

POSTPRINT

Mark Dang-Anh

Beyond parainteraction: Situatives Involvement und Footing Shifts in mobilen Livevideostreams auf *Periscope*

Mobiles Livevideostreaming ist eine Medienpraktik, bei der sich die Beteiligten in einer spezifischen Ausrichtung zueinander befinden und in der Streamer*innen und Zuschauer*innen unterschiedliche semiotische Ressourcen zur Verfügung stehen. Anhand der multimodalen Sequenzanalyse einer prägnanten Episode eines Ortswechsels im Rahmen der Berichterstattung eines Journalisten von einem politischen Ereignis auf der Livevideostreaming-Plattform *Periscope* wird die Frage bearbeitet, wie Beteiligung und *involvement* in Livevideostreams hergestellt sowie organisiert werden und dargelegt, inwiefern mobiles Livevideostreaming soziale Parainteraktion transzendiert. Es wird gezeigt, dass die Hosts der Medienpraktik ‚Livevideostreaming‘ interaktionsdominierend agieren und die Zuschauer*innen durch asymmetrische Partizipationskoordination per *footing shifts* situativ in das Geschehen involvieren.

1. Einleitung¹

Mobiles Livevideostreaming mit Smartphones ist eine alltägliche Medienpraktik. In den letzten Jahren findet auch im Journalismus das Livestreaming von Ereignissen immer mehr Verbreitung (vgl. Lünenborg 2017). Ereignisse werden per Livestreaming von mobilen Geräten wie Smartphones

- 1 Grundlage dieses Artikels ist ein Vortrag auf dem Workshop „Parainteraktion in den Medien. Multimodale Perspektiven in TV, Radio und Hypermedien“, der im Oktober 2019 an der Universität Basel stattfand. Ich danke Dorothee Meer und Martin Luginbühl für die Organisation und die Einladung zum Workshop, allen Workshopteilnehmenden und Isabell Neise für die Diskussion der Daten sowie der Deutschen Forschungsgemeinschaft (DFG) für die Förderung. Dieser Beitrag ist eine modifizierte und übersetzte Fassung von Dang-Anh (2021).

zeitsynchron dokumentiert und rezeptiv verfügbar gemacht. Charakteristisch für Livevideostreams ist ihr spezifisches *mediales* und *situatives* Arrangement: Während die Beteiligten in Livevideostreams² zeitlich quasi-synchron agieren, ist die Kommunikationssituation räumlich ‚zerdehnt‘ (vgl. Ehlich 1984), da Streamende und Zuschauende zum Zeitpunkt des Streams in der Regel physisch voneinander getrennt sind (vgl. Reeves/Greiffenhagen/Flintham et al. 2015). In einem Livevideostream befinden sich Streamer*innen und Zuschauer*innen demnach in einer asymmetrischen Relation: Während die Streamenden ihrerseits multimodal audiovisuell mit den Zuschauenden interagieren können, sind Zuschauende via Chat nur teilweise aktiv in das situative und kommunikative Geschehen, das sich auf ihren Screens abbildet, einbezogen (vgl. Licoppe/Morel 2018). Diese asymmetrische Medialität von Livevideostreams legt zunächst nahe, dass wir es bei der Kommunikation auf einer Livevideostreaming-Plattform mit integrierter Chatfunktion wie *Periscope*³ mit sozialer Parainteraktion zu tun haben.⁴ Das Ziel des Beitrags ist es erstens, vor dem Hintergrund des beschriebenen medialen und situationalen Arrangements zu erörtern, inwieweit es sich bei Livevideostreaming um soziale Parainteraktion handelt. Unter sozialer Parainteraktion verstehe ich im Anschluss an Horton und Wohl (1956) sowie Horton und Strauss (1957) technisch-medial vermittelte „Pseudo-Interaktion“ (Luginbühl 2019: 127), „scheinbare[] Interaktion“ (Holly/Baldauf 2001: 59) oder ein „simulacrum of conversational give and take“ (Horton/Wohl 1956: 215). Zweitens zielt die Analyse darauf ab, zur grundlegenden Diskussion um die wechselseitigen Relationen zwischen Medialität und Sozialität beizutragen, die vor allem in der praxeologisch orientierten Medienlinguistik geführt wird (vgl. Luginbühl/Schneider 2020).

- 2 Der Ausdruck ‚Livestream‘ kann sich auch auf Medienformate anderer Modalitätenkonstellationen beziehen, z. B. auf Audioübertragungen oder auf Textübertragungen (Live-Ticker). Im Folgenden werden jedoch ‚Livevideostream‘ und ‚Livestream‘ synonym verwendet.
- 3 *Periscope* war zwischen 2015 und 2021 eine unter der URL *periscope.tv* gehostete Plattform für mobile Livevideostreams, die in die Microblogging-Plattform *Twitter* integriert war (vgl. Kap. 3).
- 4 Horton und Wohl (1956) sprechen noch von „para-social interaction“. Ich verwende stattdessen im Anschluss an Moores (2005) den Ausdruck *soziale Parainteraktion* (bzw. vereinfacht: *Parainteraktion*), „because all human relationships, including those technologically mediated by broadcasting, are best thought of as fully ‘social’“ (75).

Ausgehend von der Annahme, dass die Ausrichtung der Beteiligten – das *Footing*, wie Goffman (1979) es nennt – sowohl medial als auch interaktional hervorgebracht wird, wird aus medien- und interaktionslinguistischer Perspektive gefragt, wie Beteiligung und Involvement in Livevideostreams durch die Teilnehmenden hergestellt und organisiert werden. Zunächst werden hierzu einleitend die theoretischen Grundlagen für die Medieninteraktionsanalyse (Kap. 2.1) und das Verständnis von Beteiligung und Involvement (Kap. 2.2) gelegt. Im Anschluss erfolgt ein Überblick über den Forschungsstand sowie die Darstellung zweier Charakteristika von Livevideostreams, ihre zerrissene Temporalität (Kap. 3.1) und die asymmetrische Relevanzsetzung (Kap. 3.2), die sich in ihnen vollzieht. Der Analyseteil gliedert sich in kurze Einführungen zu den Transkripten (Kap. 4.1) und zum Kontext der untersuchten Daten (Kap. 4.2), den analytischen Einstieg in die Einführungssequenz des untersuchten Livestreams (Kap. 4.3) und die darauf folgende Analyse einer Sequenz in drei Teilabschnitten (Kap. 4.4–4.6). Der Beitrag schließt mit einem Fazit (Kap. 5).

2. Theoretische Grundlagen

2.1 Livevideostreaming als kommunikative Medienpraktik

Livevideostreaming ist ein durch und durch soziales Unterfangen: „The 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, Herv. d. Verf.). Wie für viele andere Medienpraktiken lässt sich auch für das Livevideostreaming konstatieren: Die Sozialität einer Medienpraktik ist sowohl durch seine Medialität als auch durch seine Semiotizität bedingt. Das Verständnis von Sozialität leitet sich dabei aus sozialen Situationen ab, in der sich zwei Teilnehmende im Status gegenseitiger Wahrnehmung („Wahrnehmungswahrnehmung“ nach Luhmann 1984: 560; vgl. Hirschauer 2014: 112) und gemeinsamer Reaktionspräsenz (vgl. Goffman 1983: 2) befinden. Dadurch ist der hier angewandte Sozialitätsbegriff eng an den Begriff der Interaktion gebunden und es gilt zu fragen, *wie* Interaktion im Medium ‚Livevideostream‘ prozessiert wird. Somit ist zum einen die Medialität des Mediums für die interaktionale Konstitution des Sozialen

grundlegend. Unter Medialität eines Mediums verstehe ich mit Luginbühl und Schneider (2020) „the way it shapes the choice of signs and how we use them“ (59). Die Medialität von Medienpraktiken ist eng verknüpft mit den technischen Affordanzen (vgl. Hutchby 2014; Meredith 2017) eines medialen Arrangements, also dem Handlungspotential, das Medien zur vermittelten Interaktion nicht nur bereitstellen, sondern das auch im Rahmen von Medienpraktiken tatsächlich ausgeschöpft wird. Neben der Medialität ist zum anderen die Semiotizität (Zeichenhaftigkeit), die ihrerseits medial bedingt ist, für die interaktionale Konstitution von Sozialität entscheidend. Somit sind für die Analyse von Livevideostreams vor allem für die medialen „Verfahren der Zeichenprozessierung“ (Schneider 2017) unterschiedlicher Zeichenmodalitäten (*semiotic modes*; vgl. Bateman/Wildfeuer/Hiippala 2017: 112–123) relevant. Medienpraktiken sind wiederum als soziale Praktiken der Kommunikation zu betrachten. Mit ‚Medienpraktiken‘ lässt sich

„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/Pfeifer/Reisner et al. 2017: 12)

Medien, Praktiken und Zeichen stehen also in einer komplexen Relation zueinander, weshalb hinsichtlich des vorliegenden Gegenstands analytisch unterschieden werden muss zwischen

- dem *Medium* ‚Livevideostream‘, das als operatives, prozessierendes Arrangement von ineinander verschachtelten Apparaten und Infrastrukturen wie Smartphone, Screen, Tastatur, Internet etc. die Möglichkeitsbedingungen und Handlungspotentiale, innerhalb derer sich Medienpraktiken vollziehen, bereitstellt,
- der *kommunikativen Medienpraktik* ‚Livevideostreaming‘, im Zuge derer sich die interaktionalen Prozesse der Sozialitätskonstitution, die hier beschrieben sind, vollziehen und
- den sprachlichen und nicht-sprachlichen *Zeichen* unterschiedlicher Zeichenmodalität, die im Verlauf der multimodalen Medienpraktik ‚Livevideostreaming‘ interaktional und medial prozessiert werden.

Interaktionen sind die zwischenmenschlichen Momenta, in denen sich die Relationen von Medien, Praktiken und Zeichen synthetisch manifestieren.

Die hier angewendete Analyse des Livevideostreamings ist demnach eine semiotisch orientierte Prozessanalyse von Medieninteraktionen, die sich vor einem praktischen Hintergrund vollziehen. Als methodisches Paradigma dient daher die sequenziell vorgehende ethnomethodologische Konversationsanalyse, die in der medien- und interaktionslinguistischen Erforschung auch digitaler Kommunikation mittlerweile als etablierte Untersuchungsmethode von medial vermittelten „Vollzugswirklichkeit[en]“ (Bergmann 1981: 12)⁵ gelten kann (vgl. Arminen/Licoppe/Spagnolli 2016; Giles/Stommel/Paulus et al. 2015; die Beiträge in König/Oloff 2019; Meredith 2017; Meredith 2019), sofern sie den Medialitäten und Zeichenmodalitäten ebenso Rechnung trägt, wie den praktischen Hintergründen der untersuchten Verfertigungen (*accomplishments*).

2.2 Beteiligung und Involvement

Ermöglicht und bedingt durch die technischen Affordanzen von Livevideostreaming-Plattformen lassen sich in Livestreams unterschiedliche Beteiligungskonstellationen realisieren (vgl. Sjöblom/Törhönen/Hamari et al. 2019). Beteiligung (*participation*) wird hier interaktionstheoretisch perspektiviert und ist mit Goodwin und Goodwin unmittelbar mit *involvement* verbunden: „The term *participation* refers to actions demonstrating forms of involvement performed by parties within evolving structures of talk. [. . .] [T]he term is not being used to refer to more general membership in social groups or ritual activities“ (Goodwin/Goodwin 2004: 222). Unter ‚Involvement‘ sei hier ein aktives, beobachtbares, intelligibles und gemeinsam errungenes „achievement in conversational interaction“ (Tannen 2007: 27) zu verstehen. Im Gegensatz zum für die Linguistik zum klassischen Referenzmodell gewordenen dyadischen Sprecher-Hörer-Modell (vgl. hierzu kritisch Goffman 1979; Levinson 1988), lassen sich mit Goffman (1979) Mehrpersonenkonstellationen analysieren (vgl. Bou-Franch/Lorenzo-Dus/Blitvich 2012; Deppermann 2015), die vor allem in digitalen Medien den Regelfall darstellen dürften. Dynel charakterisiert die aus Mehrpersonenkonstellationen

5 Bergmanns Ausdrücke „Vollzug, Vollzugswirklichkeit etc.“ sind Übersetzungen von Garfinkels Konzept von „*accomplishment, accomplished objectivity, accomplished facticity* etc.“ in phänomenologischem Bezug auf Husserl, was, wie Bergmann beschreibt, Garfinkels Zusppruch fand (Bergmann/Meyer/Salomon et al. 2019: 34).

emergierenden Kommunikationssituationen als „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). Auch in Livevideostreams vollziehen sich *multi-party interactions* und es fragt sich vor allem, wie die Dynamik wechselnder Beteiligungskonstellationen interaktional und medial organisiert ist.

Goffman hat bereits vor der Popularisierung digitaler Kommunikation Beschreibungskategorien für die komplexen Beteiligungskonstellationen von Face-to-face-Interaktionen etabliert (vgl. Goffman 1979), die sich inzwischen auch für die Analyse digitaler Kommunikation durchgesetzt haben und erweitert wurden (vgl. etwa Dynel 2014; Gerhardt/Eisenlauer/Frobenius 2014; Hutchby 2014; Blommaert 2019).⁶ Er geht davon aus, dass die Art der Beteiligung an Interaktionen Gegenstand fortlaufender interaktionaler Aushandlung ist. Dabei differenziert er zwischen Sprecher- (*production format*) und Hörerrollen (*participation framework*). Sprecher*innen agieren entweder als *animator*, d. h. als die Instanz, die den Akt der Äußerung vollzieht, indem sie etwa redet oder postet. Weiterhin agieren Sprecher*innen als *author*, d. h. als die Instanz, die eine Äußerung formuliert oder als *principal*, d. h. als die Instanz, die für eine Äußerung sozial bzw. institutionell verantwortlich zeichnet, wie etwa eine Zeitungsredaktion, ein Ministerium, aber auch ein*e Politiker*in oder Redakteur*in. Neben diese drei Sprecherrollen stellt Goffman schließlich die Rolle der *figure*, einem erzählerischen Ich, das in Interaktionen als Figur dargestellt wird (vgl. Goffman 1979: 16–23). Sowohl Personen als auch Institution können die jeweiligen Beteiligungsrollen ausfüllen und oftmals fallen alle Rollen in einer Person überein (vgl. Gerhardt/Eisenlauer/Frobenius 2014).

Die Abgrenzungen zwischen *production format* und *participation framework* bilden die Grundlage für Goffmans Analyse von Wechseln (*shifts*) in der wechselseitigen Ausrichtung (*footing*) in der Interaktion (vgl. Goffman 1979: 19), welche er wie folgt beschreibt: „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“ (Goffman 1979: 5). Dass gerade diese vielzitierte Beschreibung von *footing*

6 Ich beziehe mich hier hauptsächlich auf Goffmans Aufsatz zum *Footing*, der erstmals 1979 erschienen ist und 1981 in seinem Buch *Forms of Talk* republiziert wurde (Goffman 1981).

shifts sich an einem prototypisierten und geradezu starren Gesprächsentswurf mit den Komponenten Produktion, Rezeption und Äußerung orientiert, ist der vorempirischen analytischen Trennung der Beschreibungskategorien geschuldet. Dies vermag nicht darüber hinwegzutäuschen, dass *footing* interaktional und somit hochdynamisch ausgehandelt wird. Es sind nämlich nicht die Beteiligungsstatus, die für die Analyse von Medienpraktiken zuvorderst relevant sind, sondern vielmehr deren „structural underpinnings“ (Goffman 1979: 5) sowie deren interaktionale Aushandlung. Irvine, die sich mit der Unbestimmbarkeit von Beteiligungsstatus in Interaktionen auseinandersetzt (1996), fasst diese Aushandlung als Bearbeitung eines *mapping problems* auf:

„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; d.V.]. 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.“ (Irvine 1996: 136)

Sich in der Interaktion aneinander auszurichten kann somit als ein lebhaftes Wechselspiel zwischen den Beteiligten und deren Produktions- und Rezeptionsrollen betrachtet werden. Der Begriff des ‚Beteiligungsstatus‘ geht dabei über die Frage, wer gerade spricht und wer zuhört, hinaus. ‚Footing‘ umfasst ebenso die Frage der sozialen Ausrichtung der Beteiligten zueinander, z. B. die interaktionale Aushandlung von Wissens- und Statusasymmetrien. Derartige Aushandlungen finden immer im Rahmen spezifischer Medienpraktiken, Medialitäten und Zeichenmodalitäten statt. In der Analyse (Kap. 4) wird daher hervorgehoben, inwieweit *footing shifts* mit der Medialität von Interaktionsmedien und deren Zeichenpotentialen verknüpft sind, durch die sie realisiert werden (können). Zunächst sei jedoch auf die vorliegende Forschung zu Livevideostreams eingegangen.

3. Livevideostreams

Seit ca. Mitte der 2000er Jahre gibt es Livestreaming-Plattformen. So startete 2005 *ComVu Pocket Caster*, gefolgt von *Ustream* (2006–2016; jetzt *IBM*

Watson Media)⁷ und *Bambuser* (2007; vgl. Rein/Venturini 2018: 3361). Später folgten mit *younow* (seit 2011), *Meerkat* (2015–2016) und *Periscope* (seit 2015) weitere Anbieter. Aus dem 2007 gestarteten und 2014 eingestellten *justin.tv* ging die Plattform *Twitch* hervor, auf der primär Livevideos aus dem Gaming-Bereich gestreamt werden und die mittlerweile die erfolgreichste Livestreaming-Plattform ist. Auch die großen Social Media-Plattformen wie *YouTube* und *Facebook* bieten mittlerweile eine sehr breit genutzte Livestreamfunktion an (vgl. Rein/Venturini 2018). *Twitter* hingegen kaufte *Periscope* noch vor seinem Launch im Jahr 2015 und integrierte die Livevideostreaming-Plattform in die eigene Microblog-Plattform, sodass die Streamrezeption direkt aus der *Twitter*-Oberfläche heraus möglich war. Zum 31. März 2021 wurde das Angebot von *Periscope* eingestellt.

Auch wenn Livestreaming als grundlegende Technik für Fernsehübertragungen oder Videokonferenzen betrachtet werden kann, sollen hier unter Livevideostreams über das Internet prozessierte Datenübertragungen (*stream*) von audiovisuellen Medienformaten (*video*) zum Zeitpunkt ihrer technischen Erfassung (*live*) verstanden werden. Dabei handelt es sich, im Gegensatz zu Videokonferenzen oder Videoanrufen, bezüglich des Videos um eine asymmetrische One-to-many-Kommunikation, bei der die Produzent*innen des Streams, die Streamenden, das live erstellte Video an mehrere Streamempfänger*innen übertragen. In Gegenrichtung wird kein Video übertragen, sodass alle Zuschauer die gleiche Sehfläche (vgl. Schmitz 2016), den Livestreamscreen mitsamt des Chatfensters, wahrnehmen (vgl. Abb. 1).

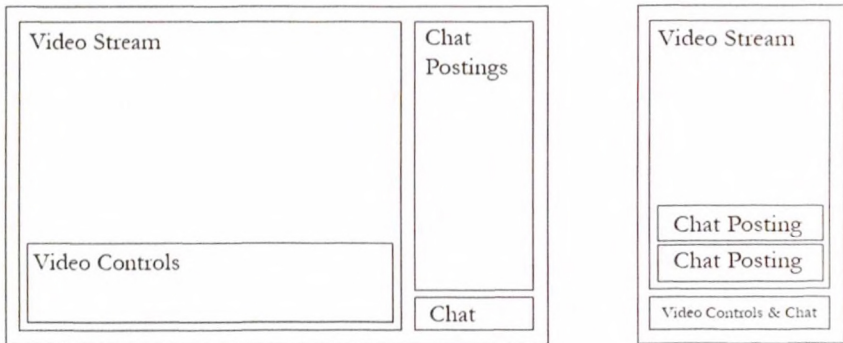


Abb. 1: Schematische Darstellung der *Periscope*-Schfläche auf stationärem Client (links) und Smartphone (rechts)

7 Vgl. [IBM: History of IBM Watson Media].

Livestreaming-Plattformen bieten nämlich in der Regel eine Chatfunktion, bei der die Videorezipient*innen untereinander und mit den Streamer*innen interagieren, so auch das hier untersuchte *Periscope*. Die Kategorisierung von Livevideostreams als One-to-many-Kommunikation lässt sich demnach nicht aufrecht erhalten (vgl. auch Dynel 2014: 38), vielmehr führt die Parallelität von Video und Chat zu einer höchst asymmetrischen Beteiligungsstruktur: „[W]hile 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“ (Licoppe/Morel 2018: 638). Reeves/Greiffenhagen/Flintham et al. (2015) zeigen anhand der Analyse eines Mixed-Reality-Spiels, bei dem sich die Spieler in einer Stadt finden müssen, während sie von Online-Spielern beobachtet werden, mit denen sie ebenfalls interagieren, dass das mediale Arrangement des Livevideostreams die Teilnehmer vor komplexe Herausforderungen stellt. Für die Spieler eröffnen sich unterschiedliche Spannungsverhältnisse zwischen ihren Körpern und der Kamera, den Anforderungen des Online-Publikums und der Situation auf der Straße, der Frontstage vor und Backstage hinter der Kamera sowie zwischen ihren Rollen als Sprecher und Performer.

3.1. Zerrissene Temporalität

Streamenden und Zuschauenden stehen im Livevideostream unterschiedliche semiotische Ressourcen zur Verfügung (audiovisuelles Video vs. textbasierter Chat), die wiederum disparaten zeitlichen Bedingungen unterliegen. Zwar werden, wie von Licoppe und Morel (2018) beschrieben, Chatbeiträge nur für kurze Zeit angezeigt und laufen, je nach Menge der aktiv Chattenden, mitunter relativ schnell durch das Chatfenster oder verblassen mit der Zeit (*fade out*). Dennoch lässt sich Flüchtigkeit („ephemeral“) vielmehr auf Seiten des Videos verbuchen, denn während die Chatbeiträge durch Streamer*in und Zuschauende immer noch ‚scrollbar‘ sind, ist das im Video Übertragene, oftmals durch die Perspektive auf den *talking head* der*des Streamenden (vgl. Licoppe/Morel 2012, 2018), insoweit vergänglich, als es während der Übertragung des Livevideostreams kontinuierlich visuell verfügbar, jedoch nicht wiederholbar, sondern flüchtig ist. Erst nach Beendigung des Livestreams

können Videos auf *Periscope* angeschaut und einzelne Stellen per Videosteuerung angesteuert werden.⁸

Überdies lässt sich bei Livevideostreams nur von quasi-synchroner Übertragung (vgl. Garcia/Jacobs 1999), nicht aber von Echtzeitkommunikation sprechen, da sowohl die Übertragung des Videosignals als auch die unterschiedlichen Übertragungen der Chatbeiträge den physikalischen und technischen Latenzen (*delay*; vgl. Schmidt/Marx/Neise 2020) von (mobilen) Internetübertragungen unterliegen. Aufgrund der größeren Datenmengen ist dabei von größeren Latenzen bei der Videoübertragung als bei der Chatübertragung auszugehen. Schmidt, Marx und Neise (2020) zeigen am Beispiel von Liveübertragungen des Videogenres *Let's Play*, bei dem das Spielen von Computerspielen live gestreamt wird, dass auf der Plattform *Twitch* der Videostream zum Zeitpunkt der Untersuchung eine Latenz von ca. 12 Sekunden aufwies. Anhand einer Sequenzanalyse legen sie dar, wie die Latenz die Teilnehmenden vor sequenzorganisatorische Herausforderungen stellt, etwa einen Chatbeitrag einer Spielaktion zuzuordnen. Überdies unterliegt die Produktion von Chatbeiträgen anderen Zeitlichkeitsbedingungen als Face-to-face-Interaktion, wie vor allem Beißwenger gezeigt hat (vgl. Beißwenger 2007, 2008). So erfordern Chatbeiträge gewisse Produktionszeiten, bevor sie versendet werden, die zudem mit Reparaturen vor ihrer Distribution verbunden sein können. Die Abfolge von Chatbeiträgen ist nicht immer eindeutig sequentiell strukturiert. Bezüge zwischen Postings sind mitunter verschachtelt und somit teils schwer zuzuordnen. Chats unterliegen also spezifischen Zeitlichkeitsbedingungen hinsichtlich der Produktion und Rezeption von Chatbeiträgen (vgl. Beißwenger 2020). Hingegen ist die im Livestream übertragende *Hic-et-nunc*-Situation der Streamenden durch ihre sequentielle Geordnetheit charakterisiert. Im Livestream sind die lokalen Situationen der Streamenden und der Chatenden, die jeweils anderen Zeitlichkeitsbedingungen unterliegen, medial auf einem Screen synthetisiert (vgl. Knorr Cetina 2009). Das in Livevideostreams synthetisierte asymmetrische Zeitlichkeitsverhältnis lässt sich als *zerrissene Temporalität* (*ruptured temporality*) beschreiben, mit der die Teilnehmenden interaktional umgehen (müssen).

8 Die technischen Beschreibungen der Livevideostreaming-Plattform *Periscope* beziehen sich auf den Stand vom Oktober 2020.

3.2 Asymmetrische Relevantsetzung

Gleichwohl werden auch Livevideostreams prinzipiell sequentiell durch die Streamenden und die Zuschauenden strukturiert, wobei die Streamenden als die Hauptdarsteller (*principal performer*) über eine sequenzsteuernde Interaktionsdominanz verfügen. Der hier untersuchte Livestream journalistischer Berichterstattung ist dadurch charakterisiert, dass der Streamer den Stream fortlaufend durch eigene Sequenzierung mittels „ritual brackets“ (Goffman 1979: 7) organisiert, etwa durch wiederholte Episodeneröffnungen (*openings*; vgl. Schegloff 2007; Frobenius 2011) und -schließungen (*closings*; vgl. Schegloff/Sacks 1973) und/oder Situations- bzw. Ortswechsel (vgl. Kap. 4). Livevideostreams sind entsprechend kontinuierliche Interaktionssituationen, die ihrerseits, vor allem bei längeren Streams, vornehmlich durch die interaktionsdominanten Streamer*innen episodisch in Interaktionssequenzen unterteilt sind. Dennoch muss festgehalten werden, dass die Beteiligten sich in den kontinuierlichen Interaktionssituationen, auch wenn eine Zeit lang nicht geredet wird, in einem fortwährenden „*state of talk*“ (Goffman 1979: 7) befinden. Interaktion kann jederzeit durch die ratifizierten Teilnehmenden im Videoscreen oder im Chat initiiert werden, wobei es vor allem der*dem streamenden Hauptdarsteller*in obliegt, inwieweit Dinge, Themen, Objekte, Akteure usw. zu Gegenständen der Interaktion gemacht werden, in die Beteiligte, zu gleichen oder unterschiedlichen Anteilen, involviert sind.

Die asymmetrische Beteiligungsstruktur bedingt spezifische, sie reproduzierende Relevantsetzungspraktiken in Livevideostreams.⁹ Hierbei ist vor allem *read aloud and response* (RAR; vgl. Licoppe/Morel 2018) als eine spezifische Praktik des *noticing* (vgl. Schegloff 2007: 217–220; vgl. zu *noticing* in Livestreams Recktenwald 2017; Schmidt/Marx/Neise 2020) zu nennen. Unter *noticing* versteht man die interaktionale Relevantsetzung von

9 Auch dieses Verhältnis zwischen asymmetrischer Beteiligungsstruktur und Relevantsetzungspraktiken im Livevideostreaming ist – typisch für Medienpraktiken – wechselseitig: In der Medialität des Mediums liegen bestimmte affordante Handlungspotentiale begründet, die in Praktiken derart genutzt und ausgestaltet werden, dass sie die Asymmetrie auf Ebene der Sozialität strukturell reproduzieren und somit verfestigen. Dass derartig strukturierende Praktiken dennoch dynamisch und variabel sind, zeigen Beispiele sich verändernder Medienpraktiken wie etwa das Liken auf Twitter, das sich seit der Einführung des Like-Buttons beträchtlich gewandelt hat (vgl. Paßmann 2018).

Merkmale (*features*), die zuvor in einer Interaktion als nicht relevant erachtet wurden und die durch das *noticing* zum Gegenstand des Gesprächs werden. Mit *RAR* bezeichnen Licoppe und Morel (2018) eine streamingtypische Praktik, bei der Streamende Chatbeiträge laut vorlesen und somit als relevant für die Interaktion setzen. Auf das Vorlesen folgt dann ein weiterer oder mehrere Turns, in denen der Streamende auf den vorgetragenen Chatbeitrag eingeht. Recktenwald beschreibt diese Praktik als *topicalising* (Recktenwald 2017: 10), was den Aspekt hervorhebt, dass Streamende unter vielen möglichen Chatbeiträgen einen bestimmten als themensetzend für die folgenden Turns auswählen. Die durch die Streamenden initiierte Selektion eines spezifischen Chatbeitrags stellt demnach eine Steuerung des Interaktionsverlaufs dar, der aufgrund des modalen Ressourcenvorteils (linear verlaufendes audiovisuelles Video vs. ‚Chunk-by-chunk‘-Sequenz des Chats) gegenüber den Chatten den Streamenden vorbehalten ist. Die um Aufmerksamkeit konkurrierenden Chatter müssen ihre Beiträge entsprechend verständlich, zielgerichtet und schnell gestalten, um sie zur selektierenden Rezitation bereitzustellen. Entsprechend verfügen die Streamenden über eine affordanzbedingte Interaktionsdominanz (vgl. Sjöblom/Törhönen/Hamari et al. 2019), die sie zur Sequenz- und Themensteuerung einsetzen.

4. Analyse

4.1 Transkription

Die hier verwendeten Transkripte orientieren sich am Basistranskript von GAT2 (Selting/Auer/Barth-Weingarten et al. 2009) sowie an den Konventionen für multimodale Transkription von Mondada (2019) und sind dem Gegenstand angepasst leicht modifiziert. Transkripte stellen immer eine zur Analyse hinreichende Selektion eines holistischen Geschehens dar, das niemals in seiner Gänze erfasst werden kann. Dennoch leiten sich die Schwerpunktsetzungen der multimodalen Analyse aus den Hervorbringungen der Beforschten ab, denn „[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 den vorliegenden Transkripten werden neben den sprachlichen

Äußerungen Fokussierungen auf die Parallelität von Videostream und Chat sowie auf die verkörperte Interaktion (*embodiment*; vgl. Mondada 2018) durch die körperliche Bewegung, den Blick und die Gesten des Streamers vorgenommen. Entsprechend wurde das Transkript um alignierte Notation von *move*, *gaze* und *hand* erweitert, die jeweils in einer eigenen Zeile vermerkt sind und alle dem Streamer zuzuordnen sind.

Dort, wo sie in der Interaktion durch die Teilnehmenden relevant gesetzt werden, sind die Chatbeiträge in einer separaten Spalte auf der rechten Seite angezeigt (vgl. Transkript 4). Durch das mehrspaltige Layout (vgl. Recktenwald 2017) werden die unterschiedlichen Zeichenmodalitäten ausgewiesen, in denen die Teilnehmenden interagieren. Zudem setzt sich das Problem der zerrissenen Temporalität zwischen Streamer*innen und Chatter*innen, das sich bereits während des Livestreamings stellt (vgl. 3.1), auch in der postsituativen Rezeption der Aufnahme des Livestreams fort: Zu unterschiedlichen Zugriffszeiten wurden bei der Wiedergabe des untersuchten Videos Chatbeiträge uneinheitlich wiedergegeben. Insofern bildet das Transkript keine zeitlich präzise Alignierung des Videos mit dem Chatfenster ab, was durch die vertikale Anordnung von Stream- und Chatbeiträgen und die gezackte Linie, die diese voneinander trennt, markiert wird (vgl. Transkript 4). Gleichwohl ist, so geht aus der Analyse hervor, die durch die Teilnehmenden hervorgebrachte Sequenzialität im Transkript hinreichend abgebildet.

4.2 Kontext & Daten

Der Ausschnitt, der in diesem Aufsatz analysiert wird, stammt aus einem Livevideostream des damaligen Reporters der deutschen überregionalen Tageszeitung *taz*, Martin Kaul (MK).¹⁰ Gesendet wurde der Livestream am 16. September 2018 von einer Versammlung in Köthen im deutschen Bundesland Sachsen-Anhalt, zu der rechtsgerichtete und rechtsextreme Organisationen anlässlich einer zuvor stattgefundenen tödlichen Auseinandersetzung

10 Seit er 2017 mit seinem Smartphone live vom G20-Gipfel in Hamburg berichtete, gilt Kaul als Vorreiter journalistischer Livevideostreaming-Berichterstattung in Deutschland. Zur Einstellung der Livestreaming-Plattform *Periscope* im März 2021 schrieb Kaul einen Abschiedstext, in dem er seine bisherige journalistische Live-Berichterstattung auf *Periscope* rekapituliert; vgl. [taz: Tschüss, Periscope!].

in der Kleinstadt aufgerufen hatten.¹¹ Bereits eine Woche zuvor hatte MK von einer ähnlichen Veranstaltung in Köthen gestreamt, dokumentierte dabei volksverhetzende Reden und wurde durch Versammlungsteilnehmende bedroht.¹² Der hier untersuchte Ausschnitt wurde auf *Periscope* am 16. September 2018 gegen 18:06 Uhr gestartet und trägt den Titel „Kaul aus #Köthen, 18.06 Uhr.“¹³ MK streamte am Tag des Versammlungsereignisses zwei Videos mit einem Smartphone an einem Stick, eines um 18:06 Uhr mit einer Länge von 1:16 Stunden und eines um 19:33 Uhr mit einer Länge von 1:17 Stunden.¹⁴

4.3 Hosting

Transkript 1: „Begrüßen“ (aus: [Periscope: Kaul aus #Köthen, 18.06 Uhr.], 00:00:00–00:00:36)

00:00 01 ((Sprechchöre, unverständlich, ca. 12 Sek.; Musik))

00:12 02 ((Sprechchöre)) a anti anticapitalista (.) a (.) anti #anticapitalista (.) a (.) anti (.) anti#capita[lista]



Abb. 2



Abb. 3

11 Vgl. [Tagesschau: Nach Todesfall in Köthen].

12 Vgl. [Deutsche Welle: Gewalt gegen Journalisten].

13 Vgl. MKs Ankündigung auf *Twitter* ([Twitter: Kaul aus #Köthen, 18.06 Uhr.]) sowie die Archivierung des Streams auf *Periscope* ([Periscope: Kaul aus #Köthen, 18.06 Uhr.]). Laut Auskunft von *Periscope* rezipierten ca. 7.000 Zuschauende das Video live zum Zeitpunkt der Ausstrahlung.

14 Vgl. [Twitter: Kaul in #Köthen, 19.33 Uhr.] und [Periscope: Kaul in #Köthen, 19.33 Uhr.].

00:22 03 MK [hallo guten TACH] herzlich
willkommen in der sächsen anhaltini-
schen kleinstadt (.) Köthen-
00:27 04 (0.42) ((Sprechchöre))
00:27 05 MK es is achtzehn uhr ähm SECHS mein name
is martin kaul,
00:30 06 (0.44) ((Sprechchöre))
00:31 07 MK das_n SERvice den die tages (-)
zeitung TAZ aus berLIN präsentiert-
00:34 08 (0.47) ((Sprechchöre))
00:35 09 MK ich erzähl #heute_n bisschen im
livestream (-) was hier LOS is.



Abb. 4

00:38 10 (2.0) ((Lautsprecherdurchsage))

Zu Beginn des Livestreams (vgl. Transkript 1) ist die Kamera auf die Straße gerichtet. Man sieht, wie auf der Straße offensichtlich Polizist*innen laufen, wobei die Kamera die uniformierten Beine bis zur Hüfthöhe filmt (vgl. Transkript 1, Abb. 2). Zu hören, aber nicht im Wortlaut zu verstehen, sind Sprechchöre und Musik. Das Setting wird bereits durch die ersten Sekunden, in denen der Streamer nicht im Bild zu sehen oder zu hören ist, aufgespannt. Straße, Polizei und Sprechchöre deuten auf Straßenproteste hin, die sich am Ort des gefilmten und gestreamten Geschehens ereignen. Die Kameraperspektive auf den Boden indiziert die Praktik, Polizeibedienstete nicht aus der Nähe ins Gesicht zu filmen, die MK im Verlaufe des Videos wiederholt vollzieht. Die Sprechchöre weisen den Protest als antikapitalistisch aus (Z. 02), wodurch eine politische Haltung ausgedrückt wird, die oftmals auch bei antifaschistisch orientierten Straßenprotesten geäußert wird.

MK steigt mit einem Eröffnungszug (*opening*; vgl. Schegloff 2007) in den Stream ein und die Kameraperspektive wechselt dabei von der rückseitigen Hauptkameraperspektive auf die der Frontkamera (vgl. Transkript 1, Abb. 3). Im Hintergrund ist im weiteren Verlauf die Vorderseite eines Demonstrationzuges mit Transparenten zu sehen, vor der MK, der im Bildvordergrund dargestellt wird, herläuft (vgl. Transkript 1, Abb. 4). Die ausgiebige Begrüßung, die nun eingenommene *talking head*-Perspektive (vgl. Transkript 1, fig. 2; vgl. Licoppe/Morel 2012) und die Tatsache, dass der Stream über seinen *Periscope*-Account gesendet wird, markieren MK als Gastgebernden (*host*) in dem Livevideostream. Bereits mit dieser Markierung werden die Ausrichtungen (*alignments*) der Beteiligten zueinander, *host* und *audience*, präfiguriert. Unter *hosting* sei die Zurichtung von medial vermittelter digitaler Interaktion zu verstehen, durch die der Host einseitig qua prospektiver Einrichtung sowie fortlaufender Administration des Livevideostreams über die medialen, multimodalen und semiotischen Rahmenbedingungen und Ressourcen einer Interaktionssituation verfügt. Somit ist *hosting* eine Medienpraktik, die digitale Interaktionssituationen hinsichtlich der Ausrichtung der Teilnehmenden präsitativ zurichtet und insituativ dominiert.¹⁵ Der Host macht sich demnach das Handlungspotential zu eigen, semiotisch agieren und Beteiligungsstatus delegieren zu können. Immer dann, wenn Zuschauende situativ in den Livevideostream involviert werden und sich somit die interaktionale Ausrichtung zueinander, vornehmlich durch den Host initiiert, verändert, wird die zuvor monologische Kommunikationssituation in eine interaktionale Kommunikationssituation transformiert. Dieses Potential für *footing shifts* als Zusammenspiel zwischen dem medialen Arrangement, den praktischen Verfertigungen durch die Beteiligten und unterschiedlichen Zeichenmodalitäten in Video und Chat charakterisiert Livevideostreams als potentiell interaktional und transzendiert somit einen Status sozialer Parainteraktion. Wenngleich

15 *Situative Kommunikation* bzw. *situative kommunikative Praktiken* seien hier analytisch differenziert in *präsitulative Kommunikation*, die Situationen zeitlich vorausgeht und zukunftsbezogen ist, *insituative Kommunikation*, die sich zeitlich simultan zu Situationen vollzieht und gegenwartsbezogen ist, *postsituative Kommunikation*, die zeitlich nach Situationen stattfindet und vergangenheitsbezogen ist und *transsituative Kommunikation*, die jederzeit stattfinden kann und situationsüberdauernde Bezüge herstellt (vgl. Dang-Anh 2019: 361–363). Situationen haben dabei unterschiedliche Reichweiten und können ineinander verschachtelt sein. So kann etwa die gesamte politische Veranstaltung, über die MK hier berichtet, eine Situation darstellen und ebenso der hier fokussierte Ortswechsel als Situation betrachtet werden.

monologisch geprägte Episoden durchaus parainteraktionale Elemente enthalten können, schließt bereits das grundsätzliche Potential zur Transformation in ‚echte‘ Interaktion aus, Livevideostreams als parainteraktional zu charakterisieren, weil vermeintlich parainteraktionale Elemente im Verlauf immer interaktionsinitiierend aufgenommen werden können (z. B. durch die explizite Rekrutierung zu Chatbeiträgen, vgl. Kap. 4.5).

Im präsentierten Ausschnitt wird durch die Eröffnungssequenz (Z. 02), die Kameraführung, die Einrichtung auf dem Account der Livevideostreaming-Plattform und auch den gewählten Zeitpunkt fokussierter Interaktion – zunächst läuft das Video 22 Sekunden lang unkommentiert, was u. a. der technischen Bedienung geschuldet ist – *hosting* als reflexiv, d. h. sich selbst als Medienpraktik anzeigend, konstituiert. Zusammengefasst: Die Ausrichtung von Streamer*innen und Zuschauer*innen zueinander wird durch das reflexive Display des Hosts hervorgebracht und im weiteren Verlauf kontinuierlich interaktional bearbeitet. Damit ist die Medienpraktiken eigentümliche wechselseitige Relation von Medialität und praktischer Verfertigung beschrieben, die bereits zu Beginn eine interaktionale Asymmetrie etabliert, die die soziale Beziehung zwischen den Streamer*innen und den Zuschauer*innen durchweg charakterisiert.

Die Situation wird im Folgenden durch MK weiter spezifiziert, indem Ort (Z. 03) und Uhrzeit (Z. 05) ebenso angegeben werden wie die handelnde Person (Z. 05). MK adressiert damit drei grundlegende Dimensionen einer Situation, die sich im Anschluss an Bühlers Begriff der Sprechsituation (Bühler 1999 [1934])¹⁶ als „die in der Interaktion praktisch vergegenwärtigte räumliche, zeitliche und personale Konstellation“ (Dang-Anh 2019: 50) definieren lässt. *Vergegenwärtigen* ist somit als situationskonstitutive Praktik zu betrachten, durch die Teilnehmende Situationen einerseits zum Ausdruck bringen, sie andererseits aber auch überhaupt erst hervorbringen. Liveübertragungen wie Videostreams sind aufgrund ihrer fortlaufenden Flüchtigkeit besonders auf kontinuierliche situative Vergegenwärtigungen angewiesen. An dieser Stelle zu Beginn des Livestreams jedoch hat die Situationsherstellung primär einführenden Charakter. Der Host informiert die Zuschauer*innen

16 Die Dimensionen *räumlich*, *zeitlich* und *personal* leiten sich aus Bühlers Sprachtheorie (1999[1934]) ab, in der er das „hier-jetzt-ich-System der subjektiven Orientierung“ (102), die Origo, als Ausgangspunkt für situationsbezogenes Sprechen mit deiktischen Ausdrücken bestimmt.

mit diesem initiierenden *grounding* (vgl. Clark 1996: 221ff.) darüber, in welcher Situation er und sie sich befinden.

Dabei agiert er wahrnehmbar als *animator* und der durchweg spontane Stil lässt vermuten, dass er auch *author* seiner Äußerungen ist. Die Erwähnung seines Arbeit- bzw. Auftraggebers, eine Tageszeitung (Z. 07), verweist hingegen auf diesen als *principal*. Damit festigt MK seine soziale Positionierung als Journalist, der im Auftrag einer Zeitung von einem Ereignis berichtet (Z. 09) und weist sein *hosting* gleichermaßen als professionell aus. MK schließt die Einführungssequenz durch eine Redepause (Z. 10), an die er eine narrative Kontextualisierung der dem berichteten Ereignis vorhergehenden Geschehnisse anschließt (nicht transkribiert).

4.4 Situieren

Im weiteren Verlauf der Analyse wird eine Episode¹⁷ detailliert untersucht, anhand derer gezeigt werden wird, wie sich ein Ortswechsel in mobilen Livevideostreamings sequentiell entfaltet und interaktional bearbeitet wird. Der Ausschnitt aus dem *Periscope*-Video „Kaul aus #Köthen, 18.06 Uhr.“ ist insgesamt 1:12 Min. lang (00:22:40–00:23:52) und bildet eine geschlossene Sequenz, die in drei Teilausschnitten analysiert wird (vgl. die Transkripte 2, 3 und 4). MK eröffnet die Sequenz durch die Partikel „soo;“ (vgl. Transkript 2, Z. 01).

17 Die in den Transkripten 2–4 angegebenen Zeiten beziehen sich auf den im Folgenden untersuchten Ausschnitt der in den Transkripten 2–4 abgebildeten Sequenz, nicht auf den gesamten Livestream.

Transkript 2: „Situieren“ (aus: [Periscope: Kaul aus #Köthen, 18.06 Uhr.], 00:22:40–00:22:56)

```
00:00  01  MK      %+soo;  +
          move   %läuft  -->>
          gaze   +blickt+vor sich
00:02  02  MK      +wir warn jetzt grade *bei #dieser
          demon-  stration da HIN* ne_#ANTIFaschistische
          ademon-  stratio-
          move   αdreht sich zur seite und zurück-->
          gaze   +blickt in die kamera-->
          hand                    *zeigt hinter sich---
          ----- *spreizt zeigefinger
          ab-->
```



Abb. 5



Abb. 6

```
00:06  03  MK      hier sind jetzt weiter son bisschen+ die
          stra-  ßen gesperrt,*α+
          move                    -->α
          gaze                    -->+blickt
          zur seite-----+
          hand                    -->*
```

(wird auf nächster Seite fortgeführt)

00:10 04 MK +ich LAUF jetzt äh (-) für alle die nich
 beSCHEID# wissen was hier gerade passiert
 (.)+nochn bisschen WEIter,
 gaze +blickt in die kamera-----
 ---+blickt vor sich-->



Abb. 7

gaze
 hand *deutet richtung kamera-----*
 zeigt vor sich-->

00:15 05 MK und +versuche mal den MARKTplatz zu
 afinden;+
 move αdreht sich nach hinten und zurück-->>
 gaze -->+blickt in die kamera-----
 -----+

Der Streamer geht, offenbar mit dem Smartphonestick in der rechten Hand, eine Straße entlang und blickt dabei in die Kamera. In Zeile 02 erfolgt eine Situationsdarstellung, die durch das Verbtempus von „warn“ ebenso wie durch die temporaldeiktische Zeitangabe „jetzt gerade“ als Retrospektive gekennzeichnet ist. „[D]ieser Demonstration“ bzw. „ne_ANTIfaschistische demonstratio“ benennen das politische Ereignis, von dem der Streamer zuvor als Journalist berichtet hat. Kombiniert ist diese Benennung mit dem Fingerzeig des Streamers, der zunächst hinter ihn, d. h. entgegen seiner Laufrichtung, weist (Abb. 5) – eine räumliche Markierung zeitlicher Retrospektive. Bei der präzisierenden Attribuierung („ne_ANTIfaschistische“) spreizt er zusätzlich zum ausgestreckten Daumen den Zeigefinger ab (Abb. 6), was als Aufzählungsgeste, die den beschriebenen Demonstrationzug als Teilereignis in einem größeren Ereigniszusammenhang kennzeichnet, verstanden werden kann. In dieser Sequenzeinleitung steht das verbale deiktische

und gestische *Zeigen* im Vordergrund. Gezeigt wird auf personale, zeitliche und räumliche Dimensionen zur Situationskonstitution. Die Situation wird in dieser Sequenz demnach neu justiert, die Szenerie gewechselt. MK möchte sich nun zur rechtsextremen Versammlung begeben, die räumlich getrennt von der Gegenversammlung stattfindet und nimmt entsprechend einen Ortswechsel vor. Die personale Konstellation adressiert MK mit dem inkludierenden Pronomen „wir“, womit er die Zuschauenden postsituativ involviert (Z. 02): „[W]arn“ bezieht sich nicht nur auf die (nun vergangene) physische Existenz im lokalen Raum vor Ort, sondern gleichsam auf die medial vermittelte Wahrnehmung der Zuschauer*innen im Livevideostream. Der Streamer ‚nimmt‘ die Adressierten nicht nur ‚mit‘ auf seinen Weg, sondern adressiert sie als teilnehmende Agenten seiner per Stream vermittelten körperlichen Bewegungshandlung und konstituiert seinen Ortswechsel als gemeinsames Projekt (*joint action*; vgl. Clark 1996: 59–91). In der anschließenden Zeile 03 vollzieht MK dann einen Wechsel in die aktuelle Situation. Durch „hier“ und „jetzt“ nimmt MK sprachlich Bezug auf zeitlich und räumlich Gegenwärtiges. Ein sehr kurzer Schwenk der Kamera auf die angesprochenen Absperungen, in der die Zuschauer sehen können, was MK sieht, komplementiert die situative Vergegenwärtigung (Z. 03). Noch deutlicher wird nun der Livecharakter der Interaktionssituation: Nach dem retrospektiven Involvement (Z. 02) erfolgt nun ein insituatives Involvement anhand der sprachlichen und medialen Konstitution eines gemeinsam erlebbaren Hier, Jetzt und Wir.

In Zeile 04 wechselt der Streamer die pronominale Perspektive und stellt in einer Handlungsbeschreibung dar, wie *er* sich bewegt: „ich lauf jetzt [...] nochn bisschen weiter“. Das Verb drückt dabei komplementär zu dem bewegten Kamerabild seine Bewegung aus. Der Einschub in Zeile 04 diversifiziert nun die Beteiligtenkonstellation: MK adressiert nun explizit „alle die nich bescheid wissen was hier gerade passiert“ und differenziert damit diese von denjenigen, die durch ihre Streamrezeption vermeintlich über situatives Wissen verfügen. Dies ist ein erster *footing shift*, ein Prozess, den Gibson (2003) im Anschluss an Goffman (1979) als „moment-by-moment shuffling of individuals between the ‚participation statuses‘ of speaker, target, and unaddressed recipient“ (1335) beschreibt. Die Differenzierung erfolgt hierbei als *grounding*, um (auch!) diejenigen durch situative Vergegenwärtigung zu involvieren, die zuvor aufgrund des Mangels an situativem Wissen („was hier gerade passiert“) nicht in dem Grad einbezogen waren, wie diejenigen Zuschauenden, die aufgrund ihrer kontinuierlicheren bzw. längeren Rezeption des Livestreams situativ ‚eingeweiht‘ waren. Jedoch ist die Unterscheidung feingliedriger als zwischen *target* und *unaddressed recipient*, denn

auch die ‚Wissenden‘ scheinen durch die Äußerung adressiert, mit der MK nicht nur *grounded*, sondern seine im Stream sichtbare körperliche Bewegung sprachlich sekundiert, um die seine Vor-Ort-Berichterstattung unterbrechende Bewegungsphase zu überbrücken. Im mobilen Livevideostreaming, das vor allem auf visuelles Involvement abzielt, etwa wie der hier analysierten journalistischen Livestream-Berichterstattung von Ereignissen, gerade bei dynamischen räumlichen Ereignissen wie z. B. Protesten, kommt es mitunter zu Phasen, in denen ‚es nichts zu sehen‘ gibt. So gibt es im gesamten Streamvideo immer mal wieder Sequenzen, in denen diese Art von *bridging* vollzogen wird: Die entstehenden ‚visuellen Leerstellen‘ überbrückt der Streamer, wie in der analysierten Sequenz, mit empraktischer Kommunikation (vgl. Bühler 1999[1934]), die seine sichtbare körperliche Bewegung sekundiert oder durch, teils vorbereite, transsituative oder nicht-situationsbezogene Kommunikation. So etwa führt MK zuvor während eines längeren Ortswechsels ein verabredetes Interview im Gehen („Kaul aus #Köthen, 18.06 Uhr.“, 00:08:03–00:15:56).

Führt MK in Zeile 04 bereits aus, dass er „WEIter“ geht, was er mit einem Fokusakzent und dem nun von ihm aus nach vorne zeigenden Finger komplementiert (Abb. 7), so vollzieht er in Zeile 05 einen weiteren Wechsel der situativen Bezugnahme. Seine Verbwahl („versuche“, „finden“) zeigt den Beginn eines indeterminierten und unsicheren Prozesses an, was mitbedeutet, dass der Versuch zukünftig gelingen oder scheitern kann. Auch zeitlich gesehen wechselt die Bezugnahme in einen präsituativen Modus, insofern eine kommende Situation auf dem „Marktplatz“ prospektiv entworfen wird. Zu beobachten ist in den Zeilen 02 bis 05 in sehr kurzer aufeinanderfolgender Reihung, wie die Situation durch vergangenheitsbezogene postsituative (Z. 02), gegenwartsbezogene insituative (Z. 03, 04) und zukunftsbezogene präsituativen Kommunikation (Z. 04, 05) konstituiert wird und die Zuschauenden situativ involviert werden. Situationsbezüge erfolgen dabei multimodal, sowohl sprachlich, als auch körperlich und medial. Erstens, hinsichtlich der *Zeit*: grammatisch durch Tempuswechsel („warn“, Z. 02 vs. „lauf“, Z. 04), temporaldeiktisch („jetzt grade“, Z. 02; „jetzt“, Z. 03; „gerade“, Z. 04), semantisch durch Handlungsverben mit prozessuellem Bedeutungsanteil („lauf“, Z. 04; „versuche [. . .] zu finden“, Z. 05), gestisch durch Zeigebewegungen des Fingers vom Streamer aus nach hinten (vergangenheitsbezogen, Z. 02) und nach vorne (zukunftsbezogen, Z. 04); zweitens, hinsichtlich des *Raums*: durch Lokaldeiktika („da HIN“, Z. 02; „hier“, Z. 03), Toponyme („MARKTplatz“, Z. 05), zudem durch Zeigegesten (nach hinten, Z. 02; nach vorne, Z. 04), Bewegung des Körpers und Wechsel der Kameraperspektive

(Schwenk auf Absperrungen); drittens, hinsichtlich der *Personen*: durch Pronomina („wir“, Z. 02; „ich“, Z. 04), die *talking head*-Kameraperspektive, Nennung von Gruppen („ANTIFaschistische demonstratio“, Z. 02) oder Adressierung („für alle die [...]“, Z. 04).

4.5 Rekrutieren

Transkript 3: „Rekrutieren“ (aus: [Periscope: Kaul aus #Köthen, 18.06 Uhr], 00:22:56–00:23:24)

```
00:18 06          %α+* (0.61) #
                move   %läuft-->
                move   odreht sich nach hinten und zurück-->
                gaze   +blickt mit körperbewegung nach hinten
                        und dann seitlich nach oben-->
                hand   *zeigt vor sich-->
```



Abb. 8

```
00:19 07 MK      schätze max (-) dass der in +DIese richtung
                  geht=%aber da könnt ihr mir mal helfen;%*
                move   -->α
                move   -->% steht und dreht sich zur seite--%
                gaze   -->+blickt in die
                        kamera-->
                hand   -->*
```

(wird auf nächster Seite fortgeführt)

00:22 08 MK %*des hier is die #MÜhlenstraße* (---)
 %die MÜhlenstraße,
 move %steht-----
 %läuft über straße-->
 hand *deutet mit zeigefinger----- *
 schräg nach oben richtung schild



Abb. 9

00:27 09 (0.6)
 00:28 10 MK und ich lauf jetzt die mühlenstraße einfach
 gradeaus (.) weiter lang,
 00:31 11 MK und +HOFfe dass+ ich dann %zum MARKTplatz
 komme.=%
 move -->%dreht sich vor
 schild-%
 gaze -->+blickt
 schräg n. oben-+blickt in die kamera-->
 00:34 12 MK #=Aber *vielleicht GIBT es ja+ jemanden(--)
 von euch (--) der #ma kurz (-) *%GOOGeln
 kann-
 move %läuft
 weiter-->>
 gaze -->+blickt unter
 die kamera
 hand *scrollt-----*



Abb. 10



Abb. 11

00:39 13 (0.7)
 00:40 14 MK +äähm-
 gaze -->+blickt in die kamera-->
 00:41 15 (0.59)
 00:42 16 MK ob DAS hier der richtige WEG is;
 00:43 17 MK *sonst lauf ich nämlich in die #FALSche
 richtung.*
 hand *zeigt vor sich-----
 -----*



Abb. 12

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

MK dreht sich nun im Gehen und blickt zunächst hinter sich und dann zur Seite, leicht nach oben (Z. 06, Abb. 8). Erneut verwendet er eine Vagheitsmarkierung („schätze ma“, Z. 07), dieses Mal mit dem epistemischen Verb ‚schätzen‘, durch die er sein Nichtwissen ob der lokalen räumlichen Gegebenheiten kennzeichnet. Beim akzentuierten „DIEse richtung“ (Z. 07) deutet sein Finger vor sich in seine Laufrichtung, was die Verschränkung von sprachlicher und gestischer Lokaldeixis belegt. In direktem Anschluss erfolgt

die Adressierung der Zuschauenden, die durch ein adversatives „aber“ eingeleitet wird, und durch die die Adressierten modalisiert („können“) zur Mithilfe („helfen“) aufgefordert werden (Z. 07). Die gesamte Zeile 07 „schätze ma (-) dass der in DIEse richtung geht=aber da könnt ihr mir mal helfen;“ stellt einen *footing shift* dar, in dem der Streamer sein Publikum aktiv in das situative Geschehen einbezieht. Die erwähnte Vagheitsmarkierung zu Beginn, gemeinsam mit diesem *recruitment* des Publikums (vgl. Drew/Couper-Kuhlen 2014; Kendrick/Drew 2016), weist einen potentiell asymmetrischen epistemischen Statusunterschied durch *epistemic stancetaking* (vgl. Heritage 2012) zwischen dem Streamer und den Zuschauenden aus. MK gibt durch seine drehende Körperbewegung, seine schnell wechselnden Blickrichtungen, seine vagheitsmarkierende Verbalisierung und seine rekrutierende Adressierung der Zuschauenden zu verstehen, dass er sich momentan im lokalen Raum orientiert und dabei auf Hilfe Dritter angewiesen ist. Indem MK die Zuschauenden direkt anspricht („aber da könnt ihr mir mal helfen“, Z. 07), erzeugt er einen ersten Teil eines Adjazenzpaares (vgl. Schegloff 2007) und evoziert somit die Folgerwartung von Chat-Posts, was auch sein anschließendes Scrollen (Z. 12) unterstreicht. Durch die derartige Konstituierung eines Orientierungsproblems wird Orientieren (*orienting*) im Livevideostream als eine situative Suchpraktik ausgewiesen, deren Erfolgchancen sich durch das situative Involvement der Zuschauenden erhöht. Obwohl der Streamer als Host interaktionsdominierend und als Journalist als Experte auftritt, markiert er hier eine Wissensasymmetrie, die dieser Statusasymmetrie entgegenzustehen scheint. Die Anzeige der Wissensasymmetrie gibt dem Streamer jedoch die Gelegenheit, sein Publikum an seiner Situation zu beteiligen. Epistemisches Stancetaking, also das Äußern von Unterschieden des epistemischen Status, dient sequenzinitiiierend „as a means of warranting conversational contributions and building expanded conversational sequences“ (Heritage 2012: 49). Dieser *footing shift* projiziert demnach Folgezüge durch das Publikum und geht somit über soziale Parainteraktion hinaus, wobei noch nicht konkretisiert ist, wie das Publikum dem Streamer in dieser Situation helfen kann.

Es kommt in der Folge zu einer objektorientierten räumlichen Konkretisierung des zuvor konstituierten Orientierungsproblems: In Zeile 08 identifiziert MK seine räumliche Situierung anhand der wiederholten Nennung des Straßennamens. Dabei dreht er sich so, dass im Hintergrund ein Straßenschild zu sehen ist und zeigt mit seinem Finger in die Richtung des Schildes (vgl. Abb. 9). Durch dieses *noticing* wird es aus der Perspektive der Zuschauenden als Bezugsobjekt im dargestellten Straßenraum identifizierbar (vgl.

Licoppe/Tuncer 2019). Während der Zeilen 09 bis 11 überquert MK die Straße und blickt dabei seitlich nach oben. Damit wiederholt er richtungsgleich den seitlich orientierten Blick aus Zeile 06, der offenbar das identische Objekt fixierte. MK wechselt pronominal („ich“) wieder in die Egoperspektive und situiert sich temporal („jetzt“) und räumlich („mühlenstraße einfach gradeaus (.) weiter lang,“), wobei die Handlungsbeschreibung nicht nur seine aktuelle Tätigkeit beschreibt, sondern per andauerndem Vorgangsverb und Richtungsangabe („lauf [. . .] einfach gradeaus (.) weiter lang,“, Z. 10) auch seine zukünftige Bewegung antizipiert. Durch das Verb „hoffen“ wird erneut die Unsicherheit ob seines Orientierungsproblems markiert (Z. 11), ebenso wie der Zielort „MARKTplatz“ wiederholt genannt wird. Sowohl der gegenwärtige Standort („Mühlenstraße“) als auch der antizipierte Zielort („MARKTplatz“) sind zu diesem Zeitpunkt mehrfach genannt und somit als räumliche Orientierungspunkte etabliert. Die beschriebene Abfolge ist somit als *direction-giving sequence* lesbar, die in mobiler Interaktion eine gewichtige Rolle spielt: „‘Direction-giving sequences’ involve practical formulations that take into account a starting point as well as a target or a destination“ (Haddington/Mondada/Nevile 2013: 20). Um seine aktuelle räumliche Situiertheit zu verankern (vgl. Hausendorf 2013), stellt sich MK vor das Schild, das zuvor nur von Weitem zu sehen war und nun im Videostream lesbar ist (Z. 11, Abb. 10). Das Objekt Straßenschild dient somit als semiotische Ressource (vgl. Nevile/Haddington/Heinemann et al. 2014), die MK als verankerndes und authentifizierendes Evidenzobjekt (vgl. Luginbühl/Hauser 2019: 364) räumlicher Orientierung und als visuellen Bezugspunkt für die Rezipierenden während seiner Bewegungspause nutzt.

In Zeile 12 adressiert MK erneut die Zuschauenden. Die Wiederaufnahme des *footing shifts* leitet er erneut mit einem adversativen „Aber“ ein, was Lösungsmöglichkeiten zum Orientierungsproblem andeutet. Dabei wählt er einen vagen generischen Adressatenzuschnitt für seine Rekrutierung im Gewand einer Existenzvermutung: „vielleicht GIBT es ja jemanden (--) von euch“ (Z. 12). Mit der Adressierungspräzisierung „der ma kurz (-) GOOgeln kann-“ macht der Streamer ein interaktionsinitiierendes (und somit Parainteraktion transzendierendes) Angebot, das insofern spezifisch auf die Teilnehmenden an der medialen Interaktionssituation zugeschnitten ist, als es deren lokale Rezeptionssituation berücksichtigt. Der Stream wird durch die Zuschauenden durch ein digitales Medium, etwa einen Rechner, ein Smartphone oder ein Tablet, rezipiert, welches ihnen gleichfalls die Möglichkeit eröffnet, parallel zur Livestreamrezeption zu „GOOgeln“. Dieser Shift ist

aber nicht nur ein Interaktionsangebot, sondern vielmehr ein *recruitment*¹⁸, durch das Partizipierende dazu gebracht werden sollen, etwas zu tun (vgl. 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)

Derartige „attempts to enlist someone’s assistance, typically with respect to an immediate, physical need, problem or wish“ (Drew/Couper-Kuhlen 2014: 17) finden in der Regel lokal statt: „Recruitment is, then, restricted to quite material, here-and-now matters; it does not involve ‚remote‘ matters“ (Kendrick/Drew 2016: 2). Im Livevideostream jedoch kommt es zu einer distanten Rekrutierung, die ja gerade dadurch charakterisiert ist, dass der rekrutierende Streamer ein lokales Orientierungsproblem hat, das die Zuschauenden *mit ihm* lösen sollen. Demnach kommt das Involvement der Rekrutierten nicht qua ‚Assistenz‘ zustande, sondern vielmehr dadurch, dass sie ihren Beitrag zum gemeinsamen Projekt des Ortswechsels leisten (vgl. Zinken/Rossi 2016). Dabei wird neuerdings vor allem betont, dass recruitment praktische Tätigkeiten betrifft, die über reine Sprechhandlungen hinaus gehen (vgl. Rossi/Floyd/Enfield 2020). Wenngleich das „GOOgeln“ selber im Livestream nicht sichtbar wird, zielt doch die Rekrutierung auf dessen praktische Verfertigung. Hieran wird deutlich, wie sich die Livevideostream-situation als medial vermittelte *synthetische Situation* (vgl. Knorr Cetina 2009) bzw. *Intersituation* (vgl. Hirschauer 2014) verstehen lässt, in der neben der lokalen Situation des Hauptdarstellers auch die medialen Rezeptionssituationen der Zuschauenden den Hintergrund des situativen Settings bilden und die Parteien dennoch ein gemeinsames Projekt verfolgen. Im Gegensatz zur monosituativen Parainteraktion bezieht sich Rekrutieren in medial vermittelten Interaktionen demnach nicht ausschließlich auf Tätigkeiten, die im gemeinsamen Hier und Jetzt durchgeführt werden, sondern wird ebenso intersituativ vollzogen.

18 Im Gegensatz zum hier dargelegten *recruitment* der Zuschauenden durch den Streamer zeigt Choe am Fall des koreanischen Livestream *mukbang*, in dem der Host vor Publikum isst, wie Chattende den Host für Aktivitäten rekrutieren (vgl. Choe 2019).

MKs eigene mediale Rezeption wird durch das offensichtliche Scrollen auf seinem Smartphone, das er parallel zu seiner rekrutierenden Adressierung durchführt, angezeigt (Z. 12, Abb. 11). Durch dieses Display der Lektüre der Chatpostings, die im Chatfenster des Screens auf dem Streaminggerät dargestellt werden, zeigt der Host nicht nur seine Wahrnehmungswahrnehmung an, sondern bringt auch die Erwartung zum Ausdruck, dass zeitgleich mit seinem Livebericht auch die Zuschauenden per Chat mit ihm oder untereinander interagieren. Das rekrutierende Interaktionsangebot wird demnach durch den Ausdruck von Interaktionserwartung komplementiert. Gleichwohl wird hierdurch kein unmittelbarer Anschluss eingefordert, denn neben der Verzögerung durch Beitragsproduktion und Latenz wurden die Zuschauenden ja dazu rekrutiert, sich situatives Wissen über die örtlichen Gegebenheiten anzueignen, was ebenfalls Zeit in Anspruch nehmen dürfte.

Während MK seinen Weg daraufhin fortsetzt, bringt er wiederholt sein unsicheres Ortswissen zum Ausdruck („ob DAS hier der richtige WEG is;“, Z. 16). Er weist seine sichtbare Bewegungshandlung zudem durch die Nennung von negativen Konsequenzen als möglicherweise kontraproduktiv aus („sonst lauf ich nämlich in die FALSche richtung.“, Z. 17). Dem Teilprojekt Ortswechsel und womöglich dem gesamten Projekt ‚Livestream von der rechten Veranstaltung‘ droht somit ein Scheitern bzw. eine Verzögerung und MK erhöht mit seinem negativen Zukunftsentwurf den Partizipationsdruck auf die Rezipierenden. Um ein Scheitern zu verhindern, ist das interaktionsaktive situative Involvement der Zuschauenden notwendig. Rekrutieren wird demnach hier punktuell zum Erwerb von situativem Wissen, als Wissen bezüglich der zeitlichen, räumlichen und personalen Situationsdimensionen, eingesetzt, bis der Zustand eines „equilibrium-for-all-practical-purposes“ (Heritage 2012: 48) eintritt, der für den Erfolg des gemeinsamen Projekts hinreichend ist.

4.6 Anerkennen

Transkript 4: „Anerkennen“ (aus: [Periscope: Kaul aus #Köthen, 18.06 Uhr.], 00:23:25–00:23:53)

00:46	19	MK	%+äähm +DA würd ich jetzt weiter HINlaufen, move %läuft--> gaze +blickt nach vorne -----+blickt unter die kamera-->	Chat
00:48	20		*(0.63) hand *scrollt-->	3
00:49	21	MK	und mal kurz die kommentare derjenigen lesen-	4 Wieviel Meter sind die voneinander entfernt?
00:52	22		(0.87)	
00:53	23	MK	die hier ZUGeschaltet sind.	5 Gerade weiter dann nach rechts
00:55	24		(1.86)	6 Nächste Kreuzung rechts 7 Pass auf dich auf!
00:57	25	MK	((hupen)) wie viele meter +sind die voneinander ent+fernt? gaze -----+blickt zur seite-----+blickt unter die kamera-->	
00:59	26	MK	((schnalzt)) das WEISS ich nicht* +weil ich die eine gruppe (--) noch +nich geSEHN hab (-) sondern nur die andere. gaze -----+blickt nach vorne-----+blickt unter die kamera--> hand -----*	8 Sind Sie da wirklich alleine? 9 Am 1. Sep waren in Chem- nitz auch Reiter- staffeln im Einsatz

Fortsetzung

01:05 27 *#(3.88) ((motorgeräusche))*
 hand *scrollt-----*



Abb. 13

01:09 28 MK grade weiter dann nach rechts sacht jemand-
 01:11 29 MK das is nett-+
 gaze -->+
 01:12 30 MK +VIElen dank für die orientierung.
 gaze +blickt nach vorne-->>

In der darauffolgenden Interaktionsphase (vgl. Transkript 4) geht MK durchgehend weiter den Bürgersteig entlang und es kommt zu einer längeren *bridging*-Phase (Z. 19–26). Er sekundiert seine Körperbewegung erneut durch einen prospektiven Handlungsentwurf in Zeile 19.¹⁹ In den Zeilen 21 bis 23 verbalisiert MK seine Rezeption des Chats, während derer er fortlaufend auf den Screen blickt und mit dem Finger scrollt. Dies ist eine Form des rezeptionsbegleitenden Sprechens, das hier neben der Überbrückung die Funktion hat, die Rezeptionshandlungen des Streamers anzuzeigen und auftretende Redepausen (Z. 20, 22, 24), die den kognitiven Fokus des Streamers auf seine Rezeptionshandlung verdeutlichen, zu kontextualisieren. Auch diese scheinbar sekundäre Handlung ist involvierend, insofern die Beiträge²⁰

- 19 Angenommen sei hier eine futuristische Lesart, die sich aus der Konjunktiv II-Entsprechung der Konstruktion *werd(en) + Infinitiv* ableitet (vgl. Fabricius-Hansen 2000), wengleich in der Situation die antizipierte Handlung zum Zeitpunkt der Äußerung bereits eingesetzt hat, wie sich aus der im Livestream sichtbaren Bewegung schließen lässt.
- 20 Die Chatbeiträge sind im Transkript zu Sequenz 4 in der rechten Spalte aufgeführt. Die Chatbeiträge 1–2, die sich zeitlich vor dem hier transkribierten Ausschnitt ereigneten, sind in den anderen Transkripten nicht abgebildet, da sie nicht relevant gesetzt werden.

der Zuschauenden („die hier ZUgeschaltet sind“, Z. 23) durch die verbalisierte und angezeigte Rezeption des Hauptdarstellers gewürdigt werden. In den Zeilen 25–26 kommt es zu einem *read aloud and response* bezüglich des Chatbeitrags 4 (s. Transkript 4), wodurch ein *footing shift* initiiert wird: Der Streamer rezitiert den interrogativen Chatbeitrag (Z. 25) und antwortet auf ihn (Z. 26). Die wiedergegebene adjazente Frage-Anwort-Sequenz (vgl. Schegloff/Sacks 1973) belegt die Interaktionalität des Livevideostreams. Durch die Überführung des Chatbeitrags in die audiovisuellen Modalitäten des Livestreamvideos macht der Streamer (als *animator*) ihn für die Rezipierenden salient und es erfolgt eine selektive Relevantsetzung des Beitrags des Chatters (als *author*).²¹

In Zeile 26 zeigt MK erneut seinen epistemischen Status an. Die Antwort auf die Frage nach dem Abstand zwischen den beiden Gruppen, zwischen denen MK den Ortswechsel vollzieht, könne er nicht beantworten, da er „die eine gruppe (-- noch nicht geSEHN“ (Z. 26) habe. Auch hiermit verweist MK auf das gemeinsame Projekt, das den Ortswechsel als notwendig für das kopräesente Livestreaming vom Zielort voraussetzt. Der gegenwärtige sowie zukunftsbezogene Verlaufscharakter wird durch die hier temporal verwendete Negation „noch nicht“ indiziert (Z. 26). Da der selektierte Beitrag nicht zur Lösung des Orientierungsproblems beiträgt, liest MK weiter (Z. 27, Abb. 13). Nach dieser längeren Rezeptionsphase kommt es in den Zeilen 28–30 zu einer weiteren *RAR*-Paarsequenz. Der in Zeile 28 rezitierte Chatbeitrag 5 „grade weiter dann nach rechts“ leistet die erwünschte Raumorientierung. In den Zeilen 29–30 erfolgt schließlich eine positive Bewertung und ein Dank für den Beitrag als Anerkennung der Rekrutierten durch den Rekrutierer (vgl. Rossi/Floyd/Enfield 2020). Die Sequenzschließung wird prosodisch durch eine tief fallende Tonhöhenbewegung angezeigt und der Blick des Streamers richtet sich vom Smartphone zurück in Laufrichtung (Z. 30). Die Sequenz findet hier ihren Abschluss in der positiven Anerkennung der Lösung des Orientierungsproblems („VIElen dank für die orientierung.“, Z. 30). Mit „orientierung“ benennt MK den Zustand, der durch die zuvor durchlaufenen Phasen des Situierens, Orientierens, Rekrutierens und Sammeln von Antworten auf das Orientierungsproblem sequentiell erreicht wurde. Mit der positiven Anerkennung („VIElen dank“, Z. 30) löst sich die in Zeile 17 ausgesprochene Bedrohung auf („sonst lauf ich nämlich in die

21 Von *topicalising* kann hier insofern nicht gesprochen werden, als der Streamer durch sein vorhergehendes *recruitment* bereits themensetzend agierte.

FALSche richtung.“), die das Orientierungsproblem und dessen interaktionale Auflösung eingeklammert hat.

5. Fazit

In Livevideostreams stehen die Teilnehmenden in einer asymmetrischen Relation zueinander. Diese Asymmetrie ist einerseits medial zugerichtet und wird andererseits im Verlauf interaktional ausgehandelt. Dabei ist die Beteiligungskonstellation dynamisch und wird hauptsächlich durch die Streamenden orchestriert, indem sie in dem Mehrpersonenensemble des Livevideostreams *footing shifts* multimodal initiieren, durch die sich die Ausrichtungen der Teilnehmenden zueinander konfigurieren. ‚Multimodal‘ bezieht sich hierbei nicht nur auf das körperliche Agieren der Streamenden im lokalen Raum, sondern auf die gesamte Medienpraktik des Livevideostreamings: Streamer*innen delegieren in ihrer Rolle als Hosts Beteiligungsstatus vor allem durch den Wechsel zwischen Zeichenmodalitäten, den die Medialität des medialen Arrangements ‚Livevideostream‘ ermöglicht. *Footing shifts* werden, wie gezeigt wurde, insofern modalitätswechselnd initiiert, als auf Relevantsetzungen von Chatbeiträgen interaktionale Erweiterungen im Videostream folgen, die den Beteiligungsmodus sozialer Parainteraktion transzendieren. Mediale und modale Asymmetrien stehen somit in enger Relation zu den sozialen Asymmetrien, die in Livevideostreams medial und praktisch verfertigt werden. Livestreamhosting ist also eine reflexive Medienpraktik, die mit dem Livestreamgeschehen kontinuierlich mitläuft. Die Wahl der Medien und Zeichenmodalitäten, die Setzung von Themen, die Initiierung von Sprecherwechseln oder die Rezipitation von Chatbeiträgen – all dies geschieht zuvorderst durch die hostenden Streamenden, befördert das situative Involvement der Zuschauenden und weist die Streamenden in der asymmetrischen Partizipationskoordination in Livevideostreams als interaktionsdominierend aus.

Mobile Livevideostreams sind in Episoden gegliedert. Die Mobilität zeigt sich insbesondere dann, wenn Ortswechsel in Livevideostreams als gemeinsam verfertigte Projekte bearbeitet und koordiniert werden. Gezeigt wurde, wie Ortswechselsequenzen durch Praktiken des Situierens, Bridgings, Orientierens, Rekrutierens, Groundings, Auflösens und Anerkennens vollzogen werden. Während des Situierens vergegenwärtigen sich die Teilnehmenden der zeitlichen, räumlichen und personalen Konstellation durch sprachliches,

gestisches und körperliches Zeigen und Verorten. Mitunter werden dabei auch Objekte einbezogen, durch die Situatives evident und authentifizierend verankert wird. Phasen der körperlichen Bewegung werden hingegen durch Involvementpraktiken überbrückt. Auch hier spielen die Streamenden die Hauptrolle. Durch die Konstitution eines Orientierungsproblems, das das gemeinsame Projekt gefährdet, und das damit verbundene Recruitment der Zuschauenden werden sowohl Interaktionsangebote als auch Partizipationsdruck erzeugt. Das Orientierungsproblem stellt sich dabei als durch epistemisches Stancetaking hervorgebrachte Wissensasymmetrie dar, die der Statusasymmetrie zwischen Host und Publikum entgegenzustehen scheint. Doch erst das Involvement der Chatter*innen führt in der beschriebenen Situation zum Grounding des Status des situativen Wissens. Mit der Auflösung löst sich auch die Bedrohung des gemeinsamen Projekts und der Host erkennt schlussendlich das situative Involvement der Zuschauenden an.

Die Analyse legt dar, dass soziale Parainteraktion nicht der grundsätzliche Beteiligungsmodus in Livevideostreams ist (vgl. Schmidt/Marx/Neise 2020), sondern vielmehr zwischen unterschiedlichen interaktionskonstitutiven Ausrichtungen der Partizipierenden zueinander durch *footing shifts* gewechselt wird (vgl. Goffman 1979; Choe 2020). Die Medialität des Livevideostreams mit der tragbaren Kamera, dem Display und einer Chatfunktion bietet die Grundlage für eine multimodale Kommunikationssituation, in der wechselseitige Interaktion *potentiell* realisiert werden kann. Im Gegensatz zur sozialen Parainteraktion, die vor allem durch einseitige Kommunikation gekennzeichnet ist, wird in der Medienpraktik des mobilen Livevideostreamings Sozialität dynamisch, gemeinsam und wechselseitig, aber auch hochgradig asymmetrisch ausgehandelt.

6. Literatur

- Arminen, Ilkka / Licoppe, Christian / Spagnolli, Anna (2016): Respecifying Mediated Interaction. In: *Research on Language & Social Interaction* 49/4, 290–309. DOI: 10.1080/08351813.2016.1234614.
- Bateman, John A. / Wildfeuer, Janina / Hiippala, Tuomo (2017): *Modality – Foundations, Research and Analysis. A Problem-Oriented Introduction*. Berlin: Mouton de Gruyter.

- Beißwenger, Michael (2007): Sprachhandlungskoordination in der Chat-Kommunikation. Berlin: de Gruyter (Linguistik – Impulse & Tendenzen 26).
- Beißwenger, Michael (2008): Situated Chat Analysis as a Window to the User's Perspective: Aspects of Temporal and Sequential Organization. In: *Language@Internet* 5.
- Beißwenger, Michael (2020): Internetbasierte Kommunikation als Textformen-basierte Interaktion: ein neuer Vorschlag zu einem alten Problem. In: Marx, Konstanze / Lobin, Henning / Schmidt, Axel (Hrsg.): *Deutsch in Sozialen Medien*. Berlin [u. a.]: de Gruyter (IDS Jahrbuch), 291–318.
- Bergmann, Jörg (1981): Ethnomethodologische Konversationsanalyse. In: Schröder, Peter / Steger, Hugo (Hrsg.): *Dialogforschung. Jahrbuch 1980 des Instituts für deutsche Sprache*. Düsseldorf: Pädagogischer Verlag Schwann (Sprache der Gegenwart 54), 9–51.
- Bergmann, Jörg / Meyer, Christian / Salomon, René 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. DOI: 10.17169/fqs-20.2.3289.
- Blommaert, Jan (2019): Formatting online actions: #justsaying on Twitter. In: *International Journal of Multilingualism* 16/2, 112–126. DOI: 10.1080/14790718.2019.1575832.
- Bou-Franch, Patricia / Lorenzo-Dus, Nuria / Blitvich, Pilar Garcés-Conejos (2012): Social Interaction in YouTube Text-Based Polylogues: A Study of Coherence. In: *Journal of Computer-Mediated Communication* 17/4, 501–521. DOI: 10.1111/j.1083-6101.2012.01579.x.
- Bühler, Karl (1999 [1934]): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Ungekürzter Neudr. der Ausg. Jena, Fischer, 1934, 3. Aufl. Stuttgart: Lucius und Lucius (UTB für Wissenschaft 1159).
- Choe, Hanwool (2019): Eating together multimodally: Collaborative eating in mukbang, a Korean livestream of eating. In: *Language in Society* 48/2, 171–208. DOI: 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. DOI: 10.1016/j.pragma.2020.02.009.
- Clark, Herbert H. (1996): *Using Language*. Cambridge, New York: Cambridge University Press.

- Dang-Anh, Mark (2021): Situational involvement and footing shifts in mobile live video streams. In: *Gesprächsforschung – Online-Zeitschrift für verbale Interaktion* 22, 304–336.
- Dang-Anh, Mark (2019): Protest twittern. Eine medienlinguistische Untersuchung von Straßenprotesten. Bielefeld: Transcript (Locating Media/ Situierete Medien 22). DOI: 10.14361/9783839448366.
- Dang-Anh, Mark / Pfeifer, Simone / Reiser, Clemens et al. (2017): Medienpraktiken. Situieren, erforschen, reflektieren. Eine Einleitung. In: *Navigationen. Zeitschrift für Medien- und Kulturwissenschaften* 17/1, 7–36. DOI: 10.25969/mediarep/1702.
- Deppermann, Arnulf (2015): Pragmatik revisited. In: Eichinger, Ludwig M. (Hrsg.): *Sprachwissenschaft im Fokus. Positionsbestimmungen und Perspektiven*. Berlin [u. a.]: de Gruyter (IDS Jahrbuch), 323–352.
- Drew, Paul / Couper-Kuhlen, Elizabeth (2014): Requesting – from speech act to recruitment. In: Drew, Paul / Couper-Kuhlen, Elizabeth (Hrsg.): *Requesting in social interaction*. Amsterdam, Philadelphia: John Benjamins Publishing Company (Studies in language and social interaction 26), 1–34.
- Dynel, Marta (2014): Participation framework underlying YouTube interaction. In: *Journal of Pragmatics* 73, 37–52. DOI: 10.1016/j.pragma.2014.04.001.
- Ehlich, Konrad (1984): Zum Textbegriff. In: Rothkegel, Anneli / Sandig, Barbara (Hrsg.): *Text – Textsorten – Semantik*. Hamburg: Buske, 9–25.
- Fabricius-Hansen, Cathrine (2000): Die Geheimnisse der deutschen Würde-Konstruktion. In: Thieroff, Rolf / Tamrat, Matthias / Fuhrhop, Nanna et al. (Hrsg.): *Deutsche Grammatik in Theorie und Praxis*. Berlin, Boston: de Gruyter.
- Frobenius, Maximiliane (2011): Beginning a monologue: The opening sequence of video blogs. In: *Journal of Pragmatics* 43/3, 814–827. DOI: 10.1016/j.pragma.2010.09.018.
- Garcia, Angela Cora / Jacobs, Jennifer Baker (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. DOI: 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. DOI: 10.1016/j.pragma.2014.08.011.

- Gibson, David. R. (2003): Participation Shifts: Order and Differentiation in Group Conversation. In: *Social Forces* 81/4, 1335–1380. DOI: 10.1353/sof.2003.0055.
- Giles, David / Stommel, Wyke / Paulus, Trena et al. (2015): Microanalysis Of Online Data. The methodological development of „digital CA“. In: *Discourse, Context & Media* 7, 45–51. DOI: 10.1016/j.dcm.2014.12.002.
- Goffman, Erving (1979): Footing. In: *Semiotica* 25/1–2, 1–30. DOI: 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 Harness (2004): Participation. In: Duranti, Alessandro (Hrsg.): *A companion to linguistic anthropology*. Malden, Mass.: Blackwell Pub (Blackwell companions to anthropology 1), 222–243.
- Haddington, Pentti / Mondada, Lorenza / Nevile, Maurice (2013): Being Mobile: Interaction on the Move. In: Haddington, Pentti / Mondada, Lorenza / Nevile, Maurice (Hrsg.): *Interaction and mobility. Language and the body in motion*. Berlin: De Gruyter (*Linguae & litterae* 20), 3–64.
- Hausendorf, Heiko (2013): On the interactive achievement of space –and its possible meanings. In: Auer, Peter / Hilpert, Martin / Stukenbrock, Anja et al. (Hrsg.): *Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives*. Berlin: de Gruyter (*Linguae & litterae* 24), 276–303.
- Heritage, John (2012): The Epistemic Engine: Sequence Organization and Territories of Knowledge. In: *Research on Language & Social Interaction* 45/1, 30–52. DOI: 10.1080/08351813.2012.646685.
- Hirschauer, Stefan (2014): Intersituativität. Teleinteraktionen und Koaktivitäten jenseits von Mikro und Makro. In: Heintz, Bettina / Tyrell, Hartmann (Hrsg.): *Interaktion – Organisation – Gesellschaft revisited. Anwendungen, Erweiterungen, Alternativen*. Stuttgart: Lucius & Lucius (*Zeitschrift für Soziologie, Sonderheft*), 109–133.
- Holly, Werner / Baldauf, Heike (2001): Grundlagen des fernsehbegleitenden Sprechens. In: Holly, Werner / Püschel, Ulrich / Bergmann, Jörg

- (Hrsg.): Der sprechende Zuschauer. Wie wir uns Fernsehen kommunikativ aneignen. Wiesbaden: Westdeutscher Verlag, 41–60.
- Horton, Donald / Strauss, Anselm (1957): Interaction in Audience-Participation Shows. In: *American Journal of Sociology* 62/6, 579–587. DOI: 10.1086/222106.
- Horton, Donald / Wohl, Richard (1956): Mass Communication and Para-Social Interaction. Observations on Intimacy at a Distance. In: *Psychiatry* 19/3, 215–229. DOI: 10.1080/00332747.1956.11023049.
- Hutchby, Ian (2014): Communicative affordances and participation frameworks in mediated interaction. In: *Journal of Pragmatics* 72, 86–89. DOI: 10.1016/j.pragma.2014.08.012.
- Irvine, Judith T. (1996): Shadow Conversations: The Indeterminacy of Participant Roles. In: Silverstein, Michael / Urban, Greg (Hrsg.): *Natural histories of discourse*. Chicago: Univ. of Chicago Press, 131–159.
- Kendrick, Kobin H. / Drew, Paul (2016): Recruitment: Offers, Requests, and the Organization of Assistance in Interaction. In: *Research on Language & Social Interaction* 49/1, 1–19. DOI: 10.1080/08351813.2016.1126436.
- Knorr Cetina, Karin (2009): The Synthetic Situation: Interactionism for a Global World. In: *Symbolic Interaction* 32/1, 61–87. DOI: 10.1525/si.2009.32.1.61.
- König, Katharina / Oloff, Florence (Hrsg.) (2019): Special Issue: Mobile Medienpraktiken im Spannungsfeld von Öffentlichkeit, Privatheit und Anonymität. In: *Journal für Medienlinguistik* 2/2.
- Levinson, Stephen C. (1988): Putting linguistics on a proper footing: Explorations in Goffman's participation framework. In: Drew, Paul / Wootton, Anthony (Hrsg.): *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. DOI: 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. DOI: 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 (2019): Mediale Durchformung. Fernsehinteraktion und Fernseh­mündlichkeit in Gesprächen im Fernsehen. In: Marx, Konstanze / Schmidt, Axel (Hrsg.): Interaktion und Medien. Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation. Heidelberg: Winter (OraLingua 17), 125–146.
- Luginbühl, Martin / Hauser, Stefan (2019): Journalistische Kulturen aus der Perspektive einer kulturanalytischen Medienlinguistik. In: Schröter, Juliane / Tienken, Susanne / Ilg, Yvonne et al. (Hrsg.): Linguistische Kulturanalyse. Berlin, Boston: de Gruyter, 347–370.
- Luginbühl, Martin / Schneider, Jan Georg (2020): Medial Shaping from the Outset. On the Mediality of the Second Presidential Debate, 2016. In: Journal für Medienlinguistik 3/1, 57–93. DOI: 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. DOI: 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. DOI: 10.1016/j.pragma.2017.03.001.
- Meredith, Joanne (2019): Conversation Analysis and Online Interaction. In: Research on Language & Social Interaction 52/3, 241–256. DOI: 10.1080/08351813.2019.1631040.
- 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. DOI: 10.1080/08351813.2018.1413878.
- Mondada, Lorenza (2019): Conventions for multimodal transcription (5.0.1). <<https://www.lorenzamondada.net/multimodal-transcription>> [zit. 30.10.2020].
- Moores, Shaun (2005): Media/theory. Thinking about media and communications. London: Routledge (Comedia).
- Nevile, Maurice / Haddington, Pentti / Heinemann, Trine et al. (2014): On the interactional ecology of objects. In: Nevile, Maurice / Haddington, Pentti / Heinemann, Trine et al. (Hrsg.): Interacting with Objects. Amsterdam: John Benjamins, 3–26.

- Paßmann, Johannes (2018): Die soziale Logik des Likes. Eine Twitter-Ethnografie. Frankfurt, New York: Campus.
- Recktenwald, Daniel (2017): Toward a transcription and analysis of live streaming on Twitch. In: *Journal of Pragmatics* 115, 68–81. DOI: 10.1016/j.pragma.2017.01.013.
- Reeves, Stuart / Greiffenhagen, Christian / Flintham, Martin et al. (2015): I'd Hide You. Performing Live Broadcasting in Public. In: Begole, Bo / Kim, Jinwoo / Inkpen, Kori et al. (Hrsg.): *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York: ACM, 2573–2582. DOI: 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. DOI: 10.1177/1461444817748954.
- Rossi, Giovanni / Floyd, Simeon / N. J. Enfield (2020): Recruitments and pragmatic typology. In: Floyd, Simeon / Rossi, Giovanni / Enfield, N. J. (Hrsg.): *Getting others to do things. A pragmatic typology of recruitments*. Berlin: Language Science Press (Studies in Diversity Linguistics 31), 1–23.
- Schegloff, Emanuel A. (2007): *Sequence organization in interaction. A primer in conversation analysis*. Cambridge: Cambridge Univ. Press.
- Schegloff, Emanuel A. / Sacks, Harvey (1973): Opening up Closings. In: *Semiotica* 8/4, 289–327. DOI: 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 (Hrsg.): *Deutsch in Sozialen Medien: de Gruyter*, 265–288.
- Schmitz, Ulrich (2016): Sehflächen. Sprache und Layout. In: Jürgen Schiewe (Hrsg.): *Angemessenheit. Einsichten in Sprachgebräuche*. Göttingen: Wallstein (Valerio 18), 24–37.
- Schneider, Jan Georg (2017): Medien als Verfahren der Zeichenprozessierung: Grundsätzliche Überlegungen zum Medienbegriff und ihre Relevanz für die Gesprächsforschung. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 18, 34–55.
- Selting, Margret / Auer, Peter / Barth-Weingarten, Dagmar et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, 353–402.

- Sjöblom, Max / Törhönen, Maria / Hamari, Juho et al. (2019): The ingredients of Twitch streaming: Affordances of game streams. In: *Computers in Human Behavior* 92, 20–28. DOI: 10.1016/j.chb.2018.10.012.
- Tannen, Deborah (2007): *Talking voices*. Cambridge: Cambridge University Press.
- Zinken, Jörg / Rossi, Giovanni (2016): Assistance and Other Forms of Cooperative Engagement. In: *Research on Language & Social Interaction* 49/1, 20–26. DOI: 10.1080/08351813.2016.1126439.

Links:

- [Deutsche Welle: Gewalt gegen Journalisten] <https://www.dw.com/de/gewalt-gegen-journalisten-wenn-twittern-zum-sicherheitsrisiko-wird/a-45435583> [zit. 30.10.2020].
- [IBM: History of IBM Watson Media] <https://blog.video.ibm.com/streaming-video-news/history-of-ibm-watson-media/> [zit. 30.10.2020].
- [Periscope: Kaul aus #Köthen, 18.06 Uhr.] <https://www.pscp.tv/w/1ZkKzNLzbjaKv> [zit. 30.10.2020].
- [Periscope: Kaul in #Köthen, 19.33 Uhr.] <https://www.pscp.tv/w/1OwxWWQmpOnxQ> [zit. 30.10.2020].
- [Tagesschau: Nach Todesfall in Köthen] <https://www.tagesschau.de/multimedia/video/video-448867.html> [zit. 30.10.2020].
- [taz: Tschüss, Periscope!] <https://taz.de/Aus-fuer-Livestream-App!/5757024/> [zit. 24.03.2021].
- [Twitter: Kaul aus #Köthen, 18.06 Uhr.] <https://twitter.com/martinkaul/status/1041357479353888768> [zit. 30.10.2020].
- [Twitter: Kaul in #Köthen, 19.33 Uhr.] <https://twitter.com/martinkaul/status/1041379614910689280> [zit. 30.10.2020].