Understanding technology use in face-to-face interaction:
A conversation analytic perspective

Florence Oloff & Joonas Råman

Prologi
– Viestinnän ja vuorovaikutuksen tieteellinen aikakauslehti

journal.fi/prologi/

Editorial

Understanding technology use in face-to-face interaction: A conversation analytic perspective

Florence Oloff

Professor
Leibniz-Institute for the German Language, Mannheim
Affiliated Researcher
University of Oulu
oloff@ids-mannheim.de

Joonas Råman

Lecturer
Centria University of Applied Sciences, Kokkola
joonas.raman@centria.fi

Technology use in face-to-face interaction

Our daily and professional lives are closely entwined with mundane technologies, such as laptops, smartphones, tablets, smart speakers, and other touch- or voice-based interfaces. More novel technologies such as virtual reality environments and social robots are becoming increasingly ubiquitous in various institutional and everyday settings as well. As our communication routines and interpersonal encounters become more reliant on technologies, our need to understand their effect on our social practices also becomes more apparent. While research and, in particular, media discourses on technologies tend to highlight technological innovations, their possible uses and potential risks (Leick, 2019, pp. 95–128), the impact of technologies, both novel and more mundane, on our communication routines in social interaction has not been fully explored. The contributions of this special issue will look at video data documenting how participants use and adapt to technologies in and for social interaction. This special issue therefore focuses on multimodal, i.e., both verbal and nonverbal, communication practices of participants in a selection of social settings, both day-to-day and institutional, where various technologies, both well-known and more innovative, are present.

To answer the central question of what the role of technology in face-to-face communication

is, the selection of studies included in this special issue addresses the variety of affordances (Hutchby, 2001) brought about by technologies to – among other things – changes in participation, novel communicative practices, and interactional problem-solving. The spectrum of roles played by technologies in the included studies is broad, ranging from participants simply using them, interacting with them, and – in the case of virtual reality – immersing themselves within them. While it cannot be denied that increasingly pervasive and immersive technologies bring new challenges and potential moments of trouble, the participants are also shown to treat these same challenges as opportunities for new ways of pursuing their interactional tasks. To say that technologies change face-to-face interaction – as the popular narrative often goes – provides a woefully one-sided perspective (with respect to, for instance, smartphones, see Avgustis, 2023; Oloff, 2021). The studies in this special issue contribute to a growing body of research interested in real-life, situated technology use.

The overall potential impact of technologies on interpersonal communication is naturally too vast to cover in a single special issue. Consequently, a set of more specific research questions was drawn. The included contributions answer, among others, the following questions:

- What kinds of new communicative practices does the use of technology foster in the context of face-to-face interactions?
- How is participation maintained, and enabled in various settings implementing technology and how are participants included or excluded from these settings?
- How are participants acquainted with new technologies, and what does this learning process look like?
- What is the role of technology in interactional problem-solving?

- How are expertise and institutional roles intertwined with technology-focused communication?

These questions are best answered by using naturally occurring data, that is, from social encounters which would have taken place even without the data collection. Furthermore, to understand how face-to-face interaction changes with the inclusion of technology, it is imperative to have a solid understanding of the basic systematics of talk-in-interaction (Stokoe et al., 2024). The requirements for naturally occurring data and an established understanding of the organization of face-to-face interaction largely dictated the selection of the research method of the contributions in this special issue; (multimodal) conversation analysis, which will be outlined in the following section.

A brief introduction to conversation analysis

This special issue contains studies that use a micro-analytic approach to social interaction, all of them inspired by conversation analysis. Conversation analysis (henceforth CA) has a distinct approach to empirical data and to how it can be used for understanding social interaction. As the journal Prologi—Journal of Communication and Social Interaction is interested in a multi-disciplinary approach to communication and social interaction, including possible foci on interpersonal relationships, public speaking, or communication training, we will now provide a short introduction to the specific framework used in this special issue.

CA has its roots in ethnomethodology, an empirical approach interested in discovering the production of social order by the participants, i.e., understanding the practices (methods) of

a specific group (ethno) that they use in their everyday private, institutional or professional lives (Garfinkel, 1967; Heritage, 1984). Instead of formulating assumptions and hypothesizing about social norms and communication patterns, ethnomethodology aims at revealing the orderliness of social conduct by identifying and describing it in natural settings. While sociological research frequently operates with pre-defined social categories and norms – or rather assumes their existence outside of and prior to interpersonal communication –, ethnomethodology requires a radically different researcher posture, as social categories and actions are seen as locally (re)produced by the participants themselves. CA has been heavily influenced by this approach, for instance, in that it also believes that social conduct is produced and understood by participants in its local context (indexicality), that context and action mutually adjust to and produce each other (reflexivity), and that participants are making their conduct intelligible and understandable for each other (accountability).

Initially a sociological endeavour, CA was founded in the 1960s and 1970s by Harvey Sacks, Emanuel A. Schegloff and Gail Jefferson. The interest of CA in uncovering the underlying "mechanisms" of social order is visible in the pioneering papers on conversational openings and closings (Schegloff, 1968; Schegloff & Sacks, 1973), turn-taking (Sacks et al., 1974), and repair in conversation (Schegloff et al., 1977). In line with its ethnomethodological roots, CA is primarily interested in naturally occurring interactions, meaning interactions that take a priori place without the researcher's involvement. CA researchers therefore prefer unconstrained, non-experimental settings and do not conduct interviews in which participants are asked to report on their communication practices, as these settings produce very different types of social conduct (Mondada, 2012). It should be noted, however, that CA scholars can work on interview and laboratory settings as well, if the focus is on how participants specifically interact in these task-driven contexts (see Kendrick, 2017). The primary data for conversation analysts are audio- and video-recordings, as they provide a convenient way to make data repeatedly available and shareable with other researchers (Sacks, 1984). Capturing the fleeting and complex nature of spoken discourse and visible conduct also allows for discovering small phenomena which, at the moment of the recording, might yet be unknown to the researcher:

> We can then come to see that a base for using close looking at the world for theorizing about it is that from close looking at the world we can find things that we could not, by imagination, assert were there (Sacks, 1984, p. 25).

Indeed, CA does not formulate hypotheses about social conduct that are then to be tested and verified. Instead, the analysis itself is inductive and data-driven, as it aims at discovering phenomena and recurrent patterns of social conduct (Sacks, 1984, p. 25). For this purpose, CA relies both on repeated viewings of the recordings and detailed transcripts of the recorded interaction. As conversation analysts assume that "[...] there is order at all points" (Sacks, 1984, p. 22), transcripts should be attentive to all details, including pauses, hesitation particles, inbreaths, vowel lengthening and repeated or recycled words, as these all meaningfully organize the interaction. While transcripts largely rely on standard orthography (Mondada, 2018a), specific transcription conventions have been developed that allow for the systematic and standardized annotation of, e.g., syntactic or prosodic features, and which furthermore take the temporal development of the conver-

sation into account (for instance, overlapping talk, see Jefferson, 1983; 2004a). The most popular transcription conventions in CA used for the transcription of talk are those developed by Gail Jefferson (2004b), while embodied conduct is frequently transcribed according to the conventions suggested by Lorenza Mondada (2018b; 2022). These conventions might differ according to different academic and linguistic traditions (see, e.g., the GAT2-conventions, Selting et al., 2009), however, they all focus on the precise rendering of the temporality of the social conduct (for an introduction to transcription practices, see, e.g., Hepburn & Bolden, 2017; Mondada, 2018b; Oloff & Hepburn, in press). CA transcripts are not simple renderings of a conversation's content, they are part and parcel of the analytic process. As analytic objects, they can therefore have different degrees of granularity, depending on the setting and phenomenon under investigation. For interactions mediated by or using technology, this means developing specific ways of representing both virtual and physical spaces, or how on-screen actions are entwined with the participants' audible and visible conduct (see, e.g., Haddington, Eilittä, et al., 2023). As all the contributions to this special issue show, different solutions for annotating technologized actions and integrating video stills have to be found, depending on the analytic focus.

By viewing the original data and transcripts, possibly adjusting and refining the latter, the researcher then looks for specific phenomena or patterns in the given interactional setting. This inductive process is not based on the researcher's subjective interpretations of the participants' conduct. Instead, it builds both on prior descriptions of systematic social conduct (such as the organization of speaker change, Sacks et al., 1974, or conversational repair, Schegloff et al., 1977), and the underlying fundamental principles of human social interaction (such as the constant aim to establish and maintain intersubjectivity, or the way actions unfold sequentially). Indeed, sequential organisation is at the heart of human social conduct (Schegloff, 2007), as actions (realized through speaking turns or embodied conduct) are formatted in a way to project a specific type of response, or, when following another turn, respond to the prior action in a specific way. A response thereby manifests how a prior action has been understood, and this understanding can then again be negotiated in the following turn. This reflexive relation between action formation and action ascription (Levinson, 2012) is based on the participants' interpretations as manifested in the subsequent turns, not on the analyst's individual assumptions on how a specific type of action should be formatted or responded to, nor on speculations about the participants' possible intentions (Deppermann, 2021). This next-turn proof procedure (cf. Sacks et al., 1974) allows to empirically anchor observations on systematic patterns in the data material, patterns that can manifest in the (non-)lexical material, syntactic structure, in prosodic format and/or embodied conduct positioned at a specific moment in relation to preceding or following actions. While observations of such interactional practices can be made based on one example (single case analysis), CA typically aims at building collections of cases, meaning examples from one or different events showing a similar type of phenomenon (see Sidnell, 2012). Collections in CA can be of varying size, depending on the specificity of the interactional practice under investigation, and the size of the data set. CA-informed studies therefore include detailed transcripts and analyses of these in the publications, as the description of a specific interactional practice should be made as comprehensible as possible. This allows for the checkability of the findings by other researchers and for their transferability

to other data sets, settings and languages. Being a micro-analytical and inductive approach, CA does not require a specific quantity of cases or statistical proof/check, and the analysis should always rely on a thorough qualitative analysis rather than simple quantifications of decontextualized phenomena (Haakana, 2002; Schegloff, 1993). However, conversation analysist can work on rather large data sets (Schegloff, 1968), and questions related to coding and quantitative methods are increasingly discussed within the field (Enfield et al., 2019; Pfeiffer et al, in press; Stivers, 2015).

CA as a method to investigate different interactional settings and phenomena

Though initially a sociological endeavour, CA has been adopted in other fields and academic communities interested in language and communication and can thus have different foci of research and even use different labels. Interactional Linguistics, for instance (Fox et al., 2012; Ochs et al., 1996; Selting & Couper-Kuhlen, 2001), aims more specifically at describing linguistic structures and grammar-in-use. Multimodal CA (or multimodal interaction analysis) works exclusively with video data and focuses on how audible and visible (i.e., multimodal) resources contribute to the organisation of social interaction, emphasizing the embodied and material dimension of social encounters (Deppermann & Streeck, 2018; Streeck et al., 2011). Especially the latter has also been informed by other scholars and studies interested in embodied conduct (for instance C. Goodwin, 1981; M.H. Goodwin, 1980; Heath, 1986; Kendon, 1967, to name but a few early inspirations, see also Heath & Luff, 2012).

Even if from its beginnings, CA has been working on technologically mediated communication as well (namely phone calls, see, e.g., Schegloff, 1968, for a general reflection on mediated interactions, see Arminen et al., 2016), the role of technologies in social interaction has been more specifically explored in institutional settings (and within the frame of "workplace studies", see Luff et al., 2000), ranging from basic technologies, such as photocopying machines (Suchman, 1987), to more complex technological set-ups, such as in aviation or subway control centres (Goodwin & Goodwin, 1996; Heath & Luff, 1996; 2000). This early focus on work settings was also linked to the fact that a lot of technologized communication environments were – and still are – initially developed and explored by bigger companies or institutions. Indeed, a lot of studies of technology use are initially motivated by an interest in assessing the technology's usability and design (see also research in related domains such as human-computer interaction and computer-supported cooperative work). Consequently, technologies tend to be explored first in complex work or lab environments, whereas their everyday, routinized uses within other settings come into focus only later. This can be illustrated, for instance, by video-mediated interaction (VMI), which has initially been investigated within company-owned media spaces (e.g., Dourish al., 1996; Heath & Luff, 1992), and the uses of which as part of more mundane communication routines have been examined only much later (e.g., Harper et al., 2017) – the more recent pandemic-related increased popularity of VMI exceptionally fuelling further research (Oloff & Ibnelkaïd, 2024).

With the improvement and fast spreading of digital infrastructure and technologies, and the increasing variety of communication platforms, applications, tools, or forms of hardware

within the last two decades, research on technology use in social interaction has become both more important and more mundane. The "mundaneisation" of this domain of research is twofold: on the one hand, it reflects the pervasiveness of technologies in all areas of our lives, in that it considers a large diversity of settings (institutional, professional, and private); on the other hand, it focusses more clearly on how basic social practices and communication with, through, and around technologies sequentially and multimodally unfold. Indeed, within the last years, the number of studies in multimodal CA interested in technologies is constantly growing, be it with respect to well-implemented technologies (such as smartphones, e.g., Avgustis, 2024; DiDomenico et al., 2020; Oloff, 2019; Råman, 2022; or tablets, e.g., Ingebrand et al., 2023; Jakonen & Niemi, 2020; Ursi & Baldauf-Quilliatre, 2021) or to newer ones (social robotics, e.g., Pelikan et al., 2022; Rollet & Clavel, 2020; virtual reality, e.g., Haddington et al., 2023; Klowait, 2023; Olbertz-Siitonen et al., 2021; AI in general, see Mlynář et al., 2024, and see also various other contributions of Ibnelkaïd & Avgustis, 2023, or Stokoe et al., 2024). It is to this emerging field that this special issue aims to make a modest contribution.

The studies of this special issue

While all articles in this special issue obviously deal with technology use, the technologies involved are quite different. On the one hand, using a desktop application, playing a video game, learning how to use a digital mobility stick or doing vocabulary training with a robot are all technology-mediated activities that build on quite different affordances (Hutchby, 2001) of the technological applications or devices. On the other hand, the role of the technology within the overall interaction can also vary, in the

sense that technology use can concern only a small task or specific participants; it can represent the main object or activity that all participants focus on during an encounter; or it can more deeply define the setting as the interaction itself is mediated through it, such as in immersive virtual reality. Therefore, the role and use of a technology within a social encounter is fundamentally linked to the task at hand. The contributions of this special issue focus on interactional moments that either reveal specific challenges with respect to the technology under investigation, or that focus on how less-expert users can be instructed and guided towards a new device or application. Consequently, we grouped the papers according to their focus, e.g. on problem-solving and assisting co-participants, instructions of specific apps and devices, and challenges in navigating between the physical and the virtual world.

Encountering and solving technology-related problems

The first set of studies examines moments where the technology itself is used as a tool to pursue a particular professional, institutional or mundane tasks, e.g. vocabulary training in a language class, compiling an exercise program during a physiotherapy consultation, or making an avatar move correctly within a game. The studies in this group address moments where the technology presents problems for the progressivity of the overall activity, and where the participants might have to negotiate to what extent they are entitled to intervene in solving the technology-related problems.

The research conducted by Jakonen, Veivo, Mutta, Maijala, Honkalammi and Johansson addresses the impact of a novel communicative and educational technology on one of the most

fundamental features of social interaction: the joint strive for the progressivity of interaction. The study *"'Am I saying it wrong?': Progressivity-related troubles and instructional opportunities in child-robot L2 interaction"* examines how students and teachers orient to moments of interactional "trouble" in robot-assisted language learning situations and provides a detailed look into how children are socialized into human-robot interaction. While the study reveals how technology can interrupt the progressivity of interaction, it also highlights the innately human ability to transform such moments of interactional trouble into moments of teaching and learning. Here, technology not only provides novel challenges, but also novel affordances for language teachers and learners.

In their paper *"Digital technology in physiotherapy consultations: Problem-solving sequences and recruitments"*, the authors Keel, Schmid and Keller investigate how troubles in handling a digital application in physiotherapy consultations are managed by both the physiotherapist and the patient. During face-to-face physiotherapy consultations, the physiotherapist compiles an exercise program with an application accessed on a desktop or laptop computer, with the aim to provide the patient with a customized training program to be carried out at home. The authors examine how problems with the app are identified and solved, by focussing on how the patient can be possibly recruited to participate in solving the problem. The task of compiling a training program with the computer is first and foremost that of the physiotherapist. However, the patients can often monitor the on-screen activity and thus perceive problems in, for instance, locating a specific exercise. These complications also become apparent in the physiotherapist's audible trouble alerts and trouble reports. If the exercise compilation is designed as a collaborative project (for instance

by verbalizing different steps of the on-screen activity for the patient and by using the inclusive pronoun "we"), the physiotherapist is more inclined to accept the patient's offer to cooperate. The authors show that within this institutional framework, the recruitment and offering of assistance for solving technological problems are linked to the specific rights and obligations that the participants hold. The analyses draw attention to the fact that these technology-related problem-solving sequences could be used to foster cooperation between physiotherapists and patients, as well as the patients' involvement in the therapeutic process.

The contribution *"Non players' embodied practices of engagement in videogaming"* by Baldauf-Quilliatre and Colón de Carvajal focusses on videogaming and reflects on different ways a non-player can engage in this activity. The setting shows a type of technology use we are probably all familiar with; while in co-presence in their private home, a couple engages in individual screen-based activities, with one participant playing a videogame and the other one using a tablet. Here, the authors seek to describe how the non-playing participant engages to different degrees with the videogame and her videogaming partner, in order to further specify the interactive construction of spectatorship. Engaging with somebody else's technology use can, in this setting, be displayed by looking at the gaming participant and the gaming screen in a brief ("noticing") or in a sustained way ("spectating"). These different ways of engaging are hinted at through a systematic annotation of the non-player's gaze shifts, and a qualitative analysis of three different types of engagement: repeated short gazes without talking display minimal engagement; longer gazes and responding to a request for assistance from the game player by touching the controller display a higher engagement; and, finally, a sustained

gaze to the game screen and offering assistance to the game player by providing instructions and accomplishing a game move on the controller corresponds to the highest degree of engagement in the gaming activity. The player, for his part, can disagree with this latter, heightened involvement in the gaming activity by the non-player. This study therefore illustrates that categories such as player and non-player are not clearly distinct, but that participants locally negotiate their mutual entitlement to engage in somebody else's "individual" technology use.

Instructing the use of technological devices or apps

The next two studies address moments where technology takes on a more central role within the interaction, i.e., as a precise techno-material object or application of which the use has to be explained and instructed. Both studies feature professional settings in which the use of a specific technology is taught to participants, who in the future will be using the technological devices in their daily lives and on their own, either for medical or professional purposes. The respective technologies are not simply "shown" to the participants. Rather, as both contributions illustrate, the role of the technology varies throughout the instructional activity, from a simple physical object handled by the instructor to an object that the participants sensorially engage with, to what could even be argued to amount to a co-participant. The multimedial nature of the types of technologies examined in this special issue is reflected in the way the use of such technologies is taught.

The study by Räisänen and Hynninen, *"Making new technology understandable through multimodal instruction: A digital mobility stick in customer training interaction"*, examines how

representatives of a health technology company introduce a novel technology, a digital mobility stick, in customer training sessions. More specifically, the focus is on how one particular aspect of the technology in question, the built-in haptic component, is illustrated and taught to the customer. The key learning outcome is to establish the relation between the customer's embodied actions – the exercise movements – and the haptic feedback provided by the technology. As the analysis shows, the outcome cannot be learned by orienting to the technology as a static object introduced by the instructor, something to be seen or talked about. Rather, as the authors argue, intersubjectivity can only be achieved by engaging the entire spectrum of task-relevant practices. The study presents an interesting examination of the role of haptic interaction in instructions and instructed actions, and at the same time sheds light on how technology itself can "guide" both embodied conduct and instructional sequences.

The contribution *"Digitaalisesta fyysiseen ja imitoituun – Uuden mobiilisovelluksen vaihtuvat roolit käyttöönottokoulutuksissa"* by Käntä uses multimodal discourse analysis to examine the way a new mobile application and device are taught to nurses in a Finnish healthcare organization. This study provides an in-depth examination on how the participants' orientation to the digital and physical technology in question is reflected in the production of their turns-at-talk. A particular focus is placed on the collaboration between speech and embodied-digital activities of the trainers. The role of the technology itself in the interaction is shown to be malleable, ranging from a simple tool to a co-participant. Furthermore, the study provides an interesting look into how technologies which are not immediately perceivable by the participants can be talked into being through multimodal means. By examining the ways in

which the trainers detach themselves from the physical and digital limitations of the taught technology, Kääntä's study emphasizes the importance of situational awareness for recipient design, that is, the trainers' ability to choose the most relevant aspects of the technology from the point of view of the nurses and their daily work.

Navigating between physical and virtual environments

While in the previous papers, different technologies are used and instructed in face-to-face or physical settings, the final set of contributions to this special issue features settings in which one or all the participants act within an immersive virtual world. Thus, in this case, technology has a stronger impact on the organisation of social interaction, in that it is mediated through a digitally created environment. The mismatch between what can be done and perceived in the physical and in the virtual world (Hindmarsh et al., 1998) leads to specific challenges for instructing novice users or for successfully engaging in a gaming activity, that the participants seek to solve by falling back on other verbal, embodied, or digital resources.

Olbertz-Siitonen and Piirainen-Marsh examine how a more experienced participant instructs a less-experienced participant on how to handle virtual reality equipment for gaming purposes through the use of talk and touch. In their paper *"Bridging physical and virtual ecologies of action: Giving and following instructions in co-located VR-gaming sessions"*, the nature of the communicational setting examined is asymmetrical. The instructed party uses the VR equipment, primarily engaging in a virtual ecology of communicational affordances with limited sensorial access to the physical

environment, while the instructing party has a more limited access to the virtual world, providing their instructions within a physical ecology of communicational affordances. The study shows how participants can overcome this type of asymmetrical or fractured ecology of affordances inherent to virtual reality settings. With the rapid increase of immersive technologies, understanding how to bridge the virtual and physical ecologies is also becoming more relevant.

In her contribution *"Environmentally coupled gestures as a communicative resource in the word explanation activity: A multimodal analysis of interaction in social VR"*, Spets looks at a setting in which two participants are both interacting within an immersive virtual environment offering different types of games. Although both players are physically co-present in a lab, their use of a head-mounted display allows them to visually perceive each other exclusively in and through the virtual space. Spets focuses on the players' use of environmentally coupled gestures within 3D Charades, a game in which one player has to explain a word to be guessed by getting their avatar to enact it or to draw objects in the air by using a virtual pen. Players can also refer to their virtual drawings by pointing at them with their avatar. However, as further excerpts from the video data show, referring to one's own body does not easily produce a correct hint for the co-player, as there is a mismatch between the movement carried out with and on the physical body, and the way it is represented by the avatar in the game. Indeed, the asymmetrical perception of both others' and one's own virtual body make it difficult to know what the pointing gestures and drawings exactly refer to, and how these are appropriately positioned and carried out so as to be understandable for the co-player. Therefore, to make gestures more intelligible in a virtual reality en-

vironment, a better correspondence between physical and virtual movements and bodies is needed. When designing their actions in this virtual environment, participants must take into account the possibilities and restrictions provided by the resources at hand, here, the avatar's appearance and mobility within the game.

Acknowledgments

References

Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction, 49*(4), 290–309. https://doi.org/10.1080/08351813.2016.1234614

Avgustis, I. (2023). Respecifying phubbing: Video-Based analysis of smartphone use in co-present interactions. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3544548.3581052

Avgustis, I. (2024). *Smartphone use by young adults. Collaborative practices of dis-/re-)engagement in face-to-face interaction* [University of Oulu]. https://urn.fi/URN:NBN:fi:oulu-202406034151

Deppermann, A. (2021). Social Actions. In M. Haugh, D. Z. Kádár, & M. Terkourafi (Eds.), *The Cambridge Handbook of Sociopragmatics* (1st ed., pp. 69–94). Cambridge University Press. https://doi.org/10.1017/9781108954105.006

Deppermann, A., & Streeck, J. (Eds.) (2018). *Time in Embodied Interaction. Synchronicity and sequentiality of multimodal resources*. John Benjamins. https://doi.org/10.1075/pbns.293

DiDomenico, S. M., Raclaw, J., & Robles, J. S. (2020). Attending to the mobile text summons: Managing multiple communicative activities across physically copresent and technologically mediated interpersonal interactions. *Communication Research, 47*(5), 669–700. https://doi.org/10.1177/0093650218803537

Dourish, P., Adler, A., Bellotti, V., & Henderson, A. (1996). Your place or mine? Learning from long-term use of audio-video communication. *Computer Supported Cooperative Work (CSCW), 5*(1), 33–62. https://doi.org/10.1007/BF00141935

Enfield, N. J., Stivers, T., Brown, P., Englert, C., Harjunpää, K., Hayashi, M., Heinemann, T., Hoymann, G., Keisanen, T., Rauniomaa, M., Raymond, C. W., Rossano, F., Yoon, K.-E., Zwitserlood, I., & Levinson, S. C. (2019). Polar answers. *Journal of Linguistics, 55*(2), 277–304. https://doi.org/10.1017/S0022226718000336

Fox, B. A., Thompson, S. A., Ford, C. E., & Couper-Kuhlen, E. (2012). Conversation Analysis and Linguistics. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 726–740). Wiley. https://doi.org/10.1002/9781118325001.ch36

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice-Hall.

Goodwin, C. (1981). *Conversational Organization. Interaction between Speakers and Hearers*. Academic Press.

Goodwin, C., & Goodwin, M. H. (1996). Seeing as a situated activity: Formulating planes. In D. Middleton & Y. Engeström (Eds.), *Cognition and Communication at Work*. Cambridge University Press. https://doi.org/10.1017/CBO9781139174077.004

Goodwin, M. H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociolgical Inquiry, 50*, 303–317. https://doi.org/10.1111/j.1475-682X.1980.tb00024.x

Haakana, M. (2002). Laughter in medical interaction: From quantification to analysis, and back. *Journal of Sociolinguistics, 6*, 207–235. https://doi.org/10.1111/1467-9481.00185

Haddington, P., Eilittä, T., Kamunen, A., Kohonen-Aho, L., Oittinen, T., Rautiainen, I.., & Vatanen, A. (Eds.) (2023). Ethnomethodological conversation analysis in motion. Emerging methods and new technologies. Routledge. https://doi.org/10.4324/9781003424888

Haddington, P., Kohonen-Aho, L., Tuncer, S., & Spets, H. (2023). Openings of interactions in immersive virtual reality: Identifying and recognising prospective participants. In P. Haddington, T. Eilittä, A. Kamunen, L. Kohonen-Aho, I. Rautiainen, & A. Vatanen (Eds.), *Complexity of Interaction* (pp. 423–456). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-30727-0_12

Harper, R., Watson, R., & Licoppe, C. (2017). Interpersonal video communication as a site of human sociality: A special issue of Pragmatics. *Pragmatics, 27*(3), 301–318. https://doi.org/10.1075/prag.27.3.01har

Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge University Press. https://doi.org/10.1017/CBO9780511628221

Heath, C., & Luff, P. (1992). Media space and communicative asymmetries: Preliminary observations of video-mediated interaction. *Human-Computer Interaction, 7*(3), 315–346. https://doi.org/10.1207/s15327051hci0703_3

Heath, C., & Luff, P. (1996). Convergent activities: Line control and passenger information on the London Underground. In Y. Engeström & D. Middleton (Eds.), *Cognition and Communication at Work* (1st ed., pp. 96–129). Cambridge University Press. https://doi.org/10.1017/CBO9781139174077.005

Heath, C., & Luff, P. (2000). *Technology in Action*. Cambridge University Press. https://doi.org/10.1017/CBO9780511489839

Heath, C., & Luff, P. (2012). Embodied action and organizational activity. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 281–307). Wiley. https://doi.org/10.1002/9781118325001.ch14

Hepburn, A., & Bolden, G. (2017). *Transcribing for social research*. SAGE.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Polity Press/Blackwell.

Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (1998). Fragmented interaction: Establishing mutual orientation in virtual environments. *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*, 217–226. https://doi.org/10.1145/289444.289496

Hutchby, I. (2001). Technologies, texts and affordances. *Sociology, 35*(2), 441–456. https://doi.org/10.1177/S0038038501000219

Ibnelkaïd, S., & Avgustis, I. (2023). Situated agency in digitally artifacted social interactions: Introduction to the special issue. *Social Interaction. Video-Based Studies of Human Sociality, 6*(1). https://doi.org/10.7146/si.v6i1.136855

Ingebrand, E., Samuelsson, C., & Hydén, L.-C. (2023). Supporting people living with dementia in novel joint activities: Managing tablet computers. *Journal of Aging Studies, 65*, 101116. https://doi.org/10.1016/j.jaging.2023.101116

Jakonen, T., & Niemi, K. (2020). Managing participation and turn-taking in children's digital activities: Touch in blocking a peer's hand. *Social Interaction. Video-Based Studies of Human Sociality, 3*(1). https://doi.org/10.7146/si.v3i1.120250

Jefferson, G. (1983). Notes on some orderlinesses of overlap onset. *Tilburg Papers in Language and Literature* (Tilburg University), *28*, 1–28.

Jefferson, G. (2004a). A sketch of some orderly aspects of overlap in natural conversation. In G. H. Lerner (Ed.), *Conversation Analysis. Studies from the first generation* (pp. 43–59). John Benjamins. https://doi.org/10.1075/pbns.125.05jef

Jefferson, G. (2004b). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis. Studies from the first generation.* (pp. 13–31). John Benjamins Publishing. https://doi.org/10.1075/pbns.125.02jef

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26,* 22–63. https://doi.org/10.1016/0001-6918(67)90005-4

Kendrick, K. H. (2017). Using conversation analysis in the lab. *Research on Language and Social Interaction, 50*(1), 1–11. https://doi.org/10.1080/08351813.2017.1267911

Klowait, N. (2023). On the multimodal resolution of a search sequence in virtual reality. *Human Behavior and Emerging Technologies*, 2023, 1–15. https://doi.org/10.1155/2023/8417012

Leick, K. (2019). *Parents, media, and panic through the years. Kids those days*. Springer. https://doi.org/10.1007/978-3-319-98319-6

Levinson, S. C. (2012). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 101–130). Wiley. https://doi.org/10.1002/9781118325001.ch6

Luff, P., Hindmarsh, J., & Heath, C. (Eds.). (2000). *Workplace studies. Recovering work practice and informing system design*. Cambridge University Press. https://doi.org/10.1017/CBO9780511628122

Mlynář, J., De Rijk, L., Liesenfeld, A., Stommel, W., & Albert, S. (2024). AI in situated action: A scoping review of ethnomethodological and conversation analytic studies. *AI & Society*. https://doi.org/10.1007/s00146-024-01919-x

Mondada, L. (2012). The Conversation analytic approach to data collection. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 32–56). Wiley. https://doi.org/10.1002/9781118325001.ch3

Mondada, L. (2018a). Transcription in linguistics. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 85–114). Continuum.

Mondada, L. (2018b). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction, 51*(1), 85–106. https://doi.org/10.1080/08351813.2018.1413878

Mondada, L. (2022). *Conventions for multimodal transcription*. https://www.lorenzamondada.net/multimodal-transcription

Ochs, E., Schegloff, E. A., & Thompson, S. A. (Eds.). (1996). *Interaction and grammar*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620874

Olbertz-Siitonen, M., Piirainen-Marsh, A., & Siitonen, M. (2021). Constructing co-presence through shared VR gameplay. *Journal Für Medienlinguistik, 4*(2), 85–122. https://doi.org/10.21248/jfml.2021.31

Oloff, F. (2019). Some systematic aspects of self-initiated mobile device use in face-to-face encounters. *Journal Für Medienlinguistik, 2*(2), https://doi.org/10.21248/jfml.2019.21

Oloff, F. (2021). New technologies – new social conduct? A sequential and multimodal approach to smartphone use in face-to-face interaction. *Bulletin VALS-ASLA, 1*, 13–34. https://oulurepo.oulu.fi/bitstream/handle/10024/31674/nbnfi-fe2021122963341.pdf

Oloff, F., & Hepburn, A. (in press). Multimodal transcription as process and analysis: Capturing the audible and visible. In: J. D. Robinson, R. Clift, K. H. Kendrick, & C. W. Raymond (Eds.). The Cambridge handbook of methods in conversation analysis. Cambridge University Press, 115–142.

Oloff, F., & Ibnelkaïd, S. (2024). 13. Video-vermittelte Interaktion. In J. Androutsopoulos & F. Vogel (Eds.), *Handbuch Sprache und digitale Kommunikation* (pp. 269–288). De Gruyter. https://doi.org/10.1515/9783110744163-013

Pelikan, H., Broth, M., & Keevallik, L. (2022). When a robot comes to life: The interactional achievement of agency as a transient phenomenon. *Social Interaction. Video-Based Studies of Human Sociality, 5*(3). https://doi.org/10.7146/si.v5i3.129915

Pfeiffer, M., König, K., Weber, K., Deppermann, A., Ehmer, O., Gipper, S., Gubina, A., Kim, K., Küttner, U., Li, X., Marmorstein, M., Maschler, Y., Ben Moshe, Y., Oloff, F. & Szczepek Reed, B. (in press). Request for confirmation sequences in ten languages. A quantitative comparison. *Open Linguistics*.

Rollet, N., & Clavel, C. (2020). "Talk to you later": Doing social robotics with conversation analysis. Towards the development of an automatic system for the prediction of disengagement. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems, 21*(2), 268–292. https://doi.org/10.1075/is.19001.roll

Råman, J. (2022). Multimodal negotiation for the right to access digital devices among elderly users and teachers. In J.-P. Alarauhio, T. Räisänen, J. Toikkanen, & R. Tumelius (Eds.), *Shaping the North Through Multimodal and Intermedial Interaction* (pp. 67–93). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-99104-3_4

Sacks, H. (1984). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action. Studies in Conversational Analysis.* (pp. 21–27). Cambridge University Press. https://doi.org/10.1017/CBO9780511665868.005

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735. https://doi.org/10.2307/412243

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist, 70*, 1075–1095. https://www.jstor.org/stable/pdf/669510.pdf

Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction, 26*(1), 99–128. https://doi.org/10.1207/s15327973rlsi2601_5

Schegloff, E. A. (2007). *Sequence organization in interaction. A primer in conversation analysis*. Cambridge University Press. https://doi.org/10.1017/CBO9780511791208

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*(2), 361–382. https://doi.org/10.2307/413107

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*, 289–327. https://doi.org/10.1515/semi.1973.8.4.289

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J. R., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., … Uhmann, S. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift Zur Verbalen Interaktion, 10*, 353–402. http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf

Selting, M., & Couper-Kuhlen, E. (Eds.). (2001). *Studies in Interactional Linguistics*. John Benjamins. https://doi.org/10.1075/sidag.10

Sidnell, J. (2012). Basic conversation analytic methods. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 77–99). Wiley-Blackwell. https://doi.org/10.1002/9781118325001.ch5

Stivers, T. (2015). Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction, 48*(1), 1–19. https://doi.org/10.1080/08351813.2015.993837

Stokoe, E., Albert, S., Buschmeier, H., & Stommel, W. (2024). Conversation analysis and conversational technologies: Finding the common ground between academia and industry. *Discourse & Communication*, 17504813241267118. https://doi.org/10.1177/17504813241267118

Streeck, J., Goodwin, C., & LeBaron, C. (Eds.) (2011). *Embodied interaction. Language and body in the material world*. Cambridge University Press.

Suchman, L. (1987). *Plans and situated actions: The problem of human machine communication*. Cambridge University Press.

Ursi, B., & Baldauf-Quilliatre, H. (2021). Toucher les écrans ça s'apprend! L'étude multimodale des gestes réalisés par des enfants au cours de sessions de jeu sur grande tablette. *Bulletin VALS-ASLA, 1*, 35–56.

OTSIKKO JA ASIASANAT SUOMEKSI:

Teknologian käytön ymmärtäminen kasvokkain tapahtuvassa vuorovaikutuksessa: keskustelunanalyyttinen näkökulma

ASIASANAT: keskustelunanalyysi, multimodaalisuus, videoaineistot, teknologian käyttö