

Social Interaction. Video-Based Studies of Human Sociality.
2021 Vol. 4, Issue 4
ISBN: 2446-3620
DOI: 10.7146/si.v4i4.121776

Social Interaction

Video-Based Studies of Human Sociality

Eye-tracking Recordings as Data in EMCA Studies: Exploring Possibilities and Limitations

Elisabeth Dalby Kristiansen & Gitte Rasmussen

University of Southern Denmark

Abstract

The article contributes to the ongoing discussion of the potential of using eye-tracking recordings in ethnomethodological conversation analysis (EMCA) by exploring to what extent and under what circumstances such recordings may be useful for EMCA studies of multimodal social interaction. For this purpose, it analyzes examples of social conduct recorded by one video camera and one set of eye-tracking glasses. The article concludes that while eye-tracking recordings may, in some specific cases, provide new analytic possibilities for studying social action, they are by no means indispensable for EMCA research in multimodal social interaction, and making use of mobile eye-tracking equipment and recordings may compromise the data as well as the analytic procedure.

Keywords: ethnomethodological conversation analysis, eye-tracking, gaze behavior

1. Introduction

The collection, analysis and presentation of video data is a well-established practice in studies of human action and interaction. The use of video has not only made it possible to capture events for repeated scrutiny (cf. Sacks, 1984) and analysis of talk in interaction (e.g., Goodwin, 1979, 1980); it has also resulted in a widespread research interest in the embodied and multimodal actions of participants in interaction (Nevile, 2015), for instance with the purpose of studying either bodies as a type of modality through which individuals produce actions which are “observable”, “witnessable”, “recognizable”, and “understandable” (e.g., Hayashi, 2005; Hazel et al., 2014; Heath, 1986; Keevallik, 2010; Mondada, 2014), or the individual’s understanding of “social meaning” through an embodied multisensorial experience (e.g., Meyer & Streeck, 2020; Streeck, 1993).

The use of video technologies for studies of interaction has proven advantageous as compared to, for instance, handwritten ethnographic notes, and researchers in the field add to their interest in social interaction research interests in investigating the uses, advantages, and disadvantages of new technological devices and features for the purpose of recording social interaction, such as camera brands, number of cameras used, mobile versus fixed cameras, etc. (McIlvenny, 2019). In methodological discussions of video technologies in Ethnomethodological Conversation Analysis (EMCA), the question of how the technology may best be used for gaining access to participants’ orientations and perspectives is central and has been subject to discussions for several years (e.g., Heath et al., 2010; Knoblauch et al., 2006; Macbeth, 1999; Mondada, 2006).

In recent years, the use of eye-tracking recordings has emerged as yet another potential tool for data collection (Brône & Oben, 2018a; Kristiansen & Rasmussen, *forthc.*; Rasmussen & Kristiansen, 2021). As this new technology is explored, an ongoing discussion of how it may contribute to EMCA research is vital. As a contribution to this discussion, this article explores the promises and problems of eye-tracking recordings, discussing the relevance of eye-tracking data for EMCA studies of multimodal everyday social interaction (Deppermann, 2013; Garfinkel, 1967; Maynard & Clayman, 1991; Mondada, 2017).

For this purpose, it provides examples of eye-tracking and video recorded actions and interactions in physical shopping environments.¹ Each example is a single case of some conduct in the shopping environments where gazing seems to play a role and is offered with the sole purpose of discussing methodological points.

¹ As opposed to digital shopping environments.

2. Prior work on the use of eye-tracking in EMCA

In recent years, it has become possible to accompany or replace video recordings with eye-tracking recordings, as, for instance, Stukenbrock and others have done (Brône & Oben, 2018a; Stukenbrock & Dao, 2019). This type of data grants researchers access to highly detailed information about participants' eye movements and fixations, and this prompts considerations of the technical possibilities as well as possible methodological and analytic implications of using eye-tracking technology.

While gaze has long been recognized and studied as a key resource for meaning making in interaction (Goodwin, 1979; Kendon, 1967; Rossano, 2012), the technology of eye tracking has recently made it possible to measure eye movement and fixations, first in experimental settings with fixed eye-tracking equipment (Pfeiffer et al., 2013), and later also in more naturalistic settings as the equipment has become more mobile (Auer, 2018; Duchowski, 2007). The use of stationary eye-tracking equipment allows researchers a large degree of control with the settings, and participants are often seated and more or less restricted in their head movements. In contrast, the use of mobile and wearable eye-tracking equipment means that participants may move around and freely turn their bodies, heads and gaze where they wish, lessening the experimental control of the setting. The use of this equipment thus makes it possible to record and study participants' gaze behavior in interaction in everyday environments, such as shopping (Stukenbrock, 2018; Stukenbrock & Dao, 2019).

The detailed information about eye movements and fixations which eye-tracking technology makes available may be of interest to researchers in e.g. customer behavior (Meißner et al., 2019) and marketing (Lohse & Wu, 2001; Pieters & Warlop, 1999; Wedel & Pieters, 2008) interested in customers' decisions to buy and the relation of the decisions to the organization of products displays in brick-and-mortar stores or of product images on the websites of internet stores. The detailed records of eye movements and fixations may also be intriguing for cognitive research with an interest in first-person perspective (Ferguson et al., 2017) and for research with a general interest in gaze behavior (Risko & Kingstone, 2011) or in social and interactional phenomena such as "joint attention" (Stukenbrock, 2020) or "addressee selection" (Auer, 2018).

A small but growing body of research is thus making use of mobile eye-tracking technology for studying gaze behavior in more naturalistic settings and situations, using conversation analytic, and interactional linguistics approaches (Auer, 2018; Holler & Kendrick, 2015; Oben, 2018; Stukenbrock, 2018, 2020; Stukenbrock & Dao, 2019; Weiß & Auer, 2016). The studies use several sets of eye-tracking glasses to record dyadic and triadic interactions. Oben's study analyses dyadic interaction in a quasi-experimental setting where both participants are wearing eye-tracking glasses, while Auer, Weiß and Auer, and Holler and Kendrick invite participants to engage in unscripted dialog in a controlled setting with fixed seats

for the participants and multiple cameras. The studies by Stukenbrock and Stukenbrock and Dao use two sets of eye-tracking glasses for studying joint attention and deixis in dyad interactions in natural surroundings while the participants are moving through space. All of the studies obtain very precise details regarding e.g. the timing of gaze movement in relation to talk, gestures, or other embodied actions with the purpose of understanding how participants' "allocation of gaze" contributes e.g. to turn allocation and speaker selection (Auer, 2018), to lexical and gestural alignment, and to establishing joint attention on objects (Stukenbrock & Dao, 2019) or referential practices (Stukenbrock, 2018).

The studies position themselves in relation to prior research studying gaze behavior in experimental settings and/or with stationary eye-tracking equipment, emphasizing the importance of studying gaze behavior in "a truly interactive manner" (Holler & Kendrick, 2015), i.e., studying people who are participating in interaction rather than people who observe others interacting, and studying natural (i.e., unscripted) interaction. Some studies also position themselves in relation to previous work on gaze and referential practices, primarily within Conversation Analysis (CA), stressing the "high degree of granularity" (Stukenbrock, 2018, p. 289) and increased precision offered by eye-tracking recordings, which allow more exact validation of previous findings and invite researchers to make use of the new and alternative sources of information which eye-tracking recordings present, including the first-person perspective and the spatial and temporal dynamics revealed in the data. Stukenbrock and Dao (2019) even state that relying exclusively on video recordings imposes serious limitations on the precision with which eye gaze is documented, adding that video analytic studies often have to rely on participants' head movements as an indicator of their gaze direction, which they believe is not sufficiently reliable. In their view, eye-tracking recordings "provide a maximal degree of precision in recording participants' gaze behavior while simultaneously preserving the ecological validity of the data" (Stukenbrock & Dao, 2019, p. 180).

However, some of the studies also mention potential methodological problems of eye-tracking data, including that measuring gaze movement and fixations in great detail does not entail gaining direct access to the conversational functions of these (Oben, 2018) and that eye-tracking data do not in fact correspond to what participants actually see, since the human peripheral vision surpasses what the cameras can capture (Stukenbrock, 2018). Stukenbrock and Dao (2019) briefly mention the problem of how to treat eye-tracking data as an analytical resource for multimodal CA: By analyzing micro-details only visible in the mobile eye-tracking data, it is possible to discover and describe hitherto unknown systematicities in gaze behavior as demonstrated by the studies reviewed above, but these systematicities are available only to the analyst, not to the co-participants in the situation. For research in interactional linguistics and CA with an *a priori* interest in specific types of gaze behavior and their role in multimodal interaction, such discoveries constitute a research breakthrough which enables

a whole new, extremely detailed level of analysis. But for EMCA research committed to studying participants' methods for understanding and achieving social actions, using eye-tracking data begs the question of whether systematicities that may be discovered on a micro-level by analyzing eye-tracking recordings has any social meaning for the participants in the local interaction (cf. Weiß & Auer, 2016), since the eye-tracking information about eye movements and fixations are not available to the participants in the interaction.

Answering Stukenbrock and Dao's call for critical and methodological discussion of how eye-tracking data may be used in EMCA analyses (Stukenbrock & Dao, 2019), this article will explore the usefulness of eye-tracking data for EMCA research on multimodal interaction, i.e., research discovering and describing how participants achieve and maintain social order, including how they make this work visible in and through their actions (in interaction), by employing available resources, including gestures, bodily movements, verbal contributions, objects, and also gaze (Hazel et al., 2014; Mondada, 2014). For this type of research, gaze is not of interest *a priori*, but only to the extent that members use and orient to it as relevant for the local organization of ordinary multimodal interaction in stores.

In our investigation of eye-tracking technology and the possibilities of the data that it yields, we are aware of a number of methodological pitfalls that may compromise the data and its analysis, and the analytic procedure which we explore in this article is developed specifically to avoid some of these. The methodological pitfalls include the practical matter of introducing relatively noticeable eye-tracking equipment, i.e., eye-tracking glasses plus an external hard drive and a wearable microphone, into the setting, which may compromise the validity of the data in terms of the participants' orientation to the eye-tracking equipment. This will be discussed in more detail in section 3 (Data and Methods). The pitfalls also include the special status which is assigned to gaze *a priori* by recording it separately by means of eye-tracking equipment. This is similar to the special status allocated to language for many years, until the fundamentally multimodal or embodied nature of social interaction was acknowledged and made a topic for research (Mondada, 2014; Nevile, 2015) and it may skew the analytic perspective in research where gaze is conceptualized as only one out of many resources that people may use for making meaning, while it may no doubt enhance research with an interest specifically in gaze such as the studies mentioned above.

Further, while we do agree that recordings of eye movements and eye fixations no doubt provide more precise and more detailed information on participants' gaze behavior than what is available from video recordings of the participants' social interaction, we will argue that in terms of the ethnomethodological commitment to describing how members of society engage in accountable actions, i.e., conduct in ways for others to observe and witness (Garfinkel, 1967), the availability *for the analyst only* of detailed information about eye movements

and fixations may be a challenge for the validity of the analysis: The question for EMCA analysts is always first and foremost whether those details are available *to the co-participants* and whether they orient to them as relevant in the interaction. For co-participants, detailed information about gaze movements and fixations is not available, except to the extent that it is visible through, for instance, head movements and bodily orientation; and while these may be less accurate or detailed than eye-tracking data (as pointed out in Stukenbrock, 2018; Stukenbrock & Dao, 2019), participants in social interaction nonetheless manage to conduct their business for all practical purposes by means of these and other publicly available resources (for examples of this, see e.g., Clark & Pinch, 2010; Rasmussen & Kristiansen, 2021). Thus, given the basic interests of EMCA, the usefulness of eye-tracking data for generating insights into social interactional phenomena may not be taken for granted but needs to be investigated empirically as we aim to do here.

In the present paper we describe one participant's methods for doing "browsing", "searching" and "examining", and for observably and witnessably doing so, with the specific aim of discussing whether and how eye-tracking recordings may contribute to such analyses. The analytical examples feature spontaneous interaction between a person and acquainted and unacquainted others who just happen to be co-present as well as interaction with the environment. Some of the examples include dialog between the customer and a salesclerk while others do not include dialog. However, in all our examples other customers are present in the shop (cf. Figure 1 and Figure 2), and therefore, even though the interaction may not be sequentially organized in the moment, any action that the participants do is observable and witnessable to co-present others and may potentially be responded to as e.g., a "pre" (Schegloff, 2007); that is, co-present others may orient to the participant's action as inviting or making possible the initiation of a face-to-face encounter, i.e., a dyad.

Figure 1. *Customers in a bookstore*



Figure 2. *Customers and staff in an interior decoration store*



This phenomenon is discussed e.g. in a study of availability, unavailability, and awareness in face-to-face encounters involving persons with dementia (Rasmussen et al., 2019) and in a study on sharing space and minimizing involvement in Danish self-service stores (Rasmussen & Kristiansen, 2021).

3. Data and methods

The aim of the Velux-funded research project RESEMINA is to describe differences and similarities in shopping practices in a range of different types of shops, including brick-and-mortar and online stores and supermarkets (Kristiansen & Rasmussen, *forthc.*; Rasmussen & Kristiansen, 2021; Rasmussen et al., *forthc.*), using insights from multimodal social semiotics (Hodge & Kress, 1998; Jewitt, 2008; Jones, 2005; Kress & Van Leeuwen, 2001; Zhu et al., 2017) and ethnomethodological conversation analysis (EMCA) of multimodal interaction (Deppermann, 2013; Fasulo & Monzoni, 2009; Haddington et al., 2014; Hazel et al., 2014; Heath & Luff, 1992; Licoppe & Morel, 2012; Mondada, 2014). Methodologically and theoretically, the RESEMINA project also aims to explore possibilities for a dialogue between the two approaches (Poulsen et al., *forthc.*).

The data collected for the RESEMINA project consist of a corpus of video as well as eye-tracking recordings of shopping interactions online and in brick-and-mortar shops in Denmark. The data were collected with the informed consent of the participants prior to the recordings, and the data are managed in accordance with Danish Law and EU regulations as sanctioned and monitored by the Data Protection Office of the University of Southern Denmark. All names in the article are pseudonyms. Talk and embodied conduct are transcribed in accordance with the Jeffersonian system of transcription (see, e.g., Transcript Notation p. ix-xvi in Atkinson and Heritage (1984)).

In this article, we draw on the video and eye-tracking recordings of shoppers in brick-and-mortar stores. The video data was recorded using a mobile video camera operated by a researcher following the shoppers as they accomplished whatever business they had in the store as illustrated in example 1.

(>>Video: Example1<<)

During the recording, the researcher constantly made methodological choices, e.g., regarding recording angle, how to move around obstacles, how to position themselves in relation to the shopping participant(s), etc. This is illustrated in example 2.

(>>Video: Example2<<)

While the shops were approached in advance, and all salesclerks had consented to participate in the recordings, customers were neither selected nor approached in advance. Instead, the researchers approached people about to enter the store,

asking if they could record them while they conducted whatever business they had in the store. In that way, the activities recorded are activities that the participants were conducting independently of the research, that is, naturally occurring (Mondada, 2012; Sacks et al., 1974).

If a pair of potential participants (De Stefani, 2013) were approached, the participants negotiated amongst themselves who would wear the eye-tracking glasses. Among other things, their negotiations took into account who initially agreed to participate; who was the primary shopper (which might change during the shopping activity); who was already wearing glasses; and who was most comfortable with the situation, including the prospect of having their gaze recorded. The latter included reflecting on what one might happen to look at. The exchange below between participant (M) and his female co-shopper (F) exemplifies this. M reflects on whether his gaze might happen to rest inappropriately on female customers in the shop:

Extract 1. Reflecting on possible inappropriate gaze behavior

- 1: F: så må vi se om der mange damer
Then we'll see if there are many women
- 2: M: ja (.) jeg håber bare der ik er nogen der pæne
yes I just hope there aren't any who are pretty
- 3: fordi ellers så går det galt
because otherwise it will go wrong

Potentially, the choices made by the researchers and participants while *preparing* the recording may affect some of the researcher's choices made while *operating* the video camera, specifically choices that are determined by the recording researcher's own member's understanding of the social situation they are recording. For instance, recording a pair of participants who were shopping together and split up during the shopping, the researcher had to decide which participant to follow. In such cases, the researcher usually followed the participant whom they understood as being in charge of the shopping activity. In Example 3, a young couple are buying ingredients for a cake in a supermarket. In the example, the video recording is represented in the main frame, and the eye-tracking recording is represented in the overlay frame in the bottom right corner. The video in the eye-tracking recording is recorded by the eye-tracking glasses, i.e., as if we are "looking through the eyes" of person wearing the glasses. The circles in the eye-tracking recording represent eye fixations, and the lines represent eye movement.

(>>Video: Example3<<)

In Example 3, the main shopper is the female participant: She is in charge of the shopping list as well as the shopping basket. The male participant, however, is wearing the eye-tracking glasses. At one point, the couple split up while moving from one area of the store to another. The researcher operating the camera follows the female participant, i.e., the main shopper.

However, in some cases, the opposite happens: The researcher follows the person with the eye-tracking glasses rather than the main shopper. In Example 4, which is recorded in a book store, the shoppers are a grandmother and her granddaughter, and the grandmother is the main shopper.

(>>Video: Example4<<)

They are in the store because she needs to buy a ledger book, and they are searching (Rasmussen et al., forthc.) for this particular item. However, the grandmother is wearing glasses, and therefore the granddaughter is wearing the eye-tracking glasses. They cannot find the item they are looking for, and at one point they split up, searching different aisles. The camera follows the granddaughter, i.e., the person with the eye-tracking glasses and not the main shopper, even though the socially natural choice would be to follow the grandmother, i.e., the person who has an errand in the shop.² In fact, as the video demonstrates (00:45-00:53), the grandmother has been talking to a member of staff, inquiring about the book, while the camera followed the granddaughter, and as the granddaughter returns, the shop assistant is showing the grandmother where to find the ledger book.

Examples like this indicate that the presence of the eye-tracking glasses may overrule the researcher's intuitive understanding of the social situation so that the camera follows the eye-tracking equipment rather than the socially most significant activity.

One way of solving this problem might be to have both participants wear eye-tracking glasses. As argued by the studies reviewed above, using multiple sets of eye-tracking glasses yields more precise details about the eye movements and fixations of both participants, which may improve the analysis. Further, it might be argued that using only one set of eye-tracking glasses introduces a skewed perspective on the interaction, since a more detailed record of gaze movements and fixations is available for one of the participants. However, we believe that introducing another set of eye-tracking glasses would limit the naturalness of the interaction, bringing the situation closer to an experimental setting, e.g., by making the eye-tracking equipment, visible on the co-participant's face, more noticeable to the participants, and by impeding the interaction of the participants with the wider environment as described in EMCA studies on e.g. museum visits (vom Lehn & Heath, 2016) or shop visits (Rasmussen & Kristiansen, 2021).

² That is not to say that it might not also be interesting to follow the "accompanying" shopper.

A key criticism of experimental studies from an EMCA point of view is that participants are keenly aware that they are serving as participants in an experiment, i.e., that they are not engaging in activities that they would otherwise engage in, which means that the interaction is not naturally occurring (Garfinkel & Wieder, 1992). Experimental studies may thus provide insights about how people interact in experimental settings, but not about how people engage in social interaction outside of experimental settings. The use of wearable eye-tracking equipment in studies of naturally occurring interaction is open to similar criticism: Even though the eye-tracking glasses used for collecting the data used in this study are relatively small, they are not entirely “unobtrusive” (Brône & Oben, 2018b, p. 7). They are noticeably different from normal glasses, which means that other customers or staff may gaze at participants, looking at the glasses; they are connected to an external hard drive by a cable, which means that the participant must carry a bag over their shoulder for the duration of the recording; and they are, at least in our study, worn by people who do not normally wear glasses. All these factors potentially render the interaction less natural. On the other hand, all participating customers in the RESEMINA project were recruited on their way into the store where they had an errand independently of the research project, to ensure that the interaction would be natural in the sense that its occurrence is not prompted or determined by the research.

Using eye-tracking equipment for collecting data thus presents a challenge for EMCA studies whose analyses are based on the researcher’s member’s understanding of the social situation (Garfinkel, 1967; Wooffitt & Hutchby, 1998), since the eye tracking equipment may skew or bypass this understanding in the recording situation. The data demonstrate that the researchers orient to the glasses in the setting and the situation of the shopping and that the participants evidently also do so. Other people in the stores may also orient to the glasses, and participants may orient to the fact that the other people in the stores orient to the glasses. Analysts working with eye-tracking data must therefore consider how this restricts the analytical possibilities of working with such data.

Further, eye-tracking data may also skew the researcher’s member’s understanding during the analytic process by providing a default focus on gaze as the primary resource for understanding the social activity as discussed above (section 2). The latter problem may be lessened by following an analytic procedure where the first step of the analysis is looking at the video recording of a shopping interaction, and the second step is to include the eye-tracking recording. In that way, the analytic procedure follows the core EMCA analytic practice of describing members’ understanding in interaction as this is publicly demonstrated in and through their actions by members for members (ten Have, 2005). Since these actions are publicly available, they are consequently also available for bystanders, including analysts, and they may also be recorded on video. By looking at the videos first, the analytic procedure thus aims to develop the analysts’ member’s understanding based on “what is publicly available” before the eye-tracking recording is introduced. In that way, the eye-tracking

recording can inform the analysis without bypassing the analyst's intuitive understanding of the situation. However, that raises the question of what kind of additional insight may be gained from using eye-tracking data if the analytic procedure introduced above is followed, which will be explored in the next section by means of three analytic examples.

4. The potential of using eye-tracking data for analyses of social conduct: Analytic examples

In this section, the analytic procedure for doing social analysis which was outlined above will be demonstrated and illustrated through analyses of three activities that one customer engages in while doing shopping. The customer is in a store which sells second-hand furniture, vases, plates, cups, etc. as well as coffee, tea, and other local specialties. In this type of store, customers know that they can never be sure what items they'll find for sale, and they know that there is not likely to be two items of a kind.

In accordance with the analytic procedure described above, each analysis begins by analyzing the video recording of the activity; next, the eye-tracking recording is analyzed; and finally, the methodological implications of these analytic steps for (multimodal) EMCA research are discussed. Thus, the first analysis informs a discussion of the potential of using details of gaze behavior captured by the eye-tracking recordings (section 4.1); the second analysis enables a discussion of the mutual support of video and eye-tracking recordings (section 4.2); and the third analysis problematizes the use of eye-tracking recordings for identifying and capturing social actions of interest to EMCA analyses (section 4.3). The aim of the analyses is thus to explore whether it is possible to gain any additional analytic insight by using eye-tracking data in EMCA analysis.

4.1 The analytic potential of using details of gaze behavior captured by the eye-tracking recordings

Prior to the beginning of Example 5, the customer is walking around the store, moving between tables and shelves and looking around at the items displayed.

(>>Video: Example5<<)

At the beginning of the clip she stops, turns her torso toward some shelves placed along the wall, and positions herself at a distance from them. Her body remains in that position while she moves her head and gaze from side to side and up and down. Her gaze direction seems to change fairly rapidly, and it spans the entire set of shelves. The gaze direction does not seem to rest on objects with a particular shape, color, or function — she seems to gaze at the yellow dishes first (00:00-00:02), then a red can (00:07-00:08), and later some of the blue porcelain (00:22-00:23). Her body posture does not change when her gaze direction

changes. As the video clip ends, she finishes looking and moves to another location in the shop.

The embodied actions of the customer indicate a low-involvement level of interest (Clark & Pinch, 2010), recognizable as doing “just looking”, that is, the customer is engaged in the activity of browsing the items on sale in the shop to see if there is anything interesting.

When the eye-tracking recording of the activity is included along with the video recording as the second step in the analysis, as it is in Example 6, it confirms that the movements of the head and eyes, i.e., the gaze direction, that were visible in the video recording do in fact correspond to the eye movements and fixations as recorded by the eye-tracking equipment.

(>>Video: Example6<<)

In the eye-tracking recording, as already mentioned, the eye movements of the person in the video are marked by a line, and eye fixations are marked by a circle. In the analyses, the terms “gaze” and “gaze direction” will be used to describe when a person is looking as indicated by the combined head and eyes movement (Kendon, 1967, 1973) as visible for co-present others and also visible for analysts on the video recording. Terms such as “eye fixation” and “eye movement” will be used to describe when a person is looking as indicated by the dots and lines on the eye-tracking recordings.

The eye-tracking recording demonstrates how the eyes do indeed move quite rapidly over a series of items. The fixations of the eyes do not remain for long on any particular item: The eyes first fixate on the red can (00:06-00:08), then a crystal bowl on a dresser (00:09-00:10), and then some black items on the bottom shelf (00:15-00:17). Further, the eyes move across shelves, up and down, and from side to side.

The eye-tracking recording thus demonstrates that the fixations of the eyes match the gaze direction which is observable from the video recording. The eye-tracking recording also confirms that the tempo of the eye movements is rather high, i.e., that the fixations are quite short. The eye-tracking recording thus supports the understanding of the social action which the video allowed us to recognize: This is a person who is browsing.

Further, the eye-tracking recording provides additional details about precisely which objects are fixated and how short the fixations are. The eye-tracking recording may also provide additional details not immediately visible on the video recording, e.g., that the fixations of the eyes move beyond the shelves to the dresser by the window (00:09-00:13).

In that way, analyses may make use of the eye-tracking recording for 1) confirming what is observable from the video, in this case the social activity of doing browsing; 2) giving a more detailed understanding of the gaze direction

than what the video provides, in this case, the duration of the fixation periods on particular “browsable” items; 3) providing information about additional details that are not clear from the video recording, i.e., that the customer may be looking further in some direction than we can see from the video, in this case moving her eyes not only to the shelves but also to the dresser next to the shelves.

The first analysis thus demonstrates how the video recording and the eye-tracking recording may reflexively elaborate each other, confirming the claims by Stukenbrock (2018) and Stukenbrock and Dao (Stukenbrock & Dao, 2019) that eye-tracking recordings yield more precise information with regard to gaze direction than video recordings.

However, it is important to consider that despite the first-person perspective, the eye-tracking recordings do not in fact correspond to what the participants see. For instance, the size of the circles representing fixations don't correspond to the specific area focused on by the participant. Rather, the information about eye movement and fixation displayed in the eye-tracking recording is only available as visual output as the result of a series of operations by software and hardware, including e.g. calibration of the data, translation of the data collected to visual output, etc. (Duchowski, 2007). And these operations are designed and implemented not by the researchers collecting the data but by technicians and engineers. In other words, researchers depend on technical decisions beyond their control or understanding for gaining access to eye-tracking data. As a consequence, it is debatable who in fact defines what researchers end up having access to on the eye-tracking recordings: the researchers or the technicians who developed the equipment. This raises the further question of how to treat any discrepancies that may be discovered between the video recording and the corresponding eye-tracking recording: Is there an actual discrepancy, or does it result from the measuring equipment or from the calculations generating the visual output?

4.2 Mutual support of video and eye-tracking recordings

The second analysis deals with another social action conducted by the same customer: searching. Again, the video recording is analyzed first, and secondly, the eye-tracking recording is included in the analysis.

The customer and the seller are at the counter, and they have just reached an agreement on the price of a piece of artwork on display in the shop. Example 7 opens with the customer initiating a new topic by stating that she also needs some tea.

(>>Video: Example7<<)

Extract 2. Searching for Earl Grey tea

- 1 C: ved du hvad så s:ka jeg os bruge noget te
you know what I also need some tea
C places the box on the counter; S takes it
- 2 S: ja (.) hva (.) ska det være for noget
yes what kind should it be
C>counter, to her left
- 3 C: .hhhh j:a
Well
C>shelves with tea
S places the box to the left
- 4 S: ska jeg lige rulle noget papir om dem /her
would you like me to wrap some paper around these
S touches the wrapping paper /S>bottles
- 5 C: ja [det må du gerne]
yes please do
- 6 S: [så de ik klasker] imod hinanden
so they don't smash against each other
S places a bottle on the paper
- 7 C: det må du gerne
please do
- 8 Ps (2.4)
S wraps the paper around the bottle-----
- 9 C: øh jam jeg vil
eh well I
C>shelves on her left

- 10 jeg sku egentlig bruge en øh
I really needed an eh
-----//
- 11 Ps (0.7)
S takes some tape
- 12 en earl grey men det må /godt være (.)
an earl grey but I'd like it to be
S finishes wrapping the bottle, takes the next one
/S>C

- 13 Ps (1.2)
S>C
C>shelves on her left
S takes a bottle, turns her head toward the shelves
- 14 med det hele jeg ka se den du har det sån en uden
with everything I can see that the one you have that's one
without
S>shelves
C>shelves
- 15 Ps (0.4)
- 16 S: jeg har en med jasmin øh: og så har [je:g]
I have one with jasmine eh and I have
S wraps the bottle-----
S>shelves
- 17 C: [og du har os]
and you also have
- 18 en /økologisk øh \earl grey [ka jeg se
an organic eh earl grey I see
/C points to shelf
\C lowers her hand

- 19 S: [ja det /har jeg] den den er ren
yes I have that that's pure
/S>bottle

- 20 og så har jeg altså den der tempeltræ (0.5) /som faktisk os er
and then I have that temple tree which is actually also
-----*//S takes some tape*
- 21 en earl grey hvor der er kornblomster og va/nilje i
an earl grey with cornflowers and vanilla in it
/S>C
- 22 Ps (0.4)
>bottles on the counter
- 23 den er sådan li/dt blid
it is kind of gentle
/S>C

As she makes the statement (line 1), the customer turns her head and torso toward the boxes with tea displayed on shelves behind the counter (00:01-00:02). Her gaze seems directed to the boxes to her left, since her head remains turned toward this part of the shelves for some time. In contrast to the shelves in the

former example, these shelves all display the same type of item: boxes with different types of tea.

The seller inquires what kind of tea the customer wants (line 2), and the customer responds by an inbreath sound (".hhh") and a hesitating "j:a" ("well") (line 3, 00:03-00:06), which indicates that she is doing being engaged in making a choice, or perhaps in doing being thinking or some other "cognitive matter" (Coulter, 1991). While producing this turn, the customer simultaneously leans forward, thus bringing her torso and head slightly closer to the boxes and aligning her body with the display of boxes, indicating a high level of involvement with them (Clark & Pinch, 2010). The customer moves her head slightly from side to side as if she is moving her gaze from one box to the next.

The clerk then initiates a side sequence (Jefferson, 1972), asking whether she should wrap the bottles that the customer is also buying (lines 4 and 6, 00:06-00:09). The customer treats the question as an insertion (Schegloff, 1972): She changes her body posture, shifting her weight to one leg, and changes her gaze direction to the bottles on the counter while responding (line 5, 00:08).

As she repeats her response to the inserted question (line 7, 00:09-00:10), the customer indicates that she is resuming the prior activity by shifting her weight back to the other leg and directing her gaze to the boxes on her left again, resuming her body alignment and high-level involvement with the boxes displayed.

After a pause (line 8), the customer describes the type of tea that she is looking for (lines 9-14, 00:13-00:22) — an Earl Grey — and the position of her head indicates a gaze direction toward the boxes to her left, moving from side to side. Her gaze seems to remain within the same limited space, which may indicate that this is where the Earl Grey teas are placed. The turn at talk in which she mentions Earl Grey (line 12, 00:17) and her gaze behavior thus reflexively make her action recognizable as doing searching, that is, looking for something specific. Note that the clerk, when looking up (line 14, 00:20), also turns toward the same boxes, thus treating that as the relevant visual field (Goodwin, 2000).

When the eye-tracking recording is added to the data material as in Example 8, it primarily serves to confirm what is evident from the video recording.

(>>Video: Example8<<)

The customer actually does look at the visual field indicated by her gaze direction as witnessable from the video, and her eyes do in fact move along the shelves as indicated by her head movement. Further, the eye-tracking recording shows that all her eye fixations are within the frame of the shelves. Her eye fixations move over a series of boxes whose exteriors are very similar, in contrast to the previous example where the items were of different shapes and colors.

Moreover, the eye-tracking recording provides details that are not available from the video: The customer's eye fixations are specifically and systematically on the labels of the boxes where the name of the tea inside each box is written. Together with the customer's talk, this pattern of eye movements and fixations shows that she remains engaged in an activity in relation to the same category — that of Earl Grey tea. Toward the end of the clip, the customer fixates her eyes on labels that match the teas mentioned in the talk, e.g., “temple tree” (line 20, 00:30-00:31).

Thus, whereas the first analytic example — Example 5 and Example 6 — showed the customer engaged in browsing in the sense of “just looking”, this example shows the customer engaged in searching in the sense of looking for something specific. It is also evident that the gaze behavior in the second analytic example — Example 7 and Example 8 — differs from the gaze behavior in the first analytic example. In that way, the analysis of the second analytic example confirms one of the methodological insights from the analysis of the first example — that the video recording and eye-tracking recordings elaborate each other — and shows that the different types of data may mutually support each other in substantiating descriptions and understandings of different kinds of social behavior, i.e., browsing versus searching. Eye-tracking recordings may thus supply information about subtle shifts in gaze that are easily missed on video recordings or possibly not available because of less-than-ideal camera angles, lighting, or other situational factors. Note, however, that while the eye-tracking recordings in our data elaborate and substantiate information available in the video recordings, this may not always be the case (Stukenbrock, 2015).

For analytic processes in EMCA, eye-tracking data may thus contribute to substantiating the analytic work with such comparisons by making available more detailed information about eye movements and fixations and their coordination with other interactional resources. More specifically, this is seen in connection with, for example, collection work that includes elements of comparisons of specific actions in relation to details in their construction; sequential environments etc. with the purpose of describing and delimiting a phenomenon — such as how restarts of turn beginnings are systematically linked to the achievement of mutual gazing between speaker and hearer (Goodwin, 1980); or how customers and sales clerks achieve jointly coordinated entry into a service encounter (Harjunpää, Mondada, & Svinhufvud, 2018).

4.3 The use of eye-tracking recordings to identify and capture members' understanding of situated actions in social interaction

The above argument that the difference between two actions does not reside only in the pattern of eye movements and fixations can be illustrated by attempting to begin the next analysis with the eye-tracking recording rather than with the video recording.

(>>Video: Example9<<)

The pattern of eye movements and fixations in the eye-tracking recording in Example 9 is clearly different from that documented in Example 6 and Example 8 which were shown in the previous analyses to be instances of "browsing" and "searching", respectively.

At the beginning of the clip (00:00-00:08), the customer's eyes move between a number of objects on or near the table. We can see that she moves closer to the table. The eye fixations then move to the black box on the table, and the rest of the clip is characterized by a series of long fixations (compared to the previous examples), all of which remain within the frame of the box. Further, we can see that the customer moves even closer to the box, to the extent that the box at the end of the example fills up almost the entire visual field on the video recording (00:17-00:18).

In Example 9, the eyes of the customer thus move and fixate within the boundaries of an object, not a shelf. This indicates that the customer must be engaged in a different action than the previous actions, "browsing" and "searching", which both involved moving the eyes across different objects, either across categories of shape, color, function, or within a certain category — in this particular case, Earl Grey tea.

It is noticeable that the eye fixations in the third analytic example are of longer duration than in the previous clips. Further, the eyes in this example also fixate on different details of the same object — in contrast to previous examples where the fixations were on different objects either of different categories or within the same category. Both observations seem to indicate a more thorough engagement with the object than in the previous examples. Nevertheless, the eye-tracking recording seen in isolation does not allow the analyst to identify the social action which the customer is engaged in, i.e., what is observable to co-present others. Rather, the understanding that the duration of the fixation seems to be significant for what is going on is inferred from a comparison with prior examples — examples which have, importantly, been described, analyzed and understood based *firstly* on social interactional analyses of the recordings and *secondly* on inclusion of the eye-tracking recordings in the analysis in order to arrive at an analysis based on both the video recordings and the eye-tracking recordings.

To provide an understanding of what kind of action the duration of fixations may possibly be indicative of, e.g., that the individual is "just looking", "staring", "looking for something specific on the object", this pattern of eye movements and fixations must therefore be analyzed in relation to how it is embedded in a bodied context, i.e., how it co-occurs with other "bodied modalities" from which specific social actions emerge that are recognizable to other co-present customers (Rasmussen, 2018).

The context for the third analytic example is provided here in terms of the corresponding video recording in Example 10.

(>>Video: Example10<<)

The video recording shows how the customer initially moves into a separate space in the store (00:00-00:01). As she turns the corner, her gaze turns toward the table located by the wall to her right, previously out of sight. She continues to move into the room, around a chair and a table, while her head remains turned in the direction of the table (00:02-00:05). Her head moves from side to side, displaying a low level of involvement with the items displayed. At one point, her head turns in the direction of the black box as she moves further into the space. First, only her head is turned in the direction of the box, and then she turns her entire body in that direction, indicating a higher level of involvement. When the customer stops moving forward (00:07), she initially positions herself at a distance from the objects toward which she seems to direct her gaze. She then moves first her head, and then her torso closer to the objects. In the video recording, her actions are thus recognizable as first “browsing”, which then becomes “examining” a particular object. The customer’s embodied action, body position, body posture, and navigation in what is recognizable as a specific direction due to her gaze behavior and stopping, render her actions observably understandable as “zooming in on” and “examining”; i.e., actions that may potentially be observed and responded to by other co-present individuals (Goffman, 1963). Immediately after Example 10 ends, the customer proceeds to pick up the object and examine it further.

(>>Video: Example11<<)

Combining the two recordings as in Example 11, a further observation can be made: The head movements which the customer makes at the beginning of the clip (00:00-00:07) and which resemble the head movements described in the previous examples, i.e., head movements that are associated with “browsing” and “searching”, stop at the same time as the customer’s eye fixations start to remain within the frame of the object focused on, i.e., the black box (00:07). In other words, eye-tracking recordings may confirm the assumptions of co-present others and analysts that eye-movement and fixations co-occur with the change in bodily posture and movement, which are in turn fitted to the material environment in which they occur. Co-present others constitute elements of the environment as well who witness and observe, etc. In this case (see Figures 3 and 4) several other customers are in the store during the recording who may see and understand her actions as just those actions (and implicit visual perception).

Figure 3. *The participant (wearing a white skirt) browsing and moving into the separate space. Another customer is moving past her in the opposite direction*

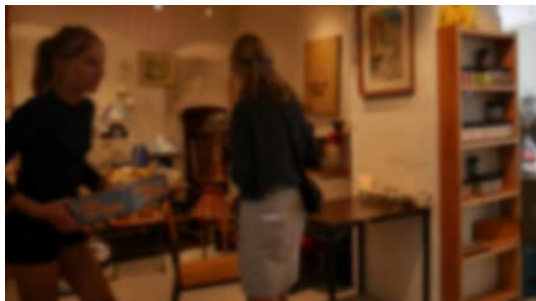


Figure 4. *The participant browsing the shelves. Note the other customers sitting in the background. They are having coffee and playing a board game*



Thus, her actions are in fact public and accountable and may be seen and understood by others *as just those actions*. They may also be taken up by others in various types of interaction, from coordinating conduct — e.g., by sharing space in self-service shops in ways that serve to minimize involvement (Rasmussen & Kristiansen, 2021) — to higher social involvement in terms of longer stretches of sequentially organized actions, which will typically also involve speech (Harjunpää et al., 2018).

Notice, however, that whereas the social action of the customer is immediately recognizable from the video recording, this is not the case for the eye-tracking recording. On the latter the eye fixations visibly move rather quickly and across a large area. But what is recorded by the eye-tracking equipment does not make the social action to which the movement of eye fixations contribute recognizable to viewers of the recording as it would be (and demonstrably is, by reference to the video recording) by co-present individuals who witness the customer's conduct (cf. Clark & Pinch, 2010; Rasmussen & Kristiansen, 2021). That is, the difference between “browsing” and “searching” does not lie in the pattern of eye movements and fixations as documented in the eye-tracking recording: “Browsing” and “searching” are embodied actions carried out in specific situations and are recognizable to other co-present people *as those actions* (Clark & Pinch, 2010). For the research field of (multimodal) EMCA, it is the kind of conduct and social actions that are witnessable and observable that fall within its analytic interests, not information, however detailed, which is unavailable to co-participants in the situated interaction.

5. Conclusion

In EMCA research, video technology has so far been used with the purpose of grasping “something” (Macbeth, 1999, p. 154), e.g., “natural” conversations in their everyday environments as they naturally unfold. The camera has primarily

been used as a means to observe (Mondada, 2006), record, i.e., account in and through recording (Macbeth, 1999) and preserve non-experimental interactions with the possibility of visually documenting the finding of “something” in an order of social practice, or to microscope (Schmidt & Wiese, 2019) sequences of embodied (and multimodal) ordinary conduct. The data is overwhelmingly produced by the researcher rather than by the participants, and work is done to minimize the impact of the researcher on the situation to avoid contaminating the “naturalness” of the data, i.e., social interaction between and amongst the observed co-present co-participants.

The use of eye-tracking recordings in combination with the camera does not, of course, remove the filmmaker from the scene. It adds yet another filmmaker to it, i.e., the co-participant(s) with eye-tracking glasses: The co-participant to be observed in everyday life situations becomes the “observer” and “documentalist” and may even become a participant observer (Kawulich, 2005), potentially with a focus on looking through the glasses and on herself looking at the situation through them as part of the social order of the situation.

Therefore, the use of eye-tracking recordings calls for as many academic discussions as video recordings have been and are subject to. This paper has aimed to point out some of the challenges and possibilities, ways of resolving some of the challenges, and what seemed unresolvable.

The paper has pointed out that the analytic procedure developed in the RESEMINA project for working with data sets combining eye-tracking recordings with video recordings enables an EMCA analysis where the detailed information on eye movement and fixations provided by the eye-tracking recordings is fitted into an already examined local social and embodied context provided by the video recordings.

Eye-tracking recordings are used in the analysis to support, substantiate and elaborate the analytic findings made on the basis of the video recordings. The eye-tracking data thus provides empirical evidence, which has previously not been available, that the gaze direction deduced by studying the participants’ publicly available head and eye movements does in fact correspond to what participants fixate with their eyes. But eye-tracking recordings do not influence the social analyses which are based on participants’ actions that are observable and witnessable to co-present others.

This analytic procedure mirrors members’ practices: Members’ social analyses of their co-participants and their social actions are not influenced by the circumstance that their co-participant cannot actually look through the eyes of the analyzing member. People construct their social actions as socially recognizable: We take it for granted that if you were in my place, you would do the same and understand the same as me, and that if I were in your place, I would do the same and understand the same as you (Garfinkel, 1967). That means that as members

and as analysts, we are taking it for granted that others can see what we see³, but we are *not* taking it for granted that they can see it *through* our eyes.

The paper also points out that eye-tracking technology is not methodologically unproblematic for EMCA interested in social action since eye-tracking recordings visualize information about eye movement and fixations, added onto first person perspective video recordings, which is specifically unavailable to co-participants and cannot for that very reason have any social significance for the participants in the interaction.

The paper argues, however, that the information available through eye-tracking recordings may still be useful for social analyses, provided that the analysis can demonstrate that the information is also oriented to by the participant wearing the eye-tracking equipment and co-present others. In an article on minimizing involvement in Danish self-service shops (Rasmussen & Kristiansen, 2021), eye-tracking recordings play an important role in the analytic process. The phenomenon discussed in the article concerns how customers get and stay out of each other's way, and the eye-tracking recordings provide access to the phenomenon because their first-person perspective allows the analyst to track when another customer enters the peripheral vision of the participant and how the participant orients to this by, e.g., moving away or stopping in their tracks to allow the other customer to pass. Importantly, the phenomenon is also recognizable in customers who are not wearing eye-tracking equipment as well as in video data where no eye-tracking equipment was used. In that way, Rasmussen's and Kristiansen's study demonstrates how eye-tracking data may indeed contribute substantially to new analytic insights in EMCA analysis of social action, provided that the information gained from the eye-tracking data is demonstrably relevant to the participants in the situation.

Moreover, the paper points out that while it is possible to use eye-tracking recordings in EMCA analysis to gain new analytic insight, adding eye-tracking recordings to a data corpus will not bring researchers any closer to the event, or to the participant's perspective. This is not possible, no matter how much or how advanced recording equipment we use for producing our data (McIlvenny, 2019). Rather than bringing us closer to understanding, the technology may instead result in methodological problems, depending on our research interests.

Consequently, gaining access to detailed records of people's eye movements and fixations is in itself unlikely to lead to new insights into people's social actions, though these records may be valuable to interactional research with an interest specifically in gaze. As the data set documents there is no clear one-to-one relationship between eye movements and fixations and social understanding: It is not possible to understand a social action by examining eye movement and

³ This does not mean that members who cannot be seen, e.g. by a blind person, do not orient to people who cannot see, cf. Due, B. L. (Ed.). (forthc.). *The Practical Accomplishment of Blind People's Ordinary Activities*. Routledge. .

fixations in isolation. In the end, neither members nor analysts can understand the action that a certain pattern of eye movements and fixations accomplishes without having access to the context to which the eye movements and fixations are fitted. One cannot extract one “resource” which, in a combination with other resources, results in the emergence of some action in order to find what emerged from the combination in that one resource (Rasmussen, 2019).

Finally, the paper points out that using eye-tracking equipment when collecting data for EMCA research projects may compromise the naturalness of the setting: We may end up turning the “natural lab” which EMCA researchers seek into an “experimental lab” by bringing the equipment from the lab into the field. For researchers interested in gaze in interaction, moving from experimental settings to mobile interaction in natural surroundings provides access to interaction which is more natural. But for researchers interested in ordinary social interaction in natural environments, introducing eye-tracking equipment, at least in its current state of technological development, into the setting may be distracting for the participants, thus compromising the validity of the data — and introducing two sets of eye-tracking glasses may increase the distraction more than what is warranted by the added value which it may contribute. Except, of course, if one is interested in studying the natural interaction of two participants wearing eye-tracking glasses.

Overall, the conclusion of this study is thus that while mobile eye-tracking equipment and the resulting recordings may provide new analytic possibilities for studying social action, they are by no means indispensable for such research, and their use in research in social interaction runs the risk of compromising both the data and the analytic procedure, which necessitates measures aimed at reducing the compromising of data and analysis. Some of these measures have been discussed in this article.

Acknowledgements

The RESEMINA project is funded by the Velux foundations (grant no.00016980).

References

- Atkinson, J. M., & Heritage, J. (Eds.). (1984). *Structures of social action: Studies in conversation analysis*. Cambridge University Press, Editions de la Maison des Sciences de l'Homme.
- Auer, P. (2018). Gaze, addressee selection and turn-taking in three-party interaction. In G. Brône & B. Oben (Eds.), *Eye-tracking in Interaction: Studies on the role of eye gaze in dialogue* (Vol. 10, pp. 197-232). John Benjamins Publishing Company.

- Brône, G., & Oben, B. (Eds.). (2018a). *Eye-tracking in Interaction: Studies on the role of eye gaze in dialogue* (Vol. 10). John Benjamins Publishing Company.
- Brône, G., & Oben, B. (2018b). Introduction. Gaze, interaction and eye-tracking: A multidisciplinary endeavour. In G. Brône & B. Oben (Eds.), *Eye-tracking in Interaction: Studies on the role of eye gaze in dialogue* (Vol. 10, pp. 1-18). John Benjamins Publishing Company.
- Clark, C., & Pinch, T. (2010). Some major organisational consequences of some 'minor', organised conduct: evidence from video analysis of pre-verbal service encounters in service encounters in a showroom retail store. In N. Llewellyn & J. Hindmarsh (Eds.), *Organisation, interaction and practice: Studies of ethnomethodology and conversation analysis* (pp. 140-171). Cambridge University Press.
- Coulter, J. (1991). Cognition: 'cognition' in an ethnomethodological mode. In G. Button (Ed.), *Ethnomethodology and the human sciences* (pp. 176-195). Cambridge University Press.
- De Stefani, E. (2013). The collaborative organisation of next actions in a semiotically rich environment: shopping as a couple. In P. Haddington, L. Mondada, & M. Nevile (Eds.), *Interaction and mobility: Language and the body in motion* (pp. 123-151). De Gruyter.
- Deppermann, A. (2013). Conversation analytic studies of multimodal interaction. Special issue of *Journal of Pragmatics*. *Journal of Pragmatics*, 46(1), 1-172.
- Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. Springer.
- Due, B. L. (Ed.). (forthc.). *The Practical Accomplishment of Blind People's Ordinary Activities*. Routledge.
- Fasulo, A., & Monzoni, C. (2009). Assessing mutable objects: A multimodal analysis. *Research on Language and Social Interaction*, 42(4), 362-376.
- Ferguson, H. J., Apperly, I., & Cane, J. E. (2017). Eye tracking reveals the cost of switching between self and other perspectives in a visual perspective-taking task. *Quarterly Journal of Experimental Psychology*, 70(8), 1646-1660. <https://doi.org/10.1080/17470218.2016.1199716>
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice Hall.
- Garfinkel, H., & Wieder, D. L. (1992). Two incommensurable, asymmetrically alternate technologies of social analysis. In G. Watson & R. M. Seiler

(Eds.), *Text in context: Contributions to ethnomethodology* (pp. 175-206). Sage Publications

- Goffman, E. (1963). *Behavior in public places. Notes on the social organization of gatherings*. The Free Press.
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97-121). Halsted Press.
- Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological inquiry*, 50(3-4), 272-302. <https://doi.org/10.1111/j.1475-682X.1980.tb00023.x>
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489-1522.
- Haddington, P., Keisanen, T., Mondada, L., & Nevile, M. (2014). *Multiactivity in social interaction: Beyond multitasking*. John Benjamins Publishing Company.
- Harjunpää, K., Mondada, L., & Svinhufvud, K. (2018). The Coordinated Entry into Service Encounters in Food Shops: Managing Interactional Space, Availability, and Service During Openings. *Research on Language and Social Interaction*, 51(3), 271-291. <https://doi.org/10.1080/08351813.2018.1485231>
- Hayashi, M. (2005). Joint turn construction through language and the body: Notes on embodiment in coordinated participation in situated activities. *Semiotica: Journal of the International Association for Semiotic Studies/Revue de l'Association Internationale de Sémiotique*, 156(1), 21-53.
- Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources - CA studies of social conduct. *Journal of Pragmatics*, 65, 1-9.
- Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge University Press/Editions de la Maison des Sciences de l'Homme.
- Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in Qualitative Research: Analysing Social Interaction in Everyday Life*. Sage.
- Heath, C., & Luff, P. (1992). Collaboration and control: Crisis management and multimedia technology in London Underground Line Control Rooms. *Computer Supported Cooperative Work (CSCW)*, 1(1-2), 69-94.

- Hodge, R., & Kress, G. R. (1998). *Social semiotics*. Polity Press.
- Holler, J., & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: optimizing reciprocity. *Frontiers in psychology, 6*(98), 76-89.
- Jefferson, G. (1972). Side sequences. In D. Sudnow (Ed.), *Studies in social interaction* (pp. 294-338). The Free Press.
- Jewitt, C. (2008). *Technology, literacy and learning: A multimodal approach*. Routledge.
- Jones, R. H. (2005). 'You show me yours, I'll show you mine': The negotiation of shifts from textual to visual modes in computer-mediated interaction among gay men. *Visual communication, 4*(1), 69-92.
- Kawulich, B. B. (2005). Participant observation as a data collection method. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*,
- Keevallik, L. (2010). Bodily Quoting in Dance Correction. *Research on Language and Social Interaction, 43*(4), 401-426.
<https://doi.org/10.1080/08351813.2010.518065>
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta psychologica, 26*, 22-63.
- Kendon, A. (1973). The role of visible behavior in the organization of social interaction. In M. v. Carnach & I. Vine (Eds.), *Social communication and movement* (pp. 29-74). Academic Press.
- Knoblauch, H., Schnettler, B., Raab, J., & Soeffner, H.-G. (Eds.). (2006). *Video analysis: methodology and methods. Qualitative Audiovisual Data Analysis in Sociology*. Peter Lang.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The Modes and Media of Contemporary Communication*. Cappelén.
- Kristiansen, E. D., & Rasmussen, G. (forthc.). Would you like a bag for that? Environmental awareness and changing practices for closing buying and selling encounters in retail shopping.
- Licoppe, C., & Morel, J. (2012). Video-in-interaction: "Talking heads" and the multimodal organization of mobile and Skype video calls. *Research on Language & Social Interaction, 45*(4), 399-429.
- Lohse, G. L., & Wu, D. (2001). Eye movement patterns on Chinese yellow pages advertising. *Electronic Markets, 11*(2), 87-96.

- Macbeth, D. (1999). Glances, trances, and their relevance for a visual sociology. In P. L. Jalbert (Ed.), *Media studies: Ethnomethodological approaches* (pp. 135-170). University Press of America.
- Maynard, D. W., & Clayman, S. E. (1991). The diversity of ethnomethodology. *Annual review of sociology*, 385-418.
- McIlvenny, P. (2019). Inhabiting spatial video and audio data: Towards a scenographic turn in the analysis of social interaction. *Social Interaction. Video-Based Studies of Human Sociality*, 2(1).
<https://doi.org/10.7146/si.v2i1.110409>
- Meißner, M., Pfeiffer, J., Pfeiffer, T., & Oppewal, H. (2019). Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research. *Journal of Business Research*, 100, 445-458. <https://doi.org/https://doi.org/10.1016/j.jbusres.2017.09.028>
- Meyer, C., & Streeck, J. (2020). Ambivalences of touch: An epilogue. In A. Cekaite & L. Mondada (Eds.), *Touch in social interaction: touch, language and body* (pp. 311-325). Routledge.
- Mondada, L. (2006). Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In H. Knoblauch, B. Schnettler, J. Raab, & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and methods. Qualitative audio-visual data analysis in sociology* (pp. 51-68). Peter Lang.
- Mondada, L. (2012). The Conversation Analytic Approach to Data Collection. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 32-56). <https://doi.org/https://doi.org/10.1002/9781118325001.ch3>
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137-156.
- Mondada, L. (2017). Walking and talking together: Questions/answers and mobile participation in guided visits. *Social science information*, 56(2), 220-253.
- Nevile, M. (2015). The embodied turn in research on language and social interaction. *Research on Language and Social Interaction*, 48(2), 121-151.
- Oben, B. (2018). Gaze as a predictor for lexical and gestural alignment. In G. Brône & B. Oben (Eds.), *Eyetracking in interaction: Studies on the role of eye gaze in dialogue* (pp. 233-262). John Benjamins Publishing Company.
- Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of

gaze in social interaction. *Neuroscience & Biobehavioral Reviews*, 37(10, Part 2), 2516-2528.

<https://doi.org/https://doi.org/10.1016/j.neubiorev.2013.07.017>

- Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of research in Marketing*, 16(1), 1-16.
- Poulsen, S. V., Rasmussen, G., & Boeriis, M. (forthc.). What's in a semiotic resource?
- Rasmussen, G. (2018). On the social constraints of having a world. *RASK: International journal of language and communication*, 7(Special Issue: The World at Your Feet), 79-107.
- Rasmussen, G. (2019). Emergence and Human Communication. In J. S. Damico & M. J. Ball (Eds.), *The SAGE Encyclopedia of Human Communication Sciences and Disorders* (pp. 681-683). SAGE Publications.
- Rasmussen, G., & Kristiansen, E. D. (2021). The sociality of minimizing involvement in Danish self-service shops: Danish customers' multi-modal practices of being, getting and staying out of the way. *Discourse & Communication, Special issue: The 'human touch' in online and offline buying and selling interaction* (Gitte Rasmussen, Theo van Leeuwen, eds.).
- Rasmussen, G., Kristiansen, E. D., & Andersen, E. M. (2019). Working out availability, unavailability and awayness in social face-to-face encounters: The case of dementia. *Discourse Studies*, 21(3), 258-279.
- Rasmussen, G., Kristiansen, E. D., & Poulsen, S. V. (forthc.). The World of Daily Life: Doing a search for (e-)shopping purposes.
- Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, 73(2), 291-296.
- Rossano, F. (2012). *Gaze behavior in face-to-face interaction*. Max Planck Institute for Psycholinguistics Series.
- Sacks, H. (1984). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 21-27). Cambridge University Press, Editions de la Maison des Sciences de l'Homme. <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: Journal of the Linguistic Society of America*, 696-735. <https://doi.org/10.2307/412243>
- Schegloff, E. A. (1972). Notes on a conversational practice: Formulating place. In J. Coulter (Ed.), *Ethnomethodological sociology* (pp. 75-119). Edward Elgar Publishing.
- Schegloff, E. A. (2007). *Sequence organization in interaction: Volume 1: A primer in conversation analysis* (Vol. 1). Cambridge University Press.
- Schmidt, R., & Wiese, B. (2019). Online Participant Videos: A New Type of Data for Interpretative Social Research. *Qualitative Social Research*, 20(2).
- Streeck, J. (1993). Gesture as communication I: Its coordination with gaze and speech. *Communications Monographs*, 60(4), 275-299.
- Stukenbrock, A. (2015). *Deixis in der face-to-face-Interaktion*. de Gruyter.
- Stukenbrock, A. (2018). Mobile dual eye-tracking in face-to-face interaction: The case of deixis and joint attention. In G. Brône & B. Oben (Eds.), *Eye-tracking in Interaction: Studies on the role of eye gaze in dialogue* (Vol. 10, pp. 265-300). John Benjamins Publishing Company.
- Stukenbrock, A. (2020). Deixis, Meta-Perceptive Gaze Practices, and the Interactional Achievement of Joint Attention [Original Research]. *Frontiers in psychology*, 11(1779). <https://doi.org/10.3389/fpsyg.2020.01779>
- Stukenbrock, A., & Dao, A. N. (2019). Joint attention in passing: What dual mobile eye tracking reveals about gaze in coordinating embodied activities at a market. In R. E. & G. C. (Eds.), *Embodied Activities in Face-to-face and Mediated Settings* (pp. 177-213). Palgrave MacMillan.
- ten Have, P. (2005). The notion of member is the heart of the matter: On the role of membership knowledge in ethnomethodological inquiry. *Historical Social Research/Historische Sozialforschung*, 30(1), 28-53.
- vom Lehn, D., & Heath, C. (2016). Action at the exhibit face: video and the analysis of social interaction in museums and galleries. *Journal of Marketing Management*, 32(15-16), 1441-1457. <https://doi.org/10.1080/0267257X.2016.1188846>
- Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. In M. N. K. (Ed.), *Review of marketing research* (Vol. 4, pp. 123-147). M. E. Sharpe.

- Weiß, C., & Auer, P. (2016). Das Blickverhalten des Rezipienten bei Sprecherhäsitationen: eine explorative Studie. *Gesprächsforschung Online Zeitschrift verbaler Interaktion*, 17, 32-167.
- Wooffitt, R., & Hutchby, I. (1998). *Conversation Analysis: Principles, Practices and Applications*. Polity Press.
- Zhu, H., Otsuji, E., & Pennycook, A. (2017). Multilingual, multisensory and multimodal repertoires in corner shops, streets and markets: introduction. *Social Semiotics*, 27(4), 383-393.