

Hubert Knoblauch/Bernt Schnettler
Jürgen Raab/Hans-Georg Soeffner
(eds.)



Video Analysis: Methodology and Methods

Qualitative Audiovisual Data Analysis in Sociology



PETER LANG

Europäischer Verlag der Wissenschaften

Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis

Introduction

The use of video is becoming a widespread practice within the social sciences interested in the real-time production of social life – and not only within specialized research areas dealing with the visual, such as visual sociology or anthropology and the studies of gesture and multimodality. Video constitutes a fundamental technique for constituting the corpora of data for analysis, as well as an important mean of rendering research results, such as in documentary films or multimedia presentations.

These uses have generated a number of *methodological* texts and handbooks giving advice on fieldwork, on technical equipment and on further analytic exploitation of video records; however, *analytical* studies focussing on video as a timed accomplishment and as a social practice are still very few. Contrary to numerous analyses of visualization practices in the exercise of science (Lynch 1985, 1991; Latour 1986), showing how the organization of images such as diagrams, photographs or maps actively constitute objects of knowledge, working as an “externalized retina” (Lynch 1988), video has not yet been investigated in this respect.

In this paper, we propose a *praxeological approach to video practices*, focussing on the way in which videos are locally and contingently produced by social scientists, as well as in other professional domains. This perspective allows us to describe some basic practices through which scholars studying social interaction actively constitute their primary data and corpora and through which they establish a first preanalytic online interpretation of the very events they document. This perspective operates a conceptual switch: it doesn't deal with video as a mere *resource*, for example in a methodological discussion, but treats video practices as a *topic* per se, within an analytic stance. It allows one to go beyond discussions of video as a methodological tool, as a source of technical bias or as a way of producing records made transparent for the description of the events they document – inviting us to consider that we see “with” the camera and not “through” it (Büscher 2004), and thus to dissipate the fallacy of an independent and pre-existing world transparently offered from “out there” to our observation.

This project thus takes seriously the invitation to develop, as Macbeth puts it, a “praxeology of seeing with a camera” (1999: 151) considering shooting as an embodied exercise of inquiry and analysis, as the “work of assembling visible social fields” (1999: 152). Camera movements, technical choices, and perspective making are an integral part of the social activities of interest here, embedded in talk-in-interaction and synchronized with it, therefore mutually elaborating each other, and further ar-

ticated with other bodily conduct, gestures, object manipulations, and material environments. These activities are constitutive of the production of the visibility, recognizability, intelligibility of the phenomena at hand and of the arrangements of phenomenal field properties which are the target of the researcher's analysis. Thus, this perspective focuses on researcher's visual perception as a social and situated action. Seeing as a situated activity (Goodwin & Goodwin 1996), professional vision (Goodwin 1994a, 1995) and the ordered production of visibility as a social accomplishment (Mondada 2005) are therefore a central object of this praxeology.

Turning video practices into a topic of analysis offers a range of possible objects of study: not only the practices and methods by which video records are *produced*, but also the practices by which they are then *edited* – going through multiple transformations thanks to digitalization, compression, cutting, reassembling, clips editing, etc. – as well as the practices by which they are *viewed*, either as films, as in the case of the ethnographic documentaries, or as data, as in the case of interactional studies: in this case, they also undergo temporal manipulations, as they are viewed frame-by-frame, in slow motion, or fast forward. Here we will focus on video as it is produced – but the same approach could/should deal with video as it is exploited and made exploitable: how its details are retrieved, how vision is enhanced by digital media, how it depends on software and their constraints, on compression modes, on types of alignment of the image with its transcripts, etc.

So, we will situate video practices within an ethnomethodologically inspired conversation analytic framework, and deal with video both as an indispensable medium for collecting data and *preserving* their relevant features in a naturalistic perspective and as a *configuring* device – taking into account the professional practices that produce it. In this sense, we focus on the detailed ways in which video recordings are reflexively produced, how they structure and arrange the very data of the analysis, shape them, give them a particular orderliness and meaning. It is therefore of fundamental importance to integrate into the analysis the practical ways in which recordings are produced, with their local contingencies and for all practical purposes. These aspects are neither marginal (so that we may ignore them) nor problematic (so that they might “distort” the phenomena at hand) with regard to the use of the resulting materials for analytic purposes. On the contrary, the very ways of producing images give us central insights into the organizational features of the recorded practices themselves, revealing their local order and intelligibility as reflexively produced by their display to and for the camera.

Documenting naturally occurring interactions

From its very beginnings, Conversation Analysis has been a pioneering movement that has explicitly recognized the use of recorded data, and even claimed that such data are embedded in its specific analytic “mentality”. In the early 60s, Harvey Sacks articulated

his topics of inquiry with the necessity of working on recordings: aiming at describing the (ethno)methods by which members organize their ordinary social life, he insisted from the start that these methods and the detailed resources they exploited could not be *imagined* but only *discovered* by a close looking at actual recordings of everyday activities.

"I started to work with tape-recorded conversations. Such materials had a single virtue, that I could replay them. I could transcribe them somewhat and study them extendedly – however long it might take. The tape-recorded materials constituted a 'good enough' record of what happened. Other things, to be sure, happened, but at least what was on the tape has happened. It was not from any large interest in language or from some theoretical formulation of what should be studied that I started with tape-recorded conversations, but simply because I could get my hands on it and I could study it again and again, and also, consequentially, because others could look at what I had studied and make of it what they could, if, for example, they wanted to be able to disagree with me" (Sacks 1984: 26; Sacks 1992, 1: 622).

Sacks' arguments concern both fundamental features of human interaction and praxeological features of its analytic practice: first of all, the way in which the tape makes it possible to organize what Garfinkel appropriately calls "another next first time" (2002: 98), but also the ways in which analysis as a social activity can be made possible, through the sharing, the reference to, the discussion and the collaborative analysis of available records. Even if Sacks worked mainly on audio, as early as the start of the 70s, Chuck and Candy Goodwin made an impressive number of videos (used also for Goodwin's dissertation, 1981), much of which got lost, but some of which eventually became "cult" fragments, such as the Auto Discussion – an excerpt of 30 minutes from tape 84, the only surviving fragment of a corpus of 3 days of continuous shots (Goodwin & Goodwin, personal communication). From the beginnings of the 70s, first in Philadelphia, then at the Summer Institute of Linguistics and in California, the Goodwins animated a number of seminars on video data. In parallel, in the UK, in about the same period, Christian Heath began to video tape medical consultations (see his dissertation, published in 1986).

Data production respects the naturalistic orientation of this framework, requiring that participants' activities be observed in their ordinary social contexts, in naturally occurring interactions, i.e. in interactions which have not been orchestrated by the researcher, which would have taken place even if she would be absent – but which represent people's ordinary business. This fundamental demand is related to a specific vision of social activities and of language developed within an ethnomethodologically inspired conversation analytic framework, which insists upon the following aspects:

– a *praxeological view of language and action*: social practices (*versus* representations, cultural beliefs, mental models, norms, etc.) are central to the constitution of the social and the grammatical order. Language is not an autonomous system, but a set of practices and resources. The natural habitat for grammatical resources is the sequential organization of social interaction (Schegloff 1996). Therefore language is

- an interactive and emergent phenomenon (Hopper 1988) for which time is an essential feature to be studied (Auer et alii 1999).
- an *endogeneous view of resources* as details oriented to and indexically exploited by participants: details of action constitute the relevant accountably phenomenal field properties which are constantly exhibited and interpreted by participants in order to coordinate their conducts. These details are often "seen but unnoticed" (Garfinkel 1967), or seen "at a glance" (Sudnow 1972) and constitute the "scenic intelligibility" (Jayyusi 1988) of conduct and participants' arrangements (Goodwin 2000). Thus, visual features, and more generally multimodal resources, are fundamental and allow a switch from a logocentric perspective on language and action to an embodied view.
 - a *situated view of social conduct*: interactions are *reflexively* structured, i.e. conduct adapt to its context and at the same time, by interpreting it in a certain way, configures it by the very fact that it adjusts to this particular feature and not another - being thus both context-shaped and context-renewing. This makes it impossible to transpose behaviour elsewhere so as to recording it more conveniently, with more sophisticated equipment or in absence of certain noise or lighting problems (e.g. by asking people to "informally chat" in a acoustic room in a academic lab).

Video as preserving relevant details of situated action

These fundamental demands concerning the empirical objects to be studied are integrated and embodied into specific practices and techniques of videotaping (cf. Goodwin 1994b, Jordan & Henderson 1995; Heath 1997; Meier 1998). Recording constitutes less a "registrierende Konservierung" than a "rekonstruierende Konservierung" as Bergmann (1985: 305) puts it. It is a paradoxical action: a dynamic fixation, a "Fixierung", which attempts to preserve the "Flüchtigkeit" of social events and their temporality. What these recordings try to reconstruct are the details to which participants orient when they produce and interpret their own and the others' conducts. If the aim is to develop an endogenous analysis of the members perspectives embedded in their practices, then the very details attended to and exploited by them have to be recorded, as well as their orientation to them (in form of gaze, of body positions, of demonstrable orientations in talk, etc.). This sets the task for a specific way of producing videotapes, consistent with a specific "analytic mentality".

This specific way adheres to an "*availability principle*" (Mondada 2003b): the analytical task of recording (and, in the same way, of digitising, anonymizing transcribing, annotating, etc.) is to provide for the availability of relevant details - which indeed makes the analysis possible. Not recording some of these details (and a part of them are not predictable) would mean not providing for the very possibility of analyzing them and therefore the action they organize methodically. Thus, the ethnomethodological conversational inquiry adopts specific ways of shooting video which preserve key dimensions such as:

- *time*. Openings as well as closings of an activity, its particular length and rhythm are respected: one continuous shot documenting the whole activity is the specific video response to this feature. This poses a series of *analytical* problems, such as the recognizable character of an opening and of pre-opening activities, or the documentation of the active relevant arrangement of bodies, objects and spaces as a precondition for the activity to begin; and also a corresponding series of *practical* problems, such as the renewal of batteries or cassettes – which can alter the very temporal organization of the activity (such as its length, but also its articulation in phases) by imposing an exogenous temporality: these problems can prompt researchers either to limit themselves to record shorter interactions or to search for technical solutions allowing a maximum of autonomy (e.g. by video recording directly onto more commodious hard disks).
- *participation framework and interactional space*. All relevant participants are considered: this produces video shots avoiding focussing too narrowly on only one participant (for example the established speaker) and, on the contrary, taking into consideration not only the recipient designed character of talk and action – which might be problematic for multiparty interactions but also the artefacts and tools manipulated by participants and their dynamic movements. This requirement implies a way of monitoring not only the ongoing action but all possible participants attending to it and produces a strong sense of the complexity of the *interactional space* – documented either with a static camera (allowing the cameraman to be absent, but then having to anticipate all possible movements done by the participants) or with a mobile camera (imposing the presence of the cameraman, who locally accompanies and projects the next possible action). Possible problems arising from the requirement to document relevant details of space and objects oriented to by participants concern the granularity of the available details (e.g. if participants read a map or look at a screen, the visual details they refer to are often barely visible on the video shot). Devices for preserving the continuity of space or participation framework as well as the complexity of objects do not only concern camera movements and focus, but also further transformations of the video records: in this sense, multiscope videos (using split screen or PIP, picture-in-picture) constitute a technical solution (adopted either during production or post-production), especially for complex spaces of action (workplace studies for example, but also ordinary conversations taking place in different rooms – cf. Zouinar, Relieu et alii, 2004, Balthasar & Mondada, in press; Mondada in press).
- *multimodal details*. Video shooting aims at documenting multimodal resources (language, gaze, gesture, body displays, facial expressions, etc.) as they are locally mobilized and attended to by participants. This means that the relevance of details is endogenously produced within courses of collective action as they are interactively and reflexively constructed moment-by-moment within the contingent unfolding of practices. Video records aim not at the production of descriptive glosses by the analyst, but at making available the ways in which participants themselves deal with these details, by methodically orienting to them and by exploiting them for the subsequent organization of action.

Video as configuring and assembling relevant details

This aim to preserve fundamental features of the events and activities is accomplished through video practices and through the cameraman's embodied analysis of the recorded events, which all contribute to the reflexive configuration of these very features – before and during the shot. In the latter case, participants can take the camera movements into account.

Setting up the video camera before the action

Choices of perspectives and spots from which to record action, choices of the beginning and the end of a recorded segment – which depend often on technical constraints (such as the length of the cassettes, the possibility of placing the cameras in difficult angles and locations, etc.) – and other technical choices – concerning the equipment, its miniaturization, angles and lenses, microphones, etc. – results from a reflexive analysis of the situation even *before* the action takes place. In this respect, fieldwork plays an essential role for the identification of expectable patterns of action to which to adjust the video shot.

For example, videorecording a work session between agronomists and computer scientists discussing about how to draw and to read farmland maps, we decided to use two cameras: one was set to point down from the ceiling, taping the working space of the table covered with maps within a vertical perspective (image A); the other was placed sideways on in order to capture all the participant's upper bodies (image B):



Image A

Image B

These two images offer very different perspectives on what happens: if the first provides a detailed view on the visible features accomplished by pointing and other gestures; the second one documents the bodily postures and gazes of the participants. In such an activity, the alternation between mutual gazes and a common focus of attention on the artefacts is a key feature that neither one nor the other image alone can capture. Analyzing the way in which turn-taking is projected by finely timed hand movements and not only by talk (Mondada 2004), we discovered that the first and the second image offer different relevant details.

(1a) (e9/agro1-47.00) (transcript based on the vertical view - image A)

- 1 PAL et puis à un autre moment,
and then at another moment
- 2 ben on va échouer, (.) en pâturage, .h
well we'll wind up (.) with pasture .h
- 3 sur l'assemblage +sans parcours. .h +je pense que+
in an assemblage without any path .h i think that
- viv +.....+moves paper+
- 4 +dans le cas du gaec du pr+adou, .h c'est tout l'un,
in the case of ((name of place)) .h it's either one
- viv +moves wrist and comes in+
- 5 tout l'autre.
or the other
- 6 VIV +.hh oui. parce que: i m'*sem+ble: eh i*- ici
.hh yes because it seems to me (that) ehm he- here
- +.....+points-->
- lau *opens booklet*
- 7 c'était s:- ce qui: ce que ça voulait représenter.
it was wh:- what what it was meant to represent.

Vivian's turn begins with a pointing gesture (line 6) which is progressively stretched along the table on the document held by Laurence – even prompting her to open the booklet in which the map pointed to by Vivian is located. A systematic feature discovered in our analysis is that pointing gestures are mobilized as multimodal resources for accomplish turn-taking, often projecting turn's begin. Here, this projection is visible in other kinds of detailed body movements, such as hand movements and displacements of objects (lines 3, 4). This movements are done methodically in a specific sequential position: they project the end of Pierre-Alain turn-constructural unit (TCU, finishing with "sans parcours/" 3) at a transition relevance point. However, at that very moment, Pierre-Alain expands his turn by an inbreath (".h" 3) and a new TCU; Viviane projects its end by coming forward with her hand. The pointing gesture line 6 is thus anticipated by other movements timely inserted within the unfolding turn organization.

Now, if we look at the other perspective, we can enrich the transcript with participant's glances:

(1b) (e9/agro1-47.00) (transcript based on the sideways view - image B)

- 1 PAL et puis à un autre moment:
- 2 ben on va échouer, . en pâturage, .h
- 3 sur l'assembla:‡ge +sans parcours, .h +je pen‡se que+
>>looks LAU->‡looks at VIV-----‡looks down->
- viv +.....+moves paper+
- 4 +dans le cas du gaec du pr+‡adou, .h c'est‡ tout l'un,
--->‡looks at LAU--‡looks at VIV-->
- viv +moves wrist and comes in+

- 5 tout l'autre.
 6 VIV +.hh oui. parçce que: i'm'*sem+ble: eh i-* ici c'était
 +.....+points-->>
 pal --->‡looks at the pointed document-->>
 lau *opens booklet*

During his description of an area of farmland, Pierre-Alain visually addresses Laurence, the computer scientist. However, even before Viviane moves her hand, Pierre-Alain moves his gaze from Laurence to her (line 3): in this way he projects the end of his explicating turn and accountably selects Viviane as the next speaker; Viviane's hand movement is responsive to this gesture, aligning with Pierre-Alain's projection. In other cases (Mondada 2004), pointing gestures allow the co-participant to self-select as possible next speaker. The difference between both methods is a matter of timed movements, whose sequential position is fundamental for deciding who is initiating and projecting the next step. The availability of these movements for the participants as well as for the camera/analyst is central too: the camera setting as well as the possibility of consulting two views during the analysis are central for the very results it can produce. In our case, the perspective view was recorded as a PIP, favouring the vertical view: however for this analysis the PIP was not precise enough, and the perspective view had to be resorted to.

Camera movements seeing and anticipating courses of action

Other kinds of reflexive online analysis take place *during* the action: camera movements display the ongoing interpretive practice of the cameraman. Videotaping is a practical accomplishment adjusted to the local contingencies of action and embodying the online interpretation the researcher makes of what happens. The continuous video record is produced within a contingent course of inquiry, where the researcher looks at social life with his camera and his look is made accountable by its record (cf. Macbeth 1999: 140). The camera sees and indeed anticipates the unfolding production of action as it is locally accomplished: the video task consists thus in the work of discovering the local interactional order, as the shot unfolds and as action unfolds.

The cameraman finds and frames the event in an embodied way, the camera being both a resource and a constraint, allowing to see in a particular way, making available a field of view narrower (or a wider, if fish-eye lenses are used) than the eye's. In this sense, as Macbeth puts it (1999) cameras can't glance but can only look or stare at social action.

Again, methods for managing turn-taking are interesting to observe as a practical problem for the participants but also for the cameraman. The classroom setting is a notoriously difficult workplace to videotape (cf. Zunger, Ford & Fassnacht 1998), providing a perspicuous setting for the analysis of the movements of a mobile camera on the selected pupils during a grammatical exercise.

(2) (Brig1-30: 5.23-5.50)

- 1 AHM *la terminaison, (.) °de l'im- (.) par- (.) # fait.°
the ending (.) °of the im- (.) per-fect°
 >>looks at ENS--->
 cam >>focussed on AHM--->
 im # im.A
- 2 CHA de l'imparfait
of the imperfect
- 3 ENS de l'imparfait, sophie.Δ (.) der*nière fois Δ hein ?
of the imperfect sophie (.) last opportunity okay ?
 ahm --->Δlooks away-----Δ
 cam --->*zoom back----->
- 4 après gare à vous* si vous ne connaissez pas tous. (.)
after that beware if you don't know (it) all
 cam --->*general view--->
- 5 Δallez *vas-y,
okay let's go
 ahm Δlooks at the front right--->
 cam ---->*moves toward the front, maintaining large view-->
- 6 SOP °alors, ° pour former le conditionnel du présentΔ
°so in order to create the conditional of the present
 ahm ---->Δ
- 7 ? ()
- 8 ENS on prend;
one takes
- 9 ? °on prend°
°one takes°
- 10 CHA on prend l'imp- (.) le:=
one takes the imp- (.) the=
- 11 VIO =sch:: tais-toi, sch::
=sch:: shut up sch::
- 12 ENS allez,
let's go
- 13 SOP xxxxx
- 14 ENS mais tu n'écoutes pas,
but you don't listen
- 15 NAD maîtresse, j'peux?
teacher, can I?
- 16 ENS euh:: saΔrina#
ehm:: sarina
 ahm Δturns quickly behind--->
 im # im B
- 17 (0.2)*(0.3)+(0.2)+ #
 cam ->*switches to the left, moving to the last row--->
 im # im C
 sar +adjust on her chair and advances her body+

- 18 SAR pour () le conditionnel
for () the conditional
- 19 ENS [pour former, il faut un p'tit
in order to create, you need a little bit
peu de vocabulaire quand même "scientifique"
of "scientific" vocabulary nevertheless
- 20 SA "pour former le conditionnel présent"
in order to create the conditional present

At the excerpt's beginning, the camera is focussed on the pupil being interrogated, Ahmed (line 1-3, image A). In the subsequent moments, camera movements will exploit Ahmed's own orientation to the ongoing action's organization for relevantly framing the interactional space of the classroom: as soon as the sequence between the teacher and Ahmed is closed and another pupil is selected (3), Ahmed looks away from the teacher. Slightly after, the camera operator zooms back, encompassing a more general view of the classroom, waiting for the identification of the next speaker. Again when Ahmed looks to his front right, the camera moves in that same direction, locating Sophie whose answer is being delayed: so, when she begins her turn, the camera is focussed in her direction (6). Interestingly, as Sophie encounters problems in continuing the exercise, another pupil is selected, Sarina (16): on the last syllable of her name uttered by the teacher, Ahmed turns visibly his head behind (image B) and this displays for the cameraman the place where Sarina sits, on the last row of the classroom (image C). During the following brief pause, the camera moves toward her and is thus able to document not only her answer but also her bodily rearrangement and adjustment to her emerging speaker position.

This excerpt shows how member's displays of the relevant interpretation and adjustment to the unfolding interactional order are exploited as resources by the cameraman following and projecting their next actions. In this sense, camera movements as visible on the tape and as transcribable as one detail among other interactional details, provide not only for the documentation of the action but for its temporally emergent accomplishment.



image A
camera focussed on AHM



image B
AHM orienting to the selected
next speaker



image C
camera moving to the last row,
focussing on the selected speaker

Interacting with the camera

We are concerned, in this paper, with the fact that video shooting actively co-produces the peculiar orderliness of the events, gestures, actions, and talk it displays and documents. Moreover, the cameraman orients to this orderliness in the very production of the visual record, contributing to its accomplishment in accordance with his practical purposes and in a recipient-designed manner, related to the co-participants, audiences, and co-workers he is engaged with:

- participants *inspect* and interpret the movements and focus of the camera, presuming possible topics of research interest; this produces an online analysis of the ongoing situation, comprising the camera, from their side too (this public availability of the camera gaze is even more explicit with a mobile camera than with a static one).
- participants also *adjust* to the camera in the organization of the course of their affairs and for their practical categorization as being filmable or not (orientation to the camera can reveal “embarrassing”, “delicate” actions or words, which are categorized as such by the very fact that at that moment the participants make a comment on/to the camera. This can work as a very useful advice for taking into account ethical issues, and for data anonymization, to be related to the local relevance of delicate matters and their emergent categorization – cf. also Speer & Hutchby 2003).
- participants can *exploit* the camera as a resource for their own activities: either for accomplishing the ongoing action (for example by using the video perspective in order to organize a queue or other arrangements of bodies) or for accomplishing another action related to the camera as inaugurating a stage or a scene. In this sense, dealing with the effects of the camera on the ongoing action as a “bias” or considering them in terms of an “observer’s paradox” (Labov 1972) misses the very entanglement of the camera and of the action and the very fact that the camera can indeed reinforce and reveal structural elements of the situation and activity (and thus can be enrolled as a resource for the production of its order and accountability).

This shows how these phenomena as well as video taping itself are embedded within the organization of the ongoing action, and have to be integrated within its analysis and not kept separated – In a methodological appendix or in the “backstage” areas of science. In these moments, what happens is a “recorded conversation” and not just a “conversation” – as a phenomenon which itself says something about the local relevancies, the participants, the organization of the ongoing activity (and not only about the way in which data are constituted). However, the relevance of the video device for the local action’s organization has not to be supposed a priori and in general, but has to be demonstrated moment-by-moment through the accountable orientations of the participants: orientation to the camera is a phenomenon that can be analyzed as a topic (cf. Heath 1986: 11-13, Lomax & Casey 1998 for good examples). Skilled ways of obtaining an informed consent by the participants and of managing the discreteness/sophistication of the recording contribute to the form these orientations might

take: in this sense, the way in which the recording is organized remains fundamental, as it enhances or diminishes the salience of the device for the participants.

Video as a professional practice

As we have seen, videos are not just data made available for research: videos are produced within specific, situated, contingent practices. However, these practices concern various kinds of social members and groups:

- not only the researchers, on which we have focussed our paper, but also
- the amateurs using video for private purposes such as videorecording a friend's wedding, snow holidays, a son's birthday, etc.;
- other professionals skilfully producing and configuring video images for all practical purpose. In this latter respect, video practices and devices are more and more deeply embedded in contemporary workplaces and institutional settings: surgeons use endoscopic cameras explore the anatomy within the course of a laparoscopic operation in order to provide for the relevant visual environment for the next surgical step (Mondada 2003c), TV cameraman and directors produce images appropriate for broadcast (Relieu 1999; Broth, 2003; Mondada, in press b), video are shot by CCTV and surveillance cameras in public spaces (Ball 2000), airport control rooms (Goodwin & Goodwin 1996; Suchman 1996), underground station operation rooms (Luff, Heath & Jirotko 2000). These settings can be studied for the way in which they produce videos as well as for the online interpreting practices by which they exploit them. In the latter case, studies focus on seeing as a professional practice that can also be documented in other settings - dealing with fixed images, computer screens or other objects (Goodwin 1994a, 2000; Heath et al 2002; Lynch 1988).

Video practices as they are organized and promoted by social actors for other than academic purposes produce video records that can be used by researchers themselves, who in this case delegate the production of data to the members they observe. In this case, the objects made available for researchers are either the products of these practices, allowing for the reconstruction of the video movements (as they are witnessable in the image) or the activity of production itself, videotaped by the researcher (for example, Schmitt, in preparation, videotapes a film team shooting a movie). This latter case is a reminder for not reducing the production to its product - this being particularly striking for media analysis (cf. the work of Relieu 1999, Broth 2003).

Working with videos produced by members within their professional activities allows to take into consideration their practical purposes and to analyze how they are materialized and embodied in particular ways of shooting. For example, surgeons operating laparoscopically are firstly interested in the endoscopic image they look at on monitors in the operating room during the procedure (im. A); however, when they broadcast the operation in an amphitheatre for teaching purposes, they add to this

endoscopic image another external view of the relevant area of the body, showing where the instruments are inserted - switching between one image and the other or using PIP in order to make both of them available (im. B) (Mondada, in press a). The researcher interested in team collaborative work can shoot her own videos within the operating room, which will produce a very different view, for other practical purposes (im. C). Although these various views can be edited in a composite image (a split screen), they are shot in order to produce very different accountabilities and intelligibilities of action (Mondada 2003a), which can well be incommensurable.



image A
interior endoscopic view used
by surgeons during the operation

image B
exterior view used for teaching
purposes, often inserted within
the interior view

image C
view of the operating room
videotaped by the researcher

Thus, video practices embedded within specific professional practices produce specific accountabilities of action. In a project aiming at documenting the experiences of persons who escaped the Holocaust, historians have recorded a series of testimonies in the form of video taped interviews of witnesses. Their video shootings are focussed on the witness, framing the upper part of his body against a curtain - accomplishing in this way the "officiality" of the settings (image A). Alternative video shootings we produced within an interactional linguistic project consist in views of the witness and the two historians interviewing him, thus constituting a wider view on the setting - which is visible as a video studio - encompassing the whole participation framework

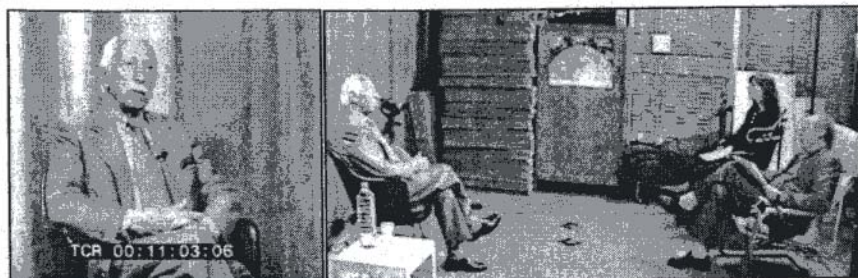


image A
historians' view on the testimony as
a monological narrative

image B
alternative view on the testimony as a collective enterprise

and the recording devices such as microphones and some cameras (image B). Whereas the historians' view makes the activity accountable as a testimony's narrative, focussing on the speaker alone, the alternative view produces the accountability of the activity as the collective elaboration of a storytelling within an interview.

Video production as a social practice – to be studied by considering symmetrically both researchers engaged in conversation analysis and other professionals or amateurs involved in the local production of videos for all practical purposes – constitutes a perspicuous setting for the study of embodied seeing practices, namely for a praxeology of seeing with a camera. Their study can show how the accountability of action is variously produced by the cameraman's embedded online choices and analyses, reflexively adjusting to the contingencies of the recorded action.

Concluding remarks

In this paper we dealt with video practices as a topic of inquiry, focussing on video production within the framework of ethnomethodology and conversation analysis. We highlighted the very fact that these videos are practical accomplishments within specific contexts and contingent courses of action, adjusting, anticipating, following the dynamics of sequential unfolding of interaction and of changing participation frameworks. Videos produced within the naturalistic requirements of ethnomethodology and conversation analysis both aim at preserving relevant details and phenomenal field features and reflexively contribute to the configuration of the very interactional order they document. This reflexive dimension of video practices can be studied in depth by also considering other skilled professional video activities, engaged in producing video documentation of events, actions or talk for various practical purposes. Thus video shooting – either by researchers, by professionals or by amateurs – can be a perspicuous setting for the observation and analysis of the way in which members develop endogenous online situated contingent analysis of social life.

Transcript conventions

Data have been transcribed according to conventions developed by Gail Jefferson. Multimodal details have been transcribed according to the following conventions (according to Mondada 2004):

- * * gestures and actions descriptions are delimited between
- + + two identical symbols (generally one symbol per participant)
- Δ Δ and are synchronized with correspondent stretches of talk
- >> gesture or action described begin before the excerpt's beginning
- >> gesture or action described continue after excerpt's end
- *---> gesture or action described continue across subsequent lines

--->*	until the same symbol is reached
....	gesture's preparation
----	gesture's apex is reached and maintained
.....	gesture's retraction
lau	participant doing gesture is identified when (s)he is not the speaker
cam	camera movements are described
im	the exact point where screen shot (image) has been taken is indicated
#	with a specific sign showing its position within turn at talk.

References

- Auer, P., Couper-Kuhlen, E., & Müller, F. (1999). *Language in Time. The Rhythm and Tempo of Spoken Interaction*. Oxford: Oxford University Press
- Ball, M. (2000). The visual availability and local organization of public surveillance systems: the promotion of social order in public space. *Sociological Research Online* (www.socresonline.org.uk), 5(1)
- Balthasar, L., Mondada, L. (in press). Multiscope videos. *Interacting Bodies. Proceedings of the 2d ISGS Conference, Lyon, June 2005*
- Bergmann, J. R. (1985). Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit: Aufzeichnungen als Daten der interpretativen Soziologie. In W. Bonss & H. Hartmann (Eds.), *Entzauberte Wissenschaft (Sonderband 3 der Sozialen Welt)*. Göttingen: Otto Schwarz
- Broth, M. (2003). Analyse de l'interaction à la télévision. *Moderna Sprak*, 2(2), 193-202
- Büscher, M. (2005). Social life under the microscope? *Sociological Research Online* (<http://www.socresonline.org.uk/10/1/buscher.htm>), 10(1)
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall
- Garfinkel, H. (2002). *Ethnomethodology's Program. Working out Durkheim's Aporism*. Lanham: Rowman and Littlefield
- Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press
- Goodwin, C. (1994a). Professional Vision. *American Anthropologist*, 96(3), 606-633
- Goodwin, C. (1994b). Recording human interaction in natural settings. *Pragmatics*, 3, 181-209
- Goodwin, C. (1995). Seeing in depth. *Social Studies of Science*, 25(2), 237-274
- Goodwin, C. (2000). Practices of Seeing, Visual Analysis: An Ethnomethodological Approach. In T. v. Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 157-182). London: Sage
- Goodwin, C., & Goodwin, M. H. (1996). Seeing as a situated activity: Formulating planes. In D. Middleton & Y. Engeström (Eds.), *Cognition and Communication at Work*. Cambridge: Cambridge University Press
- Heath, C. (1986). *Body Movement and Speech in Medical Interaction*. Cambridge: Cambridge University Press
- Heath, C. (1997). Analysing work activities in face to face interaction using video. In D. Silverman (Ed.), *Qualitative Research. Theory, Method and Practice*. London: Sage.
- Heath, C., Svensson, M. S., Hindmarsh, J., Luff, P., & vom Lehn, D. (2002). Configuring awareness. *CSCW*, 11(3-4), 317-347
- Hopper, P. (1988). Emergent grammar and the a priori grammar postulate. In D. Tannen (Ed.), *Linguistics in Context: Connecting Observation and Understanding* (pp. 103-120). Norwood: Ablex

- Jayyusi, L. (1988). Toward a socio-logic of the film text. *Semiotica*, 68(3/4), 271-296
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39-103
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Latour, B. (1986). Visualisation and cognition: Thinking with eyes and hands. *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, 6, 1-40
- Lomax, H., & Casey, N. (1998). Recording social life: Reflexivity and video methodology. *Sociological Research Online*, 3(2)
- Luff, P., Heath, C., & Jirotko, M. (2000). Surveying the scene: Technologies for everyday awareness and monitoring in control rooms. *Interacting with Computers*, 13, 193-228
- Lynch, M. (1985). Discipline and the material form of images: An analysis of scientific visibility. *Social Studies of Science*, 15, 37-66
- Lynch, M. (1988). The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. *Human Studies*, 11, 201-234
- Lynch, M. (1991). Pictures of Nothing: Visual Construals in Sociological Theory. *Sociological Theory*, 9-1(1-21)
- Macbeth, D. (1999). Glances, trances, and their relevance for a visual sociology. In P. L. Jalbert (Ed.), *Media Studies: Ethnomethodological Approaches* (pp. 135-170). Lanham: University Press of America & Int. Inst. for Ethnomethodology and Conversation Analysis
- Meier, C. (1998). Zur Untersuchung von Arbeits- und Interaktionsprozessen anhand von Videoaufzeichnungen. *Arbeit. Zeitschrift für Arbeitsforschung, Arbeitsgestaltung und Arbeitspolitik*, 7(3), 257-275
- Mondada, L. (2003a). Describing surgical gestures: the view from researcher's and surgeon's video recordings. *Proceedings of the First International Gesture Conference, Austin, July 2002* (http://www.utexas.edu/coc/cms/International_House_of_Gestures/Conferences/Proceedings/Contents/List_of_Papers.html)
- Mondada, L. (2003b). Observer les activités de la classe dans leur diversité: choix méthodologiques et enjeux théoriques. In: Perera, J., Nussbaum, L., Milian, M. (eds.). *L'educacio linguistica en situacions multicultural i multilingues*, Barcelona: ICE Universitat de Barcelona, 49-70
- Mondada, L. (2003c). Working with video: how surgeons produce video records of their actions. *Visual Studies*, 18, 1, 58-72
- Mondada, L. (2004). Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'interaction: Le pointage comme pratique de prise du tour. *Cahiers de Linguistique Française*, 26, 169-192
- Mondada, L. (2005). Visions controversées de la carte: construire le visible par les gestes et la parole en interaction. In: C. d'Alessandro et alii (eds.). *Espaces, savoirs et incertitudes*. Paris: Belin, 15-31.
- Mondada, L. (in press a). Operating together through videoconference: members' procedures accomplishing a common space of action. In: Francis, D. (ed.). *Orders of Ordinary Action*. Washington: University Press of America & Int. Institute for Ethnomethodology and Conversation Analysis
- Mondada, L. (in press b). Videorecording and the reflexive constitution of the interactional order: uses of split screen in video practices
- Relieu, M. (1999). La réalisation et la réception du produit télévisuel comme accomplissements. In J.-P. Desgoutte (Ed.), *De l'image au verbe. Le jeu interactif dans le discours audiovisuel*. Paris: L'Harmattan
- Sacks, H. (1984). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of*

- Social Action* (pp. 21-27). Cambridge: Cambridge University Press. (Edited by Gail Jefferson from various lectures)
- Sacks, H. (1992). *Lectures on Conversation [1964-72] (2 Vols.)*. Oxford: Basil Blackwell.
- Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Grammar and Interaction* (pp. 52-133). Cambridge: Cambridge University Press
- Schmitt, R. (in preparation). Das Filmset als Arbeitsplatz: Einblicke in die Struktur einer komplexen Kooperationsform
- Speer, S., Hutchby, I. (2003). From ethics to analytics: Aspects of participants' orientations to the presence relevance of recording devices. *Sociology*, 37, 2, 315-337
- Suchman, L. (1996). Constituting shared workspaces. In D. Middleton & Y. Engestrom (Eds.), *Cognition and Communication at Work*. Cambridge: Cambridge University Press.
- Sudnow, D. (1972). Temporal parameters of interpersonal observation. In D. Sudnow (Ed.), *Studies in Social Interaction* (pp. 259-279). New York: Free Press
- Zouinar, M., Relieu, M., Salembier, P., Calvet, G. (2004). Observation et capture de données sur l'interaction multimodale en mobilité. *Mobilité & Ubiquité '04*, Nice 1-3 June 2004. Nice: ACM
- Zungler, J., Ford, C., Fassnacht, C. (1998). *Analyst eyes and camera eyes: Theoretical and technological considerations in "seeing" the details of classroom interaction*. Albany: The National Research Center on English Learning and Achievement. Report Series 2.40

In the last few years we have witnessed the widespread proliferation of video camcorders as a powerful and sophisticated instrument for data collection. Video is increasingly used in broad areas of research throughout the social sciences. It allows for a rich recording of social processes and provides a completely new kind of data. Used as a "microscope of interaction", this "video revolution" is expected to exert profound impact on research practice. But despite its popularity as an instrument, the methodological discussion of video is still underdeveloped. This book gathers a selection of outstanding European researchers in the field of qualitative interpretive video analysis. The contributions discuss the crucial features of video data and present different approaches how to handle, interpret, analyse and present video data collected in a wide range of "real world" social fields. The book thereby aims at providing an overview on contemporary interpretive and qualitative approaches to video analysis.

Hubert Knoblauch is Professor of Sociology and Bernt Schnettler is Assistant at the Technical University Berlin, where they conduct research, among other topics, on communicative genres.

Jürgen Raab is Lecturer and Hans-Georg Soeffner is Professor of Sociology at the University Constance. They have developed and applied an original approach to visual data called Sociological Hermeneutics.