

Situated Organization of Video-Mediated Interaction: A Review of Ethnomethodological and Conversation Analytic Studies

JAKUB MLYNÁŘ^{1,*}, ESTHER GONZÁLEZ-MARTÍNEZ² AND DENIS LALANNE¹

¹*Human-IST Research Center, University of Fribourg, Boulevard de Pérolles 90, 1700 Fribourg, Switzerland*

²*Department of Social Sciences, Faculty of Arts, University of Fribourg, Boulevard de Pérolles 90, 1700 Fribourg, Switzerland*

*Corresponding author: jakub.mlynar@unifr.ch

Video-based communication has become a common way of interacting with remote interlocutors, whether through complex videoconferencing systems or webcams integrated into consumer technologies. Ethnomethodology and conversation analysis (EM/CA) are sociological approaches that have been influential in Human-Computer Interaction for nearly three decades due to their focus on the situated organization of practical activities. In this article, we present a state-of-the-art review of empirical research on video-mediated social interaction studied from the perspective of EM/CA. We put forward an original organization of the findings on the interplay of talk, bodily behavior and spatial and material resources. The review underscores the ways in which technology enables and constrains interaction, shaping familiar and novel social activities. We also propose directions for future research and systems design.

RESEARCH HIGHLIGHTS

- Video-mediated interaction has become ubiquitous due to dedicated technologies and devices.
- Ethnomethodology and conversation analysis have produced a number of studies on participants' practices when involved in remote synchronous video communication.
- The reviewed studies provide detailed descriptions of users' multimodal behavior when interacting with/through the technology.
- This research is moving forward, simultaneously helping to anticipate future needs and find design solutions, while concentrating on the study of evolving technologies-in-practice.

Keywords: 1.1.1.2 User studies; 1.3.2.2 Ethnographic studies; 1.3.4 Empirical studies in collaborative and social computing; 1.4.3 Ubiquitous and mobile devices; 3.2.1 Graphics systems and interfaces; 6.2.4 Sociology

Editorial Board Member: Dr. Regina Bernhaupt

Received 12 April 2017; Revised 20 August 2017; Editorial Decision 6 December 2017; Accepted 11 December 2017

1. INTRODUCTION

It is March 2017, and Robert Kelly, a professor of international relations and expert on South Korea, is being interviewed live on BBC News via an online video-call from his home office. Suddenly the door opens and a little girl marches

into the room, followed soon after by a toddler. Kelly struggles to continue the interview, and several seconds later his wife shows up and hastily drags both children out of the room. The short clip was shared by hundreds of thousands of people on social networks and became a popular item on

major news channels worldwide. In an interview following their unexpected virtual fame, the couple explained that Kelly's wife, Jung-a Kim, was watching the live interview on TV in another room, when she suddenly saw—with a delay of a few seconds—their children Marion and James on screen in her husband's home office, and acted immediately to save the situation (Johnston, 2017).

As this story shows, we are spending an increasingly large part of our lives surrounded by screens and cameras. The TV practices of remote interviewing and on-site reporting, available only to professionals for many decades, have become ubiquitous in everyday life. Whether with complex videoconferencing systems or webcams integrated into everyday technologies, video-based communication has become a common way of interacting with remote interlocutors. In this paper, we examine practices of participants involved in video-mediated interaction and their connection with the features of the technologies used. More specifically, we provide a state-of-the-art review of studies on video-mediated social interaction conducted according to two interrelated sociological approaches: ethnomethodology and conversation analysis (EM/CA).

EM/CA emerged in the 1950s and 1960s, originally based on the work of Harold Garfinkel (1967) and Harvey Sacks (1992). It investigates the methodical work that participants produce in order to accomplish the ordinary scenes of everyday life (such as common activities and situations). EM/CA research describes the social organization of practical activities in detail, as they unfold *in situ* and in real time. Researchers adopt a naturalistic approach consisting of observation and audio/video-recording of social conduct in its natural settings of occurrence. Both approaches are closely related in their historical development and epistemological presuppositions, but what distinguishes them is that EM studies the practical achievement of the organized and intelligible character of social phenomena (Lynch, 1993), while CA specifically focuses on identifying generic orders of organization of *talk-in-interaction* (Schegloff, 2007).

Over the past three decades, EM/CA has been influential in studies of technology-supported and technology-mediated social interaction (e.g. Button and Dourish, 1996; Dourish and Button, 1998; Dourish, 2001; Heath and Luff, 2000; Suchman, 2011). As early as the 1970s, Garfinkel formulated the program of 'hybrid studies of work', which substantially influenced the field of social studies of science and technology (e.g. Garfinkel, Lynch, and Livingston, 1981). Later on, following the ground-breaking study of Suchman (1987), EM/CA impacted the areas of Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW). Suchman's study devised a fruitful alternative to the dominant cognitivist understanding of action as derived from plans by focusing on the grounding of action in the contingencies of practical situations. According to Matthews (2013), one of the reasons for EM/CA's influence is the insight gained on the context of use and user practices and its utility for technology

design. Today, EM/CA continues to inform HCI, while the field provides opportunities for re-examination of EM/CA's core assumptions and earlier empirical findings in novel settings.

In this paper, we focus on EM/CA studies of situations in which participants use a specific type of technology: video-based synchronous communication systems. With these technologies, participants have mutual access to sound and image in real time, which may provide an 'illusion' (Fornel, 1996) or 'simulacra' (Rintel, 2013a) of unmediated face-to-face interaction. Our intention is not simply to recapitulate the state of the field, but also to propose an organization of the findings that is conducive to new advancements both in terms of analysis and practical applications. This organization presents participants' practices according to the main temporal phases of a video-mediated encounter: setting up, opening, maintaining/acting and closing. In the first part of the article, we summarize and discuss the main findings presented in the literature, providing an answer to the question: What have we learned so far from EM/CA studies about interactional practices in video-mediated social interaction? In the second part, we outline implications for technology design and propose directions for further research.

2. CORPUS OF LITERATURE

Video communication has progressed through several defining moments such as the introduction of the Picturephone in the 1960s (Noll, 1992), the public art installation 'Hole in Space' in 1980, which connected pedestrians in Los Angeles and New York City (Relieu, 2007), and 'media spaces' created to enhance remote collaboration in large companies (Harrison, 2009). EM/CA studies of video-mediated interaction do not show a continuous development. The first wave of research came in the early 1990s, as videophone technologies attempted an initial entry into households and more commonly into workplaces, such as videoconferencing in business meetings. Hutchby (2001) noted several years later that despite the structural support in formal organizations, these technologies have been quite slow to take off. There was a resurgence of research interest in the early 2010s, when video-mediating technologies became part of ubiquitous computational devices such as laptops, tablets and smartphones. After nearly three decades of research, the time seems right for thorough reconsideration of the progress made so far, especially since the field is growing quickly, reflecting the expansion in use of the technologies (Arminen, Licoppe, and Spagnolli, 2016; Vasilyeva, 2013; Velkovska, 2014).

Several sociological approaches, such as ethnography and Goffmanian sociology, are related to and cross-fertilize with EM/CA research while providing relevant insight on video-mediated interaction (e.g. Bernhaupt *et al.*, 2008; Carter and Mankoff, 2005; Crabtree *et al.*, 2003; Haddon, 2006; Rettie,

2009). We have nevertheless limited this review not only in terms of this specific subject, but also the specific approach, in order to ensure consistency regarding the scientific perspective and to have a chance to accurately present an already large body of literature in the limited space of a journal article. Even so, we can only point to the reviewed studies since they are hallmarked by minute descriptions of social micro-practices and furthermore focus on a diverse array of activities and technological settings.

The literature review is based on systematic searches in *ACM Library*, *Directory of Open Access Journals*, *Google Scholar*, *JSTOR*, *Science Direct*, *Scopus* and *Web of Science*. We looked for explicit references to EM/CA ('ethnomethod*' and 'conversation analy*') in co-occurrence with 'screens', 'displays', 'monitors', 'video-based communication', 'video-mediated interaction' and other combinations of relevant words. Furthermore, personal and institutional websites related to research in the field (such as *EM/CA Wiki*) were searched manually. Implementing the 'chaining technique', we also used the reference lists from the retrieved articles to find further literature and identify the most influential work. A significant part of the reviewed literature originates in French academic institutions and corporate research departments (cf. [Licoppe and Relieu, 2007](#), and the issue of *Réseaux* that they introduce); we selected the most relevant texts, opting for the ones presenting the results in English. Ultimately, we selected 63 studies for summary and discussion. These included 50 journal articles, nine book chapters and four conference proceedings.

3. PRACTICES OF VIDEO-MEDIATED INTERACTION

Research in technology-mediated interaction often takes 'unmediated' or 'face-to-face' interaction as its background. But this may not necessarily be the most appropriate way of approaching the issue, and '[m]oving away from this perspective allows us to explore a number of important, intrinsic properties of video as a communicative medium in its own right' ([Dourish et al., 1996](#), p. 34).

The EM/CA perspective follows participants' orientations and examines which features of video-mediated settings are relevant to them at any given moment, if any. The studies show 'how the characteristic opportunities for (and constraints on) actions of what we would intuitively call 'mediating technologies' accountably shape the interaction practices available or observed' ([Arminen, Licoppe, and Spagnolli, 2016](#), p. 292). [Fornel \(1996\)](#) provides an example of such orientations in his observations of participants ironically attempting practices that are impossible in video-mediated interaction, such as shaking hands or offering their interlocutor a cigarette or a piece of chocolate.

EM/CA research documents such practices and at the same time informs us about their change over time; for

instance, the above-mentioned practices may disappear as users become accustomed to the technology. It also shows the diversity of situations in which the technology has been used to date (see Fig. 1), from the street to hi-tech medical environments, from households to courtrooms. It underscores the variety of technologies involved, from videophones to smart meeting rooms, from customer applications to tailor-made professional solutions. In terms of methodology, the studies contribute ways of recruiting participants, dealing with legal and ethical considerations and collecting and presenting multimodal data that are adapted to the specificity of the studied phenomena. The following literature review, while focusing on participants' practices, also aims to report on these aspects.

3.1. Setting up, opening and closing

When initiating a video-mediated interaction, participants go through a series of phases involving distinctive practices. [Mondada \(2015\)](#) identifies and distinguishes pre-opening, opening and beginning phases in studying medical meetings. [Licoppe \(2015\)](#) analyzes the initiation of everyday-life video-mediated interactions as an approach—participants coming progressively closer to each other—during which participants show up (appear) in different modalities.



Figure 1. The diversity of video-mediated interaction—from private settings (middle left, bottom) to group meetings (top) and hi-tech professional environments (middle right).

Prior to launching the connection, participants adjust their bodily appearance, position and physical environment, for instance moving furniture around (Fornel, 1996; Pappas and Seale, 2009; Ruhleder and Jordan, 2001b). This also goes on in the pre-opening phase as they launch and try out the connection (Ibnelkaïd, 2015; Licoppe, 2015; Mondada, 2015). These initial adjustments and tests are supported by technologies that provide feedback sound and image. Such adjustments can go as far as moving to a place other than the one where the activity usually takes place or re-shaping traditional settings in considerable ways, as in the courtroom hearings analyzed by Licoppe and his colleagues (Verdier, Dumoulin, and Licoppe, 2012; Veyrier and Licoppe, 2015). They also include working on the semblance of the group in order to display each member's function and status and the hierarchical relationships among members. In the medical meetings studied by Mondada (2015), for instance, one physician indicates to another that he should move to a seat in the front of the auditorium. Of course, a working technological link is the necessary condition for video-mediated interaction, and the preparation for the interaction and the pre-opening phase are also dedicated to checking the technical aspects, such as activating the microphones during an initial exchange of try-out greetings (Mondada, 2010, 2015).

Some systems incorporate a technical notification that acts as a dedicated summons (Licoppe, 2012). Nevertheless, the appearance of a remote image on the screen, or even the screen itself, can also function as a summons (Licoppe and Dumoulin, 2007; Muñoz, 2016; Relieu, 2007). In contrast to traditional landline telephone communication, modern video-communication technologies (such as Skype) offer the called person the opportunity to decline the call in ways that display presence, but not availability, and also to choose between video and audio mode when answering the call (Ibnelkaïd, 2015). Another contrasting feature is that either the caller or the called person may speak first (Licoppe, 2015). This is due to the fact that participants rely on a series of greetings, verbal and gestural, produced during the pre-opening and opening phases as a way to display their own aural and visual appearances and confirm those of their interlocutors-to-be (Fornel, 1996; Ibnelkaïd, 2015; Licoppe, 2015; Mondada, 2015). This series of greetings is also a way to set up a proper pace and order for the sequential production of turns-at-talk (Mondada, 2015). An alternative or additional resource for the participants to check an interlocutor's availability and the proper functioning of the technology is to switch to a textual mode of communication (Ibnelkaïd, 2015).

In the opening phase of video-mediated medical meetings, Mondada observes the practice of 'roll call', in which the chair of the meeting checks the presence (i.e. successful connection) of all participants by calling their names (Mondada, 2007a). Similarly, in video-mediated courtroom hearings, the presiding judge produces a series of greetings and introduces the participants not only by their name and position, but also

by their location in the courtroom (Licoppe and Dumoulin, 2007). In both settings, the person chairing the encounter accomplishes substantial work to make sure that all the participants move forward together during the opening phase of the interaction, resulting in coordinated organized entry. This involves auditory and visual verification of the connection underway, as well as solving technical problems. In this respect, the opening phase is also the locus of socialization work into videoconferencing: novice participants being instructed in the use of technology and the order of activities. Tensions between ordinary practices and technologically constrained organization are also being resolved at this point; for instance, participants abandon the practice of standing up at the beginning of a courtroom hearing since this would put them outside the camera frame (Licoppe and Dumoulin, 2007).

Participants orient to the beginning of the activity that is the reason for the encounter (Mondada, 2015), but first they might produce talk closely associated with the technology in use and the remote nature of the encounter. For instance, interlocutors ask each other where they are, what the place that they can partially see on the screen is, what time it is and what the weather is like in the other interlocutor's location, or who else is present. If the technology allows this practice, participants may temporarily point the camera in different directions to provide answers to some of these questions (Ibnelkaïd, 2015; Veyrier and Licoppe, 2015).

Depending on the technology, video-mediated interaction might be 'hybridized' with other activities that are underway in the participants' life-spaces (Relieu, 2005). Ruhleder and Jordan (2001b) observe that opening and closing a videoconference meeting is problematic in the absence of 'dawn' and 'dusk' periods. In this case, meetings do not emerge as events delimited from the participants' previous and subsequent activities. Some videoconference settings, for instance, include wall-to-wall screens and permanently functioning video-links, which might create an illusion of a 'hyperrealistic' shared space. In one of these spaces, Bonu (2007) investigates the closing and post-closing phases of meetings, including the dispersion of the group and 're-establishment of junction' with the remote environment. He also observes trouble reconstituting distinct remote and local environments when no technical operation is needed to end the meeting. Participants recreate two physically distant environments through dedicated interactional work, including modifying their bodily orientations and conversation topics, and engaging in activities restricted to on-site members, like organizing a departure for lunch (Bonu, 2007).

The production of relevant screen-frames is a result of interpretive work by the person operating the camera: its movements do not consist simply of reflecting a situation, but they also produce it, and are accountable as such (Licoppe, Verdier, and Dumoulin, 2013; Mondada, 2007b). Acting as 'mundane video directors' (Licoppe and Morel, 2014), participants

manipulate the camera in accordance with the ongoing talk (Licoppe, 2014) and physical actions that need to be shown at any given moment, for instance a surgeon's movements during a laparoscopic surgery being broadcast by a videoconferencing system. In the words of Mondada (2003), '[c]amera 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' (p. 60).

3.2. Maintaining connection

When initiating video-mediated interaction, participants' actions are oriented in part towards the functionality of the connection. During the interaction, participants remain oriented to technological disruption as a possibility that has to be prevented, leading to routine sequences of verification (Mondada, 2007a). Indeed, technical problems can prevent reciprocity of perspectives, which is a basic assumption that makes social interaction possible (Schutz, 1962). Moreover, as Mondada (2007a) and Rintel (2013a) argue, the possibility of a technical problem is as interactionally relevant as its actual occurrence. In practical terms, this vulnerability in video-mediated interaction results in videoconference or video-call participants' reinitiating the summons-answer sequence to determine whether the communication channel is still functioning properly and also treating silences as signs of technical problems. At the same time, videoconference participants scrutinize the image on the screen with a dual purpose: not only to see what is happening, but also to monitor the functioning of the technology (Mondada, 2007a, 2015). Participants of video-mediated professional meetings may suspend the ongoing activity or put it 'on hold' by observably focusing on the screen and scanning it visually, trying to reset the software until the technical issues are resolved (Olbertz-Siitonen, 2015). Nevertheless, the interactional significance of 'trouble' (such as overlapping speech or unexpected silence) resulting from technical failure or intentional human conduct is something that has to be determined by the participants over the course of the interaction (Rintel, 2013a). For instance, Licoppe (2017) analyzes a Skype call excerpt in which a 'frozen' image is mistakenly interpreted as a showing of an object. In this case, for the participants, the actual transmission distortion—however omnipresent as a possibility—appears to be the dispreferred explanation of whatever happens on the screen.

In a study of videoconference meetings, Ruhleder and Jordan (2001a) focus on transmission delay as an inherent feature of mediating technology that causes participants not to be 'co-present to the communication in the same way' (p. 115). With detailed analysis and comparison of video recordings from both sides of the videoconference, they document that transmission delay leads to phenomena such as unintended interruptions, rephrasings, mistimed or delayed feedback and other kinds of disruptions of the turn-taking system. Moreover, 'people are unable to identify and repair trouble as

it occurs because its origin is obscured' (Ruhleder and Jordan, 2001a, p. 132). Olbertz-Siitonen (2015) works deliberately with the perspective of only one participant at a time, to stay closer to real-life conditions of professional videoconference meetings in which participants cannot compare the circumstances in the interconnected environments. She describes sequential cues used by participants as evidence of the delay that might also be referred and attended to as a source of sequential trouble, such as mismatching and mistimed contributions (Olbertz-Siitonen, 2015, p. 204).

EM/CA studies capture not only how technology limits human conduct, but also how participants exploit technological features as an interactional resource (Rintel, 2013a, 2013b, 2015). For instance, during video-mediated interactions between romantic partners, the participants can recast a lack of attention as technological trouble, for both themselves and the partner (Rintel, 2013a), or use visual distortions as resources for teasing (Rintel, 2013b).

3.3. Visual contact and attention

The very first studies on video communication systems point to the importance of being on camera and the difficulty of remaining there for the duration of the encounter (Fornel, 1992). Moreover, participants try to adjust the video-frame and their bodily position to produce a portrait-like 'head and chest' ('talking head') image on the screen. Any deviations from this 'default mode' are treated as 'noticeable and mentionable' by the participants of mobile and Skype video-calls (Licoppe and Morel, 2012). When using a mobile phone, this mode is related to interactional and technical constraints, as well as considerations of physical comfort. A close-up image of the speaker corresponds to a narrow camera angle as well as the convenience of keeping the arm flexed (Licoppe and Morel, 2009). Particularly in the case of mobile devices such as laptops or smartphones, interactional screen-frame adjustment sequences may occur anytime. Licoppe and Morel (2012) also suggest that the organized character of *video-in-interaction* derives from a single maxim: 'show the face of the current speaker on screen'. Thus, in multiparty video-mediated interactions, the person operating the mobile device makes the speaker visible on the screen by turning the camera towards them while he or she talks. By the same token, when there is something other than the face of the current speaker on screen, such as certain features of the local physical environment, the image is scrutinized by the participants of the personal mobile or Skype call for its momentary interactional relevance (Licoppe and Morel, 2012). Moreover, these maxims appear to be dropped in encounters oriented towards 'showing and talking' and producing video-as-data (Heath and Luff, 1992; Morel and Licoppe, 2009).

In some of the earliest EM/CA studies on video-mediated interaction, Heath and Luff (1991, 1992, 1993) note that the

performative significance of gaze, gestures and bodily movement appears to be decreased in video-mediated settings. As a result, the initiation of focused interaction, including securing the attention of the other participant, requires upgraded gestural practices and transformation of ordinary ways of talking, specially designed restarts, pauses and sound stretches (Heath and Luff, 1993). In a recent study of video-mediated music lessons, for instance, Duffy and Healey (2014) observe that the teacher needs to produce more extensive verbal instructions to achieve proper bodily positioning by the student.

The distance between screen and camera produces disturbing effects with respect to gaze direction. It is not possible to achieve mutual eye contact with the image on the screen. As a result, eye contact is artificially detached from attendance to the speaker. Fornel (1996) notes that early videophone users had to 'learn to face the camera even though their spontaneous reaction would be to face their interlocutor on the screen' and at the same time to 'keep an eye on the screen' (p. 55). A similar phenomenon is observed by Dourish *et al.* (1996), who also suggest that the practice improves over time, as participants gain awareness of each other's gaze patterns. In this study, office workers linked by long-term open video channels abandoned the practice of looking directly into the camera once they had associated particular gaze orientations with looking at the screen and being attentive to the speaker.

In the current design of laptops and smartphones, with the camera positioned directly above the screen, the distance between screen and camera is minimalized, yet mutual gaze is still not possible. These problems are even more acute in multiparty settings. As Hjulstad (2016) points out, participants of video-mediated classroom interaction see each other attending to the screen, but are unable to precisely distinguish at whom or what on the screen they are looking. Luff *et al.* (2016) also note that the 'Mona Lisa' effect (cf. Rogers *et al.*, 2003) applies to both gaze and pointing. Videoconference participants looking at a person on the screen who is looking and pointing forward, in their direction, will be unable to tell where exactly he or she is aiming. In a related manner, if one participant moves, he or she will have the impression of being followed by the gaze and pointing gesture of the person on the screen.

3.4. Acting in fractured ecologies

In video-mediated interaction, participants do not share the same physical environment and have asymmetrical access to visible surroundings. The mediating technology produces incongruity and incommensurability between the environment of action-production and the environment of action-reception (Heath and Luff, 1992). In this respect, Luff *et al.* (2003) have coined the notion of *fractured ecologies*, in which 'participants are unable to design their own conduct in such a way that it is

sensible and recognizable to a co-participant who has only limited access to the environment in which the action is produced. In this sense, conduct is fractured—fractured from the environment in which it is produced and from the environment in which is received.' (p. 55) As a result, the 'shared interactional zone' is highly fragile (Fornel, 1996, p. 53) and requires methodical maintenance in and through interaction. In video-mediated encounters between job-seekers and their counselors, fractured ecologies result in greater asymmetries in terms of access to relevant resources compared to face-to-face meetings; for instance, only the counselor has access to documents previously viewed in common (Velkovska and Zouinar, 2007). Nevertheless, participants might be able to re-shape their activities and adjust their communication practices to the technology at hand. For instance, they can change gestural practices of reference to material objects and manage to make sense of other participants' practices, as long as the effects remain stable (Luff *et al.*, 2016).

To influence the remote environment, participants rely on specific practices related to verbal and non-verbal referential activities (e.g. pointing) and the achievement of common orientation to an object (Luff *et al.*, 2003; Mondada, 2007b). In complex fractured ecologies involving more than two participants, these may institute new practices of gestural reference. In his study of video-mediated classroom interaction in sign language, Hjulstad (2016) observes that the '[s]igner localizes a specific spatial direction for each of the coparticipants according to the signer's own perspective' (p. 338), and thus 'points' to a different area of his or her immediate environment to refer to each of the remote participants. This practice of 'referential mapping' highlights a dependence on spatial relations to make sense of gaze and gestures in face-to-face interaction as well.

In a study of Skype calls between friends and family, Licoppe (2017) examines the practice of producing recognizable and accountable 'showings' of objects carried out by one of the participants for the benefit of the other(s). He identifies two 'interaction orders': the showing of an object that functions as a complement to talk and a showing that substitutes for talk. In the second case, Licoppe (2017) distinguishes between 'informative' and 'evocative' showing sequences. The former 'enact a recipient without any relevant knowledge with respect to the showable' (p. 81), while the latter enact a knowledgeable recipient. Apart from showing an object to a static camera, another common practice related to acting in fractured ecologies is the reorientation of the camera (Veyrier and Licoppe, 2015), which is facilitated in the case of video-mediated communication over mobile devices (Licoppe and Morel, 2009).

A further problem in fractured ecologies is the remote animation and manipulation of objects. For example, Fornel (1996) describes speech hesitation and leaning towards the screen to indicate disturbing noise coming from another room and the necessity of closing the door in the other participant's

local environment. Velkovska and Zouinar (2007) observe job search counselors who are unable to remotely guide the clients in scanning a document for them, as the hands of the clients are not visible on the screen. Spatial ‘reorganization’ of the remote environment is often achieved by talk, and in group interaction it may become the task of specific members. For example, in the telemedicine consultations investigated by Pappas and Seale (2010), nurses are responsible for physical and sensorial activities in their local environment—such as operating the camera, performing tactile examinations and evaluating symptoms—that the medical specialist ‘orchestrates’ remotely.

Connecting spaces that are remote from each other, while providing only limited access to them, brings up the distinction between *private* and *public* activities. Ruhleder and Jordan (2001b) have observed during videoconference meetings that it is not unusual for people to engage in activities that are designed to remain unnoticed by the remote participants. Hidden activities and side or parallel conversations should be taken into account since they are nevertheless related to the main activity (Tutt *et al.*, 2007). In a study of naturally occurring interactions in Google Hangouts, Rosenbaun, Rafaeli and Kurzon (2016) examine this interrelationship with the concepts of *multiactivity*, referring to two or more interwoven and co-relevant activities, and *schisming*, which is a participation framework with two parallel conversations that cannot be understood separately. Multiple engagements, with participants on the screen and bystanders outside of it but co-present in the local environment, are resolved with a variety of verbal and non-verbal interactional practices (Veyrier and Licoppe, 2015). ‘The tension between online and offline spheres is ... acknowledged and made part of the ongoing interaction’ (Rosenbaun, Rafaeli, and Kurzon, 2016, p. 307), when, for instance, physically co-present individuals outside of the screen frame are jokingly introduced and shown to the other users in multiparty Google Hangouts public sessions. With these practices, participants both establish and blur the traditional public/private and offline/online distinctions.

4. DISCUSSION

The specificity of EM/CA research is that it focuses on participants’ unfolding mutual orientations, as they are made observable and accountable in naturally occurring courses of action. It contributes detailed descriptions of participants’ practices as they happen *in situ* and in real time. This makes the findings highly relevant for technology design. On the other hand, the development of novel technologies points to new directions for study. In this section, we discuss the previously reported findings with respect to implications for future research and design. We derive these implications from the reviewed literature and our broader knowledge of the field,

both as sociologists practicing EM/CA research and as computer scientists.

4.1. Implications for future research

(a) *Generic or dependent practices?* The way people use video-mediating technology is connected to its particular ‘affordances’ (Hutchby, 2001, 2014): the actions that it enables and constrains. It is also related to the specific activities to be accomplished in the setting. This raises the question of the generalizability of findings across settings and technologies. What exactly do business videoconference and smartphone video-calls between family members have in common? In other words, do EM/CA findings refer to generic practices of video-mediated interaction, regardless of activity, setting and technology, or to context/technology-dependent forms of human behavior? Further investigations could shed more light on this subject and generate analytical and methodological contributions, such as ways of representing data, that are better adapted to the specificities of the settings under study.

(b) *Novelty and routine.* A number of EM/CA studies of video-mediated interaction investigate experimental or otherwise unusual set-ups. Research with prototypes or novel technologies (Kurvinen, Koskinen, and Battarbee, 2008; Suchman, Trigg, and Blomberg, 2002) keeps providing ‘perspicuous settings’ magnifying the work involved in interacting in/through technology (Mondada, 2015). In the words of Suchman (1987), ‘by studying what things look like when they are unfamiliar, [we can] understand better what is involved in their mastery.’ (p. 75) On the other hand, investigation of widely used video-mediating technologies will document how they blend with ordinary activities and re-shape them, once they have become an unremarkable component of everyday life.

(c) *Longitudinal studies.* EM/CA studies of video-mediated interaction tend to focus on single encounters with the technology without connecting them in temporal series. While aiming to discover the methods with which participants organize their conduct *in situ*, they may overlook how people develop these methods over time. New research could provide detailed descriptions of interactional changes over several occasions and examine the emergence of specific forms of conduct, as participants become acquainted with the technology (Pekarek Doehler, Wagner, and González-Martínez, 2018). Since the devices afford several uses, such research will also tell us how a specific form of conduct becomes the preferred one. Studies involving children or the elderly, among other new users, learning about and mastering the use of the video-mediating technology, could document the progressive development of sophisticated practices. In the process, EM/CA will develop solutions for the methodological challenges involved in comparative research focusing on the detailed organization of situated practices (Schegloff, 2009).

(d) *Transcending boundaries.* For a long time, video-mediated interaction was confined to clearly delimited spaces and moments in time. Technologies have now become ubiquitous and allow for continuous connection, even with participants on the move. A consequence of this is that the technology is now not only present in a higher number of scenes of action, but also captures more aspects of them. Future research could provide insight into new forms of interactional involvement apart from focused interaction in which pre-established participants share a common focus of attention or activity. For example, bystanders known to the participants may become momentarily involved in the video-mediated interaction (Dourish *et al.*, 1996), and strangers visible on screen can become a subject of interest, search and contact for the participants (Licoppe, 2013). The possibilities increase as technologies become interconnected and participate in human enhancement developments, the public debate about issues of privacy and confidentiality expands, and EM/CA research faces new methodological challenges, including combining data generated automatically by the technology with data produced by the researcher himself or herself (cf. Brown, McGregor, and Laurier, 2013).

4.2. Implications for design

This section reformulates recurrent findings of the reviewed literature as a set of behaviors to be supported by video-communication technologies, echoing some long-standing concerns in HCI and CSCW (cf. Finn, Sellen, and Wilbur, 1997), and relating to recent technological developments in the field.

(a) *Mutual gaze and gaze direction.* Participants expect to achieve mutual gaze and a clear understanding of gaze direction that remains constant over time (Fornel, 1996; Heath and Luff, 1992; Hjulstad, 2016). An initial solution for small groups was the use of separate devices, each of them representing one remote interlocutor, equipped with a screen, camera and audio functionalities (Gaver *et al.*, 1993). Nowadays, an increasing number of technical solutions for gaze correction are becoming available (Kuster *et al.*, 2012) and recent developments in gaze tracking also provide opportunities for new functionalities (Otsuki *et al.*, 2016).

(b) *Spatial reference.* Participants expect a clear understanding of spatial reference, including pointing, that remains constant over time (Luff *et al.*, 2003, 2016; Mondada, 2003). Remote representation of arm movements has been explored with robotics (Onishi, Tanaka, and Nakanishi, 2014), and recent progress in gesture recognition provides further potential for ordinary situations (Katsamanis *et al.*, 2017).

(c) *Camera manipulation and showings.* Participants expect video technology to be able to accommodate multiple seamlessly changing showings (Licoppe *et al.*, 2017; Mondada, 2003). Recent developments in wearable cameras combined with eye-tracking and pointing technologies could

support emergent social practices of distant communication (Kupta, Lee, and Billinghurst, 2016).

(d) *Multimodality.* Participants expect to be able to combine different communication modalities and switch between them seamlessly (Ibnelkaïd, 2015; Relieu, 2006; Sindoni, 2012). Drawing and handwriting is currently limited, especially with personal computers and mobile devices, and there are major restrictions in conveying haptic/tactile and olfactory perceptions during video-mediated interaction, although notable progress has been made in this area (Dangelmaier and Blach, 2017; Rasool and Sourin, 2016).

(e) *Awareness and control.* Participants expect awareness and control of what is being transmitted through the video (Rosenbaun, Rafaeli, and Kurzon, 2016; Ruhleder and Jordan, 2001b; Veyrier and Licoppe, 2015). Designers include aural/visual signalization of current engagement in video-mediated interaction (Mackay, 1999). Recent developments aim to offer functionalities that distinguish among people based on their participation status as well as activities produced to be public or remain private (Marlow *et al.*, 2016).

5. CONCLUSION

This article reviews video-mediated interaction research conducted over the last 30 years using the sociological approaches of ethnomethodology and conversation analysis (EM/CA). The reviewed studies focus on specific practices taking place during activities and in settings that are also very specific and describe them in great detail. The article contributes an original organization of the major findings by presenting them according to the main temporal phases of a video-mediated encounter: setting up, opening, maintaining/acting and closing. We thus put forward an array of phenomena related to the interplay of talk, bodily behavior and spatial and material resources that are relevant for the understanding of a large spectrum of video-mediated social activities. Moreover, we show that video-mediated interaction: (a) constitutes a new locus for investigating classic EM/CA phenomena (openings, closings, repairs) and discovering new ones (showings); but also (b) shows new ways of accomplishing these phenomena (via gesture instead of talk); and more importantly (c) expands the field of investigation on coordinated action, for instance to simultaneous action (mutual gaze) in addition to turn-taking organized action. Finally, we outline directions for future EM/CA research, arguing that emphasis should be given to comparative studies that follow users over time and across different settings and technologies as they transcend previous boundaries in terms of accessibility, mobility and technical interrelations.

In terms of practical implications, the review underscores the importance of detailed analysis of actual human conduct in real-life situations: (a) the practices going on in front of the screen, including—but not limited to—those related to the technology being used, but also (b) practices occurring in the vicinity of the video-captured interaction that pertain to what

is happening on the screen, in order to develop (c) context-aware technologies that embody a subtle understanding of the reflexive relationship between action and context (each of them shaping and being shaped by the other). Novel technologies are often grounded in already existing activities: ‘if the technology does not support familiar activities its actual use can become problematic’ (Crabtree *et al.*, 2009, p. 886). Studying actual practices is a basis for anticipating new ones. By reviewing and discussing EM/CA research on video-mediated interaction, we have hopefully demonstrated that it can continue to provide valuable insights for the field of HCI and technology design.

ACKNOWLEDGEMENTS

The preparation of this article was supported by the Research Fund of the University of Fribourg. The authors thank Elisabeth Lyman for her editing work.

REFERENCES

- Arminen, I., Licoppe, C. and Spagnolli, A. (2016) Respecifying mediated interaction. *Res. Lang. Soc. Interact.*, 49, 290–309. doi:10.1080/08351813.2016.1234614.
- Bernhaupt, R., Obrist, M., Weiss, A., Beck, E. and Tschelegi, M. (2008) Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *Comput. Entertain.*, 6. doi:10.1145/1350843.1350848.
- Bonu, B. (2007) Connexion continue et interaction ouverte en réunion visiophonique. *Réseaux*, 2007/5, 25–57. doi:10.3917/res.144.0025.
- Brown, B., McGregor, M. and Laurier, E. (2013). iPhone in vivo: video analysis of mobile device use. CHI ‘13: Proc. SIGCHI Conf. on Human Factors in Computing Systems, Paris, France, pp. 1031–1040. New York: ACM.
- Button, G. and Dourish, P. (1996). Technomethodology: paradoxes and possibilities. CHI ‘96: Proc. SIGCHI Conf. Human Factors in Computing Systems, Vancouver, British Columbia, Canada, pp. 19–26. New York: ACM.
- Carter, S. and Mankoff, J. (2005). When participants do the capturing: the role of media in diary studies. In CHI’05: Proc. SIGCHI Conf. Human Factors in Computing Systems, pp. 899–908. New York: ACM.
- Crabtree, A., Hemmings, T., Rodden, T., Cheverst, K., Clarke, K., Dewsbury, G., Hughes, J. and Rouncefield, M. (2003). Designing with care: Adapting cultural probes to inform design in sensitive settings. In Proc. Conf. New Directions in Interaction, Information Environments, Media, and Technology, OzCHI’03.
- Crabtree, A., Rodden, T., Tolmie, P. and Button, G. (2009). Ethnography considered harmful. In CHI ‘09: Proc. SIGCHI Conf. Human Factors in Computing Systems, pp. 879–888. New York: ACM.
- Dangelmaier, M. and Blach, R. (2017) Odor in immersive environments. In Buettner, A. (ed.), *Springer Handbook of Odor*. pp. 139–140. Springer, Cham, doi:10.1007/978-3-319-26932-0_55.
- Dourish, P. (2001) *Where the Action Is: The Foundations of Embodied Interaction*. The MIT Press, Cambridge.
- Dourish, P., Adler, A., Bellotti, V. and Henderson, A. (1996) Your place or mine? Learning from long-term use of audio-video communication. *Comput. Support. Coop. Work.*, 5, 33–62.
- Dourish, P. and Button, G. (1998) On ‘technomethodology’: foundational relationships between ethnomethodology and system design. *Hum. Comput. Interact.*, 13, 395–432.
- Duffy, S. and Healey, P.G.T. (2014) The conversational organization of musical contributions. *Psychol. Music*, 42, 888–893. doi:10.1177/0305735614545501.
- Finn, K.E., Sellen, A.J. and Wilbur, S.B. (eds) (1997) *Video-mediated Communication*. Lawrence Erlbaum, Mahwah.
- Fornel, M. de (1992) ‘Alors, tu me vois?’: Objet technique et cadre interactionnel dans la pratique visiophonique. *Cult. Tech.*, 1992, 113–120.
- Fornel, M. de (1996) The interactional frame of videophonic exchange. *Réseaux: Fr. J. Commun.*, 4, 47–72.
- Garfinkel, H. (1967) *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs.
- Garfinkel, H., Lynch, M. and Livingston, E. (1981) The work of a discovering science construed with materials from the optically discovered pulsar. *Philos. Soc. Sci.*, 11, 131–158.
- Gaver, W.W., Sellen, A., Heath, A. and Luff, C., P. (1993). One is not enough: multiple views in a media space. CHI ‘93: Proc. INTERACT ‘93 and CHI ‘93 Conf. Human Factors in Computing Systems, pp. 335–341. New York: ACM.
- Haddon, L. (2006) The contribution of domestication research to in-home computing and media consumption. *Inf. Soc. Int. J.*, 22, 195–203. doi:10.1080/01972240600791325.
- Harrison, S. (ed.) (2009) *Media Space 20+ Years of Mediated Life*. Springer, London.
- Heath, C. and Luff, P. (1991). Disembodied conduct: communication through video in a multi-media office environment. CHI ‘91: Proc. SIGCHI Conf. Human Factors in Computing Systems, New Orleans, Louisiana, USA, pp. 99–103. New York: ACM.
- Heath, C. and Luff, P. (1992) Media space and communicative asymmetries: preliminary observations of video-mediated interaction. *Hum. Comput. Interact.*, 7, 315–346. doi:10.1207/s15327051hci0703_3.
- Heath, C. and Luff, P. (1993) Disembodied conduct: interactional asymmetries in video-mediated communication. In Button, G. (ed.), *Technology in Working Order: Studies of Work, Interaction and Technology*. pp. 35–54. Routledge, London / New York.
- Heath, C. and Luff, P. (2000) *Technology in Action*. Cambridge University Press, Cambridge.
- Hjulstad, J. (2016) Practices of organizing built space in videoconference-mediated interactions. *Res. Lang. Soc. Int.*, 49, 325–341. doi:10.1080/08351813.2016.1199087.
- Hutchby, I. (2001) *Conversation and Technology: From the Telephone to the Internet*. Polity Press, Cambridge.

- Hutchby, I. (2014) Communicative affordances and participation frameworks in mediated interaction. *J. Pragmat.*, 72, 86–89. doi:10.1016/j.pragma.2014.08.012.
- Ibnelkaid, S. (2015) Scénographie d'une ouverture d'interaction vidéo. *Réseaux*, 6/2015, 125–168.
- Johnston, CH. (2017): Prof Robert Kelly: 'We were worried the BBC would never call us again'. *The Guardian*, March 15, 2017. Retrieved from <https://www.theguardian.com/media/2017/mar/14/robert-kelly-children-interrupt-live-bbc-interview-south-korea> [30/03/2017]
- Katsamanis, A., Pitsikalis, V., Theodorakis, S. and Maragos, P. (2017) Multimodal gesture recognition. In Oviatt, S., Schuller, B., Cohen, P.R., Sonntag, D., Potamianos, G. and Krüger, A. (eds), *The Handbook of Multimodal-Multisensor Interfaces*. pp. 449–487. ACM / Morgan & Claypool, New York, doi:10.1145/3015783.3015796.
- Kupta, G., Lee, G.A. and Billinghurst, M. (2016) Do you see what I see? The effect of gaze tracking on task space remote collaboration. *IEEE. Trans. Vis. Comput. Graph.*, 22, 2413–2422. doi:10.1109/TVCG.2016.2593778.
- Kurvinen, E., Koskinen, I. and Battarbee, K. (2008) Prototyping social interaction. *Des. Issu.*, 24, 46–57.
- Kuster, C., Popa, T., Bazin, J.-C., Gotsman, C. and Gross, M. (2012) Gaze correction for home video conferencing. *ACM Trans. Graph.*, 31, doi:10.1145/2366145.2366193. Article no. 174.
- Licoppe, C. (2012) Understanding mediated appearances and their proliferation: The case of the phone rings and the 'crisis of the summons'. *New Media Soc.*, 14, 1073–1091. doi:10.1177/1461444812452410.
- Licoppe, C. (2013) Merging mobile communication studies and urban research: mobile locative media, 'onscreen encounters' and the reshaping of the interaction order in public places. *Mobile Media Commun.*, 1, 122–128. doi:10.1177/2050157912464488.
- Licoppe, C. (2014) Interactions médiées et action située. *Réseaux*, 2/2014, 317–345.
- Licoppe, C. (2015) 'Apparitions', multiples salutations et 'coucou'. *Réseaux*, 6/2015, 85–124.
- Licoppe, C. (2017) Showing objects in Skype video-mediated conversations: from showing gestures to showing sequences. *J. Pragmat.*, 110, 63–82. doi:http://dx.doi.org/10.1016/j.pragma.2017.01.007.
- Licoppe, C. and Dumoulin, L. (2007) L'ouverture des procès à distance par visioconférence. *Réseaux*, 5/2007, 103–140.
- Licoppe, C., Luff, P., Heath, C., Kuzuoka, H., Yamashita, N. and Tuncer, S. (2017). Showing objects: holding and manipulating artefacts in video-mediated collaborative settings. *Proc. 2017 CHI Conf. Human Factors in Computing Systems*, pp. 5295–5306. New York: ACM.
- Licoppe, C. and Morel, J. (2009). The collaborative work of producing meaningful shots in mobile video telephony. *MobileHCI '09: Proc. 11th Int. Conf. Human-Computer Interaction with Mobile Devices and Services*, Bonn, Germany—September 15–18, 2009. Article No. 35.
- Licoppe, C. and Morel, J. (2012) Video-in-interaction: 'Talking Heads' and the multimodal organization of mobile and skype video calls. *Res. Lang. Soc. Int.*, 45, 399–429. doi:10.1080/08351813.2012.724996.
- Licoppe, C. and Morel, J. (2014) Mundane video directors in interaction: showing one's environment in Skype and mobile video calls. In Broth, M., Laurier, E. and Mondada, L. (eds), *Studies of Video Practices: Video at Work*. pp. 135–160. Routledge, London.
- Licoppe, C. and Relieu, M. (2007) Présentation. *Réseaux*, 5/2007, 9–22.
- Licoppe, C., Verdier, M. and Dumoulin, L. (2013). Courtroom Interaction as a multimedia event: the work of producing relevant videoconference frames in French Pre-Trial Hearings. *The Electronic Journal of Communication/La Revue Electronique de Communication (EJC/REC)*, 23(1–2). Retrieved from <http://www.cios.org/EJCPUBLIC/023/1/023125.HTML> [30/03/2017]
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K. and Oyama, S. (2003) Fractured ecologies: creating environments for collaboration. *Hum. Comput. Int.*, 18, 51–84. doi:10.1207/s15327051hci1812_3.
- Luff, P., Heath, C., Yamashita, N., Kuzuoka, H. and Jirotko, M. (2016) Embedded reference: translocating gestures in video-mediated interaction. *Res. Lang. Soc. Int.*, 49, 342–361. doi:10.1080/08351813.2016.1199088.
- Lynch, M. (1993) *Scientific Practice and Ordinary Action*. Cambridge University Press, Cambridge.
- Mackay, W.E. (1999) Media spaces: environments for informal multimedia interaction. In Beaudouin-Lafon, M. (ed.), *Computer-Supported Cooperative Work*. pp. 55–82. Wiley & Sons, Chichester.
- Marlow, J., van Everdingen, E. and Avrahami, D. (2016) Taking notes or playing games? Understanding multitasking in video communication. In *Proc. ACM Conf. Computer-Supported Cooperative Work & Social Computing*. pp. 1726–1737. ACM, New York.
- Matthews, B. (2013) Conversation analysis and design. In Chapelle, C.A. (ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell, Oxford.
- Mondada, L. (2003) Working with video: how surgeons produce video records of their actions. *Vis. Stud.*, 18, 58–73. doi:10.1080/1472586032000100083.
- Mondada, L. (2007a) Imbrications de la technologie et de l'ordre interactionnel: L'organisation de vérifications et d'identifications de problèmes pendant la visioconférence. *Réseaux*, 5/2007, 141–182. doi:10.3917/res.144.0141.
- Mondada, L. (2007b) Operating together through videoconference: members' procedures for accomplishing a common space of action. In Hester, S. and Francis, D. (eds), *Orders of Ordinary Action*. pp. 51–67. Ashgate, Aldershot.
- Mondada, L. (2010) Eröffnung und Vor-Eröffnung in technisch vermittelter Interaktion: Videokonferenzen. In Schmitt, R. and Mondada, L. (eds), *Situationseröffnungen: Zur multimodalen Herstellung fokussierter Interaktion*. pp. 277–334. Narr, Tübingen.

- Mondada, L. (2015) Ouverture et préouverture des réunions visio-phoniques. *Réseaux*, 6/2015, 39–84.
- Morel, J. and Licoppe, C. (2009) La vidéocommunication sur téléphone mobile. *Réseaux*, 4/2009, 165–201.
- Muñoz, A.S. (2016). Attending Multi-Party Videoconference Meetings: The Initial Problem. *Language@Internet*, 13. Retrieved from <http://www.languageatinternet.org/articles/2016/munoz> [30/03/2017]
- Noll, A.M. (1992) Anatomy of a failure: picturephone revisited. *Telecomm. Policy*, 16, 307–316. doi:10.1016/0308-5961(92)90039-R.
- Olbertz-Siitonen, M. (2015) Transmission delay in technology-mediated interaction at work. *PsychNol. J.*, 13, 203–234.
- Onishi, Y., Tanaka, K. and Nakanishi, H. (2014). PopArm: a robot arm for embodying video-mediated pointing behaviors. 2014 Int. Conf. Collaboration Technologies and Systems (CTS). doi:10.1109/CTS.2014.6867556
- Otsuki, M., Kawano, T., Maruyama, K., Kuzuoka, H. and Suzuki, Y. (2016). Representing gaze direction in video communication using eye-shaped display. *Proc. 29th Annual Symposium on User Interface Software and Technology*, pp. 65–67. New York: ACM.
- Pappas, Y. and Seale, C. (2009) The opening phase of telemedicine consultations: an analysis of interaction. *Soc. Sci. Med.*, 68, 1229–1237. doi:10.1016/j.socscimed.2009.01.011.
- Pappas, Y. and Seale, C. (2010) The physical examination in telecardiology and televascular consultations: a study using conversation analysis. *Patient. Educ. Couns.*, 81, 113–118. doi:10.1016/j.pec.2010.01.005.
- Pekarek Doehler, S., Wagner, J. and González-Martínez, E. (eds) (2018) *Longitudinal Studies on the Organization of Social Interaction*. Palgrave MacMillan, London, [in press].
- Rasool, S. and Sourin, A. (2016) Real-time haptic interaction with RGBD video streams. *Vis. Comput.*, 32, 1311–1321. doi:10.1007/s00371-016-1224-1.
- Relieu, M. (2005) Les usages des TIC en situation naturelle: une approche ethnométhodologique de l'hybridation des espaces d'activité. *Intellectica*, 2, 41–42.
- Relieu, M. (2006) Remarques sur l'analyse conversationnelle et les technologies médiatisées. *Revue française de linguistique appliquée*, 11, 17–32.
- Relieu, M. (2007) La téléprésence, ou l'autre visio-phonie. *Réseaux*, 5/2007, 183–223. doi:10.3166/Reaux.144.183-223.
- Rettie, R. (2009) Mobile phone communication: extending Goffman to mediated interaction. *Sociology.*, 43, 421–438. doi:10.1177/0038038509103197.
- Rintel, S. (2013a). Tech-tied or tongue-tied? Technological versus social trouble in relational video calling. *Proc. 46th Hawaii Int. Conf. System Sciences*, pp. 3343–3352. doi: 10.1109/HICSS.2013.512
- Rintel, S. (2013b). Video calling in long-distance relationships: The opportunistic use of audio/video distortions as a relational resource. *The Electronic Journal of Communication/La Revue Electronique de Communication (EJC/REC)*, 23(1–2). Retrieved from <http://www.cios.org/EJCPUBLIC/023/1/023123.HTML> [30/03/2017]
- Rintel, S. (2015) Omnirelevance in technologized interaction: couples coping with video calling distortions. In Fitzgerald, R. and Housley, W. (eds), *Advances in Membership Categorization Analysis*. pp. 123–150. Sage, London.
- Rogers, S., Lunsford, M., Strother, L. and Kubovy, M. (2003) The Mona Lisa effect: perception of gaze direction in real and pictured faces. In Rogers, S. and Effken, J. (eds), *Studies in Perception and Action VII*. pp. 19–24. Lawrence Erlbaum Associates, Mahwah.
- Rosenbaun, L., Rafaeli, S. and Kurzon, D. (2016) Blurring the boundaries between domestic and digital spheres: Competing engagements in public Google hangouts. *Pragmatics*, 26, 291–314.
- Ruhleder, K. and Jordan, B. (2001a) Co-constructing non-mutual realities: delay-generated trouble in distributed interaction. *Comput. Support. Cooperat. Work*, 10, 113–138. doi:10.1023/A:1011243905593.
- Ruhleder, K. and Jordan, B. (2001b) Managing complex, distributed environments: remote meeting technologies at the 'chaotic fringe'. *First Monday*, 6, <http://firstmonday.org/ojs/index.php/fm/article/view/857/766> [30/03/2017].
- Sacks, H. (1992) *Lectures on conversation I-II*. In Schegloff, E.A. and Gail Jefferson (eds.), *With introductions*. Blackwell, Oxford.
- Schegloff, E.A. (2007). *Sequence Organization in Interaction: Volume 1 – A Primer in Conversation Analysis*. Cambridge University Press, Cambridge.
- Schegloff, E.A. (2009) One perspective on conversation analysis. In Sidnell, J. (ed.), *Conversation Analysis: Comparative Perspectives*. pp. 357–406. Cambridge University Press, Cambridge.
- Schutz, A. (1962) *Collected Papers: The Problem of Social Reality*. Martinus Nijhoff, The Hague / Boston / London.
- Sindoni, M.G. (2012) Mode-switching: how oral and written modes alternate in videochats. In Cambria, M., Arizzi, C. and Cocchetta, F. (eds), *Web Genres and Web Tools: With Contributions from the Living Knowledge Project*. pp. 141–153. Ibis, Como / Pavia.
- Suchman, L.A. (1987) *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge.
- Suchman, L.A. (2011) Work practice and technology: a retrospective. In Szymanski, M.H. and Whalen, J. (eds), *Making Work Visible: Ethnographically Grounded Case Studies of Work Practice*. pp. 21–33. Cambridge University Press, New York.
- Suchman, L.A., Trigg, R. and Blomberg, J. (2002) Working artefacts: ethnomethods of the prototype. *Br. J. Sociol.*, 53, 163–179.
- Tutt, D., Hindmarsh, J., Shaikat, M. and Fraser, M. (2007). The distributed work of local action: Interaction amongst virtually collocated research teams. In L. J. Bannon, I. Wagner, C. Gutwin, R. H. R. Harper, & K. Schmidt (Eds.), *ECSCW 2007: Proc. 10th European Conf. Computer-Supported Cooperative Work, Limerick, Ireland, 24–28 September 2007*, pp. 199–218. London: Springer London.

- Vasilyeva, Z. (2013) Video-mediated communicative interaction: an analysis. *Forum Anthropol. Cult.*, 2013, 117–148.
- Velkovska, J. (2014) Ethnométhodologie des usages des TICs: Recherches françaises. *Lendemain*, 39, 40–75.
- Velkovska, J. and Zouinar, M. (2007) Interaction visiophonique et formes d'asymétries dans la relation de service. *Réseaux*, 5/2007, 225–264.
- Verdier, M., Dumoulin, L. and Licoppe, C. (2012). Les usages de la visioconférence dans les audiences judiciaires en France: les enjeux d'un protocole de recherche basé sur l'enregistrement audiovisuel des pratiques. *Ethnographiques.org*, 25 (Décembre 2012). Retrieved from <http://www.ethnographiques.org/2012/Verdier-Dumoulin-Licoppe> [18/08/2017]
- Veyrier, C.-A. and Licoppe, C. (2015) Faire apparaître un tiers à l'écran en visiocommunication. *Réseaux*, 6/2015, 169–195.