Author(s): Olbertz-Siitonen, Margarethe

Title: Transmission delay in technology-mediated interaction at work

Year: 2015

Version:

Transmission delay in technology-mediated interaction at work

Margarethe Olbertz-Siitonen[*]

University of Jyväskylä
(Finland)

ABSTRACT

This paper investigates transmission delay in technology-mediated interaction from a participant's perspective. The approach chosen here contrasts with previous studies of delay: only what is seen, heard and done at one end (i.e., what is available on screen) is considered in the analysis. It is argued that participants themselves lack access to the other sides of their interactions as they unfold at remote locations, and they thus cannot observe and deal with delay from an outside perspective. Analyzing single cases of delay-in-interaction within the framework of conversation analysis, the focus is on dispersed conversation partners' resources to detect a problem, and to make visible and relevant the experience of time lag. The data used consist of video-based mediated meetings between knowledge workers of a small Finnish company. Phenomena that may evoke reference to transmission delay in interaction include *prolonged overlap*, *gap and sequential disarray* and *missed attempt at turn-taking*. Delay in such cases is referred to as an explanation for mismatching, incongruous contributions. This, as well as participants' observable attempts to restore order, reveals fundamental expectations and orientations with respect to sequentiality as well as to technology-mediation.

1. Introduction

Transmission delay (or lag) constitutes "an unavoidable fact of life in distributed applications" (Gutwin et al., 2004, p. 503; see also Murai, Kimber, Foote, Liu and Doherty, 2005). Delay has been found to severely interfere with the structure of conversation, causing remote co-participants to interrupt one another unintentionally or to produce involuntary silences while not really becoming aware of the actual reason for problematic turn-taking or mismatching turns (e.g., Schoenenberg, Raake and

[*] Corresponding Author:
Margarethe Olbertz-Siitonen
Department of Languages, University of Jyväskylä, PO Box 35, FI-40014 University of Jyväskylä
E-mail: margarethe.olbertz-siitonen@jyu.fi

Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014). In the same vein, delay is often seen and reported on as consequential with respect to social meanings and user experience (Gutwin et al., 2004; Murai et al., 2005; Kawashima, Nishikawa and Matsuyama, 2008; Egger, Schatz, Schoenenberg, Raake and Kubin, 2012; Schoenenberg, Raake and Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014).

A closer look at people's practices and orientations in authentic situations, on the other hand, reveals that technological features such as audio and video distortions, including lag, may be actively used by co-participants as interactional resources in technology-mediated relational talk (Rintel, 2010; see also Rintel 2013a; 2013b). In this paper, following Rintel's lead, and in contrast to much previous research, I approach and analyze transmission delay not as something that happens to conversation partners, but as a phenomenon that emerges in interaction through participants' displays. In other words, delay is tackled here from an emic standpoint (see ten Have, 2007), focusing on what is made relevant in interaction. I show that and how transmission delay is occasionally referred and attended to as a plausible explanation for mismatching and mistimed contributions. The data indicate not only that certain sequential cues may work for users as evidence of lag, but also that participants are cognizant of potential technology-generated failures, which in turn provide them with the option to account for organizational trouble (e.g., pauses or overlaps that otherwise could be heard as meaningful, unrelated responses) in reference to the mediation.

This study was conducted within the framework of conversation analysis (CA). Accordingly, the analysis draws on conversation analytic notions and findings in order to account for the observed phenomena. CA is concerned with the sequential, turn-by-turn unfolding of authentic talk and allows for an in-depth examination of co-participants' procedures to mutually accomplish social interaction (e.g., Heritage, 1984a; ten Have, 2007; Deppermann, 2008; Hutchby and Wooffitt, 2008; Sidnell, 2010).

In line with the ongoing development and improvement of information and communication technology (ICT), and like other research traditions focusing on human social interaction, the interest from a CA perspective in technology-mediated (as well as computer-supported) communication has grown notably in recent years. There is, for example, a relatively long line of CA research into the organization of text-based online communication, interested particularly in aspects of turn-taking and adjacency (e.g., Garcia and Jacobs, 1999; O'Neill and Martin, 2003; Markman, 2005; 2013;

Örnberg Berglund, 2009), chat initiation and closing sequences (Rintel, Mulholland and Pittam, 2001; Markman, 2009) and repair (e.g., Schönfeldt and Golato, 2003; Jacobs and Garcia, 2013; Meredith and Stokoe, 2014). Although studies of technology-mediated communication have, in general, mostly been concerned with written online interaction (see Jenks and Firth, 2013; Brandt and Jenks, 2013), a considerable body of conversation analytic work exists, which examines particulars of audio- and video-based interaction. The central themes of many of these projects are similar to CA research on chat interaction, including the organization of turn-taking (Ruhleder and Jordan, 2001), openings of videoconferences (Mondada, 2010), as well as trouble and repair (Brandt and Jenks, 2013; Rintel, 2010; 2013a; 2013b). On the other hand, more and more conversation analysts are, for obvious reasons, concerned with multimodal aspects of video-based mediated interaction (e.g., Mondada, 2007; Licoppe and Morel, 2012; Luff et al., 2013). Similar CA-informed research interests also potentially include the organization of interaction in collaborative virtual environments. In this particular line of work, embodied activities as mediated by avatars – such as issues related to referencing (e.g., pointing) and bodily orientations – are of special relevance, and corresponding findings can be important for the design and development of virtual spaces meant for meetings and collaboration (Moore et al. 2007; Hindmarsh et al., 2000).

This short overview demonstrates that a major objective of CA studies examining technologically mediated interaction is to describe and account for practices (and their modifications) that have been found to be crucial for the interactional accomplishment of everyday face-to-face talk. The organization of turn-taking, sequential ordering, openings and closings, as well as repair constitute themes that have been central to conversation analytic inquiry since the beginning. In light of the fact that human social interaction increasingly takes place online, it is natural to ask what these conversational features look like or how they manifest in interaction that relies on technology. Attending to this question does not imply, however, that the role of technology is approached in CA as determining participants' practices. Rather, it means that the analysis should focus on parties' observable *orientations* to the possibilities (or affordances) and restrictions of the medium (see Rintel, 2013a; 2013b; Meredith and Stokoe, 2014; see also Hutchby, 2001a; 2001b).

By exploring transmission delay as it emerges and is managed in authentic situations, the goal of this paper is ultimately to provide insights into some of the ways users orient to and rely on the technology-mediated nature of their interactions. In the following

sections I first discuss the background of this study by focusing on sequentiality in the context of computer-mediated interaction and on previous research on turn-taking during transmission delay. Before turning to the analysis of a selection of several examples, which represent cases of participants explicitly attributing experiences of trouble to lag, I introduce the data and justify my methodological choices. The paper ends with concluding remarks that summarize my findings and examine their implications.

2. Background

2.1 Computer-mediated interaction and sequentiality

Classic work in CA has demonstrated that mutual understanding in ordinary conversation is tied to sequentiality, relying on the *finely tuned* succession of turns at talk (Schegloff and Sacks, 1973; Sacks, Schegloff and Jefferson, 1974; Schegloff, Jefferson and Sacks, 1977; Heritage, 1984a; Schegloff, 1984). A current speaker's talk is produced and interpreted in relation to the previous speaker's activities, and it provides a contextual framework for the next turn. However, the fact that conversation today also commonly takes place online adds a new dimension to this original assessment. Though text-based as well as audio- and video-based settings are essentially different modes of technologically mediated communication (one enabling quasi-synchronous written exchange and the other synchronous interaction through talk and, possibly, visible conduct), both show distinctive temporal features that appear to be particularly relevant for sequentiality and turn-taking.

Online written conversation, in particular, has been studied against this backdrop – not only within the framework of CA. Research on textual chat revealed what some authors have conceptualized as "disrupted turn adjacency" (Herring, 1999, p. 2), "phantom adjacency pairs" (Garcia and Jacobs, 1999, p. 354) or "false adjacency pairs" (Markman, 2005, p. 118): 'next' in the sense of the placement of a message does not necessarily mean 'next' in the sense of close sequential ordering. This is due to the circumstance that contributions (or 'turns') are typed and usually appear in the chat-window as complete messages only after they have been submitted (see Paolillo and Zelenkauskaite, 2013). Other than in face-to-face interaction ("full-duplex form of ordinary conversation" in Hutchby and Tanna, 2008, p. 146), the process of constructing a turn in most types of chat cannot be monitored (for a comparison of

structural properties evolving around two different chat systems – one providing access to message construction on a word-by-word basis, and the other relying on posting complete posts – in massively multiplayer online role playing games, see Moore et al., 2007). This means that places of transition relevance – and thus precision timing – are of no consequence for co-participants. Still, participants have been found to make use of several techniques to contextualize single contributions (e.g., Herring, 1999; Garcia and Jacobs, 1999; Hutchby, 2001b; Schönfeldt and Golato, 2003; Simpson, 2005; Nilson and Mäkitalo, 2010; Jacobs and Garcia, 2013; Markman, 2013) and thus achieve "virtual adjacency" (Schönfeldt and Golato, 2003, p. 251). Interlocutors, in other words, not only appear to fundamentally orient to and ensure coherence in chat conversation as well, they also display an awareness (or expectation) of there being no unconditional relation between successive posts, which would – as in *talk*-in-interaction – more or less automatically procure the context of conversational contributions.

Other forms of real-time computer-mediated interaction, on the other hand, have received much less attention with regard to sequential order and turn-taking, even though the question of whether (and possibly how) technological affordances become relevant in their fine-grained organization is of equal import for a more comprehensive understanding of the building blocks of social interaction in a communication-technologized world (see Jenks and Firth, 2013). Audio- and video-mediated communication is typically considered synchronous, because participants, each at their end, are simultaneously involved in the accomplishment of interaction. While this is the case also in chat, in oral computer-mediated interaction contributions are basically available to conversation partners as they unfold, which is a prerequisite for the local management of turn-taking (Sacks et al., 1974). However, the nature of such talks as mediated presupposes the more or less time-consuming transmission and processing of large amounts of data at and between remote sites (see also Schulte, Friebel and Klotzek, 2001; Friebel et al., 2003; Gutwin et al., 2004; Schoenenberg, Raake, Egger and Schatz, 2014), which means in turn that audio- and especially video-mediated interaction is inevitably subject to some temporal lag caused by the network. Tracing, for example, the path an audio-signal takes in Voice over Internet Protocol (VoIP), occurrences of temporal lag in these and similar kind of settings do not come as a surprise:

> "First, speech encoding and packetization of the recorded speech signal into VoIP packets require buffering of incoming speech samples which causes delay.

Then there is, of course, the transmission delay of packets through the network. Third, reaching the receiver side, packets need to be queued in a jitter-buffer before being unpacked, decoded and played out, to ensure a coherent speech signal for the receiver. Due to different routes that packets take through the network, packets may be missing or arrive later. The jitter-buffer tries to recollect all packets in the sent order. After a predefined time, the not yet arrived packets are usually dropped. In that case, packet-loss occurs. Some jitter-buffers are adaptive in their buffering time. Finally, the packets need to be unpacked and decoded before actual play out." (Schoenenberg, Raake, Egger and Schatz, 2014, pp. 1–2)

It is not surprising, either, that researchers both assume and find this process to cause problems for users faced with unexpected or unpredictable waiting time. Because of the finely tuned mechanism of the turn-taking machinery, the occurrence of technology-generated delay poses a number of questions concerning timing and maintaining order and, thus, achieving mutual understanding in remote interaction. Much previous research that addresses turn-taking during temporal lag, however, appears to look at delay largely from a perspective that approaches – although often implicitly – the phenomenon as a limitation or liability that determines participants' interpretations and actions as well as the course of their mediated talks. Typically such work has been either concerned with the effects of delay on communication or with the applicability and practicability of tools that make delay visible in interaction in order to improve participants' experience (e.g., Gutwin et al., 2004; Murai et al., 2005; Kawashima et al., 2008; Schoenenberg, Raake and Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014). In the following section I take a closer look at recent literature on delay in distributed audio- and video-based interaction.

2.2 Previous research on turn-taking and transmission delay

As indicated above, some work on technology-generated delay has been particularly interested in its effects on conversation, specifically on turn-taking. In two recent studies, for example, the "destructive impact delay may have on conversations" (Schoenenberg, Raake, Egger and Schatz, 2014, p. 5) has been explored in highly controlled experimental laboratory settings (Schoenenberg, Raake and Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014). In both studies the researchers added transmission delay of varying length to the connection in induced, "close-to-

natural" two- and three-party teleconferences. When co-participants' recorded streams of talk were compared with each other, it was found that the longer the delay lasted, the greater was divergence between sites (called "perceived realities"; Schoenenberg, Raake and Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014). In other words, the gap between the actual and the perceived behavior of a participant grew in proportion to increasing lag. As a consequence of this, Schoenenberg, Raake and Koeppe (2014) and Schoenenberg, Raake, Egger and Schatz (2014) note, the structure of conversations changed. According to the authors, participants experienced more mismatch between contributions, such as "unintended interruptions", during enhanced delay. Similarly, Egger et al. (2012) report an increase of the "unintended interruption rate" and a decrease of the "intended interruption rate" in delayed VoIP talks in comparison to conversations without (or with less) transmission delay. During delay, a participant starting to talk at a transition relevance place (TRP) might be heard as self-selecting in the middle of an ongoing turn ("interrupting") at the remote site, while a deliberate attempt at interrupting is likely to go unnoticed, because it arrives too late, such as at a TRP. Such effects have not only been provoked in and described for audio-only experiments, but they were also observed in naturally occurring video-mediated interaction at work. Ruhleder and Jordan (2001) studied authentic videoconferences by utilizing qualitative interaction analysis; as in the aforementioned studies by Schoenenberg, Raake and Koeppe (2014), Schoenenberg, Raake, Egger and Schatz (2014) and Egger et al. (2012), remote sites were compared, allowing the researchers to uncover the impact of delay on conversation through an objective lens. They describe "unintended interruptions" as well, with participants perceived as starting to talk in the middle of ongoing turns, and they discuss occurrences of inadvertent pauses, which caused interlocutors to produce unrelated and misplaced turns or "unnecessary additions" (p. 127). In addition, they found listener activities (or backchannels) to arrive late at the receiving site, which they argue occasionally disrupted the flow of ongoing talk or led speakers to modulate their turns in certain ways.

In addition to these apparent structural implications of delay for mediated talk, Schoenenberg, Raake and Koeppe (2014) found that participants rate others' personalities more negatively (e.g., as less friendly, less active, or less achievement-striving) when transmission delay was added to the conversation. Because participants in their study perceived no differences in the quality of the talks, Schoenenberg, Raake and Koeppe (2014) inferred that delay goes mostly unnoticed. Instead, they argue,

participants fell victim to misattributions, in which the consequences of technological shortcomings are ascribed to conversation partners' personality or disposition (Schoenenberg, Raake and Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014). Ruhleder and Jordan (2001) note as well that participants often do not become aware of delay and instead experience unexplainable trouble leading to feelings of unease and ultimately affecting trust and confidence (see also Friebel et al., 2003). Indeed, this relates to the fundamental expectation of shared experience in social interaction. According to Heath and Luff (2000), remote co-participants adhere to this "reciprocity of perspectives" with regard to the accessibility of visible conduct:

> "The technology therefore provides physically distributed individuals with incongruent environments for interaction. What I see is not what you see, and I am unable to see how you see me and the actions in which I engage. Despite this incongruity, individuals presuppose the effectiveness of their conduct and assume that their frame of reference is 'parallel' with their co-participant's. They presuppose, for the practicalities at hand and their mutually co-ordinated activity, that their image of the other is congruent with the other's image of them. [...] this presupposition of a common frame of reference, a reciprocity of perspectives, is a foundation of socially organised conduct." (p. 198)

However, some research results seem to suggest that interactants in fact do recognize technology-generated time lag. Egger et al. (2012) observed participants adapting to long transmission delay. Fischer and Tenbrink's (2003) ethnographic study of a first-time video-based meeting between researchers revealed the parties' creative ways to manage turn-taking during time lag (using a hand signal). Rintel's (2010) conversation analysis of video-mediated interactions between couples includes a case of participants detecting and playfully dealing with delay.

Problems of the kind found in the studies mentioned above (i.e., unintended interruptions or confusion between conversation partners) have been taken up by more practical work and served as impetus for developing tools that reveal delay to participants during interaction. The aim is to enable conversationalists to coordinate their turns in audio- and video-mediated talk despite time lag and thus to positively influence user experience (Gutwin et al., 2004; Murai et al., 2005; Kawashima et al., 2008). Kawashima et al. (2008), for example, have introduced and tested what they call "visual fillers" for videoconferencing, which is a tool to mark temporal gaps caused by

delay. Murai et al. (2005) as well as Gutwin et al. (2004) have also proposed and presented different tools to make transmission delay visible to participants. Through simulations, all three of these research projects found visual representations of delay to reduce difficulties in adjusting contributions.

There is, then, relatively consistent insight into how transmission delay influences mediated conversation, especially structurally. However, this knowledge represents mostly an outside perspective, and – particularly in view of the research design preferred in previous work – it should be asked how much these findings ultimately reflect co-participants' lived reality. Although others have already addressed the question of how conversationalists themselves approach and handle delay in interaction (see Rintel, 2010), this issue remains understudied. Examining delay from an emic perspective may provide insights into the role technology and mediation play for distributed participants, especially when faced with trouble.

3. Data and Methodological Considerations

Previous studies (especially on transmission delay) underline the curious fact that in technology-mediated situations co-participants interact with each other through different conversational realities. Against this backdrop, researchers maintain that in order to examine delay appropriately it is paramount to capture the unfolding talk in each of these realities and to hold respective recordings against each other (see Schulte et al., 2001). The aim is a detailed understanding of the phenomenon, as illustrated in Figure 1.
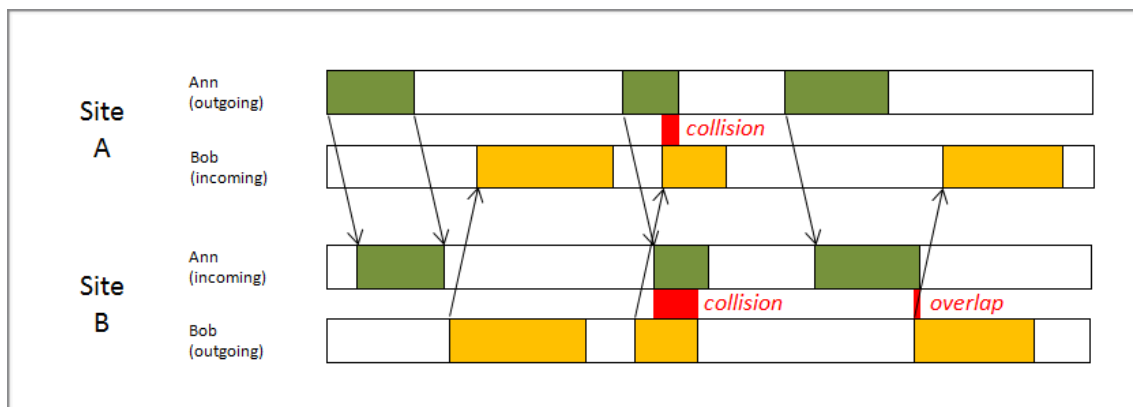

Figure 1. *Transmission delay (adapted from Murai et al., 2005; see also Friebel et al., 2003; Kawashima et al., 2008, p. 3587; Egger et al., 2012, p. 1321; Schoenenberg, Raake and Koeppe, 2014, p. 478; Schoenenberg, Raake, Egger and Schatz, 2014, p. 7; see Ruhleder and Jordan, 2001, p. 123 for authentic examples)*

This approach does, however, possess some shortcomings with respect to participants' relevancies. An external, all-inclusive examination of delay can be appreciated for uncovering its hidden sides, thereby making its impact on the timing of contributions visible and rendering studies of this kind particularly interesting for developers of communication technologies. However, such a view does not particularly promote a closer look at co-participants' work to accomplish interaction by means of mediation (and through possible transmission delay). Interactants *themselves* do not have access to what can be heard and seen (i.e., received) at the remote locations of their conversation partners. Yet they need to find ways to coordinate or match their different conversational realities in interaction in order to achieve and maintain coherence (see Hutchby, 2001b, who discusses the approach in Garcia and Jacobs, 1999, to chat-communication, including the unseen production of posts into their analysis). How do remote co-participants, then, create something close to a mutual, interactive reality – not despite but in reference to computer mediation? Attending to the problem analytically could include looking at time lag as it surfaces in interaction: What is treated as delay (or how do co-participants notice delay)? How and when is delay made relevant? How is delay dealt with?

Delay-in-interaction is at the heart of analysis in this paper, and it is studied from the recorded perspective of only one participant. In this way, information that is not available to this particular individual (i.e., what co-participants receive at the far end) cannot and will not be included in the investigation either. The objective is to enable an analysis that is close to real life and to uncover underlying expectations involved in the detection of potential transmission delay as well as the practices which are employed to address and handle the trouble.

This paper draws on authentic technology-mediated interactions between dispersed employees (and subcontractors) of a medium-sized company in Finland. The participants' work can be characterized as knowledge intensive; their remote meetings (held without the moderation of a chairperson) involve activities such as planning educational events, writing, acquiring information and programming. In collaborating (real-time) from a distance, co-workers utilize a desktop videoconferencing platform that supports audio, video, chat, simultaneous collaborative writing, and viewing and modifying documents. Figure 2 shows a distorted screenshot of the environment. In this case, the *video pod*, containing the video feeds of the participants, is on the left. The *share pod*, allowing the sharing of documents, collaborative writing and other

activities, can be seen in the middle. The *attendees pod*, which indicates who is present (sometimes not everyone participating uses a webcam), as well as the *chat pod*, enabling simultaneous text-based conversations between co-workers, are on the right. Note, however, that the positions of the pods can be changed by the hosts of the videoconference.
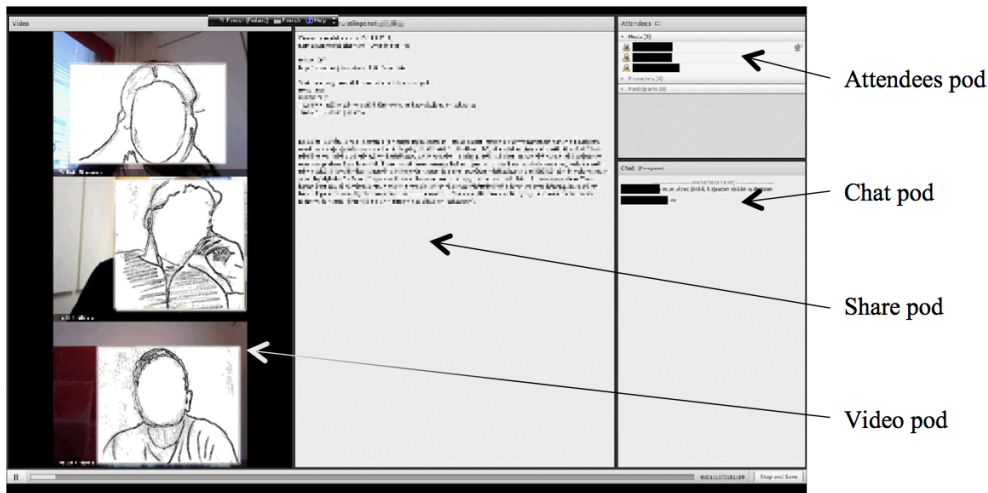
Figure 2. *Screenshot of the environment (participants, from above: "Kaarina", "Erik", and "Jyrki")*

In sum, the data consist of 34 remote work meetings recorded in April, May and June 2012. The excerpts presented here are extracted from three of these talks. The number of interlocutors ranged from two to five, and their meetings lasted from 16 minutes to approximately 2 hours. Apart from one participant ("Jyrki"), everyone used headphones.

The conferencing platform allows users to capture their interactions automatically, which was utilized in data collection: one of the participants – in the cases presented here it has been "Erik" – initiated the recording at the beginning of the meetings. All participants gave their consent prior to data collection. The automated nature of recording in this particular setting made it possible to obtain the data almost unobtrusively, because it required no physical interference in the form of, for example, setting up and operating video cameras and tape recorders (see Olbertz-Siitonen, Siitonen and Valo, 2014). It yielded data which depict the interactions exactly as they unfolded on screen. What can be observed in the recordings is equivalent to what the respective (recording) interlocutor could see and hear via the group support system, including his or her own contributions (Olbertz-Siitoen et al., 2014). Since all the

extracts presented in this article stem from talks Erik recorded, the examples portray his perspective (i.e., his conversational reality).

At the same time, the data do not provide direct insights into any side activities parallel to the virtual meetings, such as surfing or answering e-mails, nor do they include nonverbal behavior or spatial features beyond the camera angles at the different sites (as visible in the video feeds). However, because this kind of information was not available to the participants either, it bears no meaning for analysis unless it was made relevant – and thus witnessable – by the interlocutors themselves.

All excerpts have been transcribed using conventions proposed by GAT 2 (Selting et al., 2011; see appendix). A free translation is placed beneath each line of transcribed talk. The names of all participants as well as cues possibly allowing an identification of the company have been replaced.

The study presented here is exploratory in nature. In the following sections I examine several cases extracted from the data set described above, all of which illustrate mismatches similar to the kinds described in previous accounts of time lag, such as initiations of talk during ongoing contributions and long stretches of overlapping talk. However, apart from the first passage, which has been included in reference to Ruhleder and Jordan's (2001) work and functions as analytical starting point, the fragments were selected primarily on the basis of Erik's explicit displays. They constitute incidents that are observably oriented to as problematic and that are treated as being a result of transmission delay.

4. Analysis

Ruhleder and Jordan's (2001) detailed analyses not only show that lag impacts on the timing of contributions, but they also reveal the subtle ways in which these misplacements are consequential for the achievement of shared meaning. Their examples, depicting the different perspectives of both involved interlocutors, illustrate that remote participants interpret pauses, late responses and delayed feedback which are demonstrably caused by transmission delay as meaningful; interactants, then, react accordingly with, for example, additions and mitigations, treating the symptom but not its source. The underlying problem, according to Ruhleder and Jordan (2001), is that in most cases "participants cannot identify the true source of the trouble" (p. 132). Indeed,

with its focus on only one of the remote sites, the approach chosen here entails the same problem. It does not allow a definite identification of delay from the analyst's perspective, and pauses or hitches and adjustments in turn construction therefore cannot be definitively associated with lag. However, there are some indications in the data at hand that are reminiscent of Ruhleder and Jordan's (2001) observations. The following is a case in point (Extract 1). The example, involving two participants, starts with Erik specifying and demonstrating on screen (available to his colleague, Jaana, via the share pod) how the categorization of questionnaire data could be done in practice using Excel.

```
Extract 1 (H-sem-plan6 33:55-34:05 Jaana, Erik)

1    ER:  eli mä e- (-) eli m NAPpaisin
              so I e- (-) so I'd pick
2         [tÄÄltä näitä (hAvaintoja) mitä- ((cursor moves to comments
          displayed on screen))]
              [from here these (observations) which-]
3    JA:  [°hh eli SILloin tOIsin sanoin
                                   ]
              [°hh so then in other words         ]
4         me Emme kÄYttäisi- (---) ((JA suspends all movements))  joo.
              we wouldn't use- (---) yes.
5         °h sIlloin me [Emme] KÄYttäisi edes sEmmosta kÄsitettä
              °h then we [wouldn't] even use such concepts
6    ER:                [hm? ]
                        [hm?    ]
((...))
```

At line 3 Jaana begins to speak in overlap with Erik's ongoing turn (line 2). By exerting the phrase "so then in other words" and by offering a candidate understanding (see Heritage, 1984b), she marks her contribution as an understanding check, indicating some trouble with prior talk and initiating repair. After both speak simultaneously for a moment, Erik ceases to talk. However, he only stops after employing a relative pronoun ("which"), and thus before his turn reaches the next point of possible completion. Through these activities he not only shows an orientation to and restores the basic order for conversation (i.e., one speaker at a time), he also displays that he intended to go on. Jaana's turn, on the other hand, is not marked as problematic or turn-competitive, and she continues to speak without signs of experiencing overlap, until when, at line 4, she suddenly stops in the middle of her contribution for no apparent reason (i.e., as observable through the recorded perspective). Her movements come to a halt, and, while looking at the screen, her expressions as well as the position of her head remain almost steady for approximately one second. Next she produces a short

"yes" with falling intonation, and she starts to move again (line 4). This is followed by a reformulation of what she was just saying (line 5).

Jaana's activities thus include (1) no witnessable recognition of overlap while speaking simultaneously with Erik, (2) an abrupt termination of talk mid-turn, which suggests that now she experiences overlap and orients to 'one at a time', even though Erik does not talk anymore, (3) a display of listening, and (4) a reformulation, which can be seen as yet another indicator for preceding overlap on her side of the interaction (on repeats after overlap, see Schegloff, 2000, p. 34). While these features seem to match Ruhleder and Jordan's (2001) observations and therefore could be attributed to delay from an outside perspective, it also needs to be kept in mind, of course, that analyses of such incidents can only remain indicative. In any case, this example demonstrates that an examination of mediated talk through the eyes of only one participant can provide a sense of the work involved to match different conversational realities, even if time lag is not explicitly addressed.

One explanation for this occurrence not being treated as a result of transmission delay by the participants themselves may be, as Ruhleder and Jordan (2001) argue, that such hitches are not perceived as technologically induced. However, in a number of cases throughout the data at hand, similar mismatches are, in fact, treated as caused by lag, which implies that co-participants do, at least at times, consider the option of transmission delay and its possible impact on their talks. Such incidents appear to be one-sided in the sense that not everyone involved initially displays awareness of any problem at all. In other words, even though participants attend observably to the local allocation of turns on their ends, they might still generate repeated or long overlaps and long gaps or even lapse. Occasionally, these hitches are made interactively relevant and resolved by explicit, meta-communicative means. Participants account for and treat them as caused by technical trouble and, more specifically, by transmission delay.

More precisely, in the recorded meetings, phenomena that may evoke reference to transmission delay in interaction are characterized by sequentially incongruous, yet unmarked contributions. Structurally, they include *prolonged overlap*, *gap and sequential disarray*, and *missed attempt at turn-taking*. All of these constitute, in Erik's conversational reality, others' mistimed recognitions of ongoing and new activities. Building on and expanding Ruhleder and Jordan's (2001) observations, occurrences of these kinds are analyzed in detail in the following sections.

4.1 Prolonged overlap

Participants occasionally treat long stretches of overlapping talk as possibly related to technology-generated delay. The following extract illustrates such a case (Extract 2). The meeting, which is about planning an internal development workshop for the company, involves three participants, Kaarina, Jyrki, and Erik. At this point during the talk the three co-workers are concerned with scheduling the event:

```
Extract 2 (H-event-planning2 05:20-05:57 Kaarina, Jyrki, Erik)

1   JY:  ((...))
2        ((sniffles)) pUOltoist pÄIvää on ihan hyvä KOMpromissi.
             one and a half days is quite a good compromise.
3   (---)
4   JY:  MUN mIElestä siis. et-
             in my opinion at least. so-
5   (--)
6   ER:  ah (mutta) mut EIkös MEIllä ollu se Ajatus
             ah (but) but didn't we have this idea
7        että pErjaintai [on vAAn sitte nIIlle kEtkä haluaa jÄÄdä. ]
             that Friday [is only for those who want to stay.           ]
8   JY:                  [en tIIä kAnnattaaks (sitä) edellist iltaa]
                     [I don't know if it's worth to (really include) this]
9        (OIkeesti Ottaa).
             previous night.
10  (5.0) ((ER looks up and scans the screen, mouth open))
11  ER:  nYt tota mä en oo vArma VIIveistä.
             now well I'm not sure if there is a delay.
12       koska tUntuu että menee HASsusti pÄÄllekkäin=
             because it seems as if we are overlapping in a funny way.
13       KUUletsä jYrki mUa nYt.
             do you hear me now Jyrki?
14  (2.0)
15  JY:  ((nods)) (1.5) no NYT mä kUUlen.
             ((nods)) (1.5) well now I hear you.
16       mut mä rupeen hAvaitsee että mUn puhe tulee hIrveen Isolla
             but I start to notice that my talk is horribly
17       VIIvellä vAI.
             delayed or?
18  (2.0)
19  ER:  joo.
             yes.
((participants try to fix the problem: Erik lowers quality of the
tool; joking about the quality of their talk))
```

Consider lines 10 to 13, which mark the recognition of a problem and contain reference to transmission delay as its possible cause. Erik observably suspends his current activities during the pause at line 10 (by looking up, mouth slightly open, and scanning the screen), treating the current course of events as unusual. Next he initiates a meta-communicative sequence with an evaluation of what is going on. He brings up

the possibility of delay and accounts for his interpretation with reference to their talk "overlapping in a funny way" (lines 11–12).

Erik's assessment becomes tangible when examining Jyrki's actions as they unfold on Erik's side of the interaction. From the recorded perspective, the passage initially proceeds in an orderly fashion. At line 4, Jyrki displays an orientation to the option to continue his turn unless someone else self-selects at a TRP (i.e., rule 1c of the turn-taking system; Sacks et al., 1974). At line 6, Erik self-selects after Jyrki ceases talking. Erik's question is clearly directed at Jyrki's contribution, indicating a problem with and opposition to his recommended solution of "one and a half days" (which included that the event would start only Thursday afternoon and not in the morning; not shown in the transcript). However, some moments into Erik's ongoing turn (line 7), before it reaches a place of transition relevance, Jyrki begins to speak once more (line 8), which causes them to talk in overlap. Both parties continue talking for a relatively long time until they finish their turns almost simultaneously. Jyrki, however, neither marks his start as competitive nor treats the emerging overlap as problematic; instead he continues talking without signs of trouble, that is, the volume and pitch of his utterance remain unaltered (see French and Local, 1983; Selting, 1995; Schegloff, 2000), and his expression does not change. In addition, his contribution does not display any kind of relation to or recognition of Erik's question at lines 6–7, but rather appears to be a continuation of his prior talk. Jyrki had previously marked his contribution as complete (lines 4–5; on the Finnish *että/et* in turn-final position, see Koivisto, 2011). It seems as if Jyrki is not (yet) aware of Erik taking the turn at this point and is instead faced with a long pause after completion. He thus orients to rule 1c of the turn-taking system yet another time, and continues his turn after, from his perspective, none of his colleagues self-selects.

Erik's treatment of Jyrki's newly started talk at line 8 does not, interestingly, mark it as an interruption. Rather, Erik orients to Jyrki's onset as possibly unintentionally and unnoticed mid turn. This is an observation of fundamental relevance, because, as mentioned above, some authors studying transmission delay find an increase in "unintentional interrupting" (e.g., Egger et al., 2012; Schoenenberg, Raake and Koeppe, 2014; Schoenenberg, Raake, Egger and Schatz, 2014; see also Ruhleder and Jordan, 2001), which implies severe issues with regard to unwarranted moral blame. Based on the data studied here, this does not, however, reflect participants' orientations in this particular setting. Indeed, by bringing up transmission delay, blame is explicitly diverted to technology (see also Bowers, Pycock and O'Brian, 1996), which

coincides with previous findings on technological mediation as conversational resource for participants (Rintel, 2010; 2013a; 2013b; see Hutchby, 2001).

On a related note, Erik's word choice at lines 11–12 is worth paying attention to. His "I'm not sure" and "it seems" in relation to transmission delay indicate an orientation to the asymmetry of co-participants' interactional experiences from another angle. While mediation provides and is used here as a possible explanation for mismatching turns, it also prevents participants from knowing for sure. In addition, this orientation is highlighted in line 13 with Erik's approach to testing the audio connection ("do you hear me now Jyrki") and his inquiring about his co-worker's perspective. He observably attends to the fact that he has no access to the others' realities and that he cannot identify technology as the true trouble source with certainty.

After this episode of referring to transmission delay, Erik indicates that he is going to lower the current quality of the videoconference (presumably trying to increase the speed of transmission). Following this statement, the co-participants excessively joke about the ("already low") quality of their talk in general. About one-and-a-half minutes later, delay surfaces again, which is illustrated in the following example (Extract 3). Now Erik treats the occurrence of gap as well as sequential disarray as evidence that the connection remains impaired.

4.2 Gap and sequential disarray

The passage begins with Erik assessing the potential delay (line 1). However, he marks this evaluation as vague. In addition to the "apparently" in line 1, he expands with a tag question, downgrading his assertion even more and seeking confirmation ("didn't it get better", line 2; see Heritage, 2012). Again, similar to the previous example, Erik observably orients to the implications of the asymmetrical nature of technology-mediated interaction, in addition to refocusing on the quality of the connection:

```
Extract 3 (H-event-planning2 07:36-7:51 Kaarina, Jyrki, Erik)

((...))
1   ER:  ilmeisesti nyt PArani tää vI vIIve.
            apparently now the de delay got better.
2        EIkö PArantunut.
            didn't it get better.
3   (--)
4   ER:  tai siis (--) (HENkinen) [vIIve ] <<dim> eli lYheni.>
            or actually (--) that is the (mental) [delay] got shorter.
```

```
5    JY:                              [kYllä;]
                                      [yes;  ]
6    (3.0) ((ER smiles broadly))
7    ER:  jEpp. ((leans back))
          yep.
8         hEI eli [nYt tota-]
          hey so [now well-]
9    JY:          [he he   ] he he-
                  [he he   ] he he-
10   (--) ((ER lifts eyebrows and lifts himself up))
11   ER:  ei on sitä VIEläkin pAljon nÄköjään.
          no it still seems to be a lot.
((Erik asks Jyrki to reconnect; Jyrki reconnects))
```

Erik's re-assessment at line 11 ("no it still seems to be a lot") has the following trajectory: at line 4 Erik continues with a humorous remark referring to participants' prior jokes on the quality of their talk. He begins to smile broadly, highlighting his 'clarifying' addition ("or actually" in line 4) also nonverbally as a jest. However, neither Kaarina, now talking to somebody else at her location, nor Jyrki, react. After a pause of three seconds (line 6), during which his expression remains the same (i.e., smiling broadly), Erik stops smiling, he leans back and produces a discontinuity marker ("yep"). Through these activities, Erik marks the end of the delay-joke-intermezzo, and he displays an orientation to move on. This is underlined by his subsequent "hey so now well" at line 8, which indicates an incipient shift toward another topic. By continuing in this way Erik implies that the delay indeed did "get better", and at the same time he lets the joke pass, so to speak: at least Kaarina, visibly involved with side activities, can be assumed to not have heard the humorous remark; Jyrki's silence on the other hand allows for all kinds of (possibly problematic) interpretations, which are thus mitigated.

Partly in overlap with Erik's utterance at line 8, however, Jyrki leans back and begins to laugh loudly (line 9). This in turn is followed by a short pause during which Erik, who stopped talking in the middle of his turn (line 8), lifts his eyebrows and shifts (upward) in his chair (line 10). Erik then – at line 11 – resumes talking. He remarks that "it" still is, apparently, "a lot". He does not explicitly mention delay, but his previous activities (stopping to talk, lifting eyebrows, etc.) in combination with the "no" at the beginning of his statement, which does convey an answer, work to connect his assessment with prior talk (rendering this a corrected version of his previous evaluation, treating it, in light of Jyrki's laughter at line 9, as premature). Indeed, Erik next asks Jyrki to reconnect, highlighting the connection as in need of remedy.

Erik's conclusion that there is still delay reveals an orientation to or expectation of a sequentially correct place for responding laughter: next to the laughable (see Glenn,

2003). In light of this, Erik's displayed understanding of what is going on here becomes comprehensible. In effect, his assessment not only marks Jyrki's reaction as too late, but it also accounts for its – otherwise socially problematic – misplacement by reference to flawed transmission, relying on technology as an explanation.

Furthermore, Erik not only brings up time lag, but he treats it as prolonged ("a lot"; note, however, the use of the phrase "it seems") and as worth mentioning again (line 11). This is also interesting with regard to Jyrki's "yes" at line 5, to which the participants do not attend as problematic. It appears relatively close to Erik's tag question at line 2, and can be taken as a more or less timely response or at least as a sign of significantly reduced delay. Jyrki's laughter, however, now occasions considerable sequential disarray from Erik's recorded perspective. It starts after a long pause at a place that clearly marks the talk it possibly relates to as out of reach (a transition is underway), a place which, in addition, no longer conveys any irony. Again, as in the previous example, Jyrki at first does not display recognition of any problem at all, emphasizing that this kind of disorder in video-mediated interaction is a one-sided experience.

Another case of accounting for gap and sequential disarray by reference to transmission delay occurs at the beginning of a remote meeting between two participants, Erik and Loviisa (Extract 4). The purpose of their session is to plan for a meeting abroad, which both are going to attend. The excerpt begins with Erik's greeting at line 1 (*moro-*):

```
Extract 4 (H-planning meeting abroad1 00:00-00:40 Erik, Loviisa)

1    ER:  mOro-
              hiya-
2    (-)
3    ER:  pistin NAUhoituksen pÄÄlle kun tota (.) tÄÄ on tää (-)
              I put the recording on because (.) this is this (-)
4         <<smiling> VUOrovaikutustutkimushuone,> ((smiles))
              communication research room, ((smiles))
5    (5.0) ((LO continues looking at screen; expression remains same))
6    ER:  <<slightly smiling> cAn you HEAR me.> ((finnish accent))
7    (---) ((LO begins to smile slightly))
8    LO:  AH::; (-) Okay. (--) hihi. ((Erik stops smiling))
9    (---)
10   LO:  [hm ]
11   ?:   [( )]
12   (9.0)
13   ER:  YKsi- [KAKsi-     ]
              one- [two-           ]
```

```
14   LO:        [KUUluuko hY]vin. (-) vai-
               [can you hear me] well. (-) or-
15   (1.0) ((LO frowns; ER looks slightly to the side, mouth open))
16   ER:  tässä on Ilmeisesti JÄRjetön vIIve.
               there is apparently an incredible delay.
17   (4.0) ((ER looks back at the screen; video pod disappears))
18   ER:  ups- SORi. EI ollut TARkoitus.
               ups- sorry. I didn't mean to.
((recording stops))
```

Some 30 seconds into the beginning of this meeting, at line 16, Erik brings up possible transmission delay. His analysis of what is going on reveals an orientation to conversational turn-taking practices, relevant nexts, sequential order as well as the specifics of the setting (technology-mediation), and can be reconstructed as follows.

After his greeting (line 1), which is succeeded by a short pause, he continues by informing Loviisa that he started to record the meeting (lines 3–4). Towards the end of this unit he begins to smile. Following a long pause of approximately five seconds, during which he slowly reduces the smile, while Loviisa's expression, visible posture and direction of gaze remain the same, Erik begins to talk again, displaying an orientation to the option to resume after no one self-selected at a TRP (rule 1c; Sacks et al., 1974; line 6). However, there is more to his ensuing utterance: by asking Loviisa (in English), whether she can hear him, Erik suggests that Loviisa should react, that is, by way of a receipt or appreciation regarding the recording. In doing so he not only marks the previous pause as too long, as an absence, he also provides a possible explanation for why his colleague does not respond, immediately directing repair initiation at problems with the technology. More precisely, rather than turning to social trouble, technological issues are made accountable for the conversational discontinuity (see also Rintel, 2013a; 2013b): Loviisa cannot hear him due to an acoustic problem. Looking at the preceding activities, this displayed interpretation of what is likely happening here stands to reason. Up to this point Loviisa has not reacted or replied to Erik's attempts to start the meeting in any way (neither verbally nor nonverbally).

Now, at the TRP at the end of line 6, Loviisa's expression starts to change, she begins to smile slightly, and she produces a prolonged change-of-state token (*AH::;*) (see Heritage, 1984b), followed by a ratification (*Okay.*) and a short giggle. In the course of her utterance Erik now ceases to smile completely. This is followed by a pause, during which both participants look at the screen (Erik's eyes move intermittently, as if studying the screen he is looking at; line 12). After approximately 9 seconds Erik ends the pause by counting to two (*YKsi KAksi*), which marks another attempt to ascertain his audibility at Loviisa's site (line 13). His counting is reminiscent

of a sound check, which further indicates (also in light of his eye movements and this being the second time he explicitly attends to "hearing") that he has started to work on fixing the issue and that he is testing for any improvement.

In this way, Erik observably orients to sequential disarray. With his "can you hear me" at line 6, Erik produced a first pair part of an adjacency pair (see Schegloff and Sacks, 1973). By asking he made an answer relevant. This particular question, however, does not allow for a NO, as such a response would contradict itself and rule out the option of 'not hearing'. Rather, YES or no reaction at all (providing a negative answer) becomes expectable here. Loviisa's contribution at line 8, therefore, cannot be related to Erik's preceding inquiry – although indirectly her turn of course provides an answer: the reactive nature of her utterance shows that she does, in fact, hear him. Both the change-of-state token and ratification, however, make sense with reference to Erik's earlier information that this is the "communication research room".

At line 14, partly in overlap with Erik's test-counting, Loviisa takes up the issue of not hearing one another well. After that she frowns, displaying an awareness of a potential problem. At the same time Erik turns his head slightly to the side, mouth open, before he ultimately concludes that they may be dealing with severe delay (line 16) – again orienting to asymmetrical accessibility by downgrading his assessment ("apparently"). He nonverbally and verbally treats the lag as unusually long. By bringing up transmission delay now, Erik orients to Loviisa's question at line 14, which mirrors his much earlier "can you hear me" at line 6, as misplaced.

Erik's reference to lag comes up after the occurrence of several long pauses and indications of mistimed responses: Loviisa's disconnected receipt at line 8 as well as her late question at line 14. By making delay relevant Erik not only treats the issue as noteworthy, as something that needs the participants' attention, but he also provides an account for the occurrence of silences and mismatching turns (he thus ultimately ensures that both of them are *not* "hearable as 'being silent' in terms of the turn-taking machinery", Schegloff and Sacks, 1973, p. 295). Technology is not only made accountable, it is used as the first best explanation for trouble.

4.3 Missed attempt at turn-taking

In addition to long overlap, gap and considerable sequential disorder, a participant may treat recognition of newly initiated talk as late and account for such occurrences as a result of technology-generated time lag. The following instance is taken from a two-party talk (Extract 5). Jaana and Erik have met to plan a seminar for a customer

(see also Excerpt 1). At lines 1–5 Jaana continues with her current activity of exploring options for organizing a specific exercise:

```
Extract 5 (H-sem-plan6 21:28-22:02 Jaana, Erik)

1    JA:   °h ja meiän täytyy kAtoo et mitä ME tehdään tossa Omassa (-)
           °h and we have to look at what we are doing in our own (-)
2          ku (.) pUOlen tunnin osuudessa.
           ku (.) half an hour segment.
3          °hh että Onko se sItten jOtenkin sellanen että (-) et me_ei
           °hh that is it then somehow such that (-) that we don't
4          mitään nYt s Isoa mEnetelmää kEksitään-
           come up with some big method now-
5          vaan että °hhh et lAItetaan (-) pOrukat niinku pUhumaan.
           but that °hhh that the people will be (-) asked to talk.
6    (--)
7    JA:   sIEllä (.) [pÖY  ]
           there (.) [at the]
8    ER:             [siis-]
                     [so-  ]
9    JA:   pÖYdissä.
           at the tables.
10   (-)
11   ER:   mEIdän
           our
12   (-)
13   JA:   rYhmissä.
           in groups.
14   (-)  ((JA turns head slightly down))
15   JA:   mItä,
           what,
16   (---)
17   ER:   tässä on pIkku[sen VIIvettä    ] (  )
           there is a [small delay] ( )
18   JA:                 [<<f> tAI tAI> nt-]
                         [or or nt-    ]
19   (-)
20   ER:   OOTta jaana.
           wait jaana.
21        (--)
22   ER:   mä hUOmaan [että] tÄssä on pIkkusen VIIvettä.
           I notice [that] there is a small delay.
23   JA:             [jo; ]
                     [yes;]
24   ER:   niin sEn takia KOMmentit ei aina osu kOhdalle;
           and that is why our comments don't always match;
25         °h [siis]
           °h [so  ]
26   JA:      [aha.]
              [aha.]
27   ER:   mä meinasin kYsyä siitä
           I wanted to ask
((...))
```

At line 5 Jaana's multi-unit turn (see Selting, 1998) reaches possible completion. However, she now expands her last point ("but that °hhh that the people will be (-) asked to talk."), several times, first at lines 7–9, adding "there" and "at the tables", and later at line 13, by saying "in groups". She thereby marks her last utterance as possibly vague, and, in a step-by-step manner, specifies what she meant by "the people will be asked to talk". At line 8, partly in overlap with her first addition, Erik begins to speak. However, Erik does not finish his utterance (especially with regard to intonation, which only falls slightly, his talk projects continuation), orienting to Jaana still adding more information. Although his construction is incomplete, with his "so our" he also displays (almost simultaneously) some kind of trouble with regard to Jaana's turn. In Finnish the particle 'siis' is used in turn-initial position as a preface to understanding checks (for understanding checks in Finnish, see Kurhila, 2006; see also Sorjonen, 2000; Sidnell, 2010, p. 118; see Schegloff et al., 1977, p. 368).

This unfinished request for clarification is followed by a pause (line 12), after which Jaana formulates a further addition "in groups" (line 13). Tilting her head downward, now she, in turn, initiates repair: by saying "what" with slightly rising intonation (line 15), Jaana indicates a problem in hearing. Erik, however, after a pause of approximately one second, does not return to the repairable, but instead gives an account, "there is a small delay" (line 17), by which he suggests that they are facing structural problems caused by the connection. The history of this displayed understanding can be briefly traced as follows.

Although Erik initiates a new turn (marked as an understanding check) at a TRP, Jaana continues, and her "in groups" does not readily relate to Erik's talk, but as mentioned above it instead constitutes a specification of "the people will be asked to talk". Her further activities show that she hears (also visibly, namely at line 14) Erik only after she made the third addition. In other words, from Erik's perspective the first indications that Jaana perceived him talking (tilting her head and asking "what") become available at a point that has passed the reference. This is addressed in another account, at line 24, when Erik displays what his assessment of the short delay is based on: "and that is why our comments don't always match".

As in the preceding example, Erik makes estimations with respect to the quality (length) of the transmission delay. Whereas previously he recognized a long (or "incredible") delay, now he describes it as short. This is perhaps worth mentioning in the light of studies on transmission delay finding that lag, despite its impact on the conversational structure, often remains hidden from co-participants (e.g., Ruhleder and

Jordan, 2001; Friebel et al., 2003; Schoenenberg, Raake and Koeppe, 2014). While this might indeed be the case, these examples demonstrate that delay nevertheless can become an issue for remote interlocutors themselves, which indicates that technology does not entirely fade into the background. Moreover, when it is brought up, it may be treated and assessed as a dynamic phenomenon in interaction.

5. Conclusions

This paper contributes to and complements research on temporal lag in technology-mediated communication by approaching the phenomenon as delay-in-interaction. Because the study aimed at tackling participants' relevancies, the analysis focused on one reality only. This choice was based on the argument that interactants themselves need to find ways to match their conversational realities without having direct access to what is experienced at the other end. Of course, it could be argued that the sides of the other participants are equally important for a thorough understanding of delay, and that in order to provide an appropriate picture of interaction distorted by an impaired connection, we need to see when a certain activity is perceivable at distributed locations, or when exactly a smooth transition in one location constitutes an overlap in another. However, Erik could only work with what he received via the group support system. He himself had no way of knowing, in real time, when his conversation partners actually said or did something. Adopting this perspective, the recording of Erik's conversational reality functions as a window that allows a view of the participants' understandings and solutions.

The examples presented here illustrate a basic orientation to sequential order, which matches known features of interaction between co-located conversationalists. The analysis reveals expectations with respect to the proper placement of actions and reactions. This fundamental order is what Erik witnessably attended to. Accordingly, he treated temporal mismatches between turns as problematic. However, at the same time he explicitly attributed such trouble to technological distortions – particularly when several structural problems occurred shortly one after another (Extracts 3, 4, and 5). The temporary suspension of ongoing activities in favor of technological remedy (i.e., lowering the quality of the transmission and thereby improving the speed of the connection, reconnecting, but also adjusting the pace of the talk by saying, for example, "wait") further highlights the relevance of (restoring) sequentiality for

continuing a meeting. In sum, the data clearly exhibit participants' simultaneous orientation to sequential order as well as to the particularities and eventualities of computer mediation.

More precisely, phenomena that evoked reference to transmission delay in this particular data set included mistimed activities characterized by prolonged overlap, gap and sequential disarray, and missed attempt at turn-taking. None of these activities were marked as exceptional (e.g., as turn-competitive, late, or otherwise problematic) upon their initiation, which constitutes an aspect of mismatch that has been neglected in previous accounts of delay, but may be consequential for the ways in which they are perceived and treated by co-participants (i.e., as intentional or unintentional). For Erik such occurrences worked as indicators that the connection was not working properly and as a reason to deal with them explicitly, that is, to bring up technological distortion to account for their appearance. This demonstrates that lag, including its possible structural consequences, is available to participants as something that may interfere with technology-mediated communication. This finding thus does not entirely support the claim by Ruhleder and Jordan (2001) that the "disturbances in turn-taking generated by transmission delays are particularly insidious because they raise no easily identifiable trouble flags, yet disrupt the mechanisms of turn-taking and repair" (p. 119).

Erik did considerable work to not to leave sequential trouble open for speculation. Particularly prolonged overlap, lengthy gaps and sequential disarray (consider, for example, the absent/late laughter in Extract 3) can have serious social meanings and implications. Accounting for them by referring to technology constitutes an important function. Technology does not merely enable remote interaction, but enters interaction in multiple and complex ways, and may be resorted to as interactional resource: "The fact that the conversation was technologically mediated is both the reason and the resource for this disambiguation" (Rintel, 2013a, p. 6/19).

Erik further oriented observably to the fact that he and his colleagues interacted without access to each other's realities. When he brought up delay, he employed phrases such as "I'm not sure", "there is apparently", "it seems to be" and "I notice", by which he marked his evaluations as vague and accounted for the asymmetrical access to technology-mediated interaction.

Because this study focused particularly on the explicit treatment of sequential trouble as technologically induced, future work could look at more subtle ways to deal with technological distortions. Cases such as the first example, which demand a closer

inspection of other cues, constitute a worthwhile point of departure to obtain a clearer picture of co-participants accomplishing virtual mutual reality. This might also include utilizing an approach that combines the overall view (e.g., Ruhleder and Jordan, 2001) with an emic participant's perspective.

6. Acknowledgements

7. References

Bowers, J., Pycock, J., & O'Brian, J. (1996). Talk and embodiment in collaborative virtual environments. *CHI, 96*, 58-65.

Brandt, A., & Jenks, C. (2013). Computer-mediated spoken interaction: aspects of trouble in multi-party chat rooms. *Language@Internet, 10*, article 5, 21 pages. Last retrieved December 13th, 2015, from http://www.languageatinternet.org/articles/2013/Brandt/brandt.jenks.pdf.

Deppermann, A. (2008). *Gespräche analysieren. Eine Einführung.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Egger, S., Schatz, R., Schoenenberg, K., Raake, A., & Kubin, G. (2012). Same but different? – Using speech signal features for comparing conversational VoIP Quality Studies. *Proceedings of the International Conference on Communications IEEE ICC, 2012*, 1320-1324.

Fischer, K., & Tenbrink, T. (2003). Video conferencing in a transregional research cooperation: turn-taking in a new medium. In J. Döring, W. H. Schmitz, & O. Schulte (Eds.). *Connecting Perspectives. Videokonferenz: Beiträge zu ihrer Erforschung* (pp. 89-104). Aachen: Shaker Verlag.

French, P., & Local, J. (1983). Turn-competitive incomings. *Journal of Pragmatics, 7* (1), 17-38.

Friebel, M., Loenhoff, J., Schmitz, H. W., & Schulte, O. A. (2003). „Siehst Du mich?" – „Hörst Du mich?" – Videokonferenzen als Gegenstand kommunikationswissenschaftlicher Forschung. *kommunikation@gesellschaft, 4*, Beitrag 1. Last retrieved December 13th, 2015, from http://www.soz.uni-frankfurt.de/K.G/B1_2003_Friebel_Loenhoff_Schmitz_Schulte.pdf.

Garcia, A., & Jacobs, J. (1999). The eyes of the beholder: understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction, 32* (4), 337-367.

Glenn, P. (2003). *Laughter in Interaction.* Cambridge: Cambridge University Press.

Gutwin, C., Benford, S., Dyck, J., Fraser, M., Vaghi, I., & Greenhalgh, C. (2004). Revealing delay in collaborative environments. *Proceedings of the Conference on Computer Human Interaction (CHI '04)*. ACM Press, New York, 503-510.

ten Have, P. (2007). *Doing Conversation Analysis. A Practical Guide.* London: Sage Publications.

Heath, C., & Luff, P. (2000). *Technology in Action.* Cambridge: Cambridge University Press.

Heritage, J. (1984a). *Garfinkel and Ethnomethodology.* Cambridge: Polity Press.

Heritage, J. (1984b). A change-of-state token and aspects of its sequential placement. In M. J. Aktinson, & J. Heritage (Eds.). *Structures of Social Action. Studies in Conversation Analysis* (pp. 299-345). Cambridge: Cambridge University Press.

Heritage, J. (2012). Epistemics in conversation. In J. Sidnell, & T. Stivers (Eds.). *The Handbook of Conversation Analysis* (pp. 370-394). West Sussex: Wiley-Blackwell.

Herring, S. C. (1999). Interactional coherence in CMC. *Proceedings of the 32nd Hawaii International Conference on System Sciences*, IEEE Press, 13 pages.

Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (2000). Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer-Human Interaction, 7* (4), 477-509.

Hutchby, I. (2001a). Technologies, texts and affordances. *Sociology, 35* (2), 441-456.

Hutchby, I. (2001b). *Conversation and Technology.* Cambridge: Polity.

Hutchby, I., & Tanna, V. (2008). Aspects of sequential organization in text message exchange. *Discourse & Communication, 2* (2), 143-164.

Hutchby, I., & Wooffitt, R. (2008). *Conversation Analysis.* Cambridge: Polity Press.

Jacobs, J. B., & Garcia, A. C. (2013). Repair in chat room interaction. In S. C. Herring, D. Stein, & T. Virtanen (Eds.). *Pragmatics of Computer-Mediated Communication* (pp. 565-587). Berlin: Mouton de Gruyter,

Jenks, C., & Firth, A. (2013). Interaction in synchronous voice-based computer-mediated communication. In S. C. Herring, D. Stein, & T. Virtanen (Eds.). *Pragmatics of Computer-Mediated Communication* (pp. 217-241). Berlin: Mouton de Gruyter.

Kawashima, H., Nishikawa, T., & Matsuyama, T. (2008). Visual fillers: Facilitating smooth turn-taking in video conferencing with transmission delay. *Proceedings CHI EA '08 Extended Abstracts on Human Factors in Computing Systems*, 3585-3590.

Koivisto, A. (2011). *Sanomattakin selvä? "Ja", "mutta" ja "että" puheenvuoron lopussa.* [Goes without saying? Finnish conjunctions *ja*, *mutta* and *että* in turn-final position] Helsinki: Unigrafia.

Kurhila, S. (2006). *Second Language Interaction.* Amsterdam/Philadelphia: John Benjamins.

Licoppe, C., & Morel, J. (2012). Video-in-interaction. "Talking heads" and the multimodal organization of mobile and skype video calls. *Research on Language and Social Interaction, 45* (4), 399-429.

Luff, P., Jirotka, M., Yamashita, N., Kuzuoka, H., Heath, C., & Eden, G. (2013). Embedded interaction: The accomplishment of actions in everyday and video-mediated environments. *ACM Transactions on Computer-Human Interaction (TOCHI), 20* (1), 1-22.

Markman, K.M. (2005). To send or not to send: Turn construction in computer-mediated chat. *Texas Linguistic Forum 48: Proceedings of the Twelfth Annual Symposium about Language and Society – Austin April 16–18 2004*, 115-124.

Markman, K.M. (2009). "So what shall we talk about": Openings and closings in chat-based virtual meetings. *Journal of Business Communication, 46* (1), 150-170.

Markman, K.M. (2013). Conversational coherence in small group chat. In S. C. Herring, D. Stein, & T. Virtanen (Eds.). *Pragmatics of Computer-Mediated Communication* (pp. 539-564). Berlin: Mouton de Gruyter.

Meredith, J., & Stokoe, E. (2014). Repair: Comparing facebook 'chat' with spoken interaction. *Discourse & Communication, 8* (2), 181-207.

Mondada, L. (2007). Operating together through videoconference: Members procedures for accomplishing a common space of action. In S. Hester & D. Francis (Eds.). *Orders of Ordinary Action: Respecifying Sociological Knowledge* (pp. 51-67). Abington, Oxon, GBR: Ashgate.

Mondada, L. (2010). Eröffnung und Vor-Eröffnung in technisch vermittelter Interaktion: Videokonferenzen. In L. Mondada, & R. Schmitt, R. (Eds.). *Situationseröffnungen: Zur multimodalen Herstellung fokussierter Interaktion* (pp. 277-334)*.* Tübingen: Narr.

Moore, R. J., Ducheneaut, N., & Nickel, E. (2007). Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds. *Computer Supported Cooperative Work (CSCW), 16*, 265-305.

Murai, K., Kimber, D., Foote, J., Liu Q., & Doherty, J. (2005). Mediated meeting interaction for teleconferencing. *Proceedings of the 2005 IEEE International Conference on Multimedia and Expo ICME*, 1436-1439.

Nilson, M., & Mäkitalo, Å. (2010). Towards a conversational culture? How participants establish strategies for co-ordinating chat postings in the context of in-service training. *Discourse Studies, 12* (1), 90-105.

Olbertz-Siitonen, M., Siitonen, M., & Valo, M. (2014). Naturally occurring data in the study of virtual teams in working life – Challenges and opportunities. *Työelämän tutkimuspäivien konferenssijulkaisuja 5/2014.* [*Proceedings of the work life studies conference 5/2014.*] Tampere: University of Tampere, 185-192.

O'Neill, J., & Martin, D. (2003). Text chat in action. *Group, 03*, 40-49.

Örnberg Berglund, T. (2009). Disrupted turn adjacency and coherence maintenance in instant messaging conversations. *Language@Internet* 6, article 2, 25 pages. Last retrieved December 13th, 2015, from http://www.languageatinternet.org/articles/2009/2106/Berglund.pdf.

Paolillo, J.C., & Zelenkauskaite, A. (2013). Real-time chat. In S. C. Herring, D. Stein, & T. Virtanen (Eds.). *Pragmatics of Computer-Mediated Communication* (pp. 109-133), Berlin: Mouton de Gruyter.

Rintel, S. (2010). Conversational management of network trouble perturbations in personal videoconferencing. *OZCHI 2010 Proceedings*, 304-311.

Rintel, S. (2013a). Video Calling in Long-Distance Relationships: The Opportunistic Use of Audio/Video Distortions as a Relational Resource. *The Electronic Journal of Communication, 23* (1&2), 19 pages. Last retrieved December 13th, 2015, from http://www.cios.org/EJCPUBLIC/023/1/023123.HTML.

Rintel, S. (2013b). Tech-tied or tongue-tied? Technological versus social trouble in relational video calling. *Proceedings of the Forty-Sixth Hawaii International Conference on System Sciences*, 3343-3352.

Rintel, S., Mulholland, J., & Pittam, J. (2001). First things first: internet relay chat openings. *Journal of Computer-Mediated Communication, 6* (3). Last retrieved

December 13th, 2015, from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2001.tb00125.x/full.

Ruhleder, K., & Jordan, B. (2001). Co-constructing non-mutual realities: delay-generated trouble in distributed interaction. *Computer Supported Cooperative Work, 10* (1), 113-138.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696-735.

Schegloff, E. (1984). On some questions and ambiguities in conversation. In M. J. Aktinson, & J. Heritage (Eds.). *Structures of Social Action. Studies in Conversation Analysis* (pp. 28-52). Cambridge: Cambridge University Press.

Schegloff, E. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society, 29* (1), 1-63.

Schegloff, E., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53* (2), 361-382.

Schegloff, E., & Sacks, H. (1973). Opening up closings. *Semiotica, VIII* (4), 289-327.

Schoenenberg, K., Raake, A., & Koeppe, J. (2014). Why are you so slow? – Misattribution of transmission delay to attributes of the conversation partner at the far-end. *International Journal of Human-Computer Studies, 72* (5), 477-487.

Schoenenberg, K., Raake, A., Egger, S., & Schatz, R. (2014). On interaction behavior in telephone conversations under transmission delay. *Speech Communication, 63-64*, 1-14.

Schönfeldt, J., & Golato, A. (2003). Repair in chats: a conversation analytic approach. *Research on Language and Social Interaction, 36* (3), 241-284.

Schulte, O. A., Friebel, M., & Klotzek, C. (2001). Aufzeichnung technisch vermittelter Kommunikation − das Beispiel Videokonferenz. *Gesprächsforschung − Online−Zeitschrift zur verbalen Interaktion, 2*, 222-242. Last retrieved December 13th, 2015, from http://www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2001/px-schulte.pdf.

Selting, M. (1995). *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation.* Tübingen: Max Niemeyer Verlag.

Selting, M. (1998). TCUs and TRPs: The construction of units in conversational talk. *InLiSt (Interaction and Linguistic Structures), 4*, 1-50. Last retrieved December 13th, 2015, from http://www.inlist.uni-bayreuth.de/issues/4/inlist4.pdf.

Selting, M. et al. (2011). A system for transcribing talk-in-interaction: GAT 2. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion, 12*, 1-51. Last

retrieved December 13th, 2015, from http://www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2011/px-gat2-englisch.pdf.

Sidnell, J. (2010). *Conversation Analysis. An Introduction.* West Sussex: Wiley-Blackwell.

Simpson, J. (2005). Conversational floors in synchronous text-based CMC discourse. *Discourse studies, 7* (3), 337-361.

Sorjonen, M. (2000). *Responding in Conversation. A Study of Response Particles in Finnish.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

8. Appendix: Transcription conventions (adapted from Selting et al. 2011)

<u>Sequential structure</u>

`[ ]` `[ ]`	overlap and simultaneous talk
`=`	fast, immediate continuation with a new turn or segment (latching)

<u>Inbreaths</u>

`°h`	inbreaths of appr. 0.2-0.5 sec. duration
`°hh`	inbreaths of appr. 0.5-0.8 sec. duration
`°hhh`	inbreaths of appr. 0.8-1.0 sec. duration

<u>Pauses</u>

`(.)`	micro pause, estimated, up to 0.2 sec. duration appr.
`(-)`	short estimated pause of appr. 0.2-0.5 sec. duration
`(--)`	intermediary estimated pause of appr. 0.5-0.8 sec. duration
`(---)`	longer estimated pause of appr. 0.8-1.0 sec. duration
`(0.5)/(2.0)`	measured pause of appr. 0.5 / 2.0 sec. duration (to tenth of a second)

<u>Other segmental conventions</u>

`and_uh`	cliticizations within units
`:`	lengthening, by about 0.2-0.5 sec.
`::`	lengthening, by about 0.5-0.8 sec.
`:::`	lengthening, by about 0.8-1.0 sec.

<u>Laughter</u>

`hehe` `hihi`	syllabic laughter

<u>Other conventions</u>

`((coughs))`	description of non-verbal (vocal) actions and events
`<<coughing> >`	…with indication of scope
`( )`	unintelligible passage
`(may i)`	assumed wording
`((...))`	omission in transcript

<u>Accentuation</u>

`SYLlable`	focus accent
`sYllable`	secondary accent

<u>Final pitch movements of intonation phrases</u>

`?`	rising to high

`,`	rising to mid
`‐`	level
`;`	falling to mid
`.`	falling to low

<u>Loudness changes, with scope</u>

`<<f>  >`	forte, loud
`<<dim>  >`	diminuendo, increasingly softer