

Videography: An Interpretative Approach to Video-Recorded Micro-Social Interaction

Hubert Knoblauch and René Tuma

INTRODUCTION

Video technology offers expanded possibilities for looking at interaction and communication in various ‘natural’ settings: for example, workplaces and situations of informal communication or educational interaction. In the last decades international groups of scholars have developed a prominent approach to sequentially analyze visual data from social interaction. In addition to standardized ways of looking at video recorded interaction, this approach is interpretive. After a short overview of earlier research, we will briefly discuss the relevance of interpretation to video analysis. Before we describe how to sequentially analyze video-recorded social interaction, we sketch the role of data collection, coding and sampling. Because these processes are part of a focused ethnography, we use the shorthand ‘videography’ rather than video interaction analysis. ‘Sequentiality’ is a central category in fine-grained analyzes of video-recorded interaction that supports

empirically based interpretations of actions. This will be illustrated through three empirical examples, which also highlight the interrelatedness between textural and visual elements. Finally, we will discuss the relation between sequential analysis of video-recorded interaction and the wider ethnographic and social context, which is essential to videography.

BACKGROUND

During the past decades, video, i.e. audio-visual records on analogue as well as on digital technical media, not only has replaced film as a medium for private and most scientific recording but also has triggered forms of usage unknown to users of older film media. The technical feasibility of slow motion at any point, of zooming and inserting pictures as well as the digital examination of picture frames have changed the ways of studying

visual interaction. This has led to an explosive diffusion of privately produced videos using the Internet as a favorite medium of distribution. It has also opened new ways for the social sciences to produce and present data. Before the advent of the video, social scientists had been using films in quite innovative ways. In fact, the analysis of conduct was one of the first uses.¹ Throughout the twentieth century, researchers have developed ingenious ways to use film for the analysis of social conduct. Whereas these uses were restricted to an avant-garde of a few researchers, often in experimental settings, the dissemination of video as a technology has now allowed many more social scientists to use it, because it is not only relatively cheap, reliable and available but – due to its widespread private use – video has become less and less obtrusive.

To the degree that video is establishing itself in the social sciences as a technical medium for producing data, the question of how to analyze video has become an urgent topic of social scientific methodology. This chapter addresses the question of how to use video as data in the social sciences. We will examine what one may call the dominant approach in the interpretive analysis of video data. When discussing interpretive video analysis, we adopt a basic phenomenological and methodological assumption of any interpretive social science: actions cannot merely be observed; rather, actions are guided by meanings. Therefore observers must try to account for meaning not only as a general principle but also in each empirical instance (Schütz, 1971; Weber, [1921]1984). On this background (to be elaborated later), we can distinguish interpretive video analysis from standardized video analysis. Standardized video analysis derives its categories from interpreting the audiovisually recorded on the basis of theoretical assumptions. The empirical instances are then subsumed to these categories by coding, so that, for example, certain stretches of conduct are coded as ‘supportive’ or ‘non-supportive’, ‘aggressive’ or ‘non-aggressive behavior’ (Mittenecker,

1987). Note that the reasons for subsuming instances to codes, while theoretically based, are not explicated, so that the process of interpretation remains implicit. Indeed the code may be habitualized (tested by way of ‘intercoder reliability’) or even automated. In fields such as computer supported work or human–computer interaction, there are more than 40 software programs for such standardized analysis available, most of them based on predefined categories (Koch and Zumbach, 2002). Standardized methods are not restricted to experimental and quantitative studies of audio-visual conduct, but are also to be found in the so-called ‘qualitative method’ field. Thus, Dinkelaker and Herrle (2009) suggest subdividing video recorded interactions according to predefined ‘segments’ and ‘configurations’.²

As useful as standardized procedures may be for certain purposes, they differ from interpretive video analyses. For example, Heath and Luff (2007) have demonstrated that even the most standardized forms of action and interaction, such as auctions in England with their rigid pattern of bidding, exhibit interpretive elements defining the very courses of these actions. Thus, interpretation is an essential part of any video analysis that assumes conduct to carry meaning and subjects to exhibit some form of agency or subjectivity.

SUBJECTIVITY, ETHNOGRAPHY AND THE FOCUS

Starting from the assumption that actions can only be explained if we understand their meanings, Schutz (1962) suggests distinguishing between the meanings actors themselves link with their actions, ‘constructs of the first order’, and the ways that outside observers conceive of these meanings, that is ‘constructs of the second order’. With his notion of double hermeneutics, Giddens (1984: 284) added the idea that the scientific constructs themselves are not separated from

the first-order constructs, but interact with them. This interaction, however, must not be considered merely as a problem, but constitutes a resource for understanding the subject matter under investigation. Social scientific analysis, thus, rests on the assumption that the analysts themselves participate in the knowledge of a certain culture under investigation, knowledge that can be explicated in the hermeneutic circle (Soeffner, 1997). In interpreting the 'object', we are also simultaneously explicating the observer's knowledge. In this sense, the observer's and interpreter's subjectivity (based on their own competence as actors) is always involved in the process of understanding the meanings of other subject's actions.

In addition to the observer's general competence in understanding actions of others, video analysis presupposes another kind of knowledge, generally termed 'ethnographic'. Since the structure of knowledge in modern society is highly specialized and fragmented into many different settings, situations, and institutions,³ social scientists are often not familiar with many aspects of the social situations to be observed and recorded on videotape. The various methods of acquiring the situative knowledge relevant to the actors for them to be competent actors, and for the observers in order to know what is going on, is frequently referred to as ethnographic. Ethnographic methods typically include observation, some form of participation in the field, interviewing, eliciting and collecting field documents. These methods allow researchers to gather subjective knowledge of the field under investigation. Typically, it is only by doing ethnography prior to video analysis⁴ that we can understand other actors in the field. As Heath and Hindmarsh stress, for the analysis of video recordings of naturally occurring activities, 'it is critical that the researcher undertakes more conventional fieldwork' (2002: 107). For routine research practices, this implies that it is extremely useful if the researchers doing the video analysis are the same as those who do the ethnography.

One should note that definitions of ethnography differ significantly across disciplines. The type of ethnography to which we refer, focused ethnography, does not aim at encompassing large, locally distributed social structures, such as tribes, villages or cities. As opposed to such encompassing 'conventional' ethnographies (as we will call them for the sake of brevity), these ethnographies may be said to be focused in several ways.⁵

Conventional and focused ethnographies differ, first, with respect to their demands on time. The former are time extensive since they require a long, continuous period of data collection, as a rule for most students about a year. As opposed to this kind of experience-based ethnography, focused ethnographies are short term and data intensive. Even if they linger on for months and years, field contacts and recordings are rather discontinuous and selective, and fields are visited at various intervals. The standard argument against this short-ranged character is that this kind of ethnography may favor 'superficial' or 'quick and dirty' research (Hughes et al., 1994). However, this critique ignores the fact that the shorter time periods of the field contacts are compensated for by immense data intensity. As opposed to classical ethnography, which frequently reduces rich personal experiences into a restricted number of field notes and documents, focused ethnography is data intensive, collecting huge amounts of data on short stretches of actions by means of mechanical data collection devices, including audio, video and photographic recordings.

We consider videography to be a specific form of focused ethnography since the use of video has important consequences for the focusing process. First, the video as a technology entails a focus on its own. By virtue of, for example, its optical focus, the selection of a frame and the transformation into two dimensionality, data collection by video is not only 'recording' but also a sociotechnical construction of data. Although important to keep these constructive aspects of technology in mind, they have been broadly discussed elsewhere by Heath et al. (2010),

who also provide an introduction to data collection; our aim here is to address the essential role of ethnography for video analysis. For it is only by means of ethnography that situations and actions relevant to the setting or to the basic research question are identified. Preliminary ethnography allows researchers to identify the focus of observations, and it serves to collect information and data on this focus and its context. To be more specific, ethnography helps to determine exactly where, when and who is to be recorded on video.⁶

In addition to the focus on a particular situation, the use of video exhibits a particular bias. As Erickson noted, video recordings 'focus on the particular': i.e. the 'particulars of situated interaction as it occurs naturally in everyday social situations' (1988: 1083). The focus on action and interaction is linked to the specific advantages of video as a technical medium. For video is mimetic; it mimics ongoing actions to such a degree that any actor easily identifies his or her action in the medium. Moreover, video is a temporal medium. As opposed to photography, it allows the researcher to record and identify processes taking place in time. Since what are to be recorded are processes in time, there is a strong tendency to record interactions and audio-visual conduct.⁷

Video is a technology for recording as well as for the permanent re-use of audiovisual data for the sake of validation and comparison. This re-use is facilitated by a range of technical options: playing, repeating, slow motion and zooming are among the most important techniques from which researchers profit with video analysis. These techniques allow actions to be observed in a detail not even accessible to the actors themselves. For this reason, video has also been rightly compared to a 'microscope'.

The comparison to the microscope raises a third most important question which is directly related to the technical focus of the video camera: What is being focused on by the technology? Why and how is this being done?

The questions of focus are, of course, dependent on the research questions asked by the particular study, informed by theory, the area of research and the discipline. Thus, linguistics may once have emphasized the audio channel in interaction. In recent years, however, linguistics has closely examined aspects of what are termed 'modalities': next to lexical choices, codes and prosody, gestures, mimic behavior or body posture are now subject to the analysis of 'multimodality', strongly influenced either by structural linguistics or by more recent pragmatic cross-fertilizations between linguistics and neighboring disciplines (Kendon, 2004; Mondada, 2006).

In the social sciences, social actions and interactions have long been a major interest. The paradigmatic case for social scientific video analysis is what Goffman (1961) called 'focused interaction', which is the form of interaction in which the participants share a common focus of attention. In the simplest case, the interaction's focus is constituted by two actors, but focused interactions may also extend to large social occasions, such as meetings, staged events and demonstrations. The interactive focus in such situations can also be established by material items and spatial structures (such as microphones or stages), and other technological mediations of social interaction: for instance, interactivity with computer and other technologies (Rammert and Schulz-Schaeffer, 2002), computer-mediated communication in underground control rooms (Heath and Luff, 1996, 2000), or the social actions performed by medical professionals in an operating room (Schubert, 2006).

It goes without saying that the focus of the camera follows the focus of the study;⁸ on more or less static social occasions the camera itself may be static. For example, in recording PowerPoint presentations on video in various settings (to which we will refer below), the authors used fixed cameras since audiences were fixed, as were the presenters. The fact that audience and speakers were facing one another required two cameras: one

directed to the audience and the other to the speaker. If the activity is itself moving, the camera is supposed to follow the movements: for instance, during work in the operating room, when the hands of the operator are the core activity in the setting and are moving into the patient (Schubert, 2006). Under some circumstances it may therefore be useful for the camera to pan, tilt, or zoom to follow the gaze of the participating observer (Mohn, 2007).

DATA COLLECTION AND CODING

The specific demands for how to produce the video recordings are revealed through the ethnographic work preceding it. This ethnographic work is not restricted to the interactions as the primary subject of the video analysis; it also takes into account the situational and institutional context and the subjective knowledge of the actors (and of the researchers studying it). What will then be recorded depends first on what activities can be observed and, importantly, are accessible in the field. On the other hand, the production of audio-visual data depends on the distinct research question derived, the research context and one's discipline, all of which affect which actions and sequences will be considered relevant to the study. Thus, both the scientific context, as well as the field (as a frame of the research), provide the frame of the research question.

Let us summarize the steps of the research process in more detail. After the first explorative visits in the field, researchers determine which situations to focus on. One of the principal goals of this phase of preliminary scouting is also to identify common features of social situations (such as work interactions, encounters, events) to start the collection of video data. As part of a videography, data collection should not be restricted to one situation or even one type of situation. Thus, when video recording PowerPoint presentations (for example, Knoblauch, 2008), one

may want to compare them with, for example, blackboard or flip chart presentations. This ethnographic search for situations and fields can be called ethnographic sampling.

After the first hours of video data have been collected, the internal video data sampling starts: i.e. the selection of relevant sequences in the data corpus. The basic activity here consists of coding. By coding, we do not mean the application of certain fixed, theoretically deduced categories of analysis to the video data, as practiced with standardized video analysis. Rather, initial codes are developed from the 'bottom up' (as suggested by grounded theory).⁹ The first codes are informed by both ethnographic and everyday knowledge about the field, before they are verified by fine-grained sequential analysis. These codes are entered into a content log or an index describing the temporal unfolding of the interaction by reference to the video time code, so that the video fragments referred to can be found easily.¹⁰ Let us repeat that the various steps in the analysis are neither linear nor separated, but rather intertwined and iterative, as depicted in Figure 22.1.

This iterative character of the research process also holds for the codes and the log-book, which are continuously changed and corrected in the light of sequential analyses and ethnographic fieldwork. Nevertheless, it is helpful to start internal sampling with preliminary codes. As soon as initial ethnographic understanding of the situation has been achieved and some types of interactions have been established by fine-grained sequential analysis, codes start to be corroborated, and researchers should select relevant situations for further scrutiny: i.e. for further internal or ethnographic sampling. Depending on the type of situation under study, the activities and events to be coded vary in type and data extension. For example, when studying auctions, it is easy to find meaningful units: for example, bidding for a specific item will typically be marked by openings and closings.¹¹ To assign codes to certain sequences in the data, it is necessary to

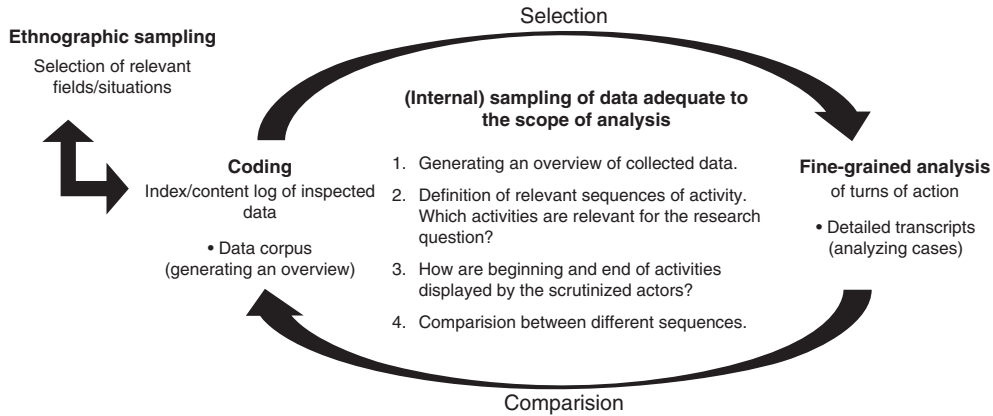


Figure 22.1 Process of analysis: © Knoblauch and Tuma (2010).

analyze sequentially: for example, to determine where and how actors mark beginnings and ends. For this purpose one has to carefully re-examine the data, scrutinizing the order of sequence (this step will be elaborated below) and comparing different situations in which these sequences occur.

When analyzing single sequences in the data, one should pay particular attention to the variation of any found pattern in the observed data corpus. This is typically accomplished by comparison. Comparison helps to identify similar and dissimilar cases according to features that are identified by coding and sequential analysis. Comparisons are associative, looking for similarities, or differentiating, looking for minimal and maximal contrasts. Both ways, the analysis will attempt to determine patterns in the sequences, varying either in terms of their internal structures, in their situative occurrence, or with respect to aspects of the setting in terms of ethnography.

INTERPRETATION AND ANALYSIS

As mentioned at the beginning, the fine-grained sequential analysis starts with interpretation. The necessity of interpretation indicates that video analysis is basically a hermeneutic

activity: interpreting not only means describing and explaining non-verbal and/or verbal behavior but also determining the knowledge that one needs to understand what is going on in a situation and identifying the visible conduct that constitutes the situation (Heath, 1997). Understanding is already involved in the basic production of data: words and sentences must be understood for transcriptions to be produced, and the description of visual processes requires seeing the directions of gazes in the recordings and knowing what people are referring to, so that all essential parts (sentence, word, movement) of a sequence make sense for observers. It is on this basic level of the everyday understanding of first-order constructs that ethnographic knowledge is indispensable. At the same time, professional, second-order constructs – i.e. theoretical, disciplinary – come into play.

Interpretation in everyday terms is the essential basis for analysis. It is involved when approaching the data, which is the tapes and transcripts; it is required for the everyday understanding of what is seen and heard, and it is needed for sequential analysis. Nevertheless, we can distinguish it from analysis. Although both processes are intertwined in practice, the latter is characterized by following certain analytical assumptions.¹² The first assumption lies in the meaningfulness of communication, reflected in the

necessity to interpret (and understand). Since meaning cannot be directly observed, the second assumption is the objectivity of what is to be understood; this objectivity is reflected in the video technology, which can by definition only record objectivated aspects of verbal and physical behavior. Methodologically, objectivity means that only those interpretations are accepted which refer to what is expressed in audio-visual conduct. Interpretations should not be based on general theoretical assumptions, or invisible factors of actors (like 'motives', 'subconscious desires' or 'attitudes'), but only on what can be observed (on the basis of understanding).

The third assumption lies in the reflexivity of communicative actions. By reflexivity, we do not mean that actions are reflected consciously, as Giddens (1984) seems to suggest. To the contrary, most video analyses address practices: i.e. the habitualized knowledge implicit in social action. The notion of reflexivity also refers to this kind of habitualized knowledge, including the ways in which actors 'indicate', 'frame' or 'contextualize' their actions.¹³ When asking a question, for example, not only do we ask the question concerning some content but also we demonstrate that we are asking a question: i.e. that we are performing an action. It is because of this reflexivity that co-actors can understand what is meant by an action. The assumption of reflexivity is reflected in the order exhibited by the objectified actions, which allows the action to be understood. In this sense, a joke or a PowerPoint presentation is not considered as something taken for granted, but as something exhibiting an order that allows participants to identify it as a joke or a Powerpoint presentation. In order to identify this order, the focus of observations lies in how the actions are being performed. By investigating the methodological resources used by participants themselves in the production of social actions and activities (Heath, 1997: 184), reflexivity is, therefore, a crucial resource for interpretation.

Sequentiality is the fourth assumption exploited in video analysis. Sequentiality is

due to the basic temporality of action and interaction – a feature that is maintained if not intensified by the technological medium of video. Like film, video is defined by the temporal sequence of frames (pictures). As a result of their temporality, pictures are watched in sequence, and sequentiality is therefore characteristic for video analysis. It is the feature of sequentiality which causes the particular focus of many video analyses – actions, reactions and interactions – since this medium preserves the time structure of these processes in a way unprecedented by earlier media (except film).

Video analysis is related to other forms of analysis based on sequentiality, including conversation analysis and hermeneutics (Soeffner, 1997; Raab and Tänzler, 2006).¹⁴ These forms share in common their consideration of sequentiality as the very structure by which social action and, thus, social order, are accomplished. For this reason, sequential analysis starts with sequences of actions or interactions as the primary subject matter. However, as opposed to structural or objective hermeneutics, the 'unit' of sequences is not defined in advance or by the technical medium as, for example, one 'frame' of the 24 frames shown on film) or one word (which in spoken language may not be the most relevant unit). Based on the notion of reflexivity, the very task of the video sequence analysis is, rather, to identify on the basis of the data what actors consider as a 'unit': i.e. anything rendered distinct by the actors, which is preceded by something else and followed by something different.

SEQUENTIAL ANALYSIS

The activities focused by sequential analysis with video are identified as sequences of social action, interaction and communication. Therefore, sequential analysis forms the core of videography, the fine-grained video analysis. In order to get a sense of this kind of analysis, we will illustrate the procedure

with some data. The notion of sequentiality may be best understood with respect to a classical empirical example from conversation analysis in which Schegloff (1968) analyzed openings of telephone calls to a ‘disaster center’. After his first analysis of the acoustic data, Schegloff had assumed that there must be a ‘distribution rule of first utterances’, i.e. the answerer speaks first. Then, however, he confronted what is called a ‘deviant case’ to this rule as follows:

#9 (Police makes call)
Receiver is lifted, and there is a one second pause
Police: Hello.
Other: American Red Cross
Police: Hello, this is Police Headquarters... er,
Officer Stratton
(Schegloff, 1968: 1079)

As Schegloff found, the rule does not hold because of the one-second pause.

The fact that the caller is talking is not really a violation of the rule; rather, the pause left by the person called is taken as a kind of answer to which the ‘Hello’ in the second turn replies. (Moreover, this ‘Hello’ is not only a greeting but also takes into account the lack of response.) The general insight is that actions are rendered meaningful in the context of other actions, and that this context is constituted by the sequence of this action.

The most interesting aspect for us in Schegloff’s example is that the ‘turn’ is not actually constituted by a ‘speech act’. In fact, it is not even direct human action but rather a

technologically mediated action – the ringing of the phone. Thus, the ‘unit’ to be considered can be a non-linguistic event if it triggers another action.

That telephone rings fulfill this function is not surprising, since acoustically coordinated action depends on the temporal ordering of sound. But how, we have to ask, can audio-visual conduct be analyzed sequentially? In order to answer this question, we discuss an audio-visual example from our own research.¹⁵ The following still is taken from an auction at a point where an auctioneer is offering a new item to the audience. The offer is interesting because the auctioneer does not reveal what it contains, but describes it as a ‘surprise’. He starts by identifying its number in the catalogue, as indicated in Table 22.1.

The auctioneer (S.) starts (1–2 in Table 22.1) with a number that, for the moment, seems to be opaque, and also his description of a ‘surprise’ object (3 in Table 22.1) remains opaque if we look only at the text. If we turn our attention to his visual conduct, we can tell that he is looking at the desk during (2 in Table 22.1) – which can easily be understood as ‘reading’. Since, as ethnographic knowledge of the situation tells us, there is one common document available in the auction – the catalogue which includes all objects on sale – actors and observers may infer that his glance is for the catalogue. In fact, we can identify a number of actors in the audience looking at a copy of the document

Table 22.1 Audio-visual example, using Gesprächsanalytisches Transkriptionssystem (GAT) transcription conventions (Selting et al., 1998)

Sequence	Original German	Translation in English
1	S:D:ie w:osechsundsiebzig*	S: number tw:oseventysix*
2	(0.6s)	(0.6s)
3	EIN Überraschungsposten	a surprise item
4	(.)	(.)
5	Reisetasche mit Inhalt;	a travel bag with ((unknown)) content
6	(0.9s)	(0.9s)
7	Fünf sind geboten; zehn sin´	Five are offered; ten are
8	geboten Fünfzehn sin´ geboten	offered fifteen are offered

Source: Degenhardt (2010).

lying on their knees (see Figure 22.2). In this case, the visual conduct is not only accompanying turn-taking (1 in Table 22.1); it helps to specify what the speaker is referring to while speaking – a specification that is made visible by his glance and the pause (2 in Table 22.1). The verbal interaction consists of a dual unit separated by a short pause (3–5 in Table 22.1). The first part is pronounced upwards – announcing the ‘surprise’, whereas the second part prosodically turns down so as to finish the turn and open the floor for the next turn. In conversation analysis, this would be called a transition relevance place (Sacks et al., 1974). Again, our analysis does not need to depend on prosodic knowledge of how the speaker inflects his voice, for the interpretation is supported sequentially: the speaker (while pronouncing ‘a surprise item’) raises his head and looks at the audience. What appears as a pause in the verbal transcript is

full of action, if we look at the video. The audience camera shows that immediately after the transition point a number of audience members raise their hands. As with the telephone ring, the raising of the auctioneer’s head can be understood as him opening the floor for the audience. The ‘reaction’ (raising of hands) of the audience, on the other hand, can be considered as a reflexive interpretation of the action by the speaker. This interpretation, again, is ratified in the next turn, for the speaker hints with his arm (as shown in Figure 22.2) towards certain actors in the audience who are raising their hands (not visible in still, highlighted with text). In the next turn he ‘interprets’ their action by identifying it with a number – ‘Five are offered’ – in response to an understanding that raising the hand is an offer. We do not want to indulge into more intricate questions, such as to how prices are fixed and negotiated in such



Figure 22.2 Split screen (auctioneer/audience) taken from auction sequence corresponding to Line 7 in the transcript reported in Table 22.1. Sourced from a video by Degenhardt (see also Degenhardt, 2010).

settings.¹⁶ Rather, we point out that the turn following the speaker's initiation is a purely visual 'unit' of action (looking up, raising the hand), and that this visual 'unit' is produced as such by the next turn of the speaker identifying it as a bid verbally and hinting in a certain direction. The visuality of these turns is emphasized if we acknowledge that none of these sequences could have been discovered from an audio recording or transcript alone. Careful study of the video not only allows the researcher to become familiar with the transcript and the spoken words but also is the concrete basis for visual observations of the structure of the sequences. Moreover, the production of the visual transcript makes understanding the speech acts easier. Familiarity with the words spoken makes it easier to identify the sequence of events and 'locate' the visual events with respect to the text.

The video sequence must be watched repeatedly to discover the order of acoustic and visual events and to identify when things are done or said.

One may argue that this could be accomplished without the transcript, but experience proves that the written transcript is very helpful as a location device when identifying the sequences. (If there are no verbal utterances, other temporal actions, such as gazing, may be transcribed into more detail.) The study of sequences thus moves from the transcribed (written) to include the visual. Hence, the visual is not treated only as an addendum to the spoken. Rather, in watching the video recordings, one frequently discovers additional sequences which allow the researcher to make sense of prior turns or sequences.

The previous example demonstrates the sequential function of 'visual conduct' since the audience only 'reacts' in a visual way (which is relevant to both the observer as well as to the speaker). One may take the 'raising' of the hand as a pure sign out of a more or less systematic sign system that is extra-situational.¹⁷ But then again, raising the hand is neither purely a reaction nor a 'sign', but is rather synchronized with the speaker's

action to make sense: the hands are only raised when the speaker turns to the audience. The audience 'reads' the speaker's body, particularly his face direction, as an indication of his focus of attention. Moreover, the indication by the speaker functions both as recognition and as a legally binding identification of the person bidding; it also opens the floor to other bidders and suggests where they are sitting. One could say that the speaker's body works as a 'mirror' to the audience, since he not only talks and turns to the audience but also indicates to the members of the audience if there are other persons acting and where they are.

Visual conduct cannot be considered as one 'mode' alongside others; the meaning of an act is constituted by visual conduct in addition to verbal utterances. In the next example we compare actions in which the visual is considered to be meaningful on its own. It is not by accident that we turn at this point to the example of PowerPoint presentations. Although there is debate concerning how much 'information' the slides of PowerPoint presentations should contain, there is no doubt that text and diagrams of the slides projected to a live audience carry meaning on the grounds of their visual features. PowerPoint presentations with speakers showing slides to an audience are, therefore, a useful example for analyzing how the visual enters social interaction. In the following fragment, which forms part of a more extended study of PowerPoint presentations,¹⁸ the speaker confronts an audience with an elaborate diagrammatic structure on his slides. While he shows the slides, he speaks the words presented in Table 22.2 (to about 50 persons in a high-profile scientific event):

The speaker starts by turning to the screen and by leaving a pause (1 in Table 22.2) – as he did quite frequently before in his speech. The new slide he refers to in this fragment opens a new topic, which is identified by the pseudonym 'lakedemons' (our pseudonym used here for a biochemical structure analyzed by his research group). As the video

Table 22.2 Visual example. Underlines indicate the time of pointing to the slides. The Gesprächsanalytisches Transkriptionssystem (GAT) transcription conventions are used (Selting et al., 1998)

Sequence	Original German	Translation in English
1	(1.0) die Auswanderung von Lakedämonen <u>aus dem Gefäß in das Gewebe</u>	(1.0) The migration of lakedemons <u>out of the vessels into the texture</u>
2	ist=relativ=gut=untersucht, man weiß die Flakomeuten brauchen=zuerst n	has=been=studied=quite well, one knows that the flamokeuts first=need 'n
3	<u>initialen Kontakt</u> , (hat ä / geringe affine) Anlagerung;	<u>Initial contact</u> , (has a reduced affine) adaption;
4	<u>des Rollen verstärkt</u> dann den Kreisbewegung Kontakt=diese=feste=Anlagerung	<u>rolling motions</u> then support the circular movements contact=this=fixed
5	<u>wandert dann raus</u> ; und wird sehr viel über <u>die</u> <u>Moleküle in</u>	<u>then migrates</u> ; and then <u>via the molecules it will</u>
6	<u>Interaktion von Lakedämonen mit</u> <u>Europolzellen</u> steuern	control <u>the interaction of lakedemons the flacocyte cells</u>

Source: Knoblauch (2008)

Ed: Pls
confirm
these
are ok?

demonstrates, he does not simply show the slide; he also relates to the slides by pointing to them with a laser pointer. Also, his introduction of the new topic is accompanied by discursive gestures illustrating the direction of movements of the 'lakedemons' ('aus' or 'out', 'in' or 'in', line 1 in Table 22.2) by two hand gestures turning outwards and turning inwards. In addition, he uses the laser pointer as a way to direct attention to features on the slide. This use of the laser pointer not only complements the structure of the talk (the time of pointing is shown by underlines in the transcript in Table 22.2); by moving the laser pointer in various ways, he also transforms the meaning of the diagram. Let us have a look at the ways in which he moves the laser pointer¹⁹ (see Figure 22.3).

As the reconstruction of the laser pointer's movements demonstrate, the speaker does not merely point at the 'lakedemons'; the movements are represented by little circles on the slide. Rather, he moves the pointer in different ways on the surface of the slides. In relating these movements to what he is speaking about, it becomes quite clear that these pointing movements and their differences are quite meaningful: the 'initial contact' (line 3 in Table 22.2) of the biochemical structures mentioned with the 'Europol-cells'

(pseudonym) are underlined by a wave-like movement, their 'rolling' by a circular movement (line 4), their 'migration' by sudden straight movements (line 5) and their interaction by loop-shaped movements (line 5f). The movements of the laser pointer seem to mimic the movements of the microscopic objects talked about. In doing so, the moving laser pointer 'illustrates' what is being said or shown. The difference between these movements corresponds to different parts of the talk ('first', 'then', 'will then').

This has serious consequences. Keeping in mind that we are talking about one slide which stays the same during these sequences, the laser pointer movements contribute to transforming the static slide into a dynamic movement. What appears as a fixed structure on the slide's diagram is thus turned into a temporal sequence of processes characterized by the movements. Of course, this meaning is accomplished by the interplay of spoken words, graphics and gestures.²⁰ It results from the speaker's sequentially fine-grained coordination, timing, ways of pointing at the slide diagram and utterances. Nonetheless, the sequence of different movements of the laser pointer takes the form of different units and thus contributes to the construction of a particular sequence for what is being talked about.

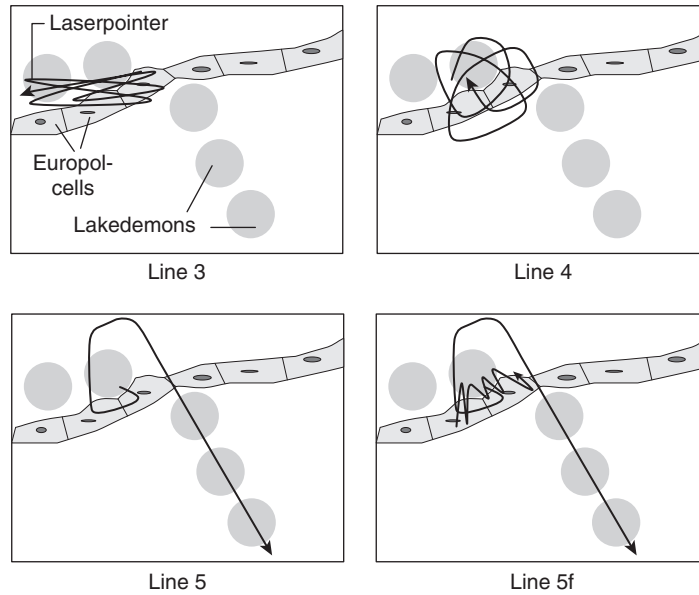


Figure 22.3 Reconstruction of the movement of the laser pointer: © Knoblauch (2008).

As this example shows, sequentiality does not only constitute a resource for the interpretation of interaction. It may also be regarded as a criterion to validate an interpretation. Or, following Schegloff (1992), what is of relevance for the analyst must be shown to be of relevance to the actors. Methodologically, it is the sequential embedding in prior and later turns that can be considered as a decisive indication of the relevance for certain visual aspects of an interactive turn.²¹ For the validation of interpretations, it is therefore critically relevant to answer detailed questions like: Do actors look at the screen before they push a button? Does a phone ring before person A walks away? Does person B point at the slide after she said X or when she says Y? Video analysis thus makes it possible to examine the sequential organization of social and human technological interactions (Heath, 1997: 184). Or, as Goodwin suggests:

Rather than wandering onto field sites as disinterested observers, attempting the impossible task of trying to catalog everything in the setting, we can

use the visible orientation of the participants as a spotlight to show us just those features of context that we have to come to terms with if we are to adequately describe the organization of their action (Goodwin 2000: 1508f).

THE VISUAL, THE SITUATION AND ETHNOGRAPHY

Sequentiality provides an essential resource for the interpretation of interaction, but it is not the only resource. Within a situation and during the time span recorded, there may just be nothing happening. Even during a series of actions happening sequentially one after another (or diachronically), there are always visual elements, such as desks, doors or projectors, which do not change at all. This may also hold for parts of the decorum of actors (such as clothing, glasses or hairstyle). These synchronic elements appear permanent on the visual representation, but it would be an error to consider them as mere 'background'.

On the contrary, as the examples in this chapter have shown, single slide representations or the location of persons in an audience may become relevant in certain situations. Goffman has suggested a helpful distinction that allows us to grasp these two aspects. Following his suggestion we can use the term *situated* for anything that occurs within the situation and ‘could not occur outside the situation’ (Goffman, 1961: 22): i.e. the *situative* character of social action (Suchman, 1987).²² *Situational*, on the other hand, are those elements of the situation that can also be present in other situations. In this sense, what is coordinated sequentially in the situation may be considered as essentially *situated*, whereas those aspects that can occur in other situations are *situational*. The distinction between *situative* and *situational* allows us to turn to the question how video may help address the synchronic aspects of the visual and *situational* elements.

One way to approach these synchronic aspects of the *situational* in video analysis has been suggested by Goodwin (2000). He assumed that semiotics might help comprehend these visual features. Thus, talk is embedded in multiple sign systems (such as graphic codes, gestures and other features of the environment). To Goodwin, actors orient to what he called ‘*semiotic fields*’, including different kinds of signs and their expressions in different kind of media. In accordance with the principle of relevance mentioned above, these *semiotic fields* may be of local relevance in that the actors demonstrably orient towards them in the situation of action. Goodwin (2000) calls this ‘*contextual configuration*’.

There is no doubt that the concept of ‘*semiotic field*’ is very useful in those cases where actors use well-developed sign systems (such as the alphabet, professional knowledge systems or even the sign elements of children playing hopscotch). We doubt, however, that semiotics remedies the problem of *visuality* in general, since structural semiotics presupposes that the signs visible are organized in a more or less systematic

way.²³ Even if one admits that words form part of a system, one may doubt that gestures can be considered as forming a sign system, let alone other visual elements (such as expressions, clothing, furniture, the order of things in space, etc.²⁴). On top of this, the meanings of visual elements are often not fixed in such a way as to be shared by any individual. Even the specialized knowledge of professionals is mostly fixed in a *de-contextualized* way and needs *re-contextualization* by ‘*situated action*’ (Suchman, 1987).

In order to grasp specialized and *situated* visual knowledge, ethnography again provides a useful set of methods. In order to clarify the meaning and signification of visual elements of the recorded data, one may, for example, use elicitation procedures. Showing the videos to the persons who have been acting in the setting and interviewing them simultaneously can retrieve their intentions, their understandings and their knowledge.²⁵ Also ‘*auto-confrontation*’, ‘*auto-ethnography*’ (of the participant observer) or ‘*video-based interviewing*’ can be useful methods to uncover various *situational* aspects of objects and events (Bayart et al., 1997).²⁶ For example, Schubert (2009) exposed medical personnel he had been video recording to recordings of their work in operating rooms, asking them to clarify actions they were involved in. These methods help researchers to reconstruct actors’ perceptions and orientations in the situations at hand. Since actors may also be able to explicate the functioning and significance of other items visible in the scene, the methods also contribute to the interpretation of the actions observed.

The importance of background knowledge elucidating visual aspects of the recordings recapitulates the importance of ethnography for video analysis referred to in the introduction. For it is by way of observation, interviews, expert interviews, etc., that researchers become familiar with the settings which are recorded by video. One must acquire local knowledge in order to understand the *audio-visual sequences*: i.e. the series of actions,

reactions and interactions (which, in many cases, involves intricate technologies). In this way, videography still incorporates the actor's perspective, yet in a specific material sense: we examine visible situations, activities and actions. This does not mean that one needs to reconstruct the whole stock of knowledge necessary to comprehend the whole field. Rather, the task of the researcher is to acquire those elements of (partly embodied) knowledge relevant to the activity on which the study is focusing. In studying technological activities, for example, only those elements of knowledge need be considered which are relevant to understanding practices involved in handling the technology under scrutiny.

As an essential part of videography, ethnography not only serves to clarify the details of interaction; by uncovering relevant knowledge, ethnography also opens the way to relate these interactions to wider social contexts and more conventional sociological concerns. Indeed, contexts such as social welfare agencies, underground stations or management offices have always been the subject of ethnography. Videography – the interpretive video analysis of social interaction – contributes to these concerns not only because it provides empirically the most reliable and 'hard' data for analysis but also because it allows reconstruction of the social construction of these contexts.

NOTES

1 Muybridge is famous for his early photographic studies of animal and human movement (Muybridge, 1901); in anthropology, Bateson and Mead (1942) used visual media for their study on the 'Balinese Character' before Birdwhistell (1952) analyzed everyday movements like smoking with the help of film. See also the film-based studies by Gregory Bateson, Frieda Fromm-Reichmann, Paul Watzlawick et al. (Bateson, 1958) at Palo Alto and the studies by Ekman and Friesen (1969a, 1969b) and Schefflen (1965) for the development of film in the social sciences.

2 The authors only analyze one type of institution (education) with wide yet not universal institutionalized standards of interaction.

3 For the general notions of knowledge, institutions and the structure of the stock of knowledge, see Berger and Luckmann (Berger and Luckmann, [1966]1991; Goodwin, 2000: 1508f).

4 For this reason, in the 1980s Corsaro (1981) proposed that there should be no video recording and analysis without completing prior ethnography – a procedure Albrecht (1985: 328f) already called 'scouting'. Also Cicourel (1992) has stressed that this focusing process presupposes prior knowledge of and prior familiarity with the field and, therefore, prior ethnography.

5 At this point, the first author draws on papers on focused ethnography, first published in German (Knoblauch, 2001b) and in English (Knoblauch, 2005).

6 Ethnography also has to clarify to what degree the use of video is 'participatory', proactive, contributing to the actions observed (Suchman and Trigg, 1991; Laurier and Philo, 2006) or 'reactive', which refers to affecting and manipulating the situation being taped. Although the technology may be obtrusive and even obstructive to the situation, there are situations in which video recording may be less distorting than the presence of an (overt or covert) observer (see, for example, vom Lehn and Heath, 2006). The use of video technology, however, may also relieve the researcher from other tasks and allow for ethnographical observations, questions and reflections while making the video records. Since the data collection is supported technically, researchers have more time to observe specific features or to enquire into aspects of the already-focused field.

7 By audio-visual conduct, we do not mean to refer to a specific sort of datum constructed by the technical medium of video recording.

8 See Heath et al. (2010: 37–60) for a very useful introduction on how to collect data in the field.

9 Cf. the methodology of the refined version of grounded theory (Glaser and Strauss, 1967; Strauss and Corbin, 1998)

10 Sophisticated software (such as Dartfish©) increasingly allows one to create visual code books with video samples and to simply insert drawings and highlight relevant movements directly in the video. NVivo™, ATLAS.ti™, HyperRESEARCH™ and Transana© allow one to code video 'on the fly'. See Chapter 27 by Raewyn Bassett in this handbook.

11 In conversation analysis, this feature of communication has been studied by Schegloff and Sacks (Schegloff, 1968; Schegloff and Sacks, 1973).

12 For a more detailed account of these theoretical assumptions, see Knoblauch (2001a).

13 This notion of reflexivity differs from notions of reflexivity that address the presentation of research (cf. Ruby, 2000).

14 Structural hermeneutics is one way to tackle the indefiniteness of synchronous visual references.

Basically it starts from the assumption that audio-visual data encompasses meaning patterns that reflect the meaning of the situations represented. It proceeds by producing as many competing interpretations of single frames as possible (in interpretation groups), which are reduced by the interpretations triggered by consecutive frames. Structural hermeneutics assumes that these meanings are reduced step by step to what becomes the 'objective meaning' (cf. Soeffner, 1997).

15 We would like to thank Felix Degenhardt (2010) for contributing this data fragment and detailed transcription. The Gesprächsanalytisches Transkriptionssystem (GAT) transcription conventions were used (Selting et al., 1998).

16 Heath and Luff (2007).

17 In fact, even the kinds of signs audience members use at auctions differ substantially, ranging from very subtle nods, short waves of the hand, to the raising of numbered signs or phone calls – the differences are not only dependent on the size of the audience but also on the 'milieu' of the auction and its formality.

18 For a more detailed analysis and transcription conventions, see Knoblauch (2008).

19 We would like to thank Marion Mackert and Sabine Petschke for the detailed reconstruction of those movements.

20 Schnettler (2006: 157) compares this interplay to 'orchestration'.

21 As seen, the turn may be observed only visually.

22 Of course video data are reducing the situated again to some aspects of the situation, but which should however cover the relevant aspects of it.

23 See Chapter 16 by Winfried Nöth on visual semiotics in this handbook.

24 Hodge and Kress (1988).

25 See Chapter 11 by Francesco Lapenta on photo-elicitation in this handbook.

26 The method dates back, of course, to Jean Rouch (Jackson, 2004).

REFERENCES

- Albrecht, G. L. (1985) 'Videotape safaris: Entering the field with the camera', *Qualitative Sociology*, 8(4): 325–344.
- Bateson, G. (1958) 'Language and psychotherapy: Frieda Fromm-Reichmann's last project', *Psychiatry*, 21(21): 96–100.
- Bateson, G. and Mead, M. (1942) *Balinese Character. A Photographic Analysis*. New York: New York Academy of Sciences.
- Bayart, D., Borzeix, A. and Lacoste, M. (1997) 'Les traversées de la gare: Filmer des activités itinerantes', *Champs visuels*, 6: 75–90.
- Berger, P. L. and Luckmann, T. [1966] (1991) *The Social Construction of Reality. A Treatise in the Sociology of Knowledge*. New York: Penguin.
- Birdwhistell, R. L. (1952) *Introduction to Kinesics. An Annotation System for the Analysis of Body Motion and Gesture*. Louisville, KY: University of Louisville.
- Cicourel, A. V. (1992) 'The interpretation of communicative contexts: Examples from medical encounters', in A. Duranti and C. Goodwin (eds), *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press. pp. 291–310.
- Corsaro, W. A. (1981) 'Something old and something new. The importance of prior ethnography in the collection and analysis of audiovisual data', *Sociological Methods and Research*, 11(2): 145–166.
- Degenhardt, F. (2010) 'Auktion, Akteurskonstellation und Struktur'. Unpublished manuscript, Technical University Berlin.
- Dinkelaker, J. and Herrle, M. (2009) *Erziehungswissenschaftliche Videographie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ekman, P. and Friesen, W. (1969a) 'The repertoire of nonverbal behavior: Categories, origins, usage and coding', *Semiotica*, 1: 63–68.
- Ekman, P. and Friesen, W. (1969b) 'A tool for the analysis of motion picture film or videotapes', *American Psychologist*, 24(3): 240–243.
- Erickson, F. (1988) 'Ethnographic description', in U. Ammon (ed.), *Sociolinguistics. An International Handbook of the Science of Language and Society*. Berlin, New York: de Gruyter. pp. 1081–1095.
- Giddens, A. (1984) *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity.
- Glaser, B. G. and Strauss, A. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine.
- Goffman, E. (1961) 'Behavior in public places. Notes on the social organization of gatherings'. Unpublished manuscript, New York.
- Goodwin, C. (2000) 'Action and embodiment within situated human interaction', *Journal of Pragmatics*, 32: 1489–1522.
- Heath, C. (1997) 'The analysis of activities in face to face interaction using video', in D. Silverman (ed.), *Qualitative Research. Theory, Method, and Practice*. London: Sage. pp. 183–200.

- Heath, C. and Hindmarsh, J. (2002) 'Analysing interaction: Video, ethnography and situated conduct', in M. Tim (ed.), *Qualitative Research in Action*. London: Sage. pp. 99–121.
- Heath, C. and Luff, P. (1996) 'Convergent activities: Line control and passenger information on the London Underground', in Y. Engeström and D. Middleton (eds), *Cognition and Communication at Work*. Cambridge: Cambridge University Press. pp. 96–129.
- Heath, C. and Luff, P. (2000) *Technology in Action*. Cambridge: Cambridge University Press.
- Heath, C. and Luff, P. (2007) 'Ordering competition: The interactional accomplishment of the sale of art and antiques at auction', *British Journal of Sociology*, 58(1): 63–85.
- Heath, C., Hindmarsh, J. and Luff, P. (2010) *Video in Qualitative Research: Analysing Social Interaction in Everyday Life*. London: Sage.
- Hodge, R. and Kress, G. (1988) *Social Semiotics*. Ithaca, New York: Cornell University Press.
- Hughes, J. A., King, V., Rodden, T. and Anderson, H. (1994) 'Moving out of the control room: Ethnography in system design', in R. Futura and C. Neuirth (eds), *Transcending Boundaries. Proceedings of the ACM Conference on Computer Supported Cooperative Work*. Chapel Hill, North Carolina. pp. 429–439.
- Jackson, J. (2004) 'An ethnographic Filmflam: Giving gifts, doing research, and videotaping the native subject/object', *American Anthropologist*, 106(1): 32–42.
- Kendon, A. (2004) *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Knoblauch, H. (2001a) 'Communication, contexts and culture. A communicative constructivist approach to intercultural communication', in A. di Luzio, S. Günthner and F. Orletti (eds), *Culture in Communication. Analyses of Intercultural Situations*. Amsterdam, Philadelphia: John Benjamins. pp. 3–33.
- Knoblauch, H. (2001b) 'Fokussierte Ethnographie', *sozialer sinn*, (1): 123–141.
- Knoblauch, H. (2005). 'Focused ethnography', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(3). [Online]. Available from: <http://nbn-resolving.de/urn:nbn:de:0114-fqs0503440> [Accessed: 10 October 2010].
- Knoblauch, H. (2008) 'The performance of knowledge: Pointing and knowledge in PowerPoint presentations', *Cultural Sociology*, 2(1): 75–97.
- Koch, S. C. and Zumbach, J. (2002) 'The use of video analysis software in behavior observation research: Interaction patterns of task-oriented small groups', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2). [Online]. Available from: <http://nbn-resolving.de/urn:nbn:de:0114-fqs0202187> [Accessed: 10 October 2010].
- Laurier, E. and Philo, C. (2006) 'Natural problems of naturalistic video data', in H. Knoblauch, B. Schnettler, J. Raab and H.-G. Soeffner (eds), *Video Analysis. Methodology and Methods. Qualitative Audiovisual Analysis in Sociology*. Frankfurt am Main, New York: Lang. pp. 183–192.
- Mittenecker, E. (1987) *Video in der Psychologie. Methoden und Anwendungsbeispiele in Forschung und Praxis*. Bern: Huber.
- Mohn, E. (2007) 'Kamera-Ethnografie: Vom Blickentwurf zur Denkbewegung', in G. Brandstetter and G. Klein (eds), *Methoden der Tanzwissenschaft. Modellanalysen zu Pina Bauschs 'Sacre du Printemps'*. Bielefeld: transcript. pp. 173–194.
- Mondada, L. (2006) 'Video recording as the reflexive preservation and configuration of phenomenal features for analysis', in H. Knoblauch, B. Schnettler, J. Raab and H.-G. Soeffner (eds), *Video Analysis. Methodology and Methods. Qualitative Audiovisual Analysis in Sociology*. Frankfurt am Main, New York: Lang. pp. 51–67.
- Muybridge, E. (1901) *The Human Figure in Motion: An Electro Photographic Investigation of Consecutive Phases in Muscular Actions*. London: Chapman and Hall.
- Raab, J. and Tänzler, D. (2006) 'Video-Hermeneutics', in H. Knoblauch, B. Schnettler, J. Raab and H.-G. Soeffner (eds), *Video Analysis. Methodology and Methods. Qualitative Audiovisual Analysis in Sociology*. Frankfurt am Main, New York: Lang. pp. 85–97.
- Rammert, W. and Schulz-Schaeffer, I. (2002) 'Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt', in W. Rammert and I. Schulz-Schaeffer (eds), *Können Maschinen denken? Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Frankfurt am Main: Campus. pp. 11–64.
- Ruby, J. (2000) *Picturing Culture: Explorations of Film and Anthropology*. Chicago, IL: UCP.
- Sacks, H., Schegloff, E. A. and Jefferson, G. (1974) 'A simplest systematics for the organization of turn-taking for conversation', *Language*, 50(4): 696–735.
- Schefflen, A. E. (1965) 'The significance of posture in communication systems', *Psychiatry*, 27: 316–331.
- Schegloff, E. (1968) 'Sequencing in conversational openings', *American Anthropologist*, 70: 1075–1095.

- Schegloff, E. (1992) 'On talk and its institutional occasions' in P. Drew and J. Heritage (eds), *Talk at Work. Interaction in Institutional Settings*. Cambridge: Cambridge University Press. pp. 101–136.
- Schegloff, E. and Sacks, H. (1973) 'Opening up closings', *Semiotica* (8): 289–327.
- Schnettler, B. (2006) 'Orchestrating bullet lists and commentaries. A video performance analysis of computer supported presentations', in H. Knoblauch, B. Schnettler, J. Raab and H.-G. Soeffner (eds), *Video Analysis. Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*. Frankfurt am Main, New York: Lang. pp. 155–168.
- Schubert, C. (2006) 'Video-analysis of practice and the practice of video-analysis', in H. Knoblauch, B. Schnettler, J. Raab and H.-G. Soeffner (eds), *Video Analysis. Methodology and Methods. Qualitative Audiovisual Analysis in Sociology*. Frankfurt am Main, New York: Lang. pp. 115–126.
- Schubert, C. (2009) 'Videographic elicitation interviews. Studying technologies, Practices and narratives in organisations', in U. T. Kissmann (ed.), *Video Interaction Analysis*. Berlin, New York: Lang. pp. 199–220.
- Schutz, A. (1962) 'Common sense and scientific interpretation of human action', in Maurice Natanson (ed.), *Collected Papers I: The Problem of Social Reality*. The Hague: Nijhoff. pp. 3–47.
- Schütz, A. (1971) *Gesammelte Aufsätze I. Das Problem der Sozialen Wirklichkeit*. Den Haag: Nijhoff.
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S. and Meier, C. (1998) 'Gesprächsanalytisches Transkriptionssystem, (GAT)', *Linguistische Berichte*, (173): 91–122.
- Soeffner, H.-G. (1997) *The Order of Rituals. The Interpretation of Everyday Life*. New Brunswick: Transaction.
- Strauss, A. and Corbin, J. (1998) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn. London: Sage.
- Suchman, L. (1987) *Plans and Situated Actions. The Problem of Human–Machine Communication*. Cambridge: Cambridge University Press.
- Suchman, L. and Trigg, R. H. (1991) 'Understanding practice: Video as a medium for reflection and design', in J. Greenbaum and M. Kyng (eds), *Design at Work. Cooperative Design of Computer Systems*. Hillsdale, NJ: Lawrence Erlbaum. pp. 65–89.
- vom Lehn, D. and Heath, C. (2006) 'Discovering exhibits: Video-based studies of interaction in museums and science centres', in H. Knoblauch, B. Schnettler, J. Raab and H. –G. Soeffner (eds), *Video Analysis. Methodology and Methods. Qualitative Audiovisual Analysis in Sociology*. Frankfurt am Main, New York: Lang. pp. 101–113.
- Weber, M. [1921] (1984) *Soziologische Grundbegriffe*. Tübingen: Mohr.