSCH ROBERT REINECKE FRANCESCA ASTRID SALVADORI STEVEN SCHOONJANS MONICA SIMONE  
RENATA GALATOLO MICK SMITH PHILIPPE SORMANI ANTONIA STEGER IOANA-MARIA STOENICA  
HANNA MAGDALENA SVENSSON BURAK S. TEKIN BIAGIO URSI CLARISSA WEIß ADRIAN YIP VERONIKA  
SCHOEB ELISABETH ZIMA GEERT BRÔNE

## Instructing and giving advices: The spectator as coach in video game interactions

HEIKE BALDAUF-QUILLIATRE & ISABEL COLON DE CARVAJAL

Lab ICAR Lyon, France

Video gaming constitutes a very popular social activity and most of gaming encounters are more or less part of larger types of interaction. If the research on video games and gaming is very large, EMCA has only been interested very recently in this specific type of interaction. And only in the last year, some analyses have focused on the role of spectators (Tekin & Reeves 2017, Baldauf-Quilliatre & Colon de Carvajal, 2017a, b). From a multimodal perspective, this aspect seems particularly important for us and we want to investigate it further in this paper by analysing one specific function of spectators' activities: Spectators act as coaches and as such they give advice, they instruct or encourage the players. We will particularly focus on instructing and advice giving. In a previous paper we did show that instructions are closely related to encouragements. Our main interest in this paper will be the relation between instruction and advice giving as multimodal actions "in the game" (Mondada 2012), related to multimodal actions "on the screen" (Reeves, Greiffenhagen & Laurier 2017): Whether the linguistic mode is essential for the instructions/advices we did observe in our data, the instructed or suggested action has primarily to be realized as game action. We therefore propose to question which modes are mobilized in the particular setting of video game interactions (axe 1). In our research, we adopted the methodology developed in conversation analysis (Sidnell & Stivers 2012) and interactional linguistics (Couper-Kuhlen & Selting 2001). Our analysis is based on two different game situations, Dragon Ball Z Budokai Tenkashi 3, a fighting game and Kinect Adventures: Reflex Ridge. In both games two participants are involved as players and several other as spectators. The two games are part of a larger corpus of 20 hours of video game interactions in French, recorded within the ICAR research lab.

# From multi-word expression to interaction-organizational resource: A multimodal analysis of 'comment on dit' (how do you say) in second language interaction over time

EVELYNE BERGER, SIMONA PEKAREK DOEHLER & KLARA SKOGMYR MARIAN

University of Neuchatel, Switzerland

CA research on second language (L2) learning has been increasingly concerned with documenting change over time in speakers' L2 practices for social interaction. To this date, a number of studies have tackled this issue by means of longitudinal investigations on the development of interactional competence (e.g. Hellermann, 2008; Pekarek Doehler & Berger, 2016). In this paper we extend this line of research by scrutinizing the role grammatical resources play in this process in concert with embodied conduct. With the use of video analysis, grammar has been reconceptualized as a resource for social interaction (cf. Ochs, Schegloff & Thompson, 1996) that is inextricably intertwined with other semiotic resources, such as gaze, gesture and posture (Goodwin 1979, 2007; Mondada 2006; inter alia). When it comes to second language development, how grammatical resources change along with other embodied conducts still needs to be explored. In this paper, we document change over time in the social-interactional purposes L2 speakers accomplish by means of the multi-word expression 'comment on dit' (how do you say) and concurrent embodied conduct. The study is based on 15 hours of video-recorded mundane conversations between L2 French speakers over twelve months. We analyze the use of 'comment on dit' by three target participants. Results indicate that the expression serves at first as an overt request for assistance; this use is typically accompanied by speaker's gaze on recipient. Over time, speakers increasingly use the expression as a marker of cognitive search and a floor-holding device; in these latter uses, speaker often deploy a "middle distance" look (Goodwin & Goodwin 1986), and the expression bares features that suggest its grammaticalization (Hopper & Traugott, 1993) into an interaction-organizational device. We consider this progressive routinization of an L2 grammar-body package for organizing social interaction as an integral part of speakers' developing L2 interactional competence.

## Measuring distance in pétanque games: Participants' practices and ways of filming

**SOFIAN A. BOUAOUINA, LORENZA MONDADA, HANNA M. SVENSSON & BURAK S. TEKIN**

University of Basel, Switzerland

This paper discusses a corpus of pétanque games video recorded in Basel from a double perspective, developed within multimodal EMCA (ethnomethodology and conversation analysis). On the one hand, we focus on a particular recurrent practice in pétanque – measuring distances. On the other hand, we discuss alternative ways of filming the game, allowing alternative accesses to that practice.

In pétanque, a recurrent issue concerns the evaluation of distances between the balls: this evaluation is crucial not only to the ongoing game for deciding who will play next but also to calculating the final result and choosing the winner. Participants engage in various practices for evaluating distances: they can glance from afar, without moving away from the throwing point, but they can also walk into the gaming field and inspect it by using their own body (feet, hands, etc.) or measuring chords as instruments. In this talk, we focus on these latter practices, further differentiating between various methodical ways of measuring, assessing and contesting distances.

Providing a systematic analysis of these measuring practices, we also discuss alternative ways of video-documenting them. In our study, we discuss various views, gathered through different cameras – sport cameras, 'traditional' cameras and drone cameras – and their adequacy for the analysis of the phenomenon at hand. In particular, we discuss the difference between a perspective view and a drone view – highlighting their advantages and challenges. For instance, the drone view allows for vertical perspectives that are at odd with participants' perspectives but document the game configuration with cartographic precision. The paper discusses how these alternative views can produce contradictory access to the phenomenon. As such the paper contributes to EMCA considering the video recording methodology as strongly embedded within conceptual and analytical issues (Heath, Hindmarsh & Luff, 2010; Mondada, 2006; Broth, Laurier & Mondada, 2014).

# On the role of eye gaze in competition for talk: Simultaneous starts in triadic interactions

GEERT BRÔNE<sup>1</sup>, CLARISSA WEIB<sup>2</sup> & ELISABETH ZIMA<sup>2</sup>

<sup>1</sup>University of Leuven, Belgium | <sup>2</sup>University of Freiburg, Germany

The recent video turn in linguistics - and most notably, the possibility to use eye tracking for interaction research – has stimulated research into the non-verbal cues that regulate turn-taking (Mondada 2007, Rossano 2012, Stivers & Rossano 2010, Holler & Kendrick 2015, Auer 2017, a.o.). This paper ties in with this development by focusing on the role of gaze for overlap resolution (Schegloff 2000) at turn beginnings, when two or more speakers simultaneously start new turns. We analyze data from six triadic conversations (both in German and Dutch, 90 minutes in total), in which the participants' gaze behavior was measured using mobile eye-tracking glasses (SMI and Tobii eye-tracking glasses). We compare the gaze behavior of all three participants in cases of: (a) simultaneous starts (defined as two turns starting within 0 - 200 milliseconds, Walker 2015) of two (or very occasionally even three) speakers that result in only one speaker finishing his/her turn while the other one(s) abandon their turn and (b) sequences in which two or more speaker start a turn simultaneously and they all finish their turn, resulting in a longer stretch of overlapping speech. The analysis reveals three dominant gaze patterns. First, speakers are more likely to finish their turn if they manage to secure their recipients' gaze. This pattern is highly robust and dominant in different conversational activities, such as storytelling, joint brainstorming activities, discussions etc. The gaze analysis hence corroborates prior claims on the influence of recipients on speakers' behavior in conversation (Lerner 2003). Furthermore, we found that if competing speakers were engaged in mutual gaze before the simultaneous start, the one that first withdraws his/her gaze is more likely to abandon the turn. However, if competing speakers did not share mutual gaze before starting to talk simultaneously, the speaker who first gazes at the competitor is more likely to drop out.

## Online text editing in pursuit of readers in news production for the web

**MATHIAS BROTH & MATTIAS ARVOLA**

Linköping University, Sweden

Statistical information about numbers of reads and unique readers of particular news items is today generally available in news production. Based on eight hours of video recordings (using five cameras) made at a Swedish crossmedia news group, this ethnomethodological and conversation analytic study takes an interest in how statistical renderings of readership per day and numbers of reads and unique readers as per the last 15 minutes inform web editors' decisions and linguistic work practices (language: Swedish). The studied media group defines success partly in terms of reaching a certain number of clicks per day. Throughout the workday, the current average number of reads is ongoingly monitored by a group of web editors, whose task it is to publish journalistic content on the web, as well as to promote reading of the different news items. When click numbers for a particular item do not reach the set target values or go down, a web editor may rework its blurb (header, brief introductory text and image combined) in an effort to boost its click rate. This amounts to a sequential response, by a web editor interacting with the journal's readership in a special and mediated way: the collective audience of readers is represented and perceived as a recurrently changing number on the web editor's screen. Different linguistic editing practices were identified for the publication and republication of blurbs on the news sites, e.g. the working in of "prospective indexicals" (Goodwin 1996) into the blurb's textual parts. Working under strict spatial constraints, web editors are further shown to orient to the sheer spatial extension of words and syntactic constructions, and may e.g. thus chose one word or construction over another just because of its length in characters. These spatial constraints of the blurb thereby reflexively configure the language being published.

## Interactional accomplishment of calculations in scrum meetings

**SAFINAZ BÜYÜKGÜZEL & JOHANNES WAGNER**

University of Southern Denmark, Denmark

In workplace interactions, people occasionally have to calculate numbers as part of their ongoing activity. This EMCA study focuses on the joint interactional accomplishment of calculations as they unfold sequentially during workplace meetings. These calculations involve talk, embodied conduct of participants, and their use of artefacts. As a mundane practice in workplace settings, calculations are an integral part of workplace interactions. Only few studies have investigated the social organization of calculations in workplaces (Anderson et al. 1989; Hughes, 2011; Clarke et al. 2006). By calculating, participants engage in an activity that favors the use of specific artefacts as e.g. white boards, magnets, calculators and other office materials that participants use to visualize tasks and to keep track of task objectives. Thus, participants treat these objects as a resource and orient to them as instrumental for their activity. For instance, in software development meetings calculations are used for negotiating technical challenges, keeping track of deadlines, backlogs and resources mainly with the use of scrum boards. The data for this study (100+ hours of video recordings) have been collected in scrum meetings in a Danish software development company. Scrum meetings are common practice in agile software development (Schwaber & Beedle, 2002). In these periodical meetings company employees meet in front of their scrum board to share information, assess finished work tasks, and plan future tasks. Scrum meetings are a prime location for observing workplace interaction in software companies since task oriented collaboration is coordinated, designed, assigned, assessed, and accomplished here.

## Filmic reflexivity in action: The interactional production of objectivity

LAURENT CAMUS

University of Basel, Switzerland

Filming and accounting for an action is not only an analytical problem. It is also, and firstly, a practical problem for participants engaged in the activity of filming a scene. This is particularly the case for TV professionals who film, edit and comment events in real time. This communication will be based on the video observation and analysis of the activity in a control room during the live broadcasting of football matches for television. It will focus in particular on the way the participants treat undefined phenomena of the game and, by the production and use of new images, achieve to qualify or requalify this action and, thus, transform an uncertain situation into an objective and tangible fact. In various situations like faults, handball, offside, new video images like replays can be used as a support for a journalistic narrative and sometimes referee's judgments. In these cases, the qualification of action is fundamental for the proceedings of the game and its TV description. This communication will be part of topic 1 of the conference. Drawing on previous work on accountability of video practices at work (Broth et al., 2014), including TV setting (Broth, 2009), medical practices (Mondada, 2014) or justice (Licoppe, Veyrier, 2017), this communication will mobilize a collection of excerpts by using multimodal analysis in a EMCA perspective. It will show how participants question the status of an action that has just taken place and, thus, produce and mobilize new video images in order to account for the situation in the game. This paper thus aims at producing an empirical inquiry into mundane uses of video in professional setting, and a theoretical reflection on the role of video in the objectivisation of reality.

## More than words can say: Embodied multimodality in an English-medium university lecture

MONICA CLUA

Universitat Internacional de Catalunya Barcelona, Spain

The expansion of English-medium instruction (EMI) in European universities has led to questions about how non-native speakers of English mobilize semiotic resources for interaction and knowledge construction in a class in which English is the academic lingua franca. This paper presents a qualitative case study that aims to describe a lecturer's deployment of embodied multimodal resources in an EMI class, and how the interplay of these modes accomplish aspects of interactional competence. The audio-visual data examined in this study were collected from the opening lecture of Immunology and Microbiology, a subject offered by one Catalan university to mostly international students undertaking the English-track Dentistry Degree program. Multimodal Interactional Analysis (Norris, 2004) was used to organize the data and focus on modal ensembles comprised of gesture, gaze, proxemics, talk and prosody. Central to the analysis was the notion of the university lecture as an interaction (Goffman, 1981). Within this institutional context, interactional competence, which is the successful collaboration between participants to construct a sphere of shared meaning (Young, 2011), and multimodality as a situated meaning-making system (Jewitt et al. 2016), provided a prism through which to examine principally how the lecturer orientated to the students to accomplish pedagogical goals. The sequences demonstrated instances where the deployment of multimodal resources served to provide emphasis and nuance to talk, elicit participation, manage dynamic interactional spaces and acknowledge student contributions to subject content construction as part of the enactment of teaching strategies. Aimed to advance the understanding of multimodality as a semiotic system in the context of higher education, this study provides evidence that multimodal resources, rather than being accessory, might be fundamental to the task of teaching in EMI settings.

# Recipients' gaze in guided tours with sign language interpreters: On the temporal dimension of spatial reference

ELWYS DE STEFANI & ISABELLE HEYERICK

KU Leuven, Belgium

Reorienting the participants' gaze toward an object of interest is a recurrent activity in guided tours and guides overwhelmingly use deictic resources to initiate such reorientations. Whereas hearing recipients may reorient their gaze to some material referent while continuing listening to the guide's explanations, this is not possible for deaf people following a tour interpreted into sign language: gaze reorientation entails more or less prolonged gaze withdrawal from the interpreter; participants hence run the risk of missing bits of the guide's explanations. Thus, gaze reorientation potentially jeopardizes deaf individuals' access to the information provided. Our analyses draw on a corpus of three tours by professional guides, which are interpreted into Flemish Sign Language. Deaf participants may respond in four different ways to the introduction of a visually accessible referent. They may a) maintain their gaze on the interpreter – thereby prioritizing access to the information provided, b) look at the referent while the interpreter continues signing – thereby prioritizing access to the material referent, c) shift their gaze toward the object of interest when the interpreter is reaching a momentary halt of a turn – thereby managing to look at the object without missing the explanations provided, or d) look at the referent as a response to the (speaking) guide's pointing gestures. In addition to describing deaf participants' available resources for coordinating face-to-face orientation with momentary object-related gaze reorientation, this contribution shows how deaf participants display their online analysis of the turn-constructive features of a turn in sign language. Furthermore, it discusses spatial deixis as interactionally accomplished in a way that is sensitive to the temporal unfolding of turns, as well as to the participants' perceptive abilities and modal properties of talk vs. signed language. Our approach to studying deixis and gaze reorientation is based on conversation analysis and multimodal interaction analysis.

## Sensibiliser les pouvoirs publics au rôle du médiateur numérique via la vidéo in situ (projet ITAC)

**CHRISTINE DEVELOTTE, HEIKE BALDAUF-QUILLIATRE & JUSTINE LASCAR**

Lab ICAR Lyon, France

Dans cette communication qui s'inscrit dans l'axe 3, nous nous intéresserons à l'utilisation de l'analyse multimodale des interactions pour la conception des formations à destination des professionnels des bibliothèques et des représentants des instances culturelles et sociales au niveau de la ville et de la collectivité territoriale. Les données étudiées, des interactions entre les membres de différentes générations d'une même famille et un médiateur de la bibliothèque autour d'une table numérique, proviennent du projet « Interactions intergénérationnelles Avec éCrans » (ITAC), une collaboration scientifique entre deux laboratoires de recherche de Lyon (ICAR et ELICO) et la Bibliothèque Municipale de Lyon. Il a pour objectif d'étudier les utilisations numériques intergénérationnelles proposées par la BML en se centrant sur les aspects relationnels, linguistiques et éducatifs associés à ces interactions et à leur médiation par les professionnels. Nous expliciterons d'abord le dispositif de recueil de données vidéos multi-sources (5 pistes vidéo et 3 pistes audio) et son traitement numérique. Ce corpus complexe de 2 heures d'animation ludique nous permet de prendre en compte un grand nombre d'aspects praxéologiques et linguistiques dans l'analyse des interactions. Dans un deuxième temps, nous montrerons à travers un extrait quelques éléments relevés dans l'analyse qui nous semblent révélateurs pour comprendre la particularité de ce type d'interaction. L'interaction étudiée plus en détail compte trois participants : le médiateur, un petit garçon et sa grand-mère. Ces particularités ayant été également pointées par les médiateurs comme étant des difficultés ressenties, nous ont amenées à réfléchir à la conception d'une formation pour les professionnels des bibliothèques. Nous préciserons donc par la suite l'intérêt que peuvent détenir de tels corpus pour la formation de ces professionnels ainsi que pour les représentants des pouvoirs publics sur le plan de l'information scientifiquement étayée propre à déboucher sur des décisions concernant les institutions culturelles.

# The interactional role of the mirror in show dance classes

JESSICA DOUGLAH

Stockholm University, Sweden

This presentation will focus on the interactional role played by the mirror during instructions in show dance classes. Previous research has highlighted the importance of the mirror in Lindy Hop dance classes with the focus on “bodily quoting” through the mirrors (Keevallik 2010). The activity, dance class rehearsals, is not talk-oriented but rather a collaboration between talk and multimodal aspects. Accordingly, multimodal interactional analysis is used for data analysis (Keevallik 2010, Broth & Keevallik 2014). My corpus consists of 18 hours of video recorded classes of show dance rehearsals. The dancers are non-professionals and their levels vary from beginners to advanced groups. For this presentation, I will show examples from one of the rehearsals with the advanced group. I will explore how the mirror functions as a tool for the leader when giving instructions. In my data, the leader faces the mirror so the students can see the leader perform the different steps in the same directions as the students should perform them, but it is also a way for the leader to check whether the students are doing the choreography in a correct way. The students both catch the instructions in the mirror, but also use the mirror to see themselves and what the rest of the group is doing. The sequential analysis of the instruction highlights how the mirror can function as a potential third person in the interaction. Based on video clips from my data I will show how the mirror plays an important interactional role. Often the meaning is not being talked into relevance, but instead the meaning is being watched into relevance.

# Interacting with a telepresence robot: Mobility, space and embodied practices

**BRIAN DUE**

University of Copenhagen, Denmark

A telepresence robot enables visual and audible access to remote settings through a video camera that can be controlled remotely. See an example from a commercial context here: <https://www.youtube.com/watch?v=R-upe-01EPQ> Recently, there have been a mobility turn in video-mediated research, paying more attention to not only the fixed computers/cameras, but also use of smart technologies, which affords moving around (Licoppe & Morel, 2014). However, a telepresence robot is quite different than the typically examined object (smartphone) because it is not carried and used in situ by a present person, but is a large self-contained moveable object (robot). Consequently, it occupies a physical position in space, it is mobile, controlled by a dislocated actor and it is seemingly oriented to as a man-machine-figuration: a cyborg (Clark, 2004). Based on multimodal interaction analysis (Streeck, Goodwin, & LeBaron, 2011; Mondada, 2014), and specifically within the mediated interaction tradition (Arminen, Licoppe, & Spagnolli, 2016), this paper reports on findings from a nursing home in Denmark, where a doctor is virtually present through a telepresence robot (BeamPro). The paper will provide an analysis of the spatial and embodied practices of the (remotely controlled) robot. We will show how 1) the object (robot) and the doctor is oriented to by situated participants as one agent in a participation framework with shifting contextual configurations (Goodwin, 2000), e.g. demonstratively paying attention to spatial distance and mutual positions in space (Goffman, 1963; Hall, 1966; Kendon, 1990; Mondada, 2009). And 2) how the man-machine-figuration (cyborg) performs embodied actions through (machine-)gaze, (machine-)head turning and (machine-)posture (cf. Goodwin, 1979, 1980, 1981; Kendrick & Holler, 2017). The paper is based on approximately 30 hours of video recordings from 5 situations within the same nursing home in Denmark. Each situation was synchronically recorded from three different locations: one GoPro on the robot, one GoPro filming the robot on a distance and one GoPro filming the doctor in his office.

## Composing terror: Analyzing video interviews of “women of ISIS”

ZOE DUPREE FINE

University of South Florida, USA

In our constantly evolving, networked world, how we communicate about terrorism through images, video, and written words, not only constitutes, but also amplifies the effects of terrorism. This paper contributes to the conference topic of multimodal analysis. The phenomenon investigated is real women’s accounts of being recruited into, and in some cases serving as members of the so-called Islamic State (IS). The video data, all in English, are three publicly accessible, online interviews, each featuring a woman telling public audiences about her experiences joining IS. One is the CNN video interview of “Khadija,” a petite 25-year-old schoolteacher turned leader of IS’s all-women’s brigade, called, “How she went from a schoolteacher to an ISIS member” (Damon & Tuysuz, 2014). Another is the New York Times interview of “Alex,” a 23-year-old Sunday school teacher and babysitter, called, “Flirting with the Islamic State,” produced by Poh Si Teng and Ben Laffin (Callimachi, 2015). The third is police interview of Tareena Shakil, a British citizen who visited to Syria with her toddler son, published by Birmingham Mail, called, “Tareena Shakil Police Interview: Entering Syria” (Larner, 2016). The analytical tools used include multimodal discourse analysis (Jewitt, 2014; Jewitt, Bezemer, & O’Halloran, 2016), visual semiotics, multimodal (inter)action analysis (Norris, Sigrid, & Daniela, 2014), and conversation analysis, the organization of turn-taking in particular. In addition to offering close analyses of the communication strategies employed in the interviews, the paper addresses the rhetorical power of international news media’s exploitation of video content. One aim of the project is to show how accounting is done through online media video interviews of real women who joined a terrorist group. Another aim is to underscore international media outlets’ use of representations like these interviews of three women to not only shape the public’s understanding of terrorism, but also tell the public stories about why and how terrorists become terrorists, and offer recommendations about how terrorism should be countered.

# Orchestrating openings: The emotional importance of the first five seconds in video calls between migrant parents and their ‘left behind’ children in China

YUMEI GAN<sup>1</sup>, CHRISTIAN GREIFFENHAGEN<sup>1</sup> & CHRISTIAN LICOPPE<sup>2</sup>

<sup>1</sup>Chinese University of Hong Kong, China | <sup>2</sup>Télécom ParisTech, France

A number of studies have investigated how people accomplish the entry into interaction, whether in telephone conversation (Schegloff, 1967) or face-to-face encounters (Pillet-Shore, 2012). Recently, a few studies have explored this in the context of video calls, pointing to the work of achieving a ‘talking head’ configuration (Licoppe & Morel, 2012) and the sequential significance of visual appearance (Licoppe, 2017). In this paper, we focus on two additional aspects of the ‘orchestration’ of openings in video calls between migrant parents and their ‘left behind’ children: firstly, how participants integrate (indeed: socialize) young children, who are not yet fully competent with these practices; secondly, the ‘emotional significance’ of these sequential practices in the interaction between a remote parent and their children (most of these parents only see their child once a year, so the achievement of proper recognition and greeting in a timely manner is important to feel ‘connected’ to their children). Our data is taken from video calls between migrant parents (who have moved to the cities for jobs) and their young ‘left-behind’ children (who are left behind to live with their grandparents in rural areas, aged 9 months to 3 years) in China. We have so far collected 37 cases, screen-capturing their mobile phone (e.g., Licoppe, 2017) as well as video-recording the interaction in their home (e.g., Sunakawa, 2012). We examine the openings in two sequential environments. Firstly, we explore the sequences, where a parent greets first and the child has to produce a timely return greeting. Secondly, where the adult participants delicately manage the opening through ‘shepherding’ work (Cekaite, 2010), so that a sequential slot is offered first to the child, making interactionally relevant with their production of the first greeting (e.g., grandparent positions the mobile phone to child’s face, and/or remote parents withholds producing an initial utterance).

## Translating positions into perspectives: A comparison of activist and journalist videos of protests in Rio de Janeiro

JACOB GEUDER

University of Basel, Switzerland

Political, social, cultural and economic positions are inevitably inscribed in videos about protests. In this presentation I pose the question how this process of translation can be identified in the traces of audiovisual representations of protesters attacking the legislative assembly, ALERJ, in Rio de Janeiro on the 17th of June 2013. By comparing two videos of the same event I aim at linking back filmic elements – such as the quality of the audiovisual footage, its editing, the perspectives of filming and the narrative created in editing – to the positioning of the respective actors who have produced these videos. On the one side, there is the Television station Globo as an actor. The TV station is part of the biggest media conglomerate in Brazil, whose rise is intrinsically linked to the times of the military dictatorship of the 1960s to 1980s and who is said to represent interests of ruling elites in Brazil until today, while simultaneously being the most popular media producer of the country. On the other side there is the “independent” media collective MIC as one of the many collectives of videoactivists that arose during the massive protests of the Jornadas de Junho in 2013. The collective claims to represent citizens’ voices and views “from below”. Members of this collective explicitly consider themselves as activists and citizens. A positionality that would characterize the Globo journalists as “professional outsiders” and the videoactivists as “involved insiders” will be critically examined in this analysis of the two videos. How does the position of the actors translate into the perspective they present in their videos? The objective of the presentation is to deepen the understanding of the links between production processes of audiovisual materials and the resulting product, the video. The presented proposal forms part of my PhD-project in which I analyze videoactivism as an urban phenomenon and how the phenomenon is being embedded in the specific contexts of Cape Town and Rio de Janeiro.

## Sequencing and unpacking powerpoint slides in university lectures

FRODE GURIBYE<sup>1</sup> & OSKAR LINDWALL<sup>2</sup>

<sup>1</sup>University of Bergen, Norway | <sup>2</sup>University of Gothenburg, Sweden

In this study, we investigate how lecturers use powerpoint slides to structure their talk. Goffman (1981) identified three distinct forms of delivering talks: memorization, reading aloud, and fresh talk. The use of powerpoint can be seen as a hybrid of these, as the lecturer often reads part of the powerpoint aloud while other parts are produced as elaborations of the words, sentences, figures, pictures or other visual material available in the slides. We have filmed 12 two hour lectures at the University of Bergen, Norway, and collected the slides used in these talks. All lectures were in Norwegian, except one given in English. Drawing on previous studies in multimodal conversation analysis, including studies of lectures and presentations (Knoblauch, 2013; Rendle-Short, 2006), we focus on sequences where the lecturer shift from one slide to another. In the work of “talking through a subject such that it may be learnt” (Sharrock & Anderson 1982, p. 171), the deck of powerpoint slides provides the lecturer with a ready-made sequence of “talking points” that can be unpacked. The transitions from one slide to another typically include the closing of one topic and opening up another, which is reflected in intonation, gaze, body orientation, movement and so on. Before a slide shift, the lecturers shift their orientation from the audience to the computer, and physically move in order to get access to the computer, while the content of the slide is summarized. The transitions are often followed by a pause, and a prospectively oriented “so”, before the title of the slide is read out loud. The sequential order of the powerpoint is thereby enacted in and through the verbal and embodied performance of the lecturer.

## La vidéo comme un outil d'auto-confrontation à l'habitus des enseignant-e-s : le cas des enseignant-e-s généralistes dans le Canton de Vaud

**MANDIRA HALDER**

Université de Genève, Suisse

L'analyse des objets enseignés dans une langue à travers les outils mobilisés par les enseignant-e-s (Schneuwly, 2000; Schneuwly & Dolz, 2009) ne permet pas de comprendre l'écart que ces enseignant-e-s constatent entre leurs conceptions des objets enseignés dans les planifications et les objets mobilisés en classe. Nous faisons l'hypothèse que cet écart est lié à l'habitus de l'enseignant-e (Perrenoud, 2010, 2012) qui est peu thématiqué dans l'analyse de son travail du point de vue des objets enseignés en classe. En revanche, l'analyse ergonomique du travail (Yvon & Garon, 2006) permet aux enseignant-e-s de se confronter à leur habitus à posteriori au moyen d'un retour réflexif sur leur action en classe. Toutefois, cette analyse menée par le biais de l'auto-confrontation simple et croisée (Clot, 2011; Yvon & Saussez, 2010) est peu reconnue au niveau institutionnel comme un outil de formation des enseignant-e-s. Afin de mieux comprendre le rôle de l'habitus des enseignant-e-s dans l'écart entre leurs conceptions des objets enseignés dans les planifications et les objets qu'ils/elles mobilisent en classe et les enjeux pour la formation des enseignant-e-s de langues, nous articulons une approche didactique des objets enseignés avec l'analyse ergonomique du travail. Ainsi, cette communication s'intéresse au rôle de la vidéo comme un outil d'auto-confrontation de 8 enseignant-e-s généralistes en formation à leur habitus dans l'enseignement de l'allemand en tant que première langue étrangère auprès des élèves débutant-e-s des écoles obligatoires vaudoises. Pour mettre en évidence le retour réflexif des enseignant-e-s sur leurs actions en classe, nous analysons le discours de ces enseignant-e-s dans les entretiens d'auto-confrontation simple d'une durée de 60 minutes et des entretiens d'auto-confrontation croisée d'une durée de 30 minutes avec leurs praticiennes formatrices. Nos données montrent que l'écart entre les conceptions des objets enseignés et leur mobilisation s'expliquent par l'absence de l'explicitation du langage non-verbal (Tellier, 2006), utilisé pour faciliter l'appropriation de l'allemand chez les élèves.

## Choosing and using instructional videos

THOMAS HILLMAN<sup>1</sup>, OSKAR LINDWALL<sup>1</sup> & BARRY BROWN<sup>2</sup>

<sup>1</sup>University of Gothenburg, Sweden | <sup>2</sup>Stockholm University, Sweden

With the advent of widely available video cameras, internet connectivity and smartphones, the once cumbersome process of producing and distributing video has become a relatively simple task. For individuals seeking to develop their competences or learn a particular skill, large scale social media platforms such as YouTube have made available a new massively prevalent social and pedagogical phenomenon: the online instructional video. There are literally millions of videos that provide step-by-step instructions on mundane and exotic skills – how to use chopsticks, calculate taxes or learn complex crafts. We have a transdisciplinary research project examining online instructional videos, including the infrastructure of distribution, their use and production, and the organisation of the videos themselves. A relevant concern for us as researchers and for the user of instructional videos is how and on what grounds a particular video amongst the millions available is selected. In the case of YouTube where the search term ‘how-to’ returns almost 555 million results, the corpus of videos available for analysis is enormous. With YouTube becoming the dominant distribution platform, there has been a shift from the earlier manual curation of websites such as Instructables and Wikihow, to algorithmic curation. This has consequences both for how we as researchers organise corpuses and for how users find videos on particular topics, since selections are now predicated on the wisdom of the crowd of millions of other users (Niederer & Van Dijck, 2010). We approach these issues by investigating metadata made available by the platform such as titles, descriptions, likes, dislikes, comments and view-counts through digital methods (cf. Lomborg & Bechmann, 2014), by interviewing users and producers of instructional videos (cf. Dubois & Ford, 2015), and by video-recording and closely analysing how users chose and make use of instructional videos in their everyday life (cf. Ayaß & Gerhardt, 2012).

# Eye-tracking and Conversation Analysis: A theoretical discussion of the approach to gaze 'from within' interactants

**KENAN HOCHULI, WOLFGANG KESSELHEIM, CHRISTOPH HOTTIGER & CHRISTINA BRANDENBERGER**

University of Zurich, Switzerland

Research with eye-tracking devices allows new insights in the analysis of social interaction, but also poses new challenges to it. Given that the important role of gaze for interaction has been demonstrated (cf. Argyle/Dean 1965; Kendon 1967; Goodwin 1980; for video-based studies cf. Kidwell 2005; Rossano 2013 and others), eye-tracking (ET) technology promises a more precise measurement of this interactional resource.

Up until now, a handful of studies have used this technology to investigate in more detail how turn transition or strategies of evading it works in dyadic or triadic conversations (cf. Weiss/Auer 2016; Holler/Kendrick 2015, Kendrick/Holler 2017; Brône et al. 2017). Setting up a quasi-experimental setting in combination with a limited space of perception of the participants allows the authors of these studies to use ET videos without an in-depth reflection of the type of data they are using. However, the pragmatic quality of proceeding in this way cannot hide the fact that the representation of an 'inner' perspective of the individuals who are wearing ET glasses poses a theoretical and methodological problem to ethnomethodological conversation analysis: So far, its full attention had been devoted to the 'outer' resources participants use to make aspects of their interactions accountable to each other. The new type of data begs the question if – when working with ET data – conversation analysis must 'look into participants' heads' and thereby give up its strict orientation towards what is "observable-and-reportable" (Garfinkel 1967: 1) in the interactional situation.

This question will serve us as a basis for a fundamental theoretical discussion about the role of eye-tracking as an instrument to record interactive phenomena in conversation analysis. By contrasting every-day situations in which participants interact in face-to-face and in side-by-side formations, we will develop an analytical assessment of the meaning of video data that is created by ET technology. We will show how ET data allow insights into the organisation of individual attention that is always tied to the interactive procedure. At the same time, we will demonstrate how distorting the gaze focus visible in ET videos can be in pointing out the relevance of the events in a situation.

## Children’s use of imitation of locomotor play

ELLIOTT M. HOEY<sup>1</sup>, DAVID DELIEMA<sup>2</sup>, RACHEL CHEN<sup>2,3</sup> & VIRGINIA J. FLOOD<sup>2</sup>

<sup>1</sup>University of Basel, Switzerland | <sup>2</sup>University of California, Berkeley, USA | <sup>3</sup>San Francisco State University, USA

Children commonly engage in social locomotor play—the horsing around, running about, and physical contact characteristic of playgrounds and neighborhood streets. To adults and outsiders, this activity may look like undisciplined chaos. However, the very recognizability of locomotor play—the fact that children observably pull it off together as a discriminate activity—points to a set of shared practices for its accomplishment. In this article, we investigate children’s methods and resources for organizing social locomotor play. Using video recordings of 5-6yo children playing during an immersive technologically-mediated science lesson, our analysis shows in fine detail some of the orderliness of social locomotor play. In particular, we demonstrate how children use imitation to organize the initiation and progression of this activity.

## Becoming the third party: Apprentices embodying a summons in hair salons

**ANNE-SYLVE HORLACHER**

University of Basel, Switzerland

Hairdressing service encounters are sometimes conceptualized as dyadic exchanges between professionals and clients (Oshima, 2009). In many cases, however, an apprentice (Billett, 2001) becomes an additional party in the encounter, hence altering the participation framework (Goodwin & Goodwin, 2004; Mondada, 2012).

This paper investigates how apprentices position themselves in spatial proximity of the hairdresser-client encounter, thereby embodying a summons. Indeed, apprentices approach the hairdresser whenever they have questions about how to pursue their own work. Their practical problem consists in (re)initiating interaction with the hairdresser while preserving the privileged relation between the client and the professional, without jeopardizing the latter's work. This presentation provides a detailed analysis of the multimodal resources apprentices call on in maximizing their visibility, thereby turning their presence into an accountable event: Apprentices may either try to identify an appropriate slot to start talking or wait for the hairdresser to invite them to formulate the reason for their presence. On occasion, the hairdresser does not disengage from his interaction with the client, thereby treating the apprentice's presence as not legitimate just now. The apprentice's summons may be treated in different ways: a) The hairdresser responds to the summons and selects the apprentice as next speaker; b) The client responds to the apprentice's summons and selects him/her as next speaker; c) A second apprentice joins the scene and helps his/her peer in upgrading his/her summons; d) The apprentice unsuccessfully leaves the scene. Drawing on Conversation Analysis (Sacks, Schegloff and Jefferson, 1974) and multimodal interaction analysis (Streeck, Goodwin and LeBaron, 2011), I analyze a collection of 15 episodes from a corpus recorded in a hair salon of the French-speaking part of Switzerland (18 sessions involving 18 clients, 1 hairdresser, 5 apprentices; 20 hours).

## Promoting research and raising awareness on social interaction and societal issues through video data

**EMILIE JOUIN-CHARDON<sup>1</sup>, JUSTINE LASCAR<sup>1</sup>, VANESSA PICCOLI<sup>1</sup>, ANNA CLAUDIA TICCA<sup>2</sup>, VÉRONIQUE TRAVERSO<sup>3</sup> & BIAGIO URSI<sup>1</sup>**

<sup>1</sup>Lab ICAR Lyon, France | <sup>2</sup>Lab LIDILEM Grenoble, France | <sup>3</sup>Institut du Proche-Orient Beyrouth, Lebanon

In the last decades, Conversation Analysis has experienced an “embodied turn” (Nevile 2015): thanks to technological advancements, naturalistic video data have become more and more available, and the research focus has moved from verbal to multimodal interactions. As a result, researchers are nowadays faced to the use of video data, to their possible applications for teaching, professional training and research’s popularisation. The ICAR Lab has a long tradition in the study of social interaction and in collecting video corpora in naturally occurring settings. Through a large variety of projects, it has shared answers to methodological and deontological questions concerning fieldwork and technical aspects of video recording (see Jouin-Chardon et al. 2010). More recently, our research group produced short films for promoting our research activity and raising awareness about societal issues (Piccoli & Ursi 2015, DISAL project, REMILAS project). This contribution will relate our experience in film creation for such purposes. In a first project, three short films issued from three different corpora were produced. Initially designed as a feedback for participants, in an attempt to improve mutual professional relationship (see Baude et al. 2006), they also turned out to be a useful tool to teach the interactional approach and to promote research to a widespread audience. In another still on-going research on multilingual communication between healthcare providers and migrants in France (the REMILAS project), video data are used not only for scientific purposes, but also for healthcare professionals and interpreters’ training, as well as artistic initiatives aimed at promoting a more informed and active citizen participation to nowadays society.

# Articulating touch and instruction in parents' and young children's accomplishment of everyday tasks

**SARA KEEL**

University of Basel, Switzerland

This paper looks at parents and children's interactive accomplishment of every day tasks, such as grating cheese, cleaning an ear with a q-tip, lighting a candle. Based on a large corpus of audiovisual recordings of French-speaking families filmed in their homes, it offers a detailed analysis of interactive moments, in which parents' instruction how to handle a knife, a q-tip, a match, are finely articulated (preceded, accompanied and/or followed up) with parents' use of touch to intervene in the child's handling of the object.

Previous studies on the social organization of touch within everyday family life and/or pre-school settings have highlighted that touch is not only essential for creating intimacy and affection (Goodwin 2017a), but also constitutes an important resource for parents and caretakers to manage a diversity of educational activities: such as shepherding children's compliance with directives (Cekaite 2010, 2015; Goodwin & Cekaite 2013), comforting young children when expressing distress (Cekaite & Holm 2017), controlling their participation and attention to particular activities (Cekaite 2016), handling their apprenticeship in everyday tasks (Goodwin 2017b).

By focusing on the ways parents' articulate instructions and touch in order to manage the child's handling of objects within the accomplishment of everyday tasks, the paper seeks on the one hand to show how parents might not only use touch to complement a verbal instruction, thus enhancing its intelligibility, but also to speed up or otherwise influence the child's compliance with it. By showing in turn, how the child might respond to distinct embodied instructions with compliance or resistance, the paper seeks on the other hand to put the spotlight on the young child's understanding of parental embodied interventions in her/his accomplishment of everyday tasks. The paper constitutes an empirical contribution and has at the same time applied implications for parents and caretakers of children.

# Multimodale Analyse von Instruktionsgesprächen im Instrumentalunterricht

**JOHANNA KEHR**

Technische Universität Dresden, Deutschland

Im Rahmen des laufenden Dissertationsprojekts wird die Kommunikation im instrumentalen Einzelunterricht an der Musikhochschule untersucht. Im Zentrum steht dabei die Frage nach den verwendeten Gesprächsressourcen für die gemeinsame Bedeutungskonstitution bei der Verständigung über musikalische Sachverhalte. Für eine qualitative Untersuchung dieses bisher in der Forschung nicht berücksichtigten Interaktionstyps wurden acht ca. sechzigminütige Videoaufnahmen von instrumentalem Einzelunterricht an verschiedenen deutschen Musikhochschulen angefertigt. Eine erste Sichtung des Materials hat ein komplexes Repertoire nonverbaler Ressourcen (Vorsingen, Vorspielen, gestische Darstellungsformen) gezeigt. Dies ist zum einen vor dem Hintergrund des in der Forschung immer wieder thematisierten Verbalisierungsproblems in Bezug auf musikalische Sachverhalte zu sehen (vgl. MAHRENHOLZ 2004, 60). Zum anderen ist hier ein Zusammenhang mit dem Imitationslernen als wichtigem Vermittlungsprinzip des Instrumentalunterrichts zu vermuten (vgl. BORK 2012, 13). Drittens scheint Metaphorik für die untersuchten Gespräche eine zentrale Rolle zu spielen, wie sie bereits für musikbezogene Textsorten gezeigt werden konnte (vgl. STÖREL 1997). Hier ist von einem komplexen Zusammenspiel verbaler und nonverbaler Ressourcen in Form von multimodaler Metaphorik auszugehen (vgl. z.B. FORCEVILLE 2009). Der gewählte gesprächsanalytische Zugang für die Beschreibung des Interaktionstyps wurde entsprechend um Methoden der multimodalen Interaktionsanalyse erweitert (vgl. FRICKE 2012). Dadurch soll eine umfassende Beschreibung der Formen und lokalen wie globalen Funktionen der multimodalen Gesprächsressourcen ermöglicht werden. Methodisch bedeutet dies neben der Anlage eines Basistranskripts nach GAT2-Konventionen (vgl. SELTING et al. 2009) zusätzlich eine abschnittsweise Annotation in ELAN (Version 5.0.0-beta 2017) für eine Aufschlüsselung und Analyse der nonverbalen Kommunikationsformen (vgl. BRESSEM et al. 2013). Während für die Gestikbeschreibungen in der Forschung bereits vielfach Verfahren erprobt wurden, stellt die Transkription von Gesang und Musizieren dabei noch eine methodische Neuheit dar, die es im Rahmen der Arbeit und des Beitrags zu diskutieren gilt. Zu klären ist auch, inwiefern Methoden der multimodalen Analyse sich auf eine breitere Datenbasis anwenden lassen.

## What is the phenomenon? The multimodal achievement of joint discoveries in science centres

**WOLFGANG KESSELHEIM, CHRISTINA BRANDENBERGER & CHRISTOPH HOTTIGER**

University of Zurich, Switzerland

Science centres are modern natural science museums which are characterised by the fact that they do not simply display a certain phenomenon. Instead, visitors first have to 'create' and 'discover' the phenomena in question with the help of so-called 'experiment stations'. If visitors jointly make use of the experiment stations, and this is the norm in science centres, the process of discovery becomes an interactive negotiation. We investigate this process from the perspective of multimodally enhanced conversation analysis.

While large-scale scientific processes of discovery have been analysed several times in ethnomethodology and conversation analysis (Garfinkel et al. 1981, 2002, Sormani et al. 2011), the local production of discoveries by means of multimodal resources has scarcely been studied yet (but cf. the first steps in Heath/vom Lehn 2008, Zimmerman et al. 2009). This is what this talk will be about.

The data base of our talk is a corpus of more than 35 hours of video and eye-tracking data which document authentic visitor interactions in the Swiss Science Center Technorama in Winterthur, Switzerland.

Based on video excerpts, we will first show what constitutes the core of the practice of 'discovering' to the visitors. Next, we will use an exemplary video to reconstruct how they use fine-tuned multimodal resources to establish a shared perceptual space (Hausendorf 2010), in which they then focus their joint attention on an object which had not been talked about before. In doing so, we will carve out the differences vis-à-vis the practice of everyday 'noticing' (cf. e.g. Goodwin/Goodwin 2012). In a final step, we will contrast video and eye-tracking data of the same process of discovery. This way, we can show that the intersubjectively produced discovery cannot be traced back to individual perceptions.

Our talk contributes to current research on the multimodal construction of knowledge in interaction (Deppermann 2015) and it addresses the question as to how eye-tracking data can be integrated into conversation analytical studies in a methodologically sound way, a point which has been largely neglected so far.

## Note-taking as recipient action

**TOM KOOLE<sup>1</sup> & PAULIEN HARMS<sup>2</sup>**

<sup>1</sup>University of Groningen, Netherlands | <sup>2</sup>University Medical Center Groningen, Netherlands

In this paper we will report an analysis of the role of note-taking in handover interaction in a hospital intensive care unit. In the interaction both information on the patient and responsibility for the patient is transferred from one doctor to another. One of the tokens by which the recipient signals reciprocity is taking notes. We will argue that different from tokens such as head nods and hmhm-s, note taking is not designed as a recipient token but we will show that it is treated as such by the transferring party. Also we will show that the positioning of note taking in the hand-over activity characterises it as responding to factual information with a steep epistemic gradient between K- and K+. In the third place, we will argue that the positioning of note-taking, either accompanying the production of a chunk of information or positioned at the end of such a chunk, sheds a new light on the distinction that Sacks (1992) made between claiming and demonstrating understanding. Our data are video-recorded simulated hand-over interactions between intensive-care specialists who were given authentic cases to hand over. The hand-over was situated around the bed of the mannequin patient and included full IC equipment and monitors. The recording were made with three static cameras , one overlooking the bed and participants from the head board of the bed and two on either end of the bed. We will include a discussion of the affordances and limits of this recording line-up and of the way we represented the three recordings in one picture frame.

## Videod displays of pain in medical consultations

JESSICA LA & ANN WEATHERALL

Victoria University of Wellington, New Zealand

Pain is commonly conceptualised as a subjective and private experience. However, as made evident by a breadth of conversation analytic research, our internal worlds become social phenomena when they are described or displayed in interaction. Existing research has shown that pain displays are fitted to the sequential and social organisation of an ongoing activity (Heath, 1989), and pain displays can be used to accomplish other interactional work (e.g in the management of getting children to eat) (Jenkins & Hepburn, 2015). The research discussed in this paper builds upon previous conversation analytic research on pain by using video data to describe and analyse the multimodal accomplishment of pain within the institutional context of medical consultations.

Our analysis documents the vocal and embodied features which make displays of pain recognisable to others as pain and its organisation with respect to turns and sequences of talk. 15 clinical encounters between general practitioners and patients make-up the data set, drawn from a larger existing corpus from the Applied Research on Communication in Health (ARCH) Corpus of Health Interactions. The 15 clinical encounters examined constitutes approximately 2.5 hours of video recordings. From the data, 22 separate instances of pain were identified and analysed.

This work describes the multimodal accomplishment of pain displays through the interplay of a range of embodied and vocal resources. For example, some elements of pain display include gasping, vocal cries, grunts, grimaces, wincing, and flinching. By examining the interactional organisation of pain, we also show how displays and responses to pain are sensitive to the local context and wider goals and activities of the medical setting. Specifically, pain plays a role in the coordination of the physical examination and collaboration with diagnostic work. This paper has relevance to the transcription of embodied action using video data. A key part of this research involved developing a transcription system that can capture the varied elements of pain displays. We discuss the multimodal transcription system developed for this research which draws on a range of well-established transcription conventions (Jefferson, 2004; Goodwin, 2000; Mondada, 2014; Hepburn, 2004) that can clearly capture the diverse ways in which pain displays unfold in social interaction.

## Advantages and challenges of the digital turn in linguistic landscape projects on educational settings

PETTERI LAIHONEN & TAMÁS PÉTER SZABÓ

University of Jyväskylä, Finland

This paper aims to advance the methodological state of the art of linguistic landscape research on educational settings (i.e. schoolscape research, see Laihonen & Szabó in press) from the point of view of the digital turn for visual methods from a critical perspective. With examples from on-going fieldwork projects using “traditional” and 360 degree videos in multilingual schools in rural Romania (Romanian, Hungarian and other languages) and urban Finland (Finnish, Swedish and other languages), we ask how co-creating such videos enhance or delimit the integration of researcher’s aims to gain analysable data, and participants’ needs to take part in meaningful activities. Visual methods have been used to generate multimodal data in order to minimize the impact that researchers have in shaping the data. Moreover, participants’ voices and agency are supposed to become foregrounded in the research narrative, too (Pietikäinen 2012). The use of digital technology further helps to engage the participants, however, the technicized fieldwork experience brings new challenges and calls for new, more critical ways to reflect on the processes of producing and analysing the digital schoolscape (see Troyer & Szabó 2017). The methods for analysing digital representations of linguistic landscape are in still in their experimental stage. We explore the challenges and new possibilities to carry out Multimodal Discourse/Interactional Analysis on such digital materials and their contexts. Finally, we will conclude by summing up our insights so far of the gains and challenges of using either a “traditional” video or a 360 degree videos from the point of views of the participants and researchers to a schoolscape project.

# Participatory watching of telecinematic fiction in an international drama viewing platform

MIRIAM A. LOCHER & THOMAS C. MESSERLI

University of Basel, Switzerland

Our research is situated in the field of the pragmatics of fiction and audio-visual translation (AVT) studies (Locher and Jucker 2017; Messerli 2017), and explores an international website that makes Asian drama series and movies accessible to an international audience by means of fan generated subtitles in over 150 languages (www.viki.com; Dwyer 2012). In addition, the site offers a social network and participatory element in that it provides viewers with different possibilities of participation. Next to producing subtitles in teams, members can comment on episodes and actors, rate shows, produce their own videos, write to each other, etc. This paper explores the possibility of viewers commenting on the episode WHILE it is being watched as a dynamic form of active reception. These comments are time-aligned with the video, which acts as the pivot of the interaction. Viewers can then read other fans' comments that pop up in a side bar and disappear again, and they can contribute their own comments and thus their own voice to this additional communal layer of drama series reception. Despite the fact that viewers typically view episodes and read/write comments asynchronously, a simultaneous viewing experience with an international community is created. We present a case study of the viewing of a first episode of a series and show how the commentators negotiate a number of issues: expectations about genre, character development, intertextuality and culture. We demonstrate that engaging with the video through written comments, engaging with each other in these comments and participating cross-linguistically is a highly interactive achievement between different modes of communication of which the video itself is the starting point.

## The camera as contextualization resource in school children's peer talk

**MARTIN LUGINBÜHL**

University of Basel, Switzerland

In this talk I will analyze direct and indirect references to the camera in children's peer talk as resource to negotiate conversational norms and to perform identities. There is a growing interest in children's peer talk in the field of CA (see Arendt 2015, Cekaite et al. 2014, Farnsworth 2012), especially with a focus on opportunities to acquire pragmatic skills (Zadunaisky-Ehrlich/Blum-Kulka 2010, Morek 2015), but also as a sphere where childhood culture can be established and negotiated. With these two planes, children's peer talk can be understood as 'double opportunity space' (Blum-Kulka et al. 2004). In my talk I will focus on Swiss German elementary school children aged 7-12 referencing to the researcher's camera in a semi-experimental setting in the context of argumentative conversations. I will draw on a corpus of 180 conversations of 2nd, 4th and 6th graders (60 conversations each, without the presence of adults), in which children in groups of four are solving a conversation task typical for school. Starting from a reflexive understanding of context (Gumperz 1982, Auer 1986), I will show in four cases analyzed in detail how the children use the camera as contextualization resource. These contextualizations take place in the context of task related sequences (and thereby as a mean to invoke certain 'conversational norms', see Grundler 2011: 166-176) or in the context of 'extracurricular' sequences (and thereby e.g. undermining the task or using them as an opportunity to 'social positioning', see Bucholtz/Hall 2005). Instead of discussing the camera's presence as a problem influencing the 'authenticity' of data, I will show how the references in question open up the opportunity to get access to different aspects of children's peer culture, especially often implicit conversational norms.

# Multimodal cueing practices for the recovery of language in aphasia speech therapy

**SARA MERLINO**

University of Basel, Switzerland

Aphasia speech and language therapy is an activity whose aim is the recovery of the client's linguistic and communicative abilities. This setting has been mainly investigated in pragmatic and discursive approaches (Silvast 1991; Cortazzi & Jin 2004) and more recently in a conversation analytic perspective (Simmons-Mackie & Damico 1999; Horton 2006; Wilkinson 2013; Merlino 2017). These studies highlight the institutional features of a communicative event that intersects health/rehabilitation with pedagogical/learning issues and that shows features observed in other instructional settings. The aphasic client is indeed involved in the accomplishment of tasks (such as naming activities) that are generally evaluated and assisted by the therapist. Our paper focuses on this cueing/scaffolding activity and on the practices used by the therapist to assist the client in the word retrieval. It particularly analyses visual phonemic cueing (that is listening and looking at the interlocutor's lips' position in order to find or repeat an item), which is a practice that makes relevant participants' mutual gaze. It describes the methods and resources through which the therapist secures the client's gaze and contextualizes the phonemic cueing: verbal instructions (such as "look at me"), prosodic changes (such as higher volume), pointing gestures and touching of the client's arm/hand. We thus show the centrality of multimodal - audible, visible and haptic - resources for the accomplishment of the therapeutic activity (traditionally investigated in a logocentric perspective - but see Wilkinson et al. 2011) and for the interactive accomplishment of cueing. The analyses are based on the video-recordings and multimodal transcriptions of a large corpus of speech and language therapy sessions in French (approx. 60 hours) that have been filmed in different therapeutic settings (hospitals, clinics and private surgeries, home settings) within the framework of a longitudinal study.

# **‘Okay blood pressure now’: Doing leadership and managing directives in obstetric emergencies**

**POLINA MESINIOTI<sup>1</sup>, JO ANGOURI<sup>1</sup>, DIMITRIOS SIASSAKOS<sup>2</sup>, STEPHEN O'BRIEN<sup>2</sup> & TIM DRAYCOTT<sup>2</sup>**

<sup>1</sup>University of Warwick, UK | <sup>2</sup>North Bristol NHS Trust, UK

The paper reports on the preliminary findings of a project on teamwork in obstetric emergencies (TREAT). It focuses on the ways senior professionals enact leadership in their context looking specifically into directives in ad hoc maternity teams from an interactional sociolinguistic perspective. Although interaction analysis has privileged the analysis of talk, recent work has placed emphasis on the importance of other semiotic modalities and the need for a more holistic analysis of social action. Mondada (2011: 290) for instance argues that ‘interaction and social action have multiple configuring effects on space, both adjusting to it and reflexively constituting its form and relevant features.’ This is particularly relevant to medical teamwork where space, room arrangement and artefacts play a role in work processes. There is, however, a dearth of studies on medical teamwork particularly in secondary care and specifically on the way directives are constituted in and through situated verbal, bodily, and spatial practices. Our paper aims to address this gap. We take an ethnographic approach and draw on four episodes of video recorded simulated obstetric emergencies. We discuss the ways in which senior doctors enact their professional role and claim control of the situation. Directives are at the heart of this process. We also pay special attention to methodological choices for representing movement in material space. Our findings show that directives are multimodally achieved; senior doctors draw on talk, gestures, body, head and gaze orientation which build on each other and converge in indexing leadership (or not). We close the paper with a discussion of our next steps and emergent theoretical framework.

# Linguistic and paralinguistic aspects of interaction with non-robotic Voice-User Interfaces

**THOMAS C. MESSERLI & STEVE DUMAN**

University of Basel, Switzerland

Voice-User Interfaces (VUIs) such as the Amazon Echo occupy an interesting embodied space. On the one hand, they process human speech and respond with human-like speech. On the other hand, they are cylindrical speakers—a significant difference from animate or robot-like entities. While several research projects have examined how interacting with robots influences human behavior and judgment (e.g., Kahn et al., 2012; Grigore et al., 2016), far fewer projects have examined how interacting with non-robot VUIs like the Amazon Echo influences conceptualization (but see Purington et al., 2017), and none that the authors are aware of analyze visual data of actual interaction with such devices. This paper discusses the video analysis of participants interacting with VUIs in a laboratory setting. Specifically, we study interactions with a non-robot VUI and the linguistic as well as in paralinguistic extensions therein. The setup compares task-oriented interaction with (1) human agents and (2) VUIs between subjects that have been primed to either focus on (a) the smooth interaction with an anthropomorphic agent or (b) on the efficient achievement of their task. Recording these interactions on video allows us to not only analyze the linguistic behavior of the participants, but also their gestures which we interpret as indicators of recipient design: Since participants know that the VUI they interact with is voice-based and can thus not interpret gestures, gesturing at the VUI is thought to indicate that participants model their utterances on human-human interaction and design the VUI as a human-like addressee. On a methodological level, this paper also discusses how the access video analysis provides to embodied participant actions is affected by researcher choices regarding recording equipment and camera placement (see also Mondada, 2006). Such choices are themselves part of the analysis and will affect the findings produced by the experiment.

## Der Einsatz multimodaler Ressourcen in der Orchesterprobe

**MONIKA MESSNER**

Universität Innsbruck, Österreich

In Orchesterproben äußert der/die Dirigierende seine/ihre Gedanken und Vorstellungen zum Stück und zur Performanz der Musiker/innen in Form von Anweisungen, die prospektiv auf noch nicht gespielte Stellen verweisen oder durch Korrekturen, die retrospektiv bereits Gespieltes verbessern sollen. Diese Anweisungen und Korrekturen werden nicht nur auf verbaler Ebene vermittelt, sondern es kommen unterschiedliche multimodale Ressourcen, wie Gestik, Mimik, Blicke, Körperbewegungen, Vor- und Nachsingen, ikonisches Vor- und Nachahmen, u.Ä. zum Einsatz. Diese Ressourcen treten häufig in Kombination auf – auch in Verbindung mit verbalen Ansagen – und können sowohl illustrativen, als auch kontrastiven oder evaluativen Charakter haben. Der Einsatz und das Zusammenspiel dieser multimodalen, bedeutungstragenden Elemente (vgl. „multimodal gestalt“ Haviland 2007; vgl. „multimodal unit“ Keevallik 2010) tragen wesentlich zur Herstellung intersubjektiven Verstehens und damit gleichzeitig zur gemeinsamen Umsetzung in der Probe bei. In dem vorliegenden Beitrag soll anhand von Videodaten und Transkripten untersucht werden, ob und wie sich die Anweisungen des/der Dirigierenden hinsichtlich ihrer multimodalen Beschaffenheit unterscheiden, d.h. es soll analysiert werden, an welchen Stellen während der Probe mehr und an welchen Stellen weniger multimodale Ressourcen eingesetzt werden. Dabei soll auch differenziert werden zwischen Ressourcen, die auf eine bestimmte Spielweise abzielen, d.h. illustrativ sind, solchen, die einen Kontrast aufzeigen und solchen, die für evaluative Zwecke genutzt werden. Außerdem soll beobachtet werden, wie die Musiker/innen auf die multimodalen Anweisungen der Dirigierenden reagieren und sich ihr Spiel verändert oder auch gleich bleibt.

## On filming and transcribing tasting: Public visibility and intersubjectivity

**LORENZA MONDADA**

University of Basel, Switzerland

Most often, taste is considered as a private cognitive-neurological-physiological and possibly culturally influenced individual matter. Within this perspective, video methodology does not have much to contribute to the study of taste. In this paper, however, tasting, rather than taste, is approached as a socially shared, intersubjective and interactionally organized experience (see Fele 2015, Liberman 2013, Mondada, 2016). In this perspective, video can bring interesting insights about the interactional organization of sensoriality.

The data on which this paper is based are video recordings of specialized tasting sessions (e.g. beer tasting session) as well as occasioned tasting practices within other activities (e.g. cheese tasting offered in a gourmet shop). In both cases, tasting happens within groups, either in a symmetrical way (all participants are involved in tasting) or in an asymmetrical way (only one participant tastes).

The question the paper raises is how to document tasting experiences through filming. It addresses limitations and pitfalls of video methodology in this respect – tasting cannot be reduced to just visual conduct – but also the advantages of video – documenting the publicly visible (and audible) cues that co-participants themselves use in displaying self and access others tasting experiences.

Video allows us to document how the co-participants (and the camera(wo)man) orient to, monitor, adjust to tasting as an intersubjective practice. Furthermore, this documentation is also achieved through multimodal transcription, which elaborates on the temporal details of the practice. The paper discusses both videorecordings and multimodal transcriptions as ways of accounting for the interactional organization of tasting as a collective activity. It shows that intersubjectivity is not only provided by the final assessments proffered by the participants, but also, and firstly, by how they orient to, monitor, adjust and respond to each other while tasting, in an embodied and often silent way.

# Capturing emotion with video: Relocating aggressive behaviors in their ecological environments

CHLOÉ MONDÉMÉ<sup>1</sup> & LAURENT CAMUS<sup>2</sup>

<sup>1</sup>CNRS Lyon, France | <sup>2</sup>University of Basel, Switzerland

In most perspectives in linguistics and social psychology, emotions are treated as behavioral demonstrations of (mental) states, and their analysis generally leads to two kind of reductions: either emotions are reduced to retrospective verbal accounts produced by the participants (i), either they are reduced to displays of bodily and/or facial expressions (ii), thus allowing for arguments in favor of the universality of emotional readability among species (Lorenz, 1950 ; Maréchal et alii, 2017). In this communication, we would like to challenge these perspectives in showing what an EMCA-informed use of video can bring to the analysis of emotional displays, by focusing on the case of aggressive behaviors. Such a framework provides adequate tools to relocate them in the unfolding of sequential relevancies (Schegloff, 2007) and normative expectancies (Garfinkel, 1967). We argue that a kinematic and multimodal analysis is able to treat aggressive displays as socially ordered interactions (although agonistic) that are a) embodied and accomplished through various multimodal resources (Mondada, 2016) and b) embedded in a dynamic and complex ecological environment (Goodwin, 2013). The contribution thus complies with questions and issues of topic 1, “multimodal analysis”. The empirical analyses will try to grasp the dynamic expression of aggressivity in two very different settings: intimidation and threatening behavior in football players and in Barbary macaques. Far from defending a reductionist approach that treats the two settings as comparable, the idea is to argue that emotions and mental states cannot be reduced to verbal accounts nor facial expressions. They are public, sometimes un-verbalized, embodied phenomena that are interpreted, and adjusted to, by others (Ruusuvuori, 2013). In that respect, video-analysis provides a fundamental tool, and the contribution’s aim is eventually to examine how, and at which price, linguistics is able take hold of that challenge.

## Adjusting mobile camera operation and participation to interactional and material contingencies in doing CA research

DAVID MONTEIRO

University of Basel, Switzerland

In the study of language and social interaction, theoretical, methodological and technological advancements have promoted the investigation of its situated and multimodal organization on the basis of video recordings of naturally-occurring activities in both ordinary and institutional settings (cf. Neville, 2015). To researchers collecting scientifically-exploitable data in complex material environments, the situated production of audiovisual recordings inescapably poses a variety of practical problems to be solved, such as the positioning of cameras, which may be fixed so to remain static throughout the interactional encounter, or operated in mobility so to reflexively adjust to the emerging transformation of the interactional space (cf. Mondada, 2012; Mondada, 2014). Adding to the complexity of coordinating static and mobile video cameras, the material ecology of the physical space in which recorded interaction takes place, as well as participants' orientation to the recording activity taking place may occasion additional constraints to the management of the recording activity (namely to the positioning of bodies and recording equipment). Based on a corpus of 22 hours of video recordings, subsequently transcribed and annotated so to preserve the sequential, temporal and multimodal organization of social action (cf. Mondada, 2018), this talk will investigate the situated and embodied production of video recordings in different settings: social work encounters taking place indoors in offices with a small number of participants concentrated in a small space, and treasure-hunting expeditions taking place in the wilderness with a large number of participants. With a focus on the interaction between researcher and other participants, as well as on the mobilization of embodied and material resources, several practical problems pertaining to the task of operating a camera will be investigated, such as: a) activating and deactivating the camera within activity junctures such as openings and closings; b) transitioning between the tasks of operating a mobile camera and participating in the interaction; c) mobility and the situated achievement of stability of the camera view in exiguous or perilous spaces. The present study aims at contributing to the promotion of an ecological approach to the study of social interaction (cf. Mondada, 2016), as well as to a growing body of studies on the production of video practices at work (cf. Broth, Mondada & Neville, 2014) and, more specifically, for doing research ethnomethodological conversation analysis (cf. Mondada, 2014).

## Inspecting objects in social interaction

**KRISTIAN MORTENSEN & JOHANNES WAGNER**

University of Southern Denmark, Denmark

In this presentation we look at how unknown or unfamiliar material and gustatory objects are publicly experienced. Using EMCA, we will focus on the sequential organization of such inspection sequences and show how they depend on the physical features of the artifact. Our study draws on data from everyday interactions, design workshops, and tasting activities. To initiate inspection sequences, participants take or receive an object. Participants display their knowledge of these objects already in the way objects are grasped, taken, received, held and used (Tolmie & Rouncefield, 2011). Alternatively, the way an object is grasped may indicate unfamiliarity with it. During the inspection proper, the object is put under scrutiny for sensorial inspection (Streeck, 1996). Earlier, we have shown how visual inspection sequences are treated as time-outs from the ongoing interaction and are typically not 'interrupted' by co-participants. The way in which this is performed depends inter alia on the objects and which of its features the inspecting participant is interested in. For example in inspecting unfamiliar material objects, participants closely coordinate visual and tactile/haptic procedures, viewing the object from different angles and experiencing its shape and substance by tactile and visual means. The way this is performed is different from for example tasting beverages, but participants typically engage in similar recognizable inspection procedures during the tasting itself (e.g. Mondada 2017a, b). The tasting is made publically available as an individual experience by drawing on a range of bodily and material resources inter alia 'chewing' the liquid, maintaining the vessel in an elevated position, and middle distance gaze (Heath, 1984). In this way, the taster redraws from the interaction and co-participants orient to this by avoiding gaze and talk during the tasting activity. As the inspection comes to a close, the participant re-engages in activities with his/her co-participants. Inspection sequences can thus be seen as a social/public practice through which the inspector momentarily disengages from the ongoing interaction while engaging in an individual sensorial experience.

# When spectators talk back: Practices of audience involvement in broadcasted video games

ISABELL NEISE & AXEL SCHMIDT

Institut für Deutsche Sprache Mannheim, Germany

Playing video games involves different participation frameworks and different personal involvements (Mondada 2012). Correspondingly, interactional moves are understandable as cues, displaying different kinds of participation (Keating/Sunakawa 2010), which have to be timely coordinated with respect to the interaction and game time (Mondada 2013). Accordingly, multimodal analysis of video gaming can tackle different levels of that activity (Reeves/Greiffenhagen/Laurier 2016). When Video games are played and presented simultaneously this causes a shift in the participation framework. This happens in so called Let's Plays where video gamers talk to their targeted audience in a moderating manner. Within Let's Plays on YouTube the interaction with an audience is delayed and restricted to written comments. In contrast, live streaming video websites (like Twitch) allow live feedback in form of an integrated chat. The latter is regulated by interaction between presenters and producers on a backstage level. Drawing on a multimodal interaction-analytical framework (Deppermann 2013) and based on a corpus of eight hours of video recordings, both from the production process and broadcast itself, we are interested in producers' practices to increase audience involvement. To this end, we consider a collection of instances where producers involve audience members during the live show. On a first level, we ask how audience participation is elicited, how it is timely coordinated during the live show, and which practices audience members use to get noticed and appreciated. On a second level, dedicated to the production, we ask how producers intervene in live shows to regulate the broadcasted communication. Our multimodal analysis contributes to the growing body of work on interaction analysis in mediated settings, focusing on the interplay between game play and the interaction between players and presenters, audience members and among producers.

## Pointing at displays: Using smartphones as social objects

FLORENCE OLOFF

University of Oulu, Finland & University of Basel, Switzerland

This paper explores how participants multimodally refer to smartphone displays when sharing visual content (especially pictures) in face-to-face interaction, more specifically, how these practices combine gesture with linguistic structure and sequential organization (conference topic 1: multimodal analysis). It has been pointed out that “mobile-supported sharing activities” (Raclaw et al. 2016) provide co-participants with expanded access to assessable referents and thus enhance co-participation (Aaltonen et al. 2014). Several studies have shown how participants manage multiple courses of action while using mobile phones (Brown et al. 2015, DiDomenico & Boase 2013, Haddington & Rauniomaa 2011, Licoppe & Figeac 2015), or how smartphones are exploited for joint orientation in space (Brown et al. 2013, Laurier et al. 2015). Despite a growing interest in object use in social interaction (e.g. Goodwin 2000, Mondada 2016, Nevile et al. 2014), only few studies have focussed on situated smartphone manipulation and gesture (e.g. Brown et al. 2013).

Based on video recordings of ordinary conversations in Czech and German and a multimodal, sequential analysis (conversation analysis), this paper investigates pointing gestures to smartphone displays while participants are jointly looking at pictures. In both languages, participants distinguish between the identification of persons or objects visible on screen (pointing with the index finger) and the reference to events that are linked to what the display is showing (pointing with the little finger, cf. Stukenbrock 2015). These two ways of pointing are systematically linked to different linguistic and action formats: index finger pointings are used with short turns containing deictic expressions and mostly carry out identifications or descriptions, while little finger pointings are accompanied by more complex and diverse turn formats, often projecting longer sequences, especially tellings. Video data can thus show how participants use smartphones as an interactional resource in systematic ways, mobilising them as social objects both through talk and embodied practices.

## Zooming into what matters: The double bind affordance of fetal ultrasound technology

ANA CRISTINA OSTERMANN & MINÉIA FREZZA

Universidade do Vale do Rio dos Sinos, Brazil

Even though images produced by ultrasound scans are normally readily understood by ultrasonographers, they are not always as obvious to lay people. On what concerns fetal ultrasonography, Mitchell (2001: 120) argues that parents are highly dependent on ultrasonographers to “see their baby amidst the swirling grey mass of echoes”. In order to understand this dependence, this study, which analyses 138 fetal ultrasound interactions video-recorded in a Brazilian hospital, investigates how intersubjectivity is pursued and affected by the presence of ‘live’ images produced with ultrasound technology. The analysis reveals that one of the factors that affects mutual understanding in these interactions relates to the ephemeral nature of the referents (Mondada 2012, 2014) projected on the screen – referents which not only move, but also disappear throughout the course of the scans. Such complex interactional ecology may result in lack of simultaneity among: (1) an information request or a description, (2) the image projected on the screen, and (3) the visual access of the participants to the projected image (i.e. when the next speaker visually orients to the image, the referent might no longer be displayed on the screen). In order to achieve intersubjectivity in such a complex ecology, some of the technological resources available to ultrasonographers originally devised as tools for fetal diagnostic investigation – in particular, zooming in and freezing – end up also serving interactional purposes, such as to build joint attention (De Stefani 2014) with the pregnant woman towards certain fetal images. By means of a multimodal analysis of the data, it is possible to observe that the professionals make the pregnant women “see their baby” by coordinating the use of the resources concerning ultrasound diagnostic technology, which are locally constrained and afforded by the setting, with pointing gestures and verbal resources (e.g. deictic expressions).

# Kohäsion und Kohärenz durch Verflechtung von Sprache und Gestik: Eine empirische Untersuchung mündlicher Erzählungen deutschsprachiger Schülerinnen und Schüler

**GÖKHAN ÖZKAYIN**

Universität Koblenz-Landau, Deutschland

Der Vortrag setzt sich vornehmlich mit folgender Hauptfragestellung auseinander: Welche Muster und syntaktische Strukturen liegen im Zusammenspiel von sprachlichen und außersprachlichen Elementen zur Herstellung von Kohäsion bzw. Kohärenz in mündlichen Erzählungen zugrunde? Die Frage, die an der Schnittstelle zwischen Syntax und Pragmatik anzusiedeln ist, wird in drei Schritten beantwortet. Diese gliedern gleichzeitig den Hauptteil des Vortrags. Das Hauptaugenmerk liegt in der Erforschung von kohäsionsstiftenden Mitteln, wie Rekurrenz, Substitution, Pro-Formen, Definitheit, Ellipsen und metakommunikative Textverknüpfungen (vgl. u. a. Linke et al. 2004; Brinker et al. 2000). Im Zuge der Beantwortung dieser Frage werden im ersten Schritt (i) sowohl die internen syntaktischen Strukturen von Diskursreferenten, die als Nominalphrasen auftreten, als auch die syntaktischen Strukturen koreferenter bzw. referenzidentischer Ausdrücke nach signifikant häufig auftretenden Mustern untersucht. Im nächsten Schritt (ii) werden die eruierten Daten, die eine Koreferentialität aufzeigen, unabhängig von ihrer Gebundenheit an das sprachliche Material analysiert, nämlich hinsichtlich der Konzepte der Textkohärenz, wie Isotopie (Greimas 1971), Frame-Theorie (Fillmore 1982) und Präsuppositionen (Levinson 1983). Abschließend (iii) wird gezeigt, dass es nicht nur für die Textkohäsion und Textkohärenz spezifische sprachliche Mittel gibt, sondern auch, dass diese von spezifischen Gesten begleitet werden. Zur Klassifizierung der Gesten werden klassische Werke zur Gestentypologie (McNeill 2000; Kendon 2004), Forschungen zur Lokaldeixis (Fricke 2007) sowie Forschungen zur face-to-face-Interaktion (Stukenbrock 2015) herangezogen. Die Grundlage der Untersuchung bildet ein eigens für das Forschungsprojekt entwickeltes Korpus aus Videodaten. Es besteht aus Erzählungen von Schülern und Schülerinnen der Altersklasse 6-18. Als Erzählanlass werden Fortsetzungsgeschichten, Bildergeschichten sowie kurze Videosequenzen, die nacherzählt werden, verwendet.

## Rendre les émotions du patient dans la traduction orale du récit : le cas des entretiens psychothérapeutiques avec demandeurs d’asile en France

VANESSA PICCOLI<sup>1</sup>, ANNA CLAUDIA TICCA<sup>2</sup> & VÉRONIQUE TRAVERSO<sup>3</sup>

<sup>1</sup>Lab ICAR Lyon, France | <sup>2</sup>Lab LIDILEM Grenoble, France | <sup>3</sup>Institut du Proche-Orient Beyrouth, Liban

Le nombre croissant de demandeurs d’asile qui arrivent en Europe confronte les institutions administratives et sanitaires au défi d’accueillir un nouveau public plurilingue, ayant des parcours et des besoins variés. Pour assurer une communication efficace dans ces contextes, la figure de l’interprète médical et social (ou interprète communautaire) devient ainsi centrale. De nombreuses études ont montré que l’interprète est bien plus qu’un « passeur linguistique » : il joue un rôle important dans la coordination de l’interaction (Wadensjö 1998, Baraldi & Gavioli 2012) et dans la co-construction de l’intercompréhension (Ticca & Traverso 2017 parmi d’autres). Dans des entretiens psychothérapeutiques avec des demandeurs d’asile, le rôle de l’interprète est spécialement délicat, d’une part à cause du lourd poids émotionnel qui peut caractériser ces interactions – notamment, durant le récit d’expériences traumatiques –, d’autre part en raison des spécificités de cette situation de soin où la communication et l’affectivité constituent une partie centrale du processus de cure (Peräkylä 2008). En adoptant la méthodologie de l’analyse conversationnelle (Sacks 1992), cette contribution s’appuie sur un important corpus audiovisuel d’entretiens entre professionnels de santé mentale (psychologues et psychiatres), demandeurs d’asile et interprètes, en France. En particulier, notre contribution propose l’analyse de séquences narratives et se concentre sur les pratiques utilisées par les interprètes pour rendre dans la traduction les émotions manifestées par le patient durant son récit d’expériences autobiographiques, de façon verbale et/ou multimodale (pleurs, soupirs, changements prosodiques, gestes, réorientations du corps, etc.). A travers la description détaillée des pratiques multimodales utilisées par les interprètes, cette étude a pour but de contribuer à mieux comprendre le fonctionnement de la communication dans ces contextes, et particulièrement la façon dont les émotions circulent en présence des interprètes au cours des narrations liées aux situations de migration.

## Address, reference and reciprocity: Participation in training of mass-casualty incidents

**KAROLA PITSCH**

University of Duisburg-Essen, Germany

Description of typical linguistic phenomena, such as reference to persons (Enfield 2013) and deixis (Mondada 2012, Stukenbrock 2015, Pitsch & Wrede 2014) gain an important level of complexity once studied using video-recordings of social interaction. Not only need verbal utterances be coordinated with bodily conduct, but they are also embedded within situations of ‘multi-activity’ (Haddington et al. 2014) and complex participation frameworks (Goodwin & Goodwin 2004, Schmitt 2012). We are interested in the phenomenon of addressing utterances to co-participants within complex participation frameworks, such as the training of a mass-casualty incident. Reacting to an alleged emergency call (15 victims), different emergency services (fire brigade, paramedics, medical doctors) are called to the location and provide aid. The setting includes 100+ participants, among which in particular two dyadic teams are responsible for coordinating the services: Team A, responsible for the overall coordination, composed by a member of the fire brigade and an emergency doctor; Team B responsible for triaging the victims and sending them into a coordinated form of medical support. We have documented this outdoor training with (i) an aerial overview, focusing on (ii) team A and (iii) team B, and (iv) documenting the patients’ treatment in the hospital. In total, time-synchronized recordings of 8 video cameras, 4 eye cams/mobile eye tracking glasses, 1 drone camera and two radio communication circuits have been obtained for the 2.5 hours’ event. Our analytic focus is on the activities involving team B triaging the victims and categorizing them into different types of patients: The victim/patient and medical doctor are surrounded by different types of bystanders who monitor the event with their respective ‘professional visions’ (Goodwin 1994). How is information addressed to one specific participant reacted to also by other participants? How do addressing, referencing and establishing reciprocity function in complex forms of participation?

# Kreativität in Interaktion: Eine multimodale Interaktionsanalyse der «Creative Economy»

ROBERT REINECKE

Universität Basel, Schweiz

Ogleich in der wissenschaftlichen und gesellschaftspolitischen Debatte ein Konsens darüber besteht, dass Kreativität im Zeitalter der Automatisierung „[...] die wissensbasierte Gesellschaft von morgen entscheidend gestalten [...]“ (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) 2000, S. 3) wird, fehlen bislang empirische Ansätze, welche die tagtägliche Herstellung, Verhandlung und Förderung von Kreativität in ebendieser Gesellschaft untersuchen (Carter 2004, S. 212). Daher wird die folgende Leitfrage ins Zentrum des Vortrages gerückt: Wie wird Kreativität im Arbeitsalltag von Werbeagenturen hergestellt und verhandelt? Um dieser Frage nachzugehen, wurden im Rahmen des von der HPSL geförderten PhD-Projekts rund 293 Stunden Videodaten von Interaktionen in Kreativräumen, also eigens zur Kreativitätsförderung geschaffenen Raumstrukturen, führender Schweizer Werbeagenturen erhoben und die darin stattfindenden Interaktionen mittels der multimodalen Interaktionsanalyse (Hausendorf et al. 2012) unter besonderer Berücksichtigung der räumlichen Komponente ausgewertet (Kesselheim und Hausendorf 2007). Auf Basis der aufbereiteten Daten soll im Vortrag anhand zweier exemplarischer Videoausschnitte aufgezeigt werden, dass Kreativität im Arbeitsalltag von Werbeagenturen interaktiv verhandelt wird und es sich bei der Verhandlung von Kreativität um ein multimodales Interaktionsgeflecht handelt, welches erst durch eine Analyse sichtbar gemacht werden kann, die „alle Ausdrucksressourcen im Prinzip als gleichwertig [...] berücksichtig[t...]“ (Hausendorf et al. 2012, S. 9). Im Vortrag wird ein Schlaglicht darauf geworfen, wie die Interaktionsteilnehmenden der untersuchten „creative economy“ (Howkins 2002) durch differenzierten Einsatz verbal-strukturierender (Sacks 1998a, 1998b, 1998c, Fina/Georgakopoulou 2012), gestischer (Stukenbrock 2015) und körperlicher (Goodwin 1984, Stukenbrock 2012), räumlicher (Hausendorf et al. 2012b) bis zu „den Beteiligten zum Zeitpunkt ihrer Interaktion perzeptorisch nicht zugänglich[er]“ (Stukenbrock 2015: 440) Ressourcen einen interaktiven Konsens über die Fähigkeit, Überraschendes, Neues und Einzigartiges hervorzubringen, etablieren. Der Vortrag versteht sich als Beitrag zur Achse «Multimodale Analyse», weshalb im Zuge der Analyseresultate zudem aufgezeigt wird, dass die multimodale Interaktionsanalyse einen methodischen Zugang bereitstellt, der mit seiner Fokussierung auf authentische Videodaten gerade hinsichtlich schwer fassbarer Phänomene wie «Kreativität» faktenbasiertes Wissen generieren kann.

# Translating feelings into actions: Communication and collaboration between motorcycle race drivers and their crew chief

FRANCESCA ASTRID SALVADORI

University of Milano, Italy

This paper explores patterns of interaction and communication between motorcycle race drivers and the crew chief. This workplace study aims to unpack issues around teamwork, and tackle those seemingly individual decisions that are actually a product of highly coordinated negotiations between team members. In particular this paper will focus on those moments of interaction during tests when the driver has limited amount of time to communicate and describe to the crew chief his feelings and experience while driving the motorcycle, in order to improve its performance. The crew chief then translates the feelings communicated by the driver, into specific engineering adjustments and communicates these decisions to the rest of the crew. This paper focuses on how these brief moments of high intensity work procedures are managed, designed and responded to; looking at how verbal and non-verbal behaviour plays into decision making. These issues are explored using a data corpus of approximately 20 hours of audio-visual recordings of naturally-occurring interactions in two separate MotoGP teams in the same pit box during tests in a race circuit. Members of the teams are not native speakers of the same language, they speak English, Italian and French with each other. Moreover, by virtue of the nature of the type of activity, the type of machines used in the pit and the presence of other teams practicing on the same circuit, the work setting explored is exceptionally loud. The crew chief uses headphones to communicate with the crew members often standing in front of him, while the driver does not use headphones. So the type of activity analysed is highly sensitive to the difficulty in hearing in this setting, and combined with issues around language use, often times team members face communication issues. One of the teams explored has just started working together and is in the process of creating a way to communicate with one another that is both efficient and clear, bringing to light interesting methodological issues around acoustics, language use and embodied actions. These recordings of video data will facilitate the consideration of a range of interactional resources used by team members to communicate, including talk, gestures, bodily orientation and also the use of artefacts and technologies such as iPads, computers and paper documents. The study's methodological approach is informed by ethnomethodology (EM) and conversation analysis (CA), using video and audio recordings reviewed and transcribed using standard CA orthographies to aid multimodal analysis. The study's approach also draws on ethnographic field observations, discussions with participants and informal interviews. Lastly, the study contributes to contemporary debates in CA, HCI, workplace studies and organisation studies; while the methodological implications of this project include the improvement of the approach to multimodal analysis, looking at new ways to collect, manage and analyse data.

## Verbale und nonverbale Abtönung im Vergleich

STEVEN SCHOONJANS

Universität Innsbruck, Österreich & KU Leuven, Belgien

Mit Waltereit (2006:62) kann ‚Abtönung‘ definiert werden als die Modifizierung (aber nicht unbedingt Abschwächung) der illokutiven Funktion einer Äußerung in Bezug auf die zu erwartende Hörerreaktion. Ein typisches Abtönungsmittel des Deutschen sind die sogenannten Abtönungspartikel oder Modalpartikeln (u. a. Thurmair 1989, Müller 2014). In Beispiel (1) (das meinem Datensatz entstammt) etwa wird durch ja signalisiert, dass der Sachverhalt als bekannt oder zumindest unstrittig betrachtet wird und dass also kein Widerspruch erwartet wird (vgl. dazu auch u. a. Rinas 2007, Reineke 2016). (1) Es ist ja nicht mehr so wie früher. Allerdings sind diese Partikeln nicht die einzigen Abtönungsmittel im Deutschen. Neben weiteren verbalen Elementen wie zum Beispiel IAW-Strukturen (Stefanowitsch 2011) können auch Gesten zu Abtönungszwecken eingesetzt werden (Schoonjans 2014). Es handelt sich sowohl um manuelle (etwa das sogenannte intersubjektive Zeigen) als auch um nonmanuelle Gesten (etwa Kopfschütteln und Nicken). Für die Deutsche Gebärdensprache (DGS) hat Herrmann (2013) ähnliche Muster mit vergleichbaren Funktionen aufgedeckt. In diesem Beitrag wird eine konstruktionsgrammatisch inspirierte kontrastive multimodale Analyse zur Abtönung im Deutschen und in DGS vorgestellt. Den Ausgangspunkt bildet eine weiterführende Analyse der Datensätze von Schoonjans (2014) und Herrmann (2013). Obwohl diese Datensätze nicht völlig vergleichbar sind (ca. 5 Stunden Fernsehgespräche verschiedenster Art fürs Deutsche, ca. 4 Stunden Nacherzählungen kurzer Geschichten im Gebärdensprachlabor für DGS), zeigt sich, dass zumindest auf der nonmanuellen Ebene nicht nur weitgehend entsprechende Formen im Deutschen und in DGS eingesetzt werden, sondern dass diese auch vergleichbare topologische Tendenzen aufweisen. Außerdem stellt sich bei einem Vergleich mit den verbalen Abtönungsformen (für den neben den Videos auch geschriebene Daten aus COSMAS in Betracht gezogen werden) heraus, dass diese topologische Tendenzen nicht nur auf nonverbale, sondern auch auf verbale Abtönungsformen zutreffen.

# Instructional practices in indoor climbing with blind and partially-sighted athletes

**MONICA SIMONE & RENATA GALATOLO**

University of Bologna, Italy

Instructional practices are key phenomena in sports training where the coach and the athlete cooperate to achieve and improve the sport performance. Despite this, analysis of instructional practices as they occur during coach-athlete interaction remains a relatively undeveloped area of research. Instructional practices routinely shape indoor climbing performed by athletes with visual impairment, both in trainings and in competitions. Indeed, while performing the climb, blind and partially sighted climbers are verbally guided by trainers, who instruct them about the position of the holds to grasp for moving upward on the route. In so doing, trainers put their visual access to the climbing route in the service of the climbers, improving their access to the climbing space and enabling the accomplishment of the sport performance. The paper aims to contribute to the conference topic “Multimodal Analysis”, exploring the ecology and the indexicality of instructional practices and the relation among verbal instructions, body movements and the environment. After providing an overview of the sequential organization of indoor climbing trainings, the paper focuses on the aspects of instructional practices that display the trainer’s orientation to the climber’s haptic access to the climbing space. We particularly demonstrate how and that the trainer produces instructions that convey timely relevant spatial references relying on the unfolding climber’s movements. The study is based on a corpus of 3 video recorded training sessions (about 8 hours). Data were collected in 2016, during a three months ethnographic work in two different gyms in Bologna (Italy). Participants are a sighted trainer and three elite climbers with visual impairment. Ethnomethodology and Conversation Analysis provide the theoretical and methodological framework of the analyses. Data were transcribed according to the conventions developed by Jefferson (Jefferson, 2004) and Mondada (2016).

## Navigating physical & epistemic landscapes: How mobility in the field structures practitioners' understanding of geologic phenomena

**MICK SMITH**

UCLA, USA

Practitioners' acquisition of knowledge is multi-dimensional: it culminates from different physical and social spaces, makes use of different forms of materiality, and originates from within a set of embodied inter-subjectivities. Embodiment and mobility are undeniably important in this regard to field geology, as the work is conducted in heterogeneous, wilderness settings that are only accessible via practitioners' movement through them. They are thus essential to the creation of the material and textual inscriptions that animates discourse in the geosciences as both a domain of knowledge and community of practice. I argue in this paper, that mobility and embodiment play a far greater role in field geology than in just providing access, however. The "moving body" constitutes what Merleau-Ponty described as the "measurement of the world" or the point of articulation where the world is joined, divided, and made sensible (2003, p. 24). In moving through a landscape, the geologist continuously (re)articulates it: What appears as phenomena at any given point in time and space is mediated by what appears at other points in time and space. Perception thus is neither a product of just the perceiver nor of just the landscape, but something that emerges in the interaction between (Carusi and Hoel, 2014, p. 212). The data come from a video ethnography of geological fieldwork conducted by senior and novice practitioners. The data are analyzed using conversation analysis (Sacks, Schegloff, & Jefferson, 1974), prioritizing embodied action and mobility (Goodwin, 2000; Mondada, 2013). I analyze in particular how talk and embodied action are coordinated with the practitioners' movement through various terrains and reveal the way in which the interactions emerge depend not just on how the participants sequence actions with one another but how they sequence different spaces as well.

## Playing and Displaying “Move 37”: The “Story of ‘Oh’”, a Multimodal Achievement In Situ

PHILIPPE SORMANI

University of Lausanne, Switzerland

For decades, playing Go at a professional level has counted among those things that “computers still can’t do” (cf. Dreyfus 1992). As it seems, this changed dramatically in early March 2016, at the Five-Star Four Seasons Hotel in Seoul, South Korea, when AlphaGo, the most sophisticated computer program in Go to date, beat Lee Sedol, an internationally top-ranked Go professional, by four games to one. A recent documentary movie has captured the unfolding drama (cf. Kohs 2017). In turn, this contribution offers a video-based interaction analysis of the second game’s “move 37,” its surprise delivery by AlphaGo, and the subsequent line of commentary by the attending experts, initiated by the news-receipt token “oh” (Heritage 1984). In the context of the analyzed game, it was impossible to enter the computer program and reconstitute its probability calculus. Yet the played out Go move was rapidly qualified as a “surprising,” “creative,” or even “cosmic” move. Why is that? How come? Then and there? In answer to the raised questions, the present contribution offers a fine-grained video analysis of “Move 37” as a locally unfolding, multi-modal achievement. Particular attention is paid to the interplay between attending experts’ game commentary and the split-screen monitoring of their local interaction, in and for the move’s assembled intelligibility. The outlined contribution offers a video-based re-specification of the “story of ‘oh’” (Macbeth et al. 2016) by homing in on how participants “index” and “animate” their encounter ad hoc and in situ.

# Flüchtige urbane Begegnungen: Eine Videoanalyse interaktiver Minimalformen im öffentlichen Raum

**ANTONIA STEGER**

Universität Zürich, Schweiz

Im Diskurs über den öffentlichen Stadtraum als sozial produzierter Raum wird oft der Begriff der ‚Begegnung‘ zwischen sich unbekanntem Personen zentral gesetzt, ohne dass jedoch empirisch präzise gefasst ist, was damit gemeint ist. Die Idealvorstellung von «zentrierten Interaktionen» (Goffman 2009) entspricht nicht der Alltagsrealität; der Grossteil an öffentlichen Begegnungen findet in flüchtigen interaktiven «Minimalformen» (Schmitt und Petrova in Vorbereitung) statt. Phänomene wie etwa das Aneinander Vorbeigehen, Grüssen, Beobachten, Nebeneinandersitzen etc. von Unbekannten, die nicht in zentrierte Interaktion miteinander treten, sind charakteristisch für die Sozialität des öffentlichen Raumes. Solche interaktiven Minimalformen sind bisher nicht systematisch empirisch untersucht und konzeptualisiert worden. Die multimodale Interaktionsanalyse (vgl. Schmitt 2007; Mondada und Schmitt 2010; Hausendorf et al. 2012; Schmitt 2013) bietet jedoch dafür einen vielversprechenden methodologischen Rahmen durch die zentrale Setzung von Videoaufnahmen (vgl. Heath et al. 2010; Tuma et al. 2013), den analytischen Einbezug multimodaler Ressourcen (vgl. Deppermann und Schmitt 2007; Streeck et al. 2013) sowie dem ethnomethodologischen Verständnis der Interaktionen als sequenziell organisierte Prozesse (vgl. Schegloff und Sacks 1973; Schegloff 1984). Ich untersuche im Kontrast exemplarisch je drei flüchtige Begegnungen auf der Strasse sowie auf einem städtischen Platz. Der Fokus der Analyse liegt auf der sequenziell-konfigurativen Organisation des Blickverhaltens, d.h. insbesondere der Steuerung der Sichtbarkeit des eigenen Sehens, sowie der Rolle von Selbstberührungen. Damit kann die spezifische Sozialität im öffentlichen Stadtraum datengeleitet diskutiert und konzeptualisiert werden. Die Untersuchung stützt sich auf ein Korpus von Videoaufnahmen auf zwei zentralen städtischen Plätzen (total rund 50 Stunden Videomaterial) sowie auf Gehsteigen (total rund 4 Stunden Videomaterial) in Zürich. In der Erhebung kamen sowohl statische wie auch mobile Aufnahmeverfahren zur Anwendung.

# A multimodal analysis of the use of relative clauses in French talk-in-interaction

IOANA-MARIA STOENICA

University of Neuchatel, Switzerland

This paper investigates the use of relative clauses (RCs) in French talk-in-interaction. RCs are traditionally defined as subordinate clauses that are introduced by a relative pronoun and that maintain a referential relationship with a noun phrase – called head of the RC – stated in the main clause (MC) (Riegel et al. 2016). In this paper, we propose a reanalysis of these constructions as part of “grammar for talk implementing action” (Schegloff 1996: 113). The data for this study comprises 10 hours of French video-recorded interactions between students during their coffee breaks, transcribed according to conversation analytic transcription conventions. Based on detailed sequential analyses, this paper shows that RCs may be used: 1) to repair a problem of referential identification, or 2) to elaborate on a given referential expression, and that these actions occur in response to interactional contingencies, such as: recipient's verbal or non-verbal display of trouble with or recognition of the referent identification. The following excerpt is an illustration of the first phenomenon:

- 01 DAV: j'ai regardé le film safe.  
02 (0.6)  
03 DAV: que Romain il nous a parlé \*une fois.=  
04 GEB: =euh avec [euh  
05 DAV: [avec eu[:h  
06 GEB: [Statham?  
07 DAV: Statham ouais=  
08 GEB: =il est bien hein?

The findings show that RCs either occur as a solution to a problem displayed by a recipient or are occasioned by recipient multimodal behavior, emphasizing thus recipient-designing and joint construction of talk. We thus argue for an integration of participants' embodied resources into the analysis of the actual use of RCs. This paper contributes to the multimodal analysis conference topic and to the growing body of research on the emergent character of grammar (Deppermann & Günthner 2015, Pekarek et al. 2015) and on its understanding as closely related to embodied semiotic resources (Keevallik 2013, 2015).

## Visible problems of hearing: correcting how to speak in public

HANNA MAGDALENA SVENSSON

University of Basel, Switzerland

Within the framework of ethnomethodology and conversation analysis, this paper discusses visual aspects of talk as consequential for how participants to large group political meetings manage emerging problems of audibility. As claimed troubles of hearing emerge, the participants attempt to solve this by corrective instructions of the current speaker's manipulation of the microphone, orienting to issues of how to speak in public.

This study is based on video recordings of political meetings in France. The large number of participants, mounting to up to 400 persons, implies a practical problem of audibility for producing public talk. That this is an issue is reflected in and reflexively established by the organizers' provision of microphones. Over and above the vocal production of linguistic resources, talking in a microphone necessitate additional embodied practice, contingent on the artifact's technical features. As acoustic issues emerge, the participants solicit solutions to this by denouncing it as a problem and eventually by engaging in corrective instructions (Weeks, 1985; Keevalik, 2010; Lindwall & Ekström, 2012) of how to use the microphone. The emerging instructions of the speaker's embodied conduct (Mondada, 2014a; 2014b) show that the participants orient to the publicly visible usage and manipulation of the microphone as an accountable course of action with regard to how to do 'speaking in public'. The corrected party's retrospective accounts for the trouble and eventual rejection of the correction, display their orientation to the instructions as doing correcting and as concerning issues of competence regarding how to speak in public, manifesting the institutionality of the setting as it is talked into being (Drew and Heritage, 1992).

The paper contributes to our understanding of talk as an embodied situated accomplishment in a context where the video recording is decisive for reconstructing the participants' orientation to visual aspects of unfolding talk as consequential for how they achieve hearing.

## On the incongruences in human machine interactions: Mutual intelligibility, responsibility and agency

**BURAK S. TEKIN**

University of Basel, Switzerland

Interactions between people and machines require similar interpretive work that characterizes interactions among people, but with different interactional resources available to parties. However, a profound discrepancy in access to the interactional resources and contingencies generates challenges for mutual intelligibility of each other's actions (Alac, 2009; Suchman, 2007).

This talk draws upon a video corpus of people playing videogames, with Microsoft Xbox One with Kinect. A congruent interaction between the player and Kinect necessitates the player to design game relevant bodily movements to be recognized by Kinect as such, and Kinect to detect and recognise the player movements done for the machine situatedly and contingently. This paper concentrates on the incongruences in Kinect videogaming interactions. More particularly, it focuses on the sequences where either player actions are not recognized appropriately by Kinect or Kinect recognizes other actions of players that are not done for the machine as game-relevant. It looks into the social and sequential organization of the ways through which participants (players, co-players, spectators, etc.) orient to the incongruences between the users and the machines. Following incongruences, players generally display puzzlement for what happened, and produce inquiries questioning themselves and their own movements. In the moments where the incongruences lead to negative game outcomes, players reject the responsibility, and deflect it to the machines through claiming that they have done what is required to be done, and that machines have failed in recognizing their movements. In these occasions, spectators who have different access refer to the normative requirements of the system, which implicitly puts the responsibility to the players and their embodied game conduct.

Based on video data and sequential analysis, this paper elucidates the collaborative practices deployed by participants in technology involved situated activities to understand what happened and to manage the challenge of what to do next, occasioned by the incongruences.

## Le refus suite à une offre pendant des repas entre ami(e)s : une étude séquentielle et multimodale

**BIAGIO URSI**

Lab ICAR Lyon, France

En analyse conversationnelle, suite à une offre ou à une proposition, le refus (versus acceptation) représente l'une des réalisations attendues de la deuxième partie de paire adjacente (Schegloff 2007). Ce type d'action réactive est considéré comme non préféré. Des indices séquentiels spécifiques sont mobilisés par les offreurs afin d'obtenir une réponse préférée (c'est-à-dire une acceptation), et d'éviter ainsi de recevoir un refus (Davidson 1984, 1990). Récemment, l'analyse multimodale a apporté de nouveaux éléments à l'étude des séquences offre – acceptation/refus. Par exemple, en s'appuyant sur des données audiovisuelles d'interactions en anglais et en finnois, Kärkkäinen & Keisanen (2012) montrent une structuration locale récurrente des offres d'objets : l'offre est réalisée en deux temps. D'abord, l'objet de l'offre est identifié comme référent, ensuite, l'offre proprement dite est accomplie, ces deux étapes pouvant être produites par le biais de ressources verbales ou gestuelles. Dans cette contribution, nous allons étudier des séquences d'offres lors d'interactions entre locuteurs francophones, en particulier pendant des repas et des apéritifs entre ami(e)s. Notre intérêt sera focalisé sur la réalisation du refus suite à une offre initiale. Le destinataire de l'offre formule son refus en deuxième position, par le biais de marques linguistiques de négation et ressources multimodales variées (gestes manuels, secouements de tête, mimiques faciales). En l'absence de réponses verbales, ces dernières jouent un rôle primordial pour annoncer ou exprimer un refus local, et elles peuvent déclencher des cycles d'insistance (cf. Traverso & Ursi 2015) : l'offreur réitère son offre plusieurs fois, afin d'obtenir une réponse préférée. En revanche, en présence de réponses verbales, les ressources multimodales peuvent accompagner, et même précéder l'expression verbale du refus, comme dans les cas de secouements de tête. La prise en compte de la multimodalité permet donc d'éclairer l'articulation séquentielle des séquences d'offres et de spécifier la temporalité des réalisations réactives.

# Das Blickverhalten des nicht-blickselektierten Sprechers während nicht-blickselektierten Turnübernahmen: Multimodale Interaktionsanalysen triadischer Gesprächssituationen mit Hilfe von Eye-Tracking

CLARISSA WEIß

Universität Freiburg, Deutschland

Das konversationsanalytische Interesse an der Erforschung des Blickverhaltens und dessen Bedeutung für das Turn-Taking hat in den letzten Jahren stark zugenommen (vgl. z. B. Kendon 1967; Goodwin z. B. 1981; Lerner 2003; Rossano & Stivers 2010; Rossano 2012; Streeck 2014; Holler & Kendrick 2015). Aus jüngeren Forschungen geht immer mehr hervor, dass der Sprecherblick nicht nur im Kontext von Fragen eine Reaktion des Rezipienten elizitieren kann (vgl. Rossano & Stivers 2010, Kendon 1967, Bavelas/Coates/Johnson 2002), sondern am Turn-Ende als next-speaker-selection-Ressource zu behandeln ist (in triadischen Gesprächen analysiert von Auer 2017). Acht Eye-Tracking-Aufnahmen von triadischen Tischgesprächen (durchschnittliche Dauer: 50 Minuten) wurden untersucht. Das Paper behandelt die folgende – potenziell problematische – Turn-Taking-Situation: Ein Sprecher (S1) blickselektiert einen nächsten Sprecher (S2). Entgegen der Hypothese, dass dies für eine Turnübernahme genügen sollte, übernimmt der nicht-blickselektierte Teilnehmer (S3) den nächsten sequenziellen Schritt. Es fanden sich 44 Beispiele des zuvor beschriebenen nicht-blickselektierten Turn-Takings. Unterschiedlichen Blickverhaltensmustern von S3 während der Turnübernahme konnten unterschiedliche interaktionale Funktionen zugeordnet werden: • S3 schaut über den Sprecherwechsel hinaus zu S1: Der Turn-Inhalt (Reparaturen oder persönliche Stellungnahmen) ist ausschließlich an S1 gerichtet, S2 wird erst nach dem informativen Kern des Turns wieder miteinbezogen (angeschaut). • S3 wendet sich zu S2 (Monitoring): S3 möchte sich ‚rückversichern‘, dass die Turnübernahme nicht zur Konkurrenz mit S2 führt. Die Nachricht wird an beide Rezipienten adressiert und weist eine andere inhaltliche Strukturierung auf als in Fall 1. • S3 schaut weg von beiden Ko-TeilnehmerInnen: S3 sieht sich als notwendiger Produzent des zweiten Paarteils aufgrund von vermeintlichen Wissensasymmetrien (vgl. Fall 1). Jeglicher visuelle Zugang zu S3 wird durch das ‚Wegschauen‘ erschwert, um sich durchzusetzen.

# “Remember what exercises we did?”: Responses to instructions during exercise training in physiotherapy

ADRIAN YIP<sup>1,2</sup> & VERONIKA SCHOEB<sup>1</sup>

<sup>1</sup>Hong Kong Polytechnic University, China | <sup>2</sup>Queen Mary University of London, UK

Introduction: Instruction is an integral and crucial component of exercise therapy during which physiotherapists need to ensure that patients are able to perform the prescribed exercises correctly. In general, the series of instructional sequences are anchored in a “request” / “compliance with request” adjacency pair (Lindwall & Ekström, 2012) – physiotherapists instruct an exercise that patients then perform. While the process of acquiring the instructed actions and various instructional/corrective sequences have been discussed in a physiotherapy context (Parry, 2004, 2005), communication of initial requests is apparently assumed to be straightforward and has not been explored. The aim of this study is to delineate physiotherapists’ verbal and embodied initiations of exercises and patients’ responses. Methods: A conversation analytic approach was adopted. Data were collected from one public and one private rehabilitation center in Hong Kong, involving Cantonese, Mandarin and English-speaking patients. A total of 47 consultations with 6 physiotherapists and 16 patients were video-recorded and transcribed (Jefferson, 2004; Mondada, 2014). An inductive multimodal analysis was conducted to examine both verbal utterances and embodied features. Findings: Initial instructional sequences were manifested in statements and questions such as invitation (e.g. “Can you...?”) and memory check (e.g. “Do you remember...?”). They carried distinct presuppositions about patients’ ability to perform and subjected them to different response frames. Moreover, physiotherapists evaluated patients’ responses and adapted their instructions accordingly. They also mobilized nonverbal resources such as gesture to facilitate patients’ comprehension. Discussion: The multimodal analysis of exercise instruction in a physiotherapeutic setting demonstrates that initiations of exercises are situated in task-relevant actions. While imperatives are commonplace in contexts characterized by a sense of urgency (Mondada, 2017), instructional sequences in the praxeological context of exercise therapy are often manifested as interrogatives. The importance of a multimodal perspective in monitoring participants’ understanding (Mondada, 2011) will also be discussed.

# On the role of gaze for eliciting and addressing listener feedback in jointly construed narratives

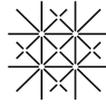
ELISABETH ZIMA<sup>1</sup> & GEERT BRÔNE<sup>2</sup>

<sup>1</sup>University of Freiburg, Germany | <sup>2</sup>University of Leuven, Belgium

In this paper, we address the question how speakers in three-party interactions manage the knowledge state of their interlocutors through different semiotic resources. More specifically, we zoom in on a particular type of interactional activity, viz. jointly construed narratives in which two (co-)narrators report on a joint experience to a third participant. In this particular setting, the current speaker needs to monitor whether (1) the third (uninformed) participant is able to reconstruct the story that is being told, and (2) the co-narrator agrees with the content and presentation of the narrated joint experience. As a result, jointly construed narratives in three-party interactions require speakers to elicit and monitor response from both co-participants (Vranjes et al. 2016), either in the form of minimal (verbal or nonverbal) response tokens (e.g. mhm or head nods) or more elaborate feedback. The focus of this contribution is mainly on the role of eye gaze as a resource used by speakers to elicit feedback from more than one interlocutor. For dyadic conversations, Bavelas et al. (2002) have shown that speakers soliciting a listener response make strategic use of the so-called gaze window, i.e. a short moment of mutual gaze during which speakers through their gaze signal that they expect feedback of some sort (cf. also Stivers & Rossano 2010). If the addressee then gives such feedback, the speaker typically averts his/her gaze again (i.e. the gaze window is terminated) and continues the ongoing turn. Our study on jointly construed narratives in triadic interactions confirms the relation between gaze windows and listener responses, as described by Bavelas and colleagues, but with more complex constellations than is the case for dyadic interactions. For instance, we observe configurations of multiple gaze windows, with speakers soliciting a response by two addressees within a single turn-constructive unit. We further show that speakers that just gave feedback do different things, i.e. shift gaze to the third party or avoid mutual gaze altogether, depending on their participant role either as recipient of the storytelling activity or as co-teller.



# SPONSORS



University  
of Basel

Department of  
Languages and Literatures

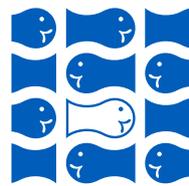


FONDS NATIONAL SUISSE  
SCHWEIZERISCHER NATIONALFONDS  
FONDO NAZIONALE SVIZZERO  
SWISS NATIONAL SCIENCE FOUNDATION



Vereinigung für Angewandte Linguistik in der Schweiz  
Association Suisse de Linguistique Appliquée  
Associazione Svizzera di Linguistica Applicata  
Associaziun Svizra da Linguistica Applitgada  
Swiss Association for Applied Linguistics

Schweizerische Akademie der Geistes- und Sozialwissenschaften  
Académie suisse des sciences humaines et sociales  
Accademia svizzera di scienze umane e sociali  
Accademia svizra da ciencias umanas e socialas  
Swiss Academy of Humanities and Social Sciences



WICKELFISCH  
ORIGINAL

