Situatedness and adaptation of practices in video-mediated interaction: reflections on research design, observed phenomena, and didactical potentials

Silke Tork
European University Institute Firenze

E-mail: silke.tork@eui.eu

Abstract. While web based video-conferencing interested researchers since its beginnings, it was the Covid 19-pandemic that led to an immense increase in the use of video-conferencing platforms as well as in related research. Not only the areas of research on video-mediated interaction expanded, but research itself was impacted by these developments. Due to the restrictions during the pandemic, researchers depending on natural data and working with research subjects were confronted with challenges regarding research design, data collection, theoretical assumptions and methodological approaches. This paper presents one of these cases, a study of multimodal interaction in task-oriented group work in higher education with two differently situated data sets: one in a classroom and one on a video-conferencing platform. Tracing the process from data collection to analysis, the paper discusses basic principles and concepts of an EMCA approach to multimodal interaction in relation to the specific settings and observed phenomena, e. g., interaction space, inter- and intra-personal embodied actions, as well as technical issues. In selected sequences, phenomena like verbal overlaps and embodied feedback are highlighted, suggesting participants' adaptation to their situatedness. Following a twofold objective, the paper concludes by relating the findings to existing and future research, as well as to didactical considerations on video-mediated interaction in educational settings.

1. Introduction

While web based video-conferencing interested researchers since its beginnings, it was the Covid 19-pandemic that led not only to an immense increase in use of video-conferencing platforms but also of related research. The safety regulations during the pandemic transformed video-conferencing from a specialised format used primarily in international cooperation with geographically dispersed teams to an omnipresent technology used by the general public in diverse contexts, such as education, health care, work and private life (cf. Suduc et al., 2023: 1-2, Due & Licoppe, 2020: 1-2). Not only did the research areas and interests in video-mediated interaction expand, but research itself was impacted by these developments. Researchers depending on natural data and working with research subjects were confronted with challenges regarding research design, data collection, theoretical assumptions, and methodological approaches.

This paper presents one of these cases, a study on multilingualism in higher education that initially didn't foresee a video-mediated setting. The research project was already on its way, with a first data set that had been collected in a seminar room at the Pedagogical University Karlsruhe (PHK). In the process of working with the data, the decision was taken to include a second data set. After gaining permission for data collection at the European University Institute (EUI) in Florence, preparations for the workshop were underway when the pandemic broke out. As hope for a quick end faded, it was discussed to collect the second data set online in 2021.

A critical revision of research design, theoretical, and methodological approach was required, as there are "obvious differences" (Due & Licoppe, 2020: 5) between task oriented group work in a seminar room or on a video-conferencing platform. Would these differences allow for meaningful results regarding the research focus – an interaction based description of multilingualism in higher education? How would the situatedness of the data sets affect multilingual practices of its participants? Would it be possible to keep the focus on multilingual practices in higher education, neither turning the project into a contrastive study of *mediated* vs. *unmediated* interaction nor ignoring the specific situatedness of the data? The foreseen multimodal

interaction analysis in the tradition of ethnomethodology and conversation analysis (EMCA) follows a phenomenological approach that starts with "unmotivated examination" (Sacks, 1984: 27) and takes the uniqueness of each interactional encounter seriously. In this sense, every single task oriented group work of the corpus was examined and described in its own right, the situatedness being part of this description but not at its centre.

This paper now shifts the focus towards observations emerging from the study of these two differently situated data sets. Tracing the process from data collection to analysis, the paper discusses basic principles and concepts of an EMCA approach to multimodal interaction in relation to the specific settings and observed phenomena (cf. Mlynář et al., 2018), e. g., interaction space, inter- and intra-personal embodied actions, as well as technical issues. Through selected sequences, this paper highlights phenomena like verbal overlaps and embodied feedback signals. The insight into questions, decisions, and findings resulting from being confronted with two differently situated data sets aims to contribute to the exchange within research on video-conferencing platforms in educational contexts and to relate the observations to didactical considerations, which are discussed in the "Outlook" of this paper.

2. Theoretical foundations and methodological approach

The critical revision of research design and methodology to address the "obvious differences" (Due & Licoppe, 2020: 5) between task oriented group work conducted in person or on a video-conferencing platform drew attention to space and mediatedness.

The spatial dimension in studies of language has been investigated from diverse perspectives such as the *space-as-container* view that understands space as a 'given' entity, or the cognitive perspective that conceptualises space as a cognitive representation. The constructivist perspective shifts the focus to the social actions by which people give sense to a place, differentiating *space* from *place*. In research on social interaction, space is understood as a setting in which people position themselves in alignment to each other and the context, which includes the institutional (or private) frame and objective of the encounter, language and socio-cultural norms, as well as the peoples' bodies and the material and spatial environment. The interaction space is conceptualised as a reflexive relationship between the action in which participants are engaged, the spatial arrangement and alteration of their bodies, and the material and spatial features of the situation (cf. De Stefani et al., 2012: 2-4, Mondada, 2016: 336-338). While the *space-as-container* perspective and the *bucket theory* of context see human action as determined and structured by the context, an EMCA approach views the interaction space as continuously being enacted by the participants *in situ* and each action is both context-shaped and context-renewing. The EMCA approach to "Video-Mediated Interaction is not based on the dualistic separation of mediation and non-mediation, but on the dogma that in each case it must be shown how the 'mediating technologies' accountably shape the available or observed interaction practices" (Due & Licoppe, 2020: 6). To do this, basic principles and concepts, such as the *emic* perspective on *natural* situations, the context and *indexicality* of social interaction, the concepts of *accountability* (Garfinkel, 1967: vii-viii, 3-11) and *order at all points* (Sacks, 1984: 22), the understanding of spatiality, sequentiality and simultaneity of multimodal interaction (Mondada, 2016: 337-341), are applicable to both, in person and video-mediated settings of interaction.

Related to the perspective on space and mediatedness is the distinction between face-to-face and video-mediated interaction. Approaches that exclude video-mediated from face-to-face interaction understand the notion as referring to a physical co-presence of participants in a shared spatial environment, while an inclusive understanding reads it as metaphor for the *emic* perspective (cf. Deppermann & Schmitt, 2007: 16) and the embodied turn in the social sciences. Focussing on how multimodal resources are used holistically and situatedly in building and organising human action, the *embodied* or *visual turn* seeks to overcome the logocentric view of social interaction by including embodied actions, their material and spatial context (cf. Mondada, 2016: 337). This does not necessarily require the participants' physical co-presence in a shared spatial environment but depends on reciprocal perceptibility. Video-conferencing platforms such as Skype, Teams, Zoom and many others enable participants to have mutual access to sound and image in real time, which allows them to synchronously build and organise a shared interaction space using language, embodied actions, objects, and their spatial environment. The specificity of the spatial dimension on video conferencing platforms is that it depends on technological transmission and is enacted within *fractured*

ecologies: "fractured from the environment in which it is produced and from the environment in which it is received" (Luff et al., 2003: 55).

Based on this perspective on interaction space, multimodality and mediatedness, this paper discusses questions and decisions regarding research design and technology, and presents observed phenomena that suggest participants adapt their practices to the shared (data set 1) or fractured (data set 2) ecology of the task oriented group work. The data was transcribed according to the "Conventions for transcribing multimodality" by Mondada (2019) and GAT 2 (Selting et al., 2009) for the verbal tier. In the EMCA tradition, the aim is not to fit data into theory but to gain data based insights into social interaction.

3. Research process and observed phenomena

As explained above, the underlying corpus consists of two data sets (15 recordings, 148.32 minutes in total) that were collected within the institutional context of higher education, within a seminar (data set 1) and a workshop (data set 2) under the title "Multilingualism: individual – societal – institutional". Both data sets show situations of task-oriented group work among peers in small groups of two to four participants. Most of the task oriented group work consisted of question-based text discussion for which the participants were given text materials on different aspects of multilingualism. Further activities included the presentation of individual language portraits and a search on google maps for linguistic landscapes.

The seminar (data set 1) took place during the winter semester 2018/2019 under pre-pandemic conditions in a classroom of the PHK. With the Covid-19 outbreak in 2020 the workshop (data set 2) was at first postponed and then held on the video conferencing platform Zoom in 2021. By the time of the workshop, the participants in the second data set already had a year of experience with that platform, its tools and functions. They also had a year's time to set up and equip their computers for video-conferencing, which is an additional factor to be considered when using and/or researching video-mediated interaction.

The following description of the research process and observed phenomena begins with reflections and decisions concerning data collection, moves on to considerations regarding the specific ecologies, participants' positioning and resources, and concludes with observations on embodied actions and verbal overlaps.

3.1. Data collection

Data collection for research on social interaction is always confronted with questions regarding research needs and interests, the research subjects' rights and interests, and technological possibilities. From the researchers' perspective it might seem advantageous to collect the broadest possible insight into the interaction space, using cameras from different angles or movement-/sound-responsive camera setups. Such a broad use of technology increases the richness of the data but can conflict with the aspiration to collect data from natural situations of interaction (Gülich & Mondada, 2008: 27-30). It can also conflict with the research subjects' interests and rights such as privacy rights or, in an educational setting, the right to learn in an undisturbed environment. The relevance of this conflict becomes obvious through the fact that not all participants gave their informed consent. If the data collection had not been perceived by some as an intrusion into their learning environment and privacy, all students would have given their informed consent.

Against this background, the data collection was conducted under the premise to keep the recording as unintrusive as possible and the recorded situation as natural as possible. For the research question of the underlying project it was central to capture the participants' use of multimodal resources in alignment to their interlocutors and so the recordings could be focused on the participants and their immediate surroundings.

To collect the first data set, where the interactions took place in person in a seminar room, the camera was placed in front of the groups with consenting students, and additionally an audio-recording device was placed on the table.

All groups were seated at tables, positioned in a side-by-side constellation, leaving only the upper part of the body visible to the interlocutors (example 1).

Data set 1: side-by-side constellation

Example 1. Data collection in classroom

With the collection of the second data set in 2021 new technological questions arose. Since then, technology has advanced and new options have been added to the used platform. Nevertheless, the underlying questions and reflections remain relevant, even more so as recording options multiply with technological developments.

While research on the participants' screen activities (Balaman & Pekarek Doehler, 2021: 54-59) or interactions with the technical interface requires screen recording and/or setting up additional cameras, the conducted research focussed on the interlocutors' use of multimodal resources in alignment to each other. Furthermore the data collection was to be kept as unintrusive as possible and so it was decided to only use the recording function provided by the video-conferencing platform. After this decision, the next, equally important step is to test recording function and evaluate its results prior to the actual data collection, as these may vary depending on view settings and specificities of the recorded situation.

In the underlying research project, one of the specificities regards the interaction type of task-oriented group work among peers, for which the video-conferencing platform Zoom offers so-called *Breakout Rooms*. It had to be tested if the recording function is available within the *Breakout Rooms* and, if so, how it can be operated. On Zoom, the recording of *Breakout Rooms* cannot be initiated from the main session; it has to be started within the *Breakout Room* itself. If the researcher is not supposed to be part of the group work, alternative recording options need to be explored. In this project it was decided to ask participants to handle the recording within their *Breakout Room* so that they were not just research subjects but became actively engaged in data collection.

The view settings impact not only how the individual participants perceive each other in their shared interaction space but also what is visible for the researcher, as the recording depends on the view settings of the recording person. At that time, the used video-conferencing platform allowed to select between a *Gallery* and a *Speaker View*. While the *Speaker View* shows the speaking participant in the centre of the video window with the other participants smaller at the top, the *Gallery View* displays all participants simultaneously and equally sized on the screen. With the focus on the participants' simultaneous and sequential use of multimodal resources in alignment to each other and the context, the *Gallery View* was chosen for the recording.

To also ensure at least a minimum level of common perception of the interaction space despite the participants' individual organisation of their screens, not only the recording person but the whole group was asked to choose the *Gallery View* as screen setting in the *Breakout Rooms*.

Except for one, all groups followed this instruction. This exception allows to demonstrate the impact of view settings on recording results (example 2).
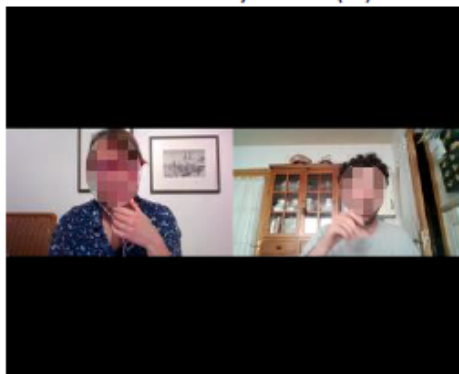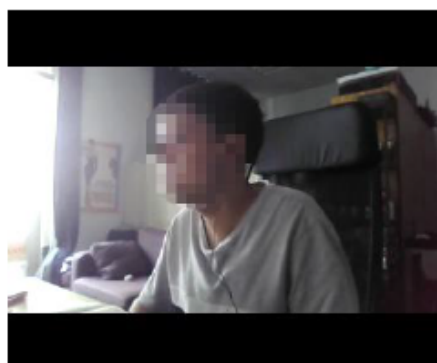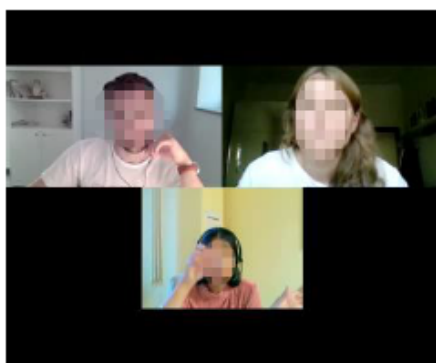
Example 2. Data collection on video-conferencing platform

Example 2 (A) shows the recording results using the *Gallery View* with all involved participants equally sized in the video window. Example 2 (B) shows the recording results using the *Speaker View* with only the current speaker visible even though in the live situation the other participants see themselves in smaller frames above the speaker. This difference between recording and live vision denies researchers access to reciprocal non-verbal actions and thus was of limited use for the original research project. However, with the focus on video-mediated interaction, the different view settings generate interesting questions of their own, in particular regarding self-perception.

In the *Gallery View* the individual participants do not only see the others and their actions but also themselves and their own actions, which isn't the case with in-person encounters. The impact of this self-perception on the participants' inter- and intra-personal actions cannot be followed up on the underlying data basis but requires a specific research design and data collection. Nonetheless, the one group that created an exception sparked the thought that the *Speaker View* might not just have been used to focus on the current speaker but also as a means to reduce self-perception.

3.2. Ecologies, positioning and resources

The task oriented group work in data set 1 takes place in a classroom that provides a common surrounding ecology, in which the participants are seated at tables in a side-by-side constellation, leaving only the upper part of the body visible to the interlocutors. In data set 2 the task oriented group work takes place on a video-conferencing platform and the participants are positioned in individual ecologies, each in front of their computer, again leaving only the upper part of the body visible to the interlocutors. The fractured visibility of the participants' bodies in both data sets is part of the ongoing activity: table talks within teaching events in the context of higher education.

However, mutual perception extends beyond visibility, it is multisensorial and the interaction space is not just a physical but a social space.

In the shared surrounding ecology of the classroom in data set 1 this includes the participants' body as well as shared access to the ecology's resources, such as chairs, tables, printed texts, pens, personal objects like bags, jackets or food and drinks. The participants make use of the shared ecology and its multimodal resources, for example, by using deictic gestures or gazes to indicate text reference, handling objects together, bodily movements towards the interlocutors to express co-orientation or interpersonal alignment etc. (example 3).
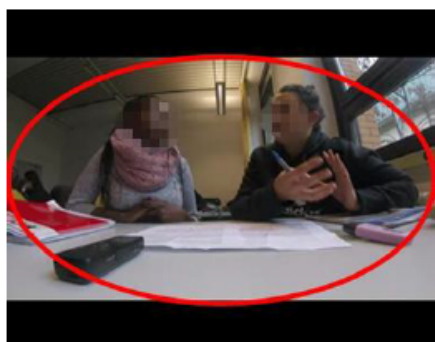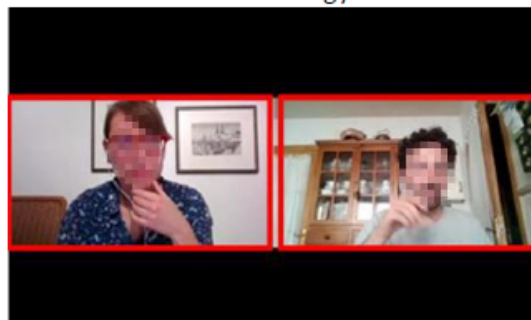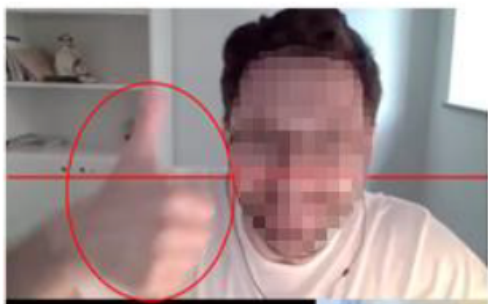

**Example 3.** Shared and fractured ecologies

In data set 2 all participants are seated with their computers placed on a table in front of them, thus creating a stable environment for the video-mediated interaction (in contrast to participants holding a mobile device in their hands). Except for one, all participants use the integrated camera of their computer and are seated in a frontal orientation to the shared interaction space. The visibility of the individuals is limited by the camera frame that includes the participants' face and upper part of the body as well as the background of their immediate environment. Some multimodal resources of the individual ecologies are visible for the interlocutors but are not equally accessible due to the fracturedness of their interaction space (example 3).

Participants adapt to this by bringing certain objects (see example 6) and embodied actions in front of the camera (example 4: A), while other actions take place in the shared perception space but are not specifically brought into focus (example 4: B).
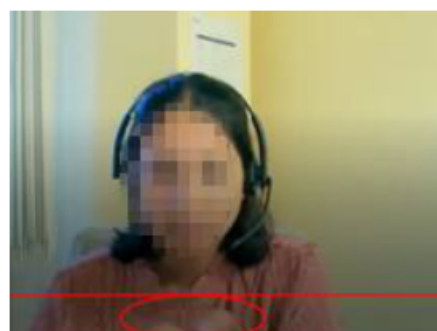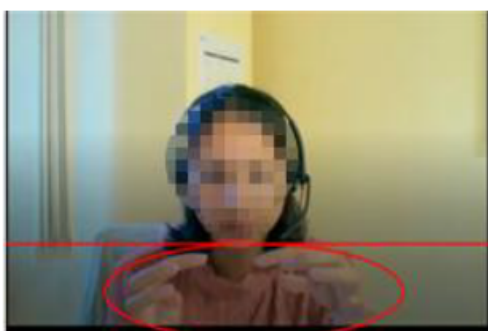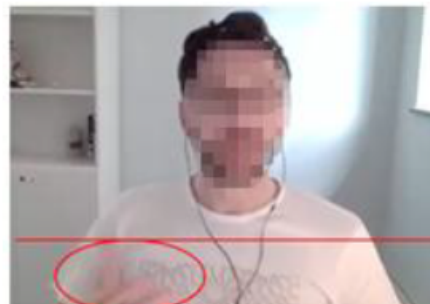
Example 4. Embodied actions

The importance of mutual visibility for multimodal interaction also becomes evident in the context of technical issues. While no technical problems like malfunctioning cameras or microphones occurred, and no questions were raised regarding the use of the platform's tools and functions, one group briefly had to deal with poor internet connection. To ease the bandwidth, the concerned participant temporarily turned off his camera, which causes his interlocutor to seek reassurance (example 5):

Example 5. Verbal compensation of visibility

The lack of visibility is immediately addressed verbally to maintain the interaction.

Apart from two brief sequences in this group, all participants kept their cameras on and the recordings show how they enact their shared interaction space through the use of multimodal resources in fractured ecologies. Some of the verbal and non-verbal actions suggest participants' adaptation to video-mediated interaction.

3.3. Embodied actions

As shown in example 3 (A), the participants in data set 1 make use of the shared ecology and its multimodal resources by gesturing or gazing towards the text, moving their bodies towards each other, etc. The embodied actions contribute to task orientation as well as to relation building and interpersonal alignment.

As participants on video-conferencing platforms are situated in individual ecologies, they need to adapt to the fractured interaction space. One example for doing so was shown in image 4 (A), where participants bring embodied actions into the camera focus. Other examples allow insight into relation building and interpersonal alignment. Instead of body movements towards each other or even physical contact, as seen in data set 1, participants in data set 2 use visualisations (example 6).
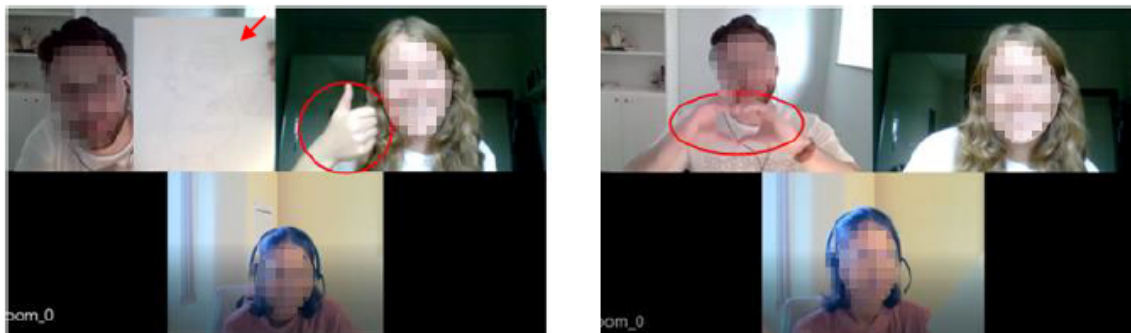
Example 6. Visualisations

Symbols from social media such as *thumbs up* or *hearts* are visualized through hand gestures in front of the camera, contributing to interpersonal alignment and relation building. Interestingly, despite Zoom offering a *Reaction* feature with these symbols, participants in data set 2 make no use of this feature but prefer manual visualisations.

Furthermore it was noted, that non-verbal feedback signals, such as nodding, facial expressions, silent laughter, gesturing and visualisations, seem to be realised very expressively and often prolonged in data set 2. This impression of intensified embodied feedback cannot be followed-up in the underlying data but hints to a compensation for the lack of bodily proximity within video-mediated interaction.

3.4. Verbal overlaps

In both data sets the verbal mode is vastly realised according to the interaction order that "[o]verwhelmingly one part talks at a time" and that "[o]ccurrences of more than one speaker at a time are common, but brief" (Sacks et al., 1974: 706). In the task oriented group work among peers, verbal overlaps are common and accepted, but there are notable differences between the two data sets.

In data set 1, verbal overlaps occur frequently, as the following sequence demonstrates (example 7):

```
Data set 1: Verbal overlaps, Group 5 (10.44.3 min. total)
Excerpt: 04.06.9-04.22.2 min.

146 B:  weil eigentlich(0.527)is ja [wie soll ich sagen]
        because in reality       is yes  how  shall i   say
147 A:                                  [das wäre         ]gut zu wissen
                                         that would be       good to know
        wie viele oder wie sind die schulen Δaufgeteilt;Δ
        how  many    or  how  are  the schools   divided up
    b:                                            Δnods      Δ
148 A:  [ist eng]lisch nur als [fach    ]
         is  english  just as   subject
149 B:  [Δhm_hmΔ]              [Δhm_hmΔ]
    b:  Δnods Δ                 Δnods Δ
150 A:  oder(0.987)und mand[arin dann dass alle schulen an          ]
        or          and  mandarin    then  that all  schools  at
151 B:                     [ich glaub die lernen die sprechen nur in] (0.64)
                            I  think they learn   they speak    just in
152     in den schulen sprechen die glaub ich englisch;
        in  the  schools speak     they believe I   english
```

Example 7. Verbal overlaps in classroom

In one sequence of just seven lines there are four overlaps occurring in turn taking moves (example 7 lines 146-147 and 150-151) and also in feedback signals, which are given verbally and non-verbally (example 7 lines 148-149).

In dataset 2, overlaps also occur for both, to give feedback and to take the turn, but they are significantly rarer. To show both phenomena two different sequences are needed (example 8).

```
Data set 2: Verbal overlap for turn-taking, Group 2.1 (11.02.2 min. total)
Excerpt: 06.29.3-06.40.7 min.

117 A:  slavic languages* und äh sie sind ein bisschen *similar* [aber ja]
                          and eh they are  a   bit        similar  but  yes
    b:     >>nods       *                              *nods    *
118 B:                                                           [ja also]
                                                                  yes so
        vielleicht wie äh spanisch und italienisch Δsie sind sehr gleich
        maybe         like eh spanish  and italian    they are very same
    a:                                             Δnods-->>

Data set 2: Verbal overlap in feedback signal, Group 2.2 (10.43.0 min. total)
Excerpt: 03.38.2-03.56.9 min.
054 B:  like one of the ideas of the text is also like this debate of
        whether languages exss exist by themself;
055     ^or [(0.507)]or it is^ also society which ^breaks them and makes differences;
    a:  ^nods       ^                             ^nods-->
056 A:      [hm:    ]
057 B:  (0.509) [(0.184)] [and he was u]sing this example of the kid; (.)
058 A:          [hm        [hm          ]
059 B:  how kids they [(0.447)]just (.) seek this- (.)
060 A:               [hmhm    ]
061 B:  or ^everything is kind of mixed;
    a:  -->^
```

Example 8. Verbal overlaps on video-conferencing platform

While lines 117-118 (example 8) of the above excerpt show one of the few overlaps for turn-taking in data set 2, the lines 54-61 (example 8) give insight into the use of verbal feedback signals. The first occurs in line

56, where it is placed within a pause of 0.507 seconds, and also in line 60 the verbal feedback signal is placed into a pause. Only in the lines 57 and 58 there is an overlap of participant B's talk and participant A's verbal feedback. The verbal feedback signal in line 58 seems to set in a little late, after B had already paused for 0.509 seconds. A's verbal feedback signal starts in the last 0.184 seconds of B's pause and then overlaps with the first words of Bs continuation, so that it could be interpreted as a misplacement, demonstrating how precise anticipation and interpersonal alignment usually are.

In the ten lines of both shown sequences from data set 2 (example 8) only two verbal overlaps occur. For feedback signals the nonverbal mode seems to be preferred and verbal feedback is often placed within pauses.

The observation, that overlaps are reduced in data set 2 occurred whilst transcribing and was checked by counting the overlaps in both data sets. To include the length of the conversations (8-14 minutes) and the quantity of verbal expressions, the total of characters within each conversation and the overlapping characters were counted and calculated into percentages with the result of 29 % of overlaps in data set 1 and only 7 % of overlaps in data set 2.

This difference in the quantity of overlaps hints to the participants' awareness of and adaptation to the video-mediatedness of their interaction space, as verbal overlaps can lead to hick-ups, delays, or gaps in the audio-transmission. After a year of experience in video-conferencing this can be assumed to be known by the participants, so that the predominantly non-verbal realisation of feedback signals in data set 2 suggests that participants adapt their feedback practices to video-mediated interaction.

4. Outlook

The given insight into questions, decisions, and findings resulting from being confronted with two differently situated data sets aims to contribute to the exchange within research on videoconferencing platforms in educational contexts and to relate the observations to didactical implications.

The comparison of the recording technologies for in classroom and video-mediated interaction highlights the complexity of underlying questions for both ecologies and the specificities to be considered. Video-conferencing platforms provide synchronous access to audio and video depending on technical transmission. Most platforms also provide an integrated recording function allowing users to extract audio and video files in standard formats and of good quality. For researchers the integrated recording function of video-conferencing platforms might seem a practical solution for data collection. However, its interconnectedness with other functions, i. e. view settings, is not to be underestimated to attain data that is useful for the specific research question. Testing the recording function and different settings is indispensable and can further additional research interests, e. g. comparative studies with groups using different view settings. Results from such studies can feed back into the planning of video-mediated teaching/learning, for example teachers on video-conferencing platforms might consider recommending a specific view setting to enhance the participants' common perception of the shared interaction space.

In the underlying research project, the *Gallery View* proofed useful to analyse the participants' reciprocal use of multimodal resources but raised questions regarding self-perception. Studies describing a video-conferencing-fatigue in the second year of the pandemic mention constant self-perception as one contributing factor and some platforms (e. g. Zoom) introduced a view setting with which participants can choose to keep their camera on, but hide the video from their own screen, so that they are no longer visible to themselves (cf. Suduc et al., 2023: 5). Prior to this technological advancement, the *Speaker View* might have been used to reduce self-perception. The impact of self-perception on inter- and intra-personal embodied actions of participants is of ongoing interest, deserving attention from researchers and practitioners alike. Within video-mediated teaching/learning events, the information of students regarding options to avoid self-perception can help them to focus on their interlocutors and reduce fatigue. Whether participants actually prefer to exclude their own video from their video window or whether, for example, younger students growing up with diverse forms of self-presentation and self-perception online develop tolerances and preferences of their own requires further research. In any case, potential impacts of self-perception on interpersonal attention, social learning, and fatigue seem worth exploring, for teaching practitioners and researchers alike.

Furthermore, the data collection of video-mediated interaction can affect the classical roles of researchers and research subjects. The use of *Breakout Rooms* for the task-oriented group work in data set 2 led to

participants becoming data collectors, which deserves discussion within the context of 'natural' situations, observers' paradox and questions of participating and non-participating observation (Schwab & Schramm, 2016: 305). Beyond theoretical debate, the active involvement of students in research opens didactical potentials, in particular within higher education. If, for example, the information on data collection is not just treated as a necessity to attain the participants' informed consent but as an opportunity to promote academic curiosity and to educate students on good research practice and ethical questions, the information on research project, data collection, and data use can become a teaching unit of its own or a sub-unit of an introductory session with the objective to interest students in and prepare them for future research of their own.

Even though multimodal interaction analysis is confronted with specificities resulting from the situatedness in a classroom or on a video-conferencing platform right at the beginning of the research process, namely during data collection, the EMCA approach doesn't start from a dualistic separation of mediation and non-mediation but sees all interaction as "mediated in some way, e.g. by language, bodies, objects, culture, history, technologies and so on (Gallagher & Zahavi, 2020). There is no unmediated interaction with which the mediated can be contrasted", thus it must be shown how context, mediatedness and spatial situatedness "accountably shape the available or observed interaction practices" (Due & Licoppe, 2020: 6) For example, the fractured visibility of the participants' bodies in the examined data is not simply a result of the opposition *mediatedness* vs. *unmediatness* but part of the ongoing activity, table talks within the context of higher education. The relation between activity and space extends beyond questions of mediatedness, it comprises institutional responsibilities regarding their students'.

In traditional classroom teaching an adequate learning environment (room, tables, chairs, boards, technical devices etc.) is provided by the institution, but with the outbreak of the pandemic this physical space was closed and a lot of students had to follow video-mediated lessons with inadequate technical devices (mobile phones) and inadequate learning environments (no individual room, no table or chair to create a stable learning environment). Now, within the process of post-pandemic normalisation of the use of video-conferencing platforms in education, these issues need to be addressed to protect educational justice and participation, especially in schools. Identifying institutional and individual responsibilities is part of this process. While the described data can exemplify the participants' agency and highly cooperative use of multimodal resources in creating their shared video-mediated interaction space, the investigated groups can also be described as privileged. In the underlying data, collected within the context of higher education, the research participants are already advanced on their educational and/or academic path (Masters' and PhD students), and all participants in data set 2 contribute to a stable learning environment within their individual ecology by using computers placed on a table and sitting on chairs instead of choosing a walking, standing or even lying position. Furthermore, all participants ensured an undisturbed learning environment without interferences from outside their group, which cannot be considered a given but might proof difficult or impossible depending on the individual learners' situation.

In addition to appropriate equipment and situatedness, the participants in data set 2 seem to adapt verbal and non-verbal practices to the fractured ecology of their shared interaction space, for example by reducing verbal overlaps and increasing embodied feedback.

The described impression of intensified and prolonged embodied feedback signals resulting from analysing natural data partly corresponds to findings from early experimental data. Experiments in which pairs of subjects performed a collaborative task in physical co-presence and remotely communicating either via video and audio links or audio links only show that both co-present and video-mediated speakers use visual cues to check for mutual understanding. In audio only communication such checks need to be conducted verbally. Despite the use of visual cues in video-mediated interaction, there is a lack of bodily proximity and the gaze is far more overall than in co-presence interaction, suggesting that "when speakers are not physically co-present they are less confident in general that they have mutual understanding, even though they can see their interlocutors, and therefore over-compensate by increasing the level of both verbal and nonverbal information" (O'Malley et al., 1996: 177). While the early experiments identify an increase in both, verbal and nonverbal information, the examined data of this paper confirms only the latter, which might be connected to technological advancement since the 1990ies and to the participants' experience with the video-conferencing platform. The perceived increase of embodied feedback in data set 2 seems to correlate with the reduction of verbal overlaps, suggesting that lack of bodily proximity and technological restraints regarding audio-transmission seem to be compensated by intensifying embodied actions. The

reduction of verbal overlaps in data set 2 (7%) compared to data set 1 (29%) and the increase and intensification of embodied feedback can be described as practices developed by the participants to contribute to interpersonal alignment, relation building and interactional achievement on video-conferencing platforms.

Whereas in the underlying data the participants themselves show a high degree of awareness and adaptation to their video-mediated learning, other educational settings might require explicit explanations to create a *common ground* and a beneficial learning environment. To develop data based guidelines for teachers in diverse educational settings, research and practice need to go hand in hand. Considering the velocity of new developments within video-conferencing, their possibilities and restraints, more insight into the process of planning (research design and/or teaching events) and more exchange regarding different experiences among researchers and teachers would be extremely helpful, not just for beginners.

References

Balaman, U., Pekarek Doehler, S. (2021). Navigating the complex social ecology of screen-based activity in video-mediated interaction. *Pragmatics*, 32(1), 54–79.

Due, B. L., Licoppe, C. (2020). Video-Mediated Interaction (VMI): Introduction to a Special issue on the multimodal accomplishment of VMI Institutional Activities. *Social Interaction. Video-Based Studies in Human Sociality*, 3(3), 1-20.

De Stefani, E., Gazin, A.-D., Ticca, A. C. (2012). Space in social interaction. An introduction. *Bulletin VALS-ASLA 96 - Bulletin suisse de linguistique appliquée 96*, 1-14.

Deppermann, A., Schmitt, R. (2007). Koordination. Zur Begründung eines neuen Forschungsgegenstandes. In R. Schmitt (Ed.), *Koordination. Analysen zur multimodalen Interaktion* (pp. 15-54). Narr: Tübingen.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. New Jersey: Polity Press.

Gülich, E., Mondada, L. (2008). *Konversationsanalyse. Eine Einführung am Beispiel des Französischen.* Tübingen: Niemeyer.

Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., Oyama, S. (2003). Fractured Ecologies: Creating Environments for Collaboration. *Human-Computer Interaction*, 18(1), 51-84.

Mlynář, J., González-Martínez, E., Lalanne, D. (2018). Situated organization of video-mediated interaction: A Review of Ethnomethodological and Conversation Analytic Studies. *Interacting with Computers*, 30(2), 73-84.

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of sociolinguistics*, 20(3), 336-366.

Mondada, L. 2019. *Conventions for transcribing multimodality.* https://www.lorenzamondada.net/multimodal-transcription. Accessed 5 January 2025.

O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., Bruce, V. (1996). Comparison of face-to-face and video-mediated interaction. *Interacting with Computers* 18(2), 177-192.

Sacks, H. 1984. Notes on Methology. In J. M. Atkinson, J. Heritage (Eds.), *Structures of Social Action* (pp. 21-27). Boston, Cambridge: University Press.

Sacks, H., Schegloff, E. A., Jefferson, G. 1974. *A Simplest Systematics for the Organization of Turn-Taking for Conversation*. In Language 50(4.1), pp. 696-735.

Schwab, G., Schramm, K. 2016: Diskursanalytische Auswertungsmethoden. In D. Caspari, F. Kippel, M. K. Legutke, K. Schramm (Eds.), *Forschungsmethoden in der Fremdsprachendidaktik: Ein Handbuch* (pp. 302-319). Tübingen, Narr, Francke, Attempto.

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., 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., Stukenbrock, A., Uhlmann, S. (2009).

Gesprächsanalytisches Transkriptionssystem 2 (GAT 2*). Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 10 – 2009, 353-402. http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf. Accessed 5 January 2025.

Suduc, A.M., Bizoi, M., Filip, F.G. 2023. Status, challenges and trends in videoconferencing platforms. *International Journal Of Computers Communications & Control*, 18(3), 1-10.