Situational involvement and footing shifts in mobile live video streams

Mark Dang-Anh

English abstract

Mobile live video streaming with smartphones is an everyday media practice in which the participants are in a specific multimodal constellation and streamers and viewers have access to various semiotic resources for interactionally establishing alignment. Based on the multimodal sequence analysis of a concise episode of a journalist's livestream coverage of a political event on the streaming platform *Periscope*, I will address the question of how participation and involvement in live video streams are achieved and organised by the participants. I will show that hosts in the media practice of live video streaming act in an interaction-dominant manner and involve the viewers in the situation through asymmetrical participation coordination via footing shifts.

Keywords: live video stream – media practices – mediated interaction – footing shifts – situational involvement – Periscope – media linguistics.

German abstract

Mobiles Livevideostreaming mit Smartphones ist eine alltägliche Medienpraktik, bei der sich die Beteiligten in einer spezifischen multimodalen Konstellation zueinander befinden und in der Streamer*innen und Zuschauer*innen unterschiedliche semiotische Ressourcen zur interaktionalen Ausrichtung zur Verfügung stehen. Anhand der multimodalen Sequenzanalyse einer prägnanten Episode im Rahmen der Berichterstattung eines Journalisten von einem politischen Ereignis auf der Streamingplattform *Periscope* wird die Frage bearbeitet, wie Beteiligung und Involvement in Livevideostreams durch die Teilnehmenden hergestellt und organisiert werden. Es wird gezeigt, dass die Hosts bei der Medienpraktik des Livevideostreamings interaktionsdominierend agieren und die Zuschauer*innen durch asymmetrische Partizipationskoordination per Footing Shifts situativ in das Geschehen involvieren.

Keywords: Livevideostream – Medienpraktiken – Medieninteraktion – Footing Shifts – situatives Involvement – Periscope – Medienlinguistik.

1. Introduction[1]

Mobile live video streaming with smartphones is an everyday media practice. In recent years, livestreaming of events has also become increasingly common in journalism (cf. Lünenborg 2017). Events are documented synchronously via livestreaming from mobile devices such as smartphones and made available for 'real-time' reception. A characteristic feature of live video streams[2] is their specific medial and situational arrangement: while the participants in live video streams act quasi-synchronously in terms of time, the communication situation is spatially 'stretched' (cf. Ehlich 1984), since the streamers and the viewers are usually physically separated from each other at the time of streaming (cf. Reeves et al. 2015). In a live video stream, therefore, streamers and viewers are situated in an asymmetrical relationship: while the streamers can interact multimodally in the semiotic modes of the audiovisual video stream, viewers are only partially actively involved in the situational and communicative events displayed on their screens via chat (cf. Licoppe/Morel 2018). The aim of the multimodal sequence analysis of the communicative media practice of live video streaming conducted here is to contribute to the discussion about the interrelation between mediality and sociality, which is fundamental for praxeologically oriented media linguistics (cf. Luginbühl/Schneider 2020).

---

[1] This article is based on a presentation at the workshop "Parainteraktion in den Medien. Multimodale Perspektiven in TV, Radio und Hypermedien" ('Parainteraction in the Media. Multimodal Perspectives in TV, Radio, and Hypermedia'), which took place in October 2019 at the University of Basel. I would like to thank Martin Luginbühl and Dorothee Meer for the organisation of and invitation to the workshop, all workshop participants and Isabell Neise for the discussion of the data, the German Research Foundation (DFG) and the Leibniz Association for funding this research, and the anonymous reviewers for their helpful comments and suggestions. A slightly modified German version of this article will appear in Luginbühl/Meer (forthcoming).

[2] The term 'livestream' can also refer to media formats of other modality constellations, e.g. audio transmissions or text transmissions (ticker tape). In the following, however, 'live video stream' and 'livestream' are used synonymously.

To gain an impression of what is happening in the live video stream being discussed and how the described arrangement appears on the screen, we will briefly consider a scene from the analysed data material. In this excerpt, a journalist (MK) is streaming live from a demonstration event via the streaming platform *Periscope* (cf. transcript 1).[3] Fig. 5 shows the screenshot of the recording of the live stream in the subsequent playback via the browser.

Transcript 1: "Aligning"
(from: "Kaul aus #Köthen, 18.06 Uhr.", 00:28:35–00:28:45)

```
28:35 01 MK   +*ich latsch jetzt mal %weiter #hier die straße  %
              lang Oder?+
              i keep walking down that road here now right
        move +turns to the walking direction and
             back------+
        gaze *looks underneath the
             camera-->>
        hand                         %points upwards/behind him%
```


#fig.1

```
28:38 02      +%##(2.8)
        move +walks-->>
        hand %scrolls-->
```


#fig.2   #fig.3

```
28:41 03 MK   soo;%
              well
        hand -->%
28:43 04 MK   nochn stück weiter dann rechts;
              a little further on then right
28:44 05 MK   %SO_#machen wirs-%
              this is how we do it
        hand %raises index     %
             finger
```

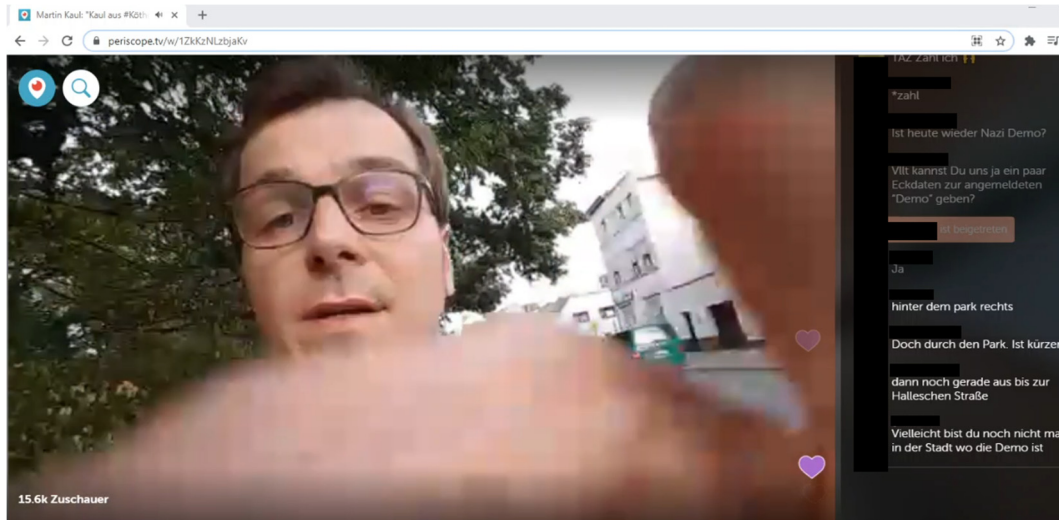
#fig.4

---

3   For details of the situational context of the data, see section 4.3.

Fig. 5.: Screenshot of the live video stream playback "Kaul aus #Köthen, 18.06 Uhr." from the *Periscope* website (https://www.periscope.tv/w/1ZkKzNLzbjaKv [23.03.2021])

In this episode of his live video stream, the streamer engages with the audience while he changes location, which is a common and recurrent activity in mobile live video broadcasts of political events or street protests. While the streamer is on foot at the reporting location with his smartphone camera attached to a mobile phone stick, he turns to his viewers and asks them if he is on the right track. His physical activity is accompanied by a movement-describing utterance (01). By means of the question tag *oder*, he establishes an orientation problem, which he apparently seeks to solve together with the spectators. We arrive at this assumption through his rough rendering of a chat post (04), which provides the answer to his question (01). What can be observed here in a nutshell is how the streamer aligns himself with the viewers and thus interactively involves them in a previously rather monologic communication situation.

Based on the assumption that the alignment of the participants – the footing, as Goffman (1979) terms it – is produced both medially and interactively, the question of how participation and involvement in live video streams are achieved and organised by the participants is addressed from a media linguistics and interactional linguistics point of view. First, the theoretical foundations for media interaction analysis (section 2.1.) are laid and an explanation of participation and involvement (section 2.2.) is presented. Subsequently, an overview of the current state of research and the delineation of two characteristics of live video streams – their ruptured temporality (section 3.1.) and the asymmetrical establishment of relevance (section 3.2.) – are given. The methods section opens with methodological notes on the analysis of live video streams (section 4.1.) and provides brief introductions to the transcriptions (section 4.2.) and the context of the data being examined (section 4.3.). The analysis section is divided into four parts: a discussion of the establishment of hosting in the opening part of the livestream (section 5.1.), and the analysis of another episode in three subsections, focusing on the practices of situating, recruiting, and acknowledging (sections 5.2. – 5.4.). The article closes with some concluding remarks (section 6.).

2. Theoretical foundations

2.1. Live video streaming as communicative media practice

Live video streaming is a thoroughly social endeavour: "[t]he combination of real-time video and chat promotes high levels of engagement and constitutes a key characteristic of *social* live-streaming" (Rein/Venturini 2018:3361; emphasis added). As for many other media practices, the following applies to live video streaming: the sociality of a media practice is conditioned both by its mediality and its semioticity. An understanding of sociality is derived from social situations in which two participants find themselves in a position of mutual perception (*Wahrnehmungswahrnehmung*, 'perception of perception' according to Luhmann 1984:560; cf. Hirschauer 2014:112) and shared "response presence" (Goffman 1983:2). Hence, the concept of sociality applied here is closely tied to the concept of interaction and it is essential to investigate *how* interaction is processed in the medium 'live video stream'. Thus, the mediality of the medium is fundamental for the interactional constitution of the social. According to Luginbühl and Schneider (2020), the mediality of a medium refers to "the ways it shapes the choice of signs and how we use them" (59). The mediality of media practices is closely linked to the technical affordances (cf. Hutchby 2014; Meredith 2017) of a medial arrangement, where 'affordances' refers not only to the potential for mediated interaction that is available, but also to that which is actually exploited in media practices. In addition to mediality, semioticity, which in turn is medially conditioned, is also essential for the interactional constitution of sociality. Thus, for the analysis of live video streams, the medial 'procedures of sign processing'[4] (Schneider 2017) of different semiotic modes (cf. Bateman et al. 2017:112ff.) are particularly relevant. Media practices, in turn, are to be regarded as social practices of communication. 'Media practices' can be described as (cf. Dang-Anh et al. 2017:12)

> 'the manifold interplay between people and media as actors of media practices. In this sense, media only ever become media in the accomplishment of media practices – practices are grasped in relation to their medial processing. The notion of media practices encompasses [...] different scopes of practices, i.e. practices within practices and media within media.'[5]

Media, practices, and signs thus stand in a complex relationship to each other, which is why, with regard to the subject matter at hand, an analytical distinction must be made between

- the medium 'live video stream', which, as an operative, processing arrangement of intertwined apparatuses and infrastructures such as smartphone, screen, keyboard, Internet, etc., provides the conditions of possibility and potential for agency and action within which media practices take place,

---

[4] Orig.: "Verfahren der Zeichenprozessierung" (Schneider 2017). Text passages from non-English literature are paraphrased and marked with single quotation marks.

[5] Orig.: "das vielfältige Wechselspiel zwischen Menschen und Medien als jeweilige Akteure von Medienpraktiken beschreiben. So verstanden werden Medien immer nur zu Medien im Vollzug von Medienpraktiken – Praktiken werden erfasst in Bezug auf ihre mediale Verfertigung. Dabei umfasst der Begriff der Medienpraktiken […] unterschiedliche Reichweiten von Praktiken, d. h. Praktiken in Praktiken und Medien in Medien" (Dang-Anh et al. 2017:12).

- the communicative media practice 'live video streaming', in the course of which the interactional processes of the constitution of sociality take place, and

- the linguistic and non-linguistic signs of different semiotic modes, which are interactively and medially processed in the course of the multimodal media practice 'live video streaming'.

Interactions are the interpersonal momenta in which the relations of media, practices, and signs are synthetically manifested. The analysis of live video streaming applied here is thus a semiotically oriented process analysis of media interactions that unfold against a practical background. Sequentially conducted ethnomethodological conversation analysis (CA) therefore serves as a methodological paradigm, which in the media linguistic and interaction-focused research of digital communication can be regarded as an established method of investigating mediated *Vollzugswirklichkeiten* ('accomplished objectivities',[6] Bergmann 1981:12; cf. Arminen et al. 2016; Giles et al. 2015; the contributions in König/Oloff 2019; Meredith 2017), as long as it takes into account the medialities and semiotic modes as well as the practical backgrounds of the accomplishments being analysed.

2.2. Participation and involvement

Enabled and conditioned by the technical affordances of live video streaming platforms (cf. Sjöblom et al. 2019), different participation constellations can be realised in livestreams. Here, 'participation' is theorised from an interactional perspective and is, according to Goodwin and Goodwin (2004:222), closely related to 'involvement':

> The term participation refers to actions demonstrating forms of involvement performed by parties within evolving structures of talk. […] the term is not being used to refer to more general membership in social groups or ritual activities.

'Involvement' is to be understood here as an active, observable, intelligible, and jointly accomplished "achievement in conversational interaction" (Tannen 2007: 27). In contrast to the dyadic speaker-listener model, which has become the classic reference model for linguistics (cf. critically Goffman 1979; Levinson 1988), with Goffman (1979), multi-person constellations, which are common for digital media, can be analysed (cf. Bou-Franch et al. 2012; Deppermann 2015). Dynel characterises the communication situations emerging from multi-person constellations as: "multi-party interaction[s] which may involve many producers of turns, one (or more in the case of choral production) taking the floor one at a time, and many individuals at the reception end, who can be classified as various hearers/listeners to an interactional turn" (Dynel 2014:38). Multi-party interactions also take place in live video streams and the main question is how the dynamics of changing participation constellations are organised in terms of interaction and media.

Even before the popularisation of digital communication, Goffman established description categories for the complex participation constellations of face-to-face

---

[6] Bergmann's terms "*Vollzug, Vollzugswirklichkeit* etc." are themselves translations of Garfinkel's notions of "*accomplishment, accomplished objectivity, accomplished facticity* etc." in phenomenological reference to Husserl, which, as Bergmann states, appealed to Garfinkel (Bergmann et al. 2019:34).

interactions (cf. Goffman 1979), which have since become established and been expanded for the analysis of digital communication (cf. for example Dynel 2014; Gerhardt et al. 2014; Hutchby 2014; Blommaert 2019).[7] He assumes that the modes of participation in interactions are the subject of ongoing interactive negotiation. Thereby, he analytically differentiates between the roles of speaker (production format) and listener (participation framework). Speakers act as *animators*, i.e. the instance that carries out the act of utterance by talking or posting, *authors*, i.e. the instance that formulates a statement, or *principals*, i.e. the instance that is socially or institutionally responsible for a statement, such as a newspaper editorial office, a ministry, a politician, or editor. In addition to these three speaker roles, Goffman posits the role of the *figure*, for example a narrative ego, which is represented as a character in interactions (cf. Goffman 1979:16-23). Both individuals and institutions can fill the various participation roles and often all roles coincide in one person (cf. Gerhardt et al. 2014).

The demarcations between production format and participation framework form the basis for Goffman's analysis of shifts in the mutual alignment (*footing*) in interaction (cf. Goffman 1979:19), which he describes as follows (Goffman 1979:5):

> A change in footing implies a change in the alignment we take up to ourselves and the others present as expressed in the way we manage the production or reception of an utterance.

The fact that precisely this much-cited description of footing shifts is based on a prototypical and almost rigid draft of a conversation with the components production, reception, and utterance is due to the pre-empirical analytical separation of the descriptive categories. It cannot disguise the most important aspect that footing is negotiated interactively and thus highly dynamically. For it is not the statuses of participation that are primarily relevant for the analysis of media practices, but rather their "structural underpinnings" (Goffman 1979:5) and their interactional negotiations. Irvine, who examines the "indeterminacy of participation roles" in interactions (1996), conceives of this negotiation as a solution to a *mapping problem* (Irvine 1996:136):

> The best approach to the mapping problem, then, may be to resist the temptation to try to arrive at a single, all-purpose solution that thrusts the problem into the background as merely the rationale for a scheme of PRs [participation roles]. Instead, the mapping problem itself—the process by which participation structures are constructed, imagined, and socially distributed—is what should come to the fore. It is not only an analytical problem; it is also a participants' problem, to which there are creative, if often evanescent, solutions.

Alignment in interaction can thus be seen as a lively interplay between the participants and their production and reception roles. The concept of 'participation role' goes beyond the question of who is speaking and who is listening. 'Footing' also includes the question of the social alignment of the participants to each other, e.g. the interactional negotiation of knowledge and status asymmetries. Such negotiations always take place within the framework of specific media practices, medialities, and semiotic modes. The analysis (section 5.) therefore emphasises the extent to which footing shifts are linked to the mediality of interactional media and the

---

7   I refer here mainly to Goffman's essay on footing, first published in 1979 and republished in 1981 in his book *Forms of Talk* (Goffman 1981).

semiotic potentials through which they can be realised. However, the current state of research on live video streams will be discussed first.

3. Live video streams

Livestreaming platforms first appeared around the middle of the 2000s. For example, *ComVu Pocket Caster* was launched in 2005, followed by *Ustream* (2006-2016; now *IBM Watson Media*)[8] and *Bambuser* (2007; cf. Rein/Venturini 2018:3361). Later, other providers, such as *younow* (since 2011), *Meerkat* (2015-2016) and *Periscope* (since 2015) followed. *Justin.tv*, which was launched in 2007 and discontinued in 2014, gave rise to the *Twitch* platform, on which primarily live videos from the gaming sector are streamed and which has since become the most successful livestreaming platform. The major social media platforms such as *YouTube* and *Facebook* now also offer a very widely used livestream function (cf. Rein/Venturini 2018). *Twitter Inc.*, on the other hand, bought *Periscope* before its launch in 2015 and integrated the live video platform into its own microblog platform, allowing stream reception directly from the *Twitter* interface. In late 2020, *Periscope* announced that the service would be discontinued from 31 March 2021 and that *Twitter* would offer an integrated streaming function.[9]

Even if livestreaming can be regarded as basic technology for television broadcasts or video conferences, live video streams over the Internet are understood here to mean processed data transmissions (*stream*) of audiovisual media formats (*video*) at the time of their technical recording (*live*). In contrast to video conferences or video calls, this is an asymmetrical one-to-many form of communication in which the producers of the stream – the streamers – transmit the live video to several stream recipients. No video is transmitted in the opposite direction, with the result that all viewers see the same livestream screen together with the chat window (cf. fig. 5; cf. fig. 6 to compare stationary and mobile clients).
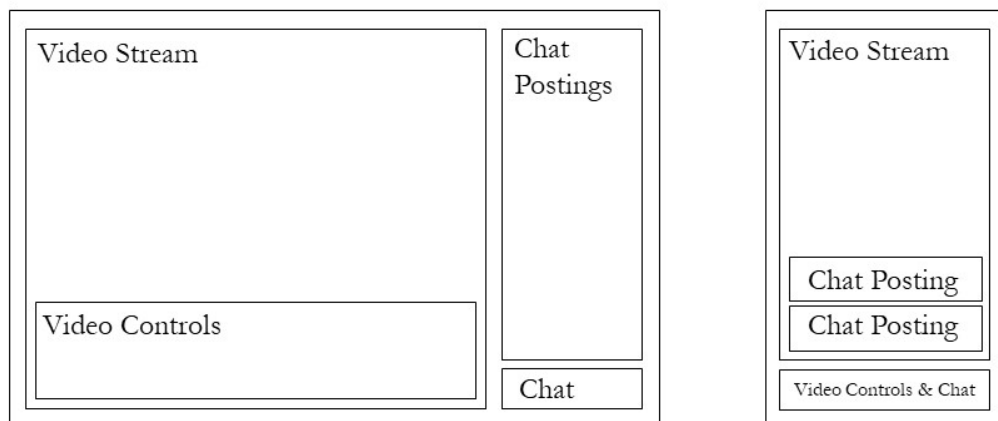

Fig. 6: Schematic illustration of the *Periscope* viewing area
on stationary client (left) and smartphone (right)

---

8   Cf. https://blog.video.ibm.com/streaming-video-news/history-of-ibm-watson-media/ [06.10.2020].
9   Cf. https://periscope.medium.com/farewell-periscope-164db2742b7c [17.03.2021].

Livestreaming platforms usually offer a chat function, which enables the video recipients to interact with each other and with the streamers, as in the case of *Periscope*. The categorisation of live video streams as a one-to-many form of communication cannot therefore be maintained (cf. also Dynel 2014:38); rather the parallelism of video and chat leads to a highly asymmetrical participation structure (Licoppe/Morel 2018:638):

> while streamers may gesture and talk, the audience's involvement is mostly mediated by text messages, which, moreover, are ephemeral, for they only remain visible for a few seconds and cannot easily be archived or retrieved.

As Reeves et al. (2015) show by analysing a mixed reality game in which players must find each other in a city while being watched by online players with whom they also interact, the medial arrangement of the live video stream poses complex challenges for participants. The players find themselves in various tensions between their bodies and the camera, the demands of the online audience and the scene on the streets, the frontstage in front of and backstage behind the camera, as well as between their roles as speakers and performers.

3.1. Ruptured temporality

During live video streaming, different semiotic resources are available to streamers and viewers (audiovisual video vs. text-based chat), which in turn are subject to disparate temporal conditions. Indeed, as described by Licoppe and Morel, chat posts are only displayed for a short time and, depending on the number of active chat users, sometimes run through the chat window relatively quickly or fade out after a while. Nevertheless, ephemerality is more likely to be attributed to the video, as, while the chat postings by the streamer and viewers are still 'scrollable', what is transmitted in the video, often through the perspective of the streamer's talking head (cf. Licoppe/Morel 2012, 2018), is transient to the extent that it is visually available continuously during the transmission of the live video stream, but is not repeatable. Only after the livestream has ended can videos be viewed on *Periscope* and selected parts accessed via video control.[10]

Moreover, live video streams can only be described as quasi-synchronous transmissions (cf. Garcia/Jacobs 1999), rather than real-time communication, as both the transmission of the video signal and the different transmissions of chat postings are subject to the physical and technical latencies (*delay*; cf. Schmidt et al. 2020) of (mobile) Internet transmissions. Due to the larger amounts of data, it can be assumed that there are greater latencies in video transmission than in chat transmission. Schmidt et al. (2020) use the case of live broadcasts of the video genre *Let's Play*, in which the playing of computer games is streamed live, to show that on the *Twitch* platform the video stream had a latency of about 12 seconds at the time of the study. Using sequence analysis, they show how latency presents the participants with challenges in terms of sequence organisation, such as how to assign a chat posting to a game action. In addition, the production of chat postings is subject to temporality conditions that are different from those of face-to-face interaction, as

---

[10] The technical descriptions of the live video stream platform *Periscope* refer to its status in October 2020.

Beißwenger has shown (cf. Beißwenger 2007). Chat postings require a certain amount of production time before they are sent and may also involve repairs before distribution. The succession of chat posts is not always clearly sequential. References between postings are at times interlaced and therefore difficult to assign. Chats are thus subject to specific temporality conditions regarding the production and reception of chat contributions (cf. Beißwenger 2020). On the other hand, the *hic-et-nunc* situation of the livestreamers is characterised by its sequential orderliness. In the livestream, the local situations of those streaming and chatting, which are each subject to different temporality conditions, are synthesised in the media on a screen (cf. Knorr Cetina 2009). The asymmetrical temporality synthesised in live video streams can be described as *ruptured temporality*, which the participants (must) deal with interactionally.

3.2. Asymmetrical establishment of relevance

Nevertheless, live video streams are also structured sequentially by the streamers and the viewers, with the streamers as the principal performers dominating the episodes and controlling the interaction. The livestream of journalistic reporting analysed here is characterised by the fact that the streamer organises the streams continuously by sequencing them by means of "ritual brackets" (Goffman 1979:7), for example by repeated openings (cf. Schegloff 2007; Frobenius 2011) and closings (cf. Schegloff/Sacks 1973) and/or changes of situation or location (cf. section 5.). Accordingly, live video streams are continuous interaction situations, which for their part, especially in the case of longer streams, are episodically divided into interaction sequences primarily by the interaction-dominant streamers. It must be noted, however, that in the continuous situations of interaction, even if there is no talking for a while, the participants are in a constant "state of talk" (Goffman 1979:7). Interaction can be initiated at any time by the ratified participants in the video screen or chat, while it is primarily up to the streaming host to decide to what extent things, topics, objects, actors, etc. are rendered subjects of interaction in which participants are involved, in equal or different proportions.

The asymmetrical participant structure requires specific practices for establishing relevance in live video streams that reproduce this configuration.[11] *Read aloud and respond* (*RAR*; cf. Licoppe/Morel 2018) as a specific practice of noticing (cf. Schegloff 2007:217-220; on noticing in livestreams Recktenwald 2017; Schmidt et al. 2020) should be mentioned here. 'Noticing' refers to the interactive establishment of relevance of features that were previously not considered relevant in an interaction and which become the subject of the conversation through noticing. Licoppe and Morel (2018) describe *RAR* as a typical streaming practice in which streamers read chat posts aloud and thus make them relevant for interaction. This is then followed by one or more turns in which the streamer responds to the chat posting.

---

[11] The relationship between asymmetrical participation structure and practices of establishing relevance in live video streaming is mutual, as is typical for media practices: the mediality of the medium provides the basis for certain affordant potentials for action, which are employed and implemented in practices in such a way that they structurally reproduce the asymmetry at the level of sociality and thus consolidate it. The fact that such structuring practices are nevertheless dynamic and variable is evidenced by cases of changing media practices, such as liking on *Twitter*, which has changed considerably since the introduction of the like button (cf. Paßmann 2018).

Recktenwald describes this practice as 'topicalising' (2017), which emphasises the fact that the streamer selects a specific topic for subsequent turns from many possible chat postings. The selection of a specific chat posting initiated by the streamer therefore represents a control of the course of interaction which, due to the modal resource advantage (linear audiovisual video vs. 'chunk-by-chunk' sequence of the chat), is reserved for the streamer as opposed to the chat participants. This requires chatters, who are competing for attention, to design their contributions intelligibly, purposefully, and quickly in order to provide them for recitation. Accordingly, the streamers have a dominance of interaction due to medial affordance (cf. Sjöblom et al. 2019), which they use for sequence and topic control.

4. Data & methods

4.1. Analysing live video streams: methodological remarks

Live video streaming is characterised by long passages in which the streamers exclusively hold the interactional floor. It may initially appear contradictory that although live video streaming seems to unfold primarily in a mode of monologic talk, it is analysed as a thoroughly social activity driven by interactional dynamics. Furthermore, in contrast to more dyadic communication situations, the analysis of live video streaming reveals a fundamental methodological issue of sequence analysis: while CA is primarily conducted along the participant's understandings displayed in *n*-th positions in interaction sequences (*next turn proof procedure;* cf. Hutchby/ Wooffitt 1998:15ff.), such displays are often absent in rather monologic turns in live streams. During the extended passages in which the streamers hold the floor, understanding is not necessarily documented by the displays of further participants in sequential turns. As will be shown, streamers tend to anticipate and elicit subsequent actions of viewers, depending on the practice being carried out. The present study aims primarily to analyse such interaction-initiating and involving media practices of the participants against the background of the journalistic practice of broadcasting and reporting via live video streams. In doing so, the analysis takes an *ex situ* perspective available via the documented playback video of the live video stream, which might differ from the *in situ* perspective via the streamer's and recipient's devices and clients at the time of streaming. However, I take the praxeological view that practices are reflexive, whereby this reflexivity, according to Garfinkel, is to be regarded as a "self-explicating property of ordinary actions" (Ten Have 2004:20). Thus, as intelligible and accountable semiotic practices, they are also accessible to an analytical *ex situ* perspective that does not draw exclusively on the participant's interactional displays (in the sense of documentations of understanding) but also on the analyst's interpretations (which in turn emerge from practices of intersubjective negotiation in data sessions). [12]

---

[12] The methodological critique cannot be fully developed here. Reference should therefore be made to Deppermann, who states that the concept of *display*, which forms the methodological basis of conversation analysis (cf. Hutchby/Wooffitt 1998:15ff.), makes the 'interpretative constitutive performance of the analyst' methodologically invisible in a positivist and structuralist manner (orig.: "*interpretationskonstitutive Leistung des Analytikers*", Deppermann 2001:56, original emphasis). The sequence-analytical premise that 'every activity can only be adequately understood by us as analysts on the basis of its procedural consequences' (orig.: "jede Aktivität für uns

4.2. Transcription

The transcripts used here are oriented towards the basic transcript of GAT2 (Selting et al. 2011) and the conventions for multimodal transcription by Mondada (2019) and are slightly modified to suit the subject matter. Transcripts always represent a sufficient selection for analysis of a holistic event that can never be captured in its entirety. Nevertheless, the emphasis of multimodal analysis is derived from the accomplishments of the participants, because "[t]he prioritization of one resource over another is not a matter that can be decided a priori but is an empirical issue that depends on the type of situated activity and how participants format it" (Mondada 2018:87). In addition to the linguistic utterances, the present transcripts focus on the parallelism of video stream and chat as well as on the embodied interaction (cf. Mondada 2018) through the physical movement, gaze, and gestures of the streamer. Accordingly, the transcript was extended by aligned notation of *move*(ment), *gaze*, and *hand*, each of which is noted in a separate line and all of which are to be assigned to the streamer. The streamer consistently holds the smartphone camera with a stick in his right hand in the clips viewed. Since the streamer maintains the talking head perspective almost entirely in the analysed excerpts, the camerawork is not transcribed, with one exception in the opening turn (notation: *cam*, cf. transcript 2). The streamer's camerawork is aligned with his movements, i.e. body movements, for example when the streamer turns, imply a camera movement. Hence, the co-ordination of camerawork and body movement establishes local relevance (cf. Broth et al. 2014; Reeves et al. 2015).

Where they are made relevant to the interaction by the participants, the chat postings are displayed in a separate column on the right-hand side (cf. transcript 5). The multi-column layout (cf. Recktenwald 2017) indicates the different semiotic modes

als Analytiker nur anhand ihrer prozeduralen Konsequenzen adäquat zu verstehen", Deppermann 2001:55) leads to a 'temporal regress' (Deppermann 2001:55), according to which an interactional phenomenon can only be understood after its completion. For live video streams, this meant that situated interpretations could be carried out neither by the participants *in situ* nor by the analysts *ex situ* on the basis of situational excerpts. Deppermann derives from his critique the demand for *conversation analysis as an explicative construction*, which consists of showing in the analysis 'which explicative-argumentative interpretations of empirical interaction events can be gained under the condition of certain background assumptions' (orig.: "*welche explikativ-argumentativen Interpretationen empirischer Interaktionsereignisse unter der Voraussetzung bestimmter Hintergrundannahmen zu gewinnen sind*", Deppermann 2001:66f., original emphasis). From a praxeologically oriented media-linguistic perspective, of particular relevance here is the extent to which a) the practice-bound nature of the phenomenon under investigation, b) ethnographically obtained background assumptions and c) (media) linguistic expertise guide the analysis interpretatively. While the streamers make relevant selected displays from the chat by *noticing* (cf. section 3.2), the present analysis additionally rests on the constructions of the analyst, who incorporates a) the journalistic practice of live video reporting, b) his ethnographic experience in the field of street protests (cf. Dang-Anh 2019) and c) intersubjective expertise from data sessions into the interpretation. Deppermann's résumé can thus also be applied here: 'We do not know whether the participants in the interaction share these constructions - and in my opinion it is not decisive for many studies.' (orig.: "Ob die Interaktionsteilnehmer diese Konstruktionen teilen, wissen wir nicht - und ist m.E. für viele Untersuchungen auch nicht entscheidend", Deppermann 2001:67). The analysis rather achieves the following: 'general, i.e., typologically transferable, interpretative principles and practices are explicatively constructed using the example of particular conversations' (orig.: "am Beispiel einzelner Gespräche werden allgemeine, d.h. typologisch in andere Kontexte transferierbare Interpretationsprinzipien und -praktiken explikativ konstruiert", Deppermann 2001:67).

in which the participants interact. In addition, the problem of ruptured temporality between streamers and chatters, which already arises during livestreaming (see above), also continues in the *ex situ* reception of the recording of the livestream: at different access times, chat messages were displayed inconsistently during playback of the video. The chat messages are aligned with the streamer's turns based on the documented video of the stream. In this respect, the transcript does not represent a precise temporal alignment of the video with the chat window as it appeared to the participants at the time of the livestream, which is marked by the vertical arrangement of stream and chat postings and the double line separating them (cf. transcript 5). However, the analysis indicates that the sequentiality, as produced by the participants, is sufficiently documented in the transcript.

4.3. Context and data

The excerpt analysed in this essay is taken from a live video stream by Martin Kaul (MK), then reporter for the German national daily newspaper *taz*.[13] The livestream was broadcast on 16 September 2018 from a gathering in Köthen in the German federal state of Saxony-Anhalt, which was organised by far-right and right-wing extremist groups following an altercation in the town after which a man died.[14] A week before, MK had streamed from a similar event in Köthen, where he documented hate speech and was threatened by those attending the gathering.[15] The clip examined here was launched on *Periscope* on 16 September 2018 at about 18:06 hrs and is entitled "Kaul aus #Köthen, 18.06 Uhr.".[16] MK streamed two videos with a smartphone on a stick, one at 18:06 with a length of one hour 16 minutes and one at 19:33 with a length of one hour 17 minutes.[17]

---

[13] Kaul pioneered live video streaming coverage in Germany after reporting live from the G20 summit in Hamburg with his smartphone in 2017. Upon the discontinuation of the livestreaming platform *Periscope*, Kaul wrote an 'obituary' in which he recapitulates his past journalistic live reporting on *Periscope* (cf. "Tschüss, Periscope!", taz),
https://taz.de/Aus-fuer-Livestream-App/!5757024/ [24.03.2021].

[14] Cf. "Nach Todesfall in Köthen: Hunderte folgen Demonstrationsaufruf rechter Gruppen", Tagesschau (video)
https://www.tagesschau.de/multimedia/video/video-448867.html [09.10.2020].

[15] Cf. "Gewalt gegen Journalisten: Wenn Twittern zum Sicherheitsrisiko wird", Deutsche Welle,
https://www.dw.com/de/gewalt-gegen-journalisten-wenn-twittern-zum-sicherheitsrisiko-wird/a-45435583 [09.10.2020].

[16] Cf. the announcement on *Twitter* at
https://twitter.com/martinkaul/status/1041357479353888768 [09.10.2020]
and the archiving of the stream on *Periscope* at
https://www.pscp.tv/w/1ZkKzNLzbjaKv [09.10.2020].
According to information from *Periscope*, around 7,000 viewers were watching the video live at the time of broadcast. The number at the bottom left of the browser view screenshot (fig. 5) indicates the total number of viewers of the live video and the documented playback video.

[17] Cf. https://twitter.com/martinkaul/status/1041379614910689280
and https://www.pscp.tv/w/1OwxWWQmpOnxQ [09.10.2020].

5. Analysis

Four transcripts are discussed below. In the first excerpt, the practice of *hosting* is examined by looking at the streamer's opening turn (section 5.1., transcript 2). Transcripts 3 to 5 cover a longer episode over three parts, discussing the practices of *situating*, *recruiting*, and *acknowledging*. (sections 5.2.-5.4.).

5.1. Hosting

Transcript 2: "Greeting"
(from: "Kaul aus #Köthen, 18.06 Uhr.", 00:00:00–00:00:36)

```
00:00 01    §((Sprechchöre, unverständlich, ca. 12 Sek.; Musik))
            ((chants, incomprehensible, approx. 12 sec.; music))
       cam §shot diagonally downwards-->
00:12 02    ((Sprechchöre)) a anti anticapitalista (.) a (.) anti
            #anticapitalista (.) a (.) anti (.)§ anti#cap-
            ita[lista]
            ((chants)) a anti anticapitalista a anti anticapital-
            ista a anti anticapitalista

       cam                             -->§camera switches
```



```
            #fig.7                        #fig.8
00:22 03 MK         [hallo guten TACH] herzlich willkommen in
            der sächsen anhaltinischen kleinstadt (.) KÖthen-
            hello good day welcome to the small town of Köthen in
            Saxony-Anhalt
00:27 04    (0.42) ((Sprechchöre))
                   ((chants))
00:27 05 MK es is achtzehn uhr ähm SECHS mein name is martin
            kaul,
            it's six minutes past six my name is martin kaul
00:30 06    (0.44) ((Sprechchöre))
                   ((chants))
00:31 07 MK das_n SERvice den die tages (-) zeitung TAZ aus ber-
            LIN präsentiert-
            this is a service presented by the daily newspaper
            taz from berlin
00:34 08    (0.47) ((Sprechchöre))
                   ((chants))
00:35 09 MK ich erzähl #heute_n bisschen im livestream (-) was
            hier LOS is.
            i will tell you today a little bit in the livestream
            what is going on here
```

#fig.9

00:38 10   (2.0) ((Lautsprecherdurchsage))
*   ((loudspeaker announcement))*

At the beginning of the livestream (cf. transcript 2) the camera is pointing at the street. There appear to be police officers walking in the street, with the camera filming their uniformed legs up to hip height (cf. fig. 7). What can be heard, but not understood in words at first, are chants and music. The setting is already apparent in the first few seconds when the streamer is not visible or audible in the picture. Street, police, and chanting indicate street protests taking place at the site of the events being filmed and streamed. The camera's position on the ground reflects the practice of not filming police officers' faces at close range, something MK repeatedly avoids doing throughout the video. The chants identify the protest as anti-capitalist (line 02), conveying a political stance that is often also expressed in anti-fascist-oriented street protests.

MK enters the stream with an opening turn (cf. Schegloff 2007) and the camera view switches from rear main camera to front camera (cf. fig. 8). In the background, a demonstration march with banners can be seen, in front of which MK, who is shown in the foreground, is walking (cf. fig. 9). The extensive greeting, the talking head shot that has now been adopted (cf. fig. 8; cf. Licoppe/Morel 2012), and the fact that the stream is being sent via his *Periscope* account identify MK as the host in the live video stream. This marking already determines the alignments of the participants to each other: *host* and *audience*. Hosting is to be understood as the arrangement of mediated digital interaction, through which the host unilaterally conditions the medial, multimodal, and semiotic framework and resources of an interaction situation by both prospectively setting it up and continuously administrating it. Hosting is thus a media practice that presituatively prepares digital interaction situations with regard to the orientation of the participants and dominates them insituatively.[18] The host thus adopts the potential to act semiotically and delegate participation status. In the excerpt, the opening sequence (line 02), the camerawork, the set-up on the account of the live video stream platform, and also the point in time chosen for focused interaction – initially the video runs without commentary for 22 seconds, which is partly for technical reasons – constitute hosting as reflexive, i.e. displaying itself as a media practice. To summarise: the alignment of streamers and viewers to each other is established by the host's reflexive display

---

[18] *Situative communication* or *situative communicative practices* can be analytically differentiated here into *presituative communication*, which precedes situations in time and is related to the future, *insituative communication*, which takes place simultaneously with situations and is related to the present, *postsituative communication*, which takes place after situations and is related to the past, and *transsituative communication*, which can take place at any time and establishes relations that last beyond the situation (cf. Dang-Anh 2019:361ff.). Situations have different ranges and can be nested within each other. For example, the entire political event on which MK is reporting here can represent a situation and the change of location focused here can also be regarded as a situation.

and is then continuously processed interactively. This describes the media practice's peculiar mutual relation of mediality and practical accomplishment, which from the very beginning establishes an interactional asymmetry that consistently characterises the social relationship between the streamers and the viewers.

MK further specifies the situation in the following by providing information about place (line 03) and time (line 05) as well as those taking part (line 05). MK thus addresses three fundamental dimensions of a situation which, following Bühler's notion of the speech situation (Bühler 1999 [1934]),[19] can be defined as 'the spatial, temporal, and personal constellation practically made present in interaction' (Dang-Anh 2019:50).[20] 'Making-present' (*Vergegenwärtigen*) is thus to be seen as a situation-constitutive practice through which participants express situations, but also create them in the first place. Live broadcasts such as video streams are particularly dependent on ongoing situational practices of 'making-present' due to their ongoing ephemerality. At this point at the beginning of the livestream, however, the constitution of the situation primarily has an introductory character. With this initial grounding (cf. Clark 1996:221ff.), the host informs the audience about the situation in which he and they find themselves.

He perceptibly acts as an animator and the consistently spontaneous style suggests that he is also the author of his statements. When he mentions his employer, a daily newspaper (line 07), however, he refers to this as the principal. MK thus consolidates his social positioning as a journalist who is reporting on an event on behalf of a newspaper (line 09), while at the same time displaying his hosting as professional. MK closes the introductory sequence by pausing (line 10), which is followed by a narrative contextualisation of the events preceding the reported event (not transcribed).

5.2. Situating

In the following, an episode[21] will be analysed in detail to show how a change of location in mobile live video streaming unfolds sequentially and is processed interactively. The excerpt from the *Periscope* video "Kaul aus #Köthen, 18.06 Uhr." is one minute 13 seconds long in total (00:22:40–00:23:53) and forms a closed sequence, which is analysed in three partial excerpts (cf. transcripts 3, 4, and 5). MK opens with the particle *soo* (cf. transcript 3, line 01).

---

[19] The spatial, temporal, and personal dimensions are derived from Bühler's theory of language (1999 [1934]), in which he defines the 'here-now-me system of subjective orientation' (orig.: "hier-jetzt-ich-System der subjektiven Orientierung", 102), the *origo*, as the starting point for situation-related speaking with deictic expressions.

[20] Orig.: "die in der Interaktion praktisch vergegenwärtigte räumliche, zeitliche und personale Konstellation" (Dang-Anh 2019:50).

[21] The times given in transcripts 3-5 refer to the section of the episode shown in transcripts 3-5 examined below, not to the entire livestream.

Transcript 3: "Situating"
(from: "Kaul aus #Köthen, 18.06 Uhr.", 00:22:40–00:22:56)

```
00:00   01   MK       %+soo;+
                      well
             move     %walks-->>
             gaze      +look+s straight ahead
00:02   02   MK       +wir warn jetzt grade *bei #dieser demonst-
                      ration da HIN* ne_#ANTIfaschistische
                      αdemonstratio-
                      we just were at this demonstration back
                      there an antifascist demonstratio
             move     α turns to the side and back-->
             gaze     +looks into the camera-->
             hand                          *points behind him----
                      -------------*extends index finger-->
```



```
             #fig.10          #fig.11
00:06   03   MK        hier sind jetzt weiter son bisschen+ die
                      straßen gesperrt,*α+
                      here the roads are now closed a little bit
             move                      -->α
             gaze                             -->+looks to
                      the side-----------+
             hand                -->*
00:10   04   MK       +ich LAUF jetzt äh (-) für alle die nich
                      *beSCHEID# wissen was hier gerade passiert*
                      (.)+nochn bisschen WEIter,
                      i'm walking now to all those who don't know
                      what's going on here right now a little fur-
                      ther
             gaze     +looks into the camera---------------------
                      ---+looks straight ahead-->
             hand     * points in the direction of the camera----*
                                            points forwards-->>
```



```
             #fig.12
00:15   05   MK       und +versuche mal den MARKTplatz zu αfin-
                      den;+
                      and try to find the market place
             move                                      αturns
                      back and forth-->>
             gaze      -->+looks into the camera-----------------
                      ----+
```

The streamer, apparently with the smartphone stick in his right hand, walks down a street and looks into the camera. In line 02 the situation is presented as a retrospective, which is characterised by the past tense of the verb *warn* ('were') as well as by

the temporal deictic indication of time *jetzt gerade* ('right now'). *[D]ieser Demonstration* ('this demonstration') and *ne_ANTIfaschistische demonstratio* ('an anti-fascist demonstratio') denote the political event that the streamer has just reported on as a journalist. This naming is combined with the streamer's finger pointing behind him, i.e. against his direction of movement (fig. 10) – a spatial marking of a temporal retrospective. During the more precise attribution (*ne_ANTI-fascist*) he extends his index finger in addition to his outstretched thumb (fig. 11), which can be understood as an enumerative gesture that marks the described demonstration march as a partial event in a larger event context. In this sequence opening, the emphasis is on verbal deictic and gestural pointing to personal, temporal, and spatial dimensions of the constitution of the situation. The situation is thus readjusted in this sequence; the scenery is changed. MK now wants to go to the far-right gathering, which is being held separately from the counter-demonstration, and changes location accordingly. The personal constellation is addressed by MK using the inclusive pronoun *wir* ('we'), which he employs to involve the viewers postsituatively (line 02): *warn* ('were') refers not only to the (now past) physical existence in the local space on site, but also to the mediated perception of the viewers in the live video stream. The streamer not only 'takes' the people he is addressing 'with him' on his journey, but also addresses them as participating agents of his physical movement action mediated by the stream, thereby constructing his change of location as a joint project (*joint action*; cf. Clark 1996:59-91). In line 03 MK then shifts to the current situation. Through *hier* ('here') and *jetzt* ('now') MK linguistically refers to the present in time and space. A very short panning shot of the barriers, which enables the audience to see what MK sees, complements the situational 'making-present' (line 03). The live character of the interaction situation now becomes even more distinct: after the retrospective involvement (line 02), an insituative involvement now takes place through the linguistic and medial construction of a jointly experienced here, now, and us.

In line 04 the streamer changes the pronominal perspective and describes how he is moving: *ich lauf jetzt [...] nochn bisschen weiter* ('i'm walking now [...] a little further'). The verb expresses his movement complementary to the moving camera image. The insertion in line 04 diversifies the constellation of participants: MK now explicitly addresses *alle die nich bescheid wissen was hier gerade passiert* ('all those who don't know what's going on here right now') and thus differentiates them from those who are supposed to know about the situation due to their stream reception. This is a first footing shift, a process that Gibson, following Goffman (1979), describes as "moment-by-moment shuffling of individuals between the 'participation statuses' of speaker, target, and unaddressed recipient" (Gibson 2003:1335). The differentiation is carried out as *grounding* in order to involve through situational 'making-present' those who were previously not included to the same extent due to a lack of situational knowledge (*was hier gerade passiert*, 'what's going on here right now') as the viewers who are aware of the situation due to their more continuous or longer reception of the livestream. However, the distinction is subtler than between target and unaddressed recipient, because even the 'knowing ones' seem to be addressed by the utterance with which MK not only grounds, but also linguistically seconds his physical movement, which is visible in the stream, in order to bridge the movement phase interrupting his on-site reporting. In mobile live

video streaming, the primary aim of which is visual participation, such as the journalistic livestream coverage of events analysed here, especially in the case of dynamic spatial events such as protests, there are sometimes phases in which there is 'nothing to see'. Throughout the video there are repeated sequences in which this kind of *bridging* is carried out: the resulting 'visual voids' are bridged by the streamer, as in the analysed sequence, with empractical communication (cf. Bühler 1999 [1934]), which is secondary to his visible physical movement, or by means of partly prepared transsituative or non-situational communication. For example, during a longer change of location, MK conducts an arranged interview while walking ("Kaul aus #Köthen, 18.06 Uhr.", 00:08:03– 00:15:56).

While MK already states in line 04 that he is going *WEIter* ('further'), which he complements with a focus accent and his finger now pointing forwards away from him (fig. 12), he makes a further shift in situational reference in line 05. His choice of verbs (*versuche, finden*, 'try', 'to find') indicates the beginning of an indeterminate and uncertain process, which implies that the attempt may succeed or fail in the future. In terms of time, too, the reference changes into a presituative mode, insofar as a future situation on the *Marktplatz* ('market place') is outlined prospectively. Lines 02 to 05 show in very short consecutive order how the situation is constituted by past-related postsituative communication (line 02), present-related insituative communication (lines 03, 04), and future-related presituative communication (lines 04, 05), and how the audience is involved in the situation. Situational references are made linguistically, physically, and medially, i.e. multimodally: firstly, with regard to *time*: grammatically by changing the tense (*warn*, 'were', line 02 vs. *lauf*, 'walk', line 04), temporal-deictically (*jetzt grade*, line 02; *jetzt*, line 03; *gerade*, line 04; 'right now', 'now'), semantically through action verbs with processual meaning (*lauf*, 'walk', line 04; *versuche…zu finden*, 'try to find', line 05), and gesturally by pointing the finger backwards from the streamer (past-related, line 02) and forwards (future-related, line 04); secondly, as regards *space*: through local deictics (*da HIN*, 'back there', line 02; *hier*, 'here', line 03), toponyms (*MARKTplatz*, 'market place', line 05), and also by pointing gestures (backwards, line 02; forwards, line 04), moving the body and changing the camera angle (panning along barriers, line 02); thirdly, as regards *persons*: by pronouns (*wir*; 'we', line 02; *ich*, 'i', line 04), the talking head camera angle, naming of groups (*ANTIfaschistische demonstratio*, 'anti-fascist demonstratio', line 02), and addressing the viewers (*für alle die* […], 'to all those who […]', line 04).

5.3. Recruiting

Transcript 4: "Recruiting"
(from: "Kaul aus #Köthen, 18.06 Uhr.", 00:22:56–00:23:24)

```
00:18   06            %α+*(0.61)#
              move    %walks-->
              move     αturns back and forth-->
              gaze      + looks backwards with body movement and
                      then laterally upwards-->
              hand       *points forwards-->
```


```
                      #fig.13
```
```
00:19   07   MK       schätze maα (-) dass der in +DIEse richtung
                      geht=%aber da könnt ihr mir mal helfen;%*
                      guess that it is in this direction but you
                      can help me there
              move               -->α
              move      -->%stands and turns to the side-----%
              gaze                          -->+looks into the
                      camera-->
              hand                                    -->*
00:22   08   MK       %*des hier is die #MÜHlenstraße* (---)
                      %die MÜHlenstraße,
                      this here is the mühlenstraße the mühlen-
                      straße
              move    %stands-------------------------------
                      %walks across the street-->
              hand      *points with his index finger diagonally
                      upwards in the direction of the sign--------
                      -----------------------------*
```


```
                      # fig.14
```
```
00:27   09            (0.6)
00:28   10   MK       und ich lauf jetzt die mühlenstraße einfach
                      gradeaus (.) weiter lang,
                      and i am now simply walking straight ahead
                      along mühlenstraße
00:31   11   MK       und +HOFfe dass+ ich dann %zum MARKTplatz
                      komme.=%
                      and hope that i will then come to the market
                      place
              move                          -->% turns before the
                      sign---%
              gaze      -->+ looks diagonally upwards--------------
                      --------------+looks into the camera-->
```

```
00:34   12   MK        #=Aber *vielleicht GIBT es ja+ jemanden(--)
                       von euch (--) der #ma kurz (-) %*GOOgeln
                       kann-
                       perhaps there is someone among you who can
                       briefly google
             move                                    %walks on-->>
             gaze                            -->+looks un-
                       derneath the camera
             hand            *scrolls----------------*
```

fig.15   # fig.16

```
00:39   13        (0.7)
00:40   14   MK        +äähm-
                       eehm
             gaze   -->+looks into the camera-->
00:41   15        (0.59)
00:42   16   MK   ob DAS hier der richtige WEG is;
                  whether this is the right way
00:43   17   MK   *sonst lauf ich nämlich in die #FALsche
                  richtung.*
                  otherwise i'm walking in the wrong direction
             hand   *points forwards---------------------------
                       ---------*
```

fig.17

```
00:45   18        (0.68)+
             gaze      -->+
```

MK now turns while walking and looks first behind him and then to the side, slightly upwards (line 06, fig. 13). Again, he uses a vagueness marker, this time with the epistemic verb *schätze[n]* ('[to] guess', line 07), by which he marks his lack of knowledge of the local spatial conditions. In the accentuated *DIEse richtung* ('this direction', line 07) his finger points in front of him in the direction in which he is walking, which documents the intertwining of linguistic and gestural local deictics. This is followed directly by the addressing of the viewers, which is initiated by an adversative *aber* ('but') and by means of which the viewers are requested to help (*helfen*, 'help', line 07) using the modal verb *können* ('can'). The whole of line 07 *schätze ma (-) dass der in DIEse richtung geht=aber da könnt ihr mir mal helfen;* ('guess that it is in this direction but you can help me there') constitutes a footing shift by which the streamer actively involves his audience in the situational events. The aforementioned vagueness marker at the beginning, together with this audience *recruitment* (cf. Drew/Couper-Kuhlen 2014; Kendrick/Drew 2016), indicates a potentially asymmetrical epistemic status difference by means of epistemic stancetaking (cf. Heritage 2012) between the streamer and the audience. With his rotating body movement, his rapidly changing directions of gaze, his vagueness-

marking verbalisation, and his addressing of the audience, MK indicates that he is currently orienting himself in the local space and is dependent on the help of third parties. Highlighting the viewer's potential ability to help MK (line 07) are first pair parts that, as further evidenced by his subsequent scrolling (line 12), aim to elicit chat posts as second pair parts in an adjacency pair (cf. Schegloff 2007). By constituting an orientation problem in this way, orienting in the live video stream is identified as a situational search practice whose chances of succeeding are increased by the situational involvement of the viewers. Although the streamer is interaction-dominant as a host and an expert as a journalist, here he marks a knowledge asymmetry that seems to oppose these status asymmetries. However, displaying the knowledge asymmetry gives the streamer the opportunity to involve his audience in his situation. Epistemic stancetaking, i.e. the expression of differences in epistemic status, serves to initiate sequences "as a means of warranting conversational contributions and building expanded conversational sequences" (Heritage 2012:49). This footing shift therefore projects subsequent turns taken by the audience, although it is not yet clear how the audience can help the streamer in this situation.

Thereafter, an object-oriented spatial concretisation of the previously constituted orientation problem takes place: in line 08, MK identifies his spatial situation by repeatedly mentioning the street name. He turns so that a road sign can be seen in the background and points his finger in the direction of the sign (see fig. 14). Through this noticing, the sign becomes identifiable from the perspective of the spectators as a reference object in the depicted urban space (cf. Licoppe/Tuncer 2019). During lines 09 to 11 MK crosses the street and looks up sideways. He thus repeats in the same direction the laterally oriented view from line 06, which apparently fixed the identical object. MK changes pronominally (*ich*, 'I') back to the ego perspective and situates himself temporally (*jetzt*, 'now') and spatially (*mühlenstraße eingeradeaus (.) weiter lang*, 'mühlenstraße straight ahead further along'), whereby the description of the action not only describes his current activity, but also anticipates his future movement by means of a continuous process verb and indication of direction (*lauf […] einfach geradeaus (.) weiter lang*, 'just walk [...] straight ahead further along', line 10). The verb *hoffen* ('hope') again marks the uncertainty concerning his orientation problem (line 11), just as the destination *MARKTplatz* ('market place') is repeatedly mentioned. Both the current location (*mühlenstraße*) and the anticipated destination (*MARKTplatz*) are mentioned several times at this point and are thus established as reference markers for spatial orientation. The sequence described can thus be read as a *direction-giving sequence*, which plays an important role in mobile interaction: "'Direction-giving sequences' involve practical formulations that take into account a starting point as well as a target or a destination" (Haddington et al. 2013:20). To anchor his actual spatial situation (cf. Hausendorf 2013), MK places himself in front of the sign, which was previously only visible from a distance and can now be read in the video stream (line 11; fig. 15). The object 'street sign' thus serves as a semiotic resource (cf. Nevile et al. 2014) for the now shared practice of orienting: MK uses the sign as a locally anchoring and authenticating evidence object of spatial orientation and as a visual reference point for the recipients during his pause in movement.

In line 12 MK again addresses the viewers. He introduces the resumption of the footing shift again with an adversarial *Aber* ('but'), hinting at the possibility of solving the orientation problem. He chooses a vague recipient design for his recruitment, disguised as a presumption of existence: *vielleicht GIBT es ja jemanden (--) von euch* ('perhaps there is someone among you', line 12). By specifying the address with *der ma kurz (-) GOOgeln kann* ('who can briefly google'), the streamer makes an interaction-initiating offer which is specifically tailored to the participants in the mediated interaction situation in that it takes into account their local reception situation. The stream is received by the viewers via a digital medium, such as a computer, a smartphone, or a tablet, which also gives them the opportunity to 'google' alongside the livestream reception. However, this shift is not only an offer of interaction, but rather a *recruitment*[22] through which participants are to be made to do something (cf. Drew/Couper-Kuhlen 2014; Kendrick/Drew 2016):

> Recruitment covers the various ways in which one person can ask for, seek, or solicit help from another, including giving indirect and perhaps embodied indications of their need for assistance, as well as another's anticipation of someone's need for help and their offering or giving that help without being asked, without their help having been solicited. (Kendrick/Drew 2016:2)

Such "attempts to enlist someone's assistance, typically with respect to an immediate, physical need, problem or wish" (Drew/Couper-Kuhlen 2014:17) are usually performed locally: "Recruitment is, then, restricted to quite material, here-and-now matters; it does not involve 'remote' matters" (Kendrick/Drew 2016:2). In the live video stream, however, there is distant recruitment, which is characterised precisely by the fact that the recruiting streamers have a local problem of orientation that the viewers are supposed to solve *with them* according to their local situations, which differ from those of the streamers. Accordingly, the involvement of the recruits does not come about through 'assistance', but rather through their contribution to the joint project of changing location (cf. Zinken/Rossi 2016). Recently, it has been emphasised that recruitment concerns practical activities that go beyond mere speech acts (cf. Rossi et al. 2020). Although the 'googling' itself is not visible in the livestream, the aim of the recruitment is its practical accomplishment. This clearly indicates how the live video stream situation can be understood as a mediated *synthetic situation* (cf. Knorr Cetina 2009) or *intersituation* (cf. Hirschauer 2014), in which, in addition to the local situation of the main actor, the media reception situations of the viewers also form the background to the situational setting and the parties nevertheless pursue a joint project. Recruitment in mediated interactions does not exclusively refer to activities that are carried out in the shared here and now but is also accomplished intersituatively.

MK's own media reception is indicated by the obvious scrolling on his smartphone, which he carries out alongside his recruitment address (line 12, fig. 16). By displaying the reading of the chat postings, which are shown in the chat window of the screen on the streaming device, the host not only signals his perception of perception (*Wahrnehmungswahrnehmung*, see above), but also expresses the expectation that at the same time as his live report, viewers will also interact with him or

---

[22]   In contrast to the recruitment of viewers by the streamer described here, Choe shows in the case of the Korean livestream *mukbang*, in which the host eats in front of an audience, how chatters recruit the host for activities (cf. Choe 2019).

with each other via chat. The recruiting interaction offer is thus complemented by the expression of interaction expectation. However, not only because of the ruptured temporality of the live video stream, the aim of this display is not necessarily to elicit an immediate response, since, in addition to the delay caused by the production of chat postings and latency, the viewers were recruited to acquire situational knowledge about local conditions, which is also likely to take time.

While MK continues on his way, he repeatedly expresses his uncertain knowledge of the place (*ob DAS hier der richtige WEG is*, 'whether this is the right way', line 16). Moreover, by mentioning negative consequences while walking, he indicates his visible movement action as being possibly counterproductive (*sonst lauf ich nämlich in die FALsche richtung*, 'otherwise i'm walking in the wrong direction', line 17, fig. 17). The sub-project 'change of location' and possibly the entire project 'live video stream from the right-wing event' is thus threatened with failure or delay and MK increases the pressure on the recipients to participate by presenting a negative outline of the future. To prevent failure, the interactional situational involvement of the audience is necessary. Recruitment is therefore used here to acquire situational knowledge, as knowledge concerning the temporal, spatial, and personal dimensions of situations, until the state of an "equilibrium-for-all-practical-purposes" (Heritage 2012:48) is achieved, which is sufficient for the success of the joint project of changing location.

5.4. Acknowledging

Transcript 5: "Acknowledging"
(from: "Kaul aus #Köthen, 18.06 Uhr.", 00:23:25–00:23:53)

```
00:46  19  MK    %+äähm +DA würd ich jetzt weiter        Chat
                 HINlaufen,
                 eehm i would keep on walking
                 there
           move  %walks-->>
           gaze   +looks straight ahead
                  ------+looks underneath the cam-
                 era-->
00:48  20        *(0.63)                             3  👆
           hand  *scrolls-->
00:49  21  MK    und mal kurz die kommentare der-    4  Wieviel Meter
                 jenigen lesen-                         sind die von-
                 and briefly read the comments of       einander ent-
                 those                                  fernt?
00:52  22        (0.87)                                 How many me-
                                                        ters apart
                                                        are they from
                                                        each other?
00:53  23  MK    die hier ZUgeschaltet sind.         5  Gerade weiter
                 who are tuned in here                   dann nach
                                                        rechts
                                                        Straight
                                                        ahead then to
                                                        the right
```

```
00:55  24          (1.86)                        6  Nächste
                                                    Kreuzung
                                                    rechts
                                                    Next crossing
                                                    right

                                                 7  Pass auf dich
                                                    auf!
                                                    Take care of
                                                    yourself!
00:57  25  MK      ((hupen)) wie viele meter
                       +sind die voneinander
                   +entfernt?
                   ((honking)) how many meters apart
                   are they from each other?
           gaze    -->+looks to the side-----------
                   +looks underneath the camera-->
00:59  26  MK      ((schnalzt)) das WEISS ich nicht*   8  Sind Sie da
                       +weil ich die eine gruppe (--)     wirklich al-
                   noch +nich geSEHN hab (-) sondern      leine?
                   nur die andere.                        Are you re-
                                                          ally alone
                   ((clicks)) i don't know yet be-        there?
                   cause i haven't seen one of the     9  Am 1. Sep wa-
                   groups only the other                  ren in Chem-
                                                          nitz auch
                                                          Reiter-staf-
                                                          feln im Ein-
                                                          satz
                                                          On Sep 1st in
                                                          Chemnitz
                                                          there were
                                                          also mounted
                                                          police on
                                                          duty
           gaze    -->+looks straight ahead---------
                   -----+looks underneath the  cam-
                   era-->
           hand                                -->*
01:05  27          *#(3.88) ((motorgeräusche))*
                   (3.88) ((engine sounds))

           hand    *scrolls-------------------*
```


#fig.18

```
01:09  28  MK      grade weiter dann nach rechts
                   sacht jemand-
                   straight ahead then to the right
                   says someone
01:11  29  MK      das is nett-+
                   this is nice
           gaze              -->+
01:12  30  MK      +VIElen dank für die orientie-
                   rung.
                   many thanks for the orientation
           gaze    +looks straight ahead-->>
```

In the subsequent interaction phase (transcript 5) MK continues to walk along the pavement and a longer *bridging* phase occurs (19-26). He seconds his body movement again by a prospective plan of action in line 19.[23] In lines 21 to 23 MK verbalises his reception of the chat, during which he continuously looks at the screen and scrolls with his finger. This is a form of reception-accompanying talk, which, in addition to bridging, has the function of indicating the streamer's reception actions and contextualising recurrent pauses (lines 20, 22, 24), which emphasise the streamer's cognitive focus on the chat postings. This seemingly secondary action also involves the viewers, insofar as the postings[24] of the spectators (*die hier zugeschaltet sind*, 'who are tuned in here', line 23) are appreciated by the verbalised and displayed reception of the main actor. In lines 25-26 MK produces a *read aloud and respond* regarding chat post 4 (transcript 5), which initiates a footing shift: the streamer reads out the interrogative chat posting (line 25) and responds to it (line 26). The adjacent question-and-answer sequence (cf. Schegloff/Sacks 1973) substantiates the interactionality of the live video stream. By transforming the written chat posting into the audio-visual semiotic mode of the livestream video, the streamer makes it salient and thus relevant for the recipient.[25] The shift in modality entails a footing shift in speaker roles, with the streamer acting as the animator of the contribution authored by the chatter.

In line 26 MK now displays his epistemic status again. As he states, he could not answer the question about the distance between the two groups, between which MK is moving, because he had not 'seen one of the groups' (line 26) yet. Here, too, MK refers to the joint project, which presupposes that the change of location is necessary for the co-present livestreaming from the target location. The present-related as well as future-related character of the progression is indicated by the temporal negation *noch nich* ('not yet', line 26). However, MK's first choice did not meet the recruitment requirements as it did not sufficiently contribute to solving the orientation problem. After a longer phase of reception (line 27, fig. 18), a further *RAR* pair sequence occurs in lines 28-30. Chat posting 5 *grade weiter dann nach rechts* ('straight ahead then to the right'), which is read out in line 28, provides the desired spatial orientation. Finally, in lines 29-30 a positive evaluation and thanks for the contribution are given as an acknowledgement of the recruitees by the recruiter (cf. Rossi et al. 2020). The sequence closing is indicated prosodically by a deep falling pitch movement and the streamer's gaze is directed back from the smartphone to the direction in which he is walking (line 30). The sequence ends here with the positive recognition of the resolution of the orientation problem (*vielen dank für die orientierung*, 'many thanks for the orientation', line 30). MK uses the term *orientierung* to describe the state that has been sequentially achieved by the previous phases of situating, orienting, recruiting, and collecting solutions to the orientation

---

[23] Assumed here is a futuristic reading derived from the subjunctive II equivalent of the construction *werd(en) + infinitive* (cf. Fabricius-Hansen 2000), although in the situation the anticipated action has already begun at the time of utterance, as can be concluded from the movement visible in the livestream.

[24] The chat postings in transcript 5 are given in the right-hand column. The chat messages are aligned with the streamer's turns based on the documented video of the stream (cf. section 4.2.). Chat postings 1-2, which occurred before the section transcribed here, are not shown in the other transcripts as they are not made relevant by the participants.

[25] However, the term 'topicalising' cannot be applied here as the streamer has already set the topic by means of the previous recruitment.

problem. By means of the positive recognition (*vielen dank*, line 30) the threat expressed in line 17 (*sonst lauf ich nämlich in die falsche richtung,* 'otherwise i'm walking in the wrong direction'), which has bracketed the problem of orientation and its interactive resolution, dissolves and the situational involvement of the viewers is acknowledged.

6. Conclusions

In live video streams the participants are in an asymmetrical relationship to each other. This asymmetry is both medially arranged and interactively negotiated. The participation constellation is dynamic and is orchestrated mainly by the streamers, who initiate footing shifts multimodally in the multi-party interaction of the live video stream. 'Multimodal' here refers not only to the physical action of the streamers in the local space, but to the entire media practice of live video streaming: streamers, in their role as hosts, delegate participation statuses above all by switching between semiotic modes, made possible by the mediality of the medium 'live video stream'. Footing shifts are, as has been shown, initiated by shifting modalities in that the establishment of relevance of chat postings is followed by interactive extensions in the live video stream. Medial and modal asymmetries are thus closely related to the social asymmetries that are processed medially and practically in live video streams. Livestream hosting is therefore a reflexive media practice that runs continuously along with livestream events. The choice of media and semiotic modes, the setting of topics, the initiation of turn-taking, and the reading out of chat postings – all this is done first and foremost by the hosting streamers, promotes the situational involvement of the viewers, and reveals the streamers as interaction-dominant in the asymmetrical coordination of participation in live video streams.

Mobile live video streams are divided into episodes. Mobility is particularly evident when changes of location are processed and coordinated in live video streams as jointly accomplished projects. It has been shown that sequences of changing places are carried out through practices of situating, bridging, orienting, recruiting, grounding, dissolving, and acknowledging. While situating, the participants 'make-present' themselves in the temporal, spatial, and personal constellation through linguistic, gestural, and bodily pointing and locating. Occasionally, objects are also included, by means of which the situation is anchored in an evident and authenticating manner. Phases of bodily movement, on the other hand, are bridged by involvement practices. Here too, the streamers take the leading role. The constitution of an orientation problem which endangers the joint project, and the related recruitment of the audience, generates both offers of interaction and pressure to participate. The problem of orientation presents itself as an asymmetry of knowledge brought about by epistemic stancetaking, which seems to oppose the status asymmetry between host and audience. But it is the involvement of the chatters in the situation described that leads to the grounding of the status of situational knowledge. With the dissolution, the threat to the joint project is also resolved and the host finally acknowledges the situational involvement of the audience.

The analysis shows that participants in live video streams change between different interaction-constitutive alignments by means of footing shifts (cf. Goffman 1979; Choe 2020). The mediality of the live video stream with the portable camera, the display, and a chat function provides the basis for a multimodal communication

situation in which mutual interaction can potentially be realised. In the media practice of mobile live video streaming sociality is negotiated dynamically, jointly, and mutually, but also highly asymmetrically.

7. References

Arminen, Ilkka / Licoppe, Christian / Spagnolli, Anna (2016): Respecifying Mediated Interaction. In: Research on Language & Social Interaction 49 (4), 290-309, https://doi.org/10.1080/08351813.2016.1234614.

Bateman, John / Wildfeuer, Janina / Hiippala, Tuomo (2017): Multimodality – Foundations, Research and Analysis. A Problem-Oriented Introduction. Berlin: Mouton de Gruyter.

Beißwenger, Michael (2007): Sprachhandlungskoordination in der Chat-Kommunikation. Berlin: De Gruyter.

Beißwenger, Michael (2020): Internetbasierte Kommunikation als Textformen-basierte Interaktion: ein neuer Vorschlag zu einem alten Problem. In: Marx, Konstanze / Lobin, Henning / Schmidt, Axel (eds.): Deutsch in Sozialen Medien. Berlin: De Gruyter, 291-318,
https://doi.org/10.1515/9783110679885-015.

Bergmann, Jörg (1981): Ethnomethodologische Konversationsanalyse. In: Schröder, Peter / Steger, Hugo (eds.): Dialogforschung. Jahrbuch 1980 des Instituts für deutsche Sprache. Düsseldorf: Pädagogischer Verlag Schwann, 9-51.

Bergmann, Jörg et al. (2019): Garfinkel folgen, heißt, die Soziologie vom Kopf auf die Füße zu stellen. Jörg Bergmann & Christian Meyer im Gespräch mit René Salomon & Hannes Krämer. In: Forum Qualitative Sozialforschung / Forum Qualitative Social Research 20 (2), 1-55,
https://doi.org/10.17169/fqs-20.2.3289.

Blommaert, Jan (2019): Formatting online actions: #justsaying on Twitter. In: International Journal of Multilingualism 16 (2), 112-126,
https://doi.org/10.1080/14790718.2019.1575832.

Bou-Franch, Patricia / Lorenzo-Dus, Nuria / Blitvich, Pilar (2012): Social Interaction in YouTube Text-Based Polylogues: A Study of Coherence. In: Journal of Computer-Mediated Communication 17 (4), 501-521,
https://doi.org/10.1111/j.1083-6101.2012.01579.x.

Broth, Mathias (2014): Pans, Tilts, and Zooms: Conventional Camera Gestures in TV Production. In: Broth, Mathias / Laurier, Eric / Mondada, Lorenza (eds.): Studies of Video Practices. Video at Work. New York: Routledge, 63-96.

Broth, Mathias / Laurier, Eric / Mondada, Lorenza (eds.) (2014): Studies of Video Practices. Video at Work. New York: Routledge.

Bühler, Karl (1999 [1934]): Sprachtheorie. Die Darstellungsfunktion der Sprache. Ungekürzter Neudr. der Ausg. Jena, Fischer, 1934, 3. Aufl, Erstausgabe 1934. Stuttgart: Lucius und Lucius.

Choe, Hanwool (2019): Eating together multimodally: Collaborative eating in mukbang a Korean livestream of eating. In: Language in Society 48 (2), 171-208, https://doi.org/10.1017/S0047404518001355.

Choe, Hanwool (2020): Talking the cat: Footing lamination in a Korean livestream of cats mukbang. In: Journal of Pragmatics 160, 60-79,
https://doi.org/10.1016/j.pragma.2020.02.009.

Clark, Herbert (1996): Using Language. Cambridge, New York: Cambridge University Press.

Dang-Anh, Mark (2019): Protest twittern. Eine medienlinguistische Untersuchung von Straßenprotesten. Bielefeld: Transcript, https://doi.org/10.14361/9783839448366.

Dang-Anh, Mark et al. (2017): Medienpraktiken. Situieren, erforschen, reflektieren. Eine Einleitung. In: Navigationen. Zeitschrift für Medien- und Kulturwissenschaften 17 (1), 7-36, https://doi.org/10.25969/mediarep/1702.

Deppermann, Arnulf (2001): Gesprächsanalyse als explikative Konstruktion – Ein Plädoyer für eine reflexive ethnomethodologische Konversationsanalyse. In: Iványi, Zsuzsanna/Kertész, András (eds.): Gesprächsforschung. Tendenzen und Perspektiven. Frankfurt am Main, New York: P. Lang, 43-73.

Deppermann, Arnulf (2015): Pragmatik revisited. In: Eichinger, Ludwig (eds.): Sprachwissenschaft im Fokus. Positionsbestimmungen und Perspektiven. Berlin et al.: De Gruyter, 323-352.

Drew, Paul / Couper-Kuhlen, Elizabeth (2014): Requesting – from speech act to recruitment. In: Drew, Paul / Couper-Kuhlen, Elizabeth (eds.): Requesting in social interaction. Amsterdam, Philadelphia: John Benjamins Publishing Company, 1-34.

Dynel, Marta (2014): Participation framework underlying YouTube interaction. In: Journal of Pragmatics 73, 37-52, https://doi.org/10.1016/j.pragma.2014.04.001.

Ehlich, Konrad (1984): Zum Textbegriff. In: Rothkegel, Annely / Sandig, Barbara (eds.): Text – Textsorten – Semantik. Hamburg: Buske, 9-25.

Fabricius-Hansen, Cathrine (2000): Die Geheimnisse der deutschen würde-Konstruktion. In: Thieroff, Rolf et al. (eds.): Deutsche Grammatik in Theorie und Praxis. Berlin, Boston: De Gruyter, https://doi.org/10.1515/9783110933932.83.

Frobenius, Maximiliane (2011): Beginning a monologue: The opening sequence of video blogs. In: Journal of Pragmatics 43 (3), 814-827, https://doi.org/10.1016/j.pragma.2010.09.018.

Garcia, Angela / Jacobs, Jennifer (1999): The Eyes of the Beholder: Understanding the Turn-Taking System in Quasi-Synchronous Computer-Mediated Communication. In: Research on Language & Social Interaction 32 (4), 337-367, https://doi.org/10.1207/S15327973rls3204_2.

Gerhardt, Cornelia / Eisenlauer, Volker / Frobenius, Maximiliane (2014): Participation framework revisited: (New) media and their audiences/users. In: Journal of Pragmatics 72, 1-4, https://doi.org/10.1016/j.pragma.2014.08.011.

Gibson, David (2003): Participation Shifts: Order and Differentiation in Group Conversation. In: Social Forces 81 (4), 1335-1380, https://doi.org/10.1353/sof.2003.0055.

Giles, David et al. (2015): Microanalysis of Online Data. The methodological development of "digital CA". In: Discourse, Context & Media 7, 45-51, https://doi.org/10.1016/j.dcm.2014.12.002.

Goffman, Erving (1979): Footing. In: Semiotica 25 (1-2), 1-30, https://doi.org/10.1515/semi.1979.25.1-2.1.

Goffman, Erving (1981): Forms of Talk. Philadelphia: University of Pennsylvania Press (University of Pennsylvania publications in conduct and communication).

Goffman, Erving (1983): The Interaction Order. American Sociological Association, 1982 Presidential Address. In: American Sociological Review 48 (1), 1-17.

Goodwin, Charles / Goodwin, Marjorie (2004): Participation. In: Duranti, Alessandro (eds.): A companion to linguistic anthropology. Malden, Mass.: Blackwell Pub, 222-243.

Haddington, Pentti / Mondada, Lorenza / Nevile, Maurice (2013): Being Mobile: Interaction on the Move. In: Haddington, Pentti / Mondada, Lorenza / Nevile, Maurice (eds.): Interaction and mobility. Language and the body in motion. Berlin: De Gruyter, 3-64.

Hausendorf, Heiko (2013): On the interactive achievement of space – and its possible meanings. In: Auer, Peter et al. (eds.): Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives. Berlin: De Gruyter, 276-303.

Heritage, John (2012): The Epistemic Engine: Sequence Organization and Territories of Knowledge. In: Research on Language & Social Interaction 45 (1), 30-52,
https://doi.org/10.1080/08351813.2012.646685.

Hirschauer, Stefan (2014): Intersituativität. Teleinteraktionen und Koaktivitäten jenseits von Mikro und Makro. In: Heintz, Bettina / Tyrell, Hartmann (eds.): Interaktion – Organisation – Gesellschaft revisited. Anwendungen, Erweiterungen, Alternativen. Stuttgart: Lucius & Lucius (Zeitschrift für Soziologie, Sonderheft), 109-133.

Hutchby, Ian (2014): Communicative affordances and participation frameworks in mediated interaction. In: Journal of Pragmatics 72, 86-89,
https://doi.org/10.1016/j.pragma.2014.08.012.

Hutchby, Ian / Wooffitt, Robin (1998): Conversation Analysis. Principles, Practices and Applications. Cambridge: Polity Press.

Irvine, Judith (1996): Shadow Conversations: The Indeterminacy of Participant Roles. In: Silverstein, Michael / Urban, Greg (eds.): Natural histories of discourse. Chicago: Univ. of Chicago Press, 131-159.

Kendrick, Kobin / Drew, Paul (2016): Recruitment: Offers, Requests, and the Organization of Assistance in Interaction. In: Research on Language & Social Interaction 49 (1), 1-19, https://doi.org/10.1080/08351813.2016.1126436.

Knorr Cetina, Karin (2009): The Synthetic Situation: Interactionism for a Global World. In: Symbolic Interaction 32 (1), 61-87,
https://doi.org/10.1525/si.2009.32.1.61.

König, Katharina / Oloff, Florence (eds.) (2019): Special Issue: Mobile Medienpraktiken im Spannungsfeld von Öffentlichkeit, Privatheit und Anonymität (Journal für Medienlinguistik – Journal for Media Linguistics 2 (2)),
https://doi.org/10.21248/jfml.2019.9.

Levinson, Stephen (1988): Putting linguistics on a proper footing: Explorations in Goffman's participation framework. In: Drew, Paul / Wootton, Anthony (eds.): Goffman: Exploring the interaction order. Oxford: Polity Press, 161-227.

Licoppe, Christian / Morel, Julien (2012): Video-in-Interaction. "Talking Heads" and the Multimodal Organization of Mobile and Skype Video Calls. In: Research on Language & Social Interaction 45 (4), 399-429,

https://doi.org/10.1080/08351813.2012.724996.

Licoppe, Christian / Morel, Julien (2018): Visuality, text and talk, and the systematic organization of interaction in Periscope live video streams. In: Discourse Studies 20 (5), 637-665,
https://doi.org/10.1177/1461445618760606.

Licoppe, Christian / Tuncer, Sylvaine (2019): The initiation of showing sequences in video-mediated communication. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion (20), 545-571.

Luginbühl, Martin / Meer, Dorothee (eds.) (forthcoming): Parainteraktion in den Medien – Multimodale Perspektiven in TV, Radio und Hypermedien. Frankfurt am Main: Peter Lang.

Luginbühl, Martin / Schneider, Jan (2020): Medial Shaping from the Outset: On the Mediality of the Second Presidential Debate, 2016. In: Journal für Medienlinguistik – Journal for Media Linguistics 3 (1), 57-93,
https://doi.org/10.21248/jfml.2020.34.

Luhmann, Niklas (1984): Soziale Systeme. Grundriss einer allgemeinen Theorie. Frankfurt am Main: Suhrkamp.

Lünenborg, Margreth (2017): Von Mediengattungen zu kontingenten Hybriden: Konstruktivistische und performativitätstheoretische Perspektiven für die Journalistik. In: Medien & Kommunikationswissenschaft 65 (2), 367-384,
https://doi.org/10.5771/1615-634X-2017-2-367.

Meredith, Joanne (2017): Analysing technological affordances of online interactions using conversation analysis. In: Journal of Pragmatics 115, 42-55,
https://doi.org/10.1016/j.pragma.2017.03.001.

Mondada, Lorenza (2018): Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. In: Research on Language & Social Interaction 51 (1), 85-106,
https://doi.org/10.1080/08351813.2018.1413878.

Mondada, Lorenza (2019): Conventions for multimodal transcription.
https://www.lorenzamondada.net/multimodal-transcription.

Nevile, Maurice et al. (2014): On the interactional ecology of objects. In: Nevile, Maurice et al. (eds.): Interacting with Objects. Amsterdam: Benjamins, 3-26.

Paßmann, Johannes (2018): Die soziale Logik des Likes. Eine Twitter-Ethnografie. Frankfurt am Main, New York: Campus.

Recktenwald, Daniel (2017): Toward a transcription and analysis of live streaming on Twitch. In: Journal of Pragmatics 115, 68-81,
https://doi.org/10.1016/j.pragma.2017.01.013.

Reeves, Stuart/Greiffenhagen, Christian/Flintham, Martin/Benford, Steve/Adams, Matt/Row Farr, Ju/Tandavantij, Nicholas (2015): I'd Hide You. Performing Live Broadcasting in Public. In: Begole, Bo/Kim, Jinwoo/Inkpen, Kori/Woo, Woontack (eds.): Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. New York: ACM, 2573-2582,
https://doi.org/10.1145/2702123.2702257.

Rein, Katharina / Venturini, Tommaso (2018): Ploughing digital landscapes: How Facebook influences the evolution of live video streaming. In: New Media & Society 20 (9), 3359-3380,
https://doi.org/10.1177/1461444817748954.

Rossi, Giovanni / Floyd, Simeon / Enfield, N. J. (2020): Recruitments and pragmatic typology. In: Floyd, Simeon / Rossi, Giovanni / Enfield, N. J. (eds.): Getting others to do things. A pragmatic typology of recruitments. Berlin: Language Science Press (Studies in Diversity Linguistics, 31), 1-23, https://doi.org/10.5281/zenodo.4018370.

Schegloff, Emanuel (2007): Sequence organization in interaction. A primer in conversation analysis. Cambridge: Cambridge University Press.

Schegloff, Emanuel / Sacks, Harvey (1973): Opening up Closings. In: Semiotica 8 (4), 289-327,
https://doi.org/10.1515/semi.1973.8.4.289.

Schmidt, Axel / Marx, Konstanze / Neise, Isabell (2020): Produktion – Produkt – Rezeption? Medienketten in audiovisuellen Webformaten am Beispiel von Let's Plays. In: Marx, Konstanze / Lobin, Henning / Schmidt, Axel (eds.): Deutsch in Sozialen Medien. Berlin: De Gruyter, 265-288,
https://doi.org/10.1515/9783110679885-014.

Schneider, Jan (2017): Medien als Verfahren der Zeichenprozessierung: Grundsätzliche Überlegungen zum Medienbegriff und ihre Relevanz für die Gesprächsforschung. In: Gesprächsforschung 18, 34-55.

Selting, Margret et al. (2011): A system for transcribing talk-in-interaction: GAT 2. Translated and adapted for English by Elizabeth Couper-Kuhlen and Dagmar Barth-Weingarten In: Gesprächsforschung 12, 1-51.

Sjöblom, Max / Törhönen, Maria / Hamari, Juho / Macey, Joseph (2019): The ingredients of Twitch streaming: Affordances of game streams. In: Computers in Human Behavior 92, 20-28,
https://doi.org/10.1016/j.chb.2018.10.012.

Tannen, Deborah (2007): Talking voices. Cambridge: Cambridge University Press, https://doi.org/10.1017/CBO9780511618987.

Ten Have, Paul (2004): Understanding qualitative research and ethnomethodology. London: SAGE.

Zinken, Jörg / Rossi, Giovanni (2016): Assistance and Other Forms of Cooperative Engagement. In: Research on Language & Social Interaction 49 (1), 20-26, https://doi.org/10.1080/08351813.2016.1126439.

8. Appendix

Transcription conventions (cf. section 4.2.)

Talk is transcribed based on GAT 2 (Selting et al. 2011), movement and gaze are transcribed based on Mondada's conventions for multimodal transcription (2019), and camerawork is transcribed in accordance with Broth (2014).

+ + * * % % α α	Descriptions of actions other than talk are delimited between two identical symbols (one symbol per action) that are synchronised with corresponding stretches of talk or time indications.
§ §	Delimit descriptions of camerawork.
-->	The action described continues across subsequent lines until the same symbol is reached.
-->+	
-->>	The action described continues after the end of the excerpt.
----	Action is maintained.
fig.	The exact moment at which a screen shot was taken is indicated with a sign (#) showing its position within the turn/a time measure.
#	
‖	A vertical double line separates chat postings by the viewers from talk and embodied actions by the streamer. Chat posts are aligned with the documented video stream, which does not necessarily correspond to their temporal alignment at the time of the stream.

Dr. Mark Dang-Anh
Leibniz Institute for the German Language
Department of Lexical Studies
R 5, 6-13
68161 Mannheim
Germany

dang@ids-mannheim.de