



## **Smartphone-assistierte Geschichtenerzählungen**

Eine konversationsanalytische Untersuchung des Gebrauchs mobiler Endgeräte  
in alltäglichen Gesprächen

David Suderland

Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Philosophie  
der Universität Mannheim

Dekan: Prof. Dr. Philipp Gassert

Gutachterin: Prof. Dr. Angela Keppler

Gutachterin: Dr. Florence Oloff

Datum der Abgabe: 13.01.2021

## **Danksagung**

Zuerst möchte ich mich bei meiner Betreuerin Angela Keppler dafür bedanken, dass sie mir die Möglichkeit gegeben hat, diese Arbeit zu schreiben, mir alle Freiheiten gewährt hat, das Projekt eigenständig umzusetzen, und mich auf diesem Weg stets mit inspirierenden fachlichen Gesprächen, konstruktiven Hilfestellungen und viel menschlichem Feingefühl begleitet hat. Außerdem danke ich Florence Oloff für die anregenden Diskussionen in Datensitzungen, wertvolles inhaltliche Feedback zu Manuskriptentwürfen und die Möglichkeit, wichtige Vortragserfahrungen zu sammeln. Beiden danke ich, dass sie mir ihre Sichtweise auf und Neugierde für die Details alltäglicher Interaktionen vermittelt haben.

Diese Arbeit wäre ohne die Kooperationsbereitschaft der TeilnehmerInnen dieses Projekts nicht zustande gekommen – ich danke ihnen für die Offenheit, mir Einblicke in ihre Praxis des Smartphonegebrauchs und alltäglichen Gespräche zu gewähren. Ein Dank gilt außerdem den BetreiberInnen und MitarbeiterInnen der Cafés, die es mir ermöglicht haben, in ihren Räumlichkeiten zu filmen. Diese Arbeit hat zudem zum einen von dem lebhaften, fachlichen Austausch mit den TeilnehmerInnen der „Forschungswerkstatt Ethnomethodologische Konversationsanalyse“ an der Universität Bielefeld profitiert: Für die Organisation dieses Forums möchte ich Prof. Ruth Ayaß und Dr. Sarah Hitzler danken. Zum anderen möchte ich mich bei den Mitgliedern der „Institutsschreibgruppe“ – Dr. Frank Schneider, Annabell Halfmann und Sarah Lutz – für die regelmäßige ‚externe‘ Motivation zum Schreiben und die vielen hilfreichen Ratschläge bedanken.

Ich danke allen meinen KollegInnen, die das Promovieren am Institut für Medien- und Kommunikationswissenschaft der Universität Mannheim zu einem stets kollegialen, abwechslungsreichen und unterhaltsamen Unterfangen gemacht haben. Neben meinen Lehrstuhlkolleginnen Dr. Anja Peltzer, Maike Kohler, Madeline Dahl und Katrin Jacobi möchte ich mich dabei besonders bei zwei Personen bedanken: Zum einen bei Dr. Andreas Wagenknecht für das stets offene Ohr und die anregenden Gespräche, aber auch für die „Schreibstube“, die während der Lockdowns einen unschätzbaren Wert für mich hatte. Zum anderen bei Carolina Fenner für ihre intensive Lektüre und kritischen Kommentare in der finalen Schreibphase.

Von Herzen bedanken möchte ich mich bei meiner Familie, Maja, Detlev und Jakob, und meinen FreundInnen, die mir auf dem ganzen Weg eine große moralische Unterstützung waren. Der größte Dank geht an meine Frau Laura, die mir auch in den schwierigen Phasen Kraft gegeben und dabei stets an mich geglaubt hat – selbst wenn bei mir die Zweifel überwogen haben. Danke!

# **Inhaltsverzeichnis**

<b>1 Einleitung: Smartphone-assistierte Geschichtenerzählungen .....</b>	<b>1</b>
<b>1.1 Smartphonegebrauch in sozialen Situationen als Untersuchungsgegenstand.....</b>	<b>1</b>
<b>1.2 Aufbau der Arbeit.....</b>	<b>11</b>
<b>2 Theoretischer Hintergrund und Forschungsstand.....</b>	<b>14</b>
<b>2.1 Theoretische Ansätze der Erforschung mobiler Kommunikationstechnologien .....</b>	<b>15</b>
2.1.1 Mediensoziologische Zugänge zu mobilen Kommunikationstechnologien .....	15
2.1.2 Der medien- und kommunikationswissenschaftliche Mediatisierungsansatz .....	17
<b>2.2 Forschungsstand – Der Gebrauch mobiler Kommunikationstechnologien in sozialen     Situationen .....</b>	<b>20</b>
2.2.1 Der Gebrauch von Handys im öffentlichen Raum .....	22
2.2.2. Der individuelle Gebrauch mobiler Endgeräte in F2F-Interaktionen.....	24
2.2.3. Konvergente Handy- und Smartphonennutzung in F2F-Interaktionen .....	30
2.2.4 Zusammenfassung: konvergenter Smartphonegebrauch in F2F-Interaktionen.....	36
<b>3 Multimodale Konversationsanalyse als Forschungsansatz zur Untersuchung des Smartphonegebrauchs in F2F-Interaktionen .....</b>	<b>39</b>
<b>3.1 Die ethnomethodologischen Grundlagen der multimodalen KA.....</b>	<b>41</b>
<b>3.2 Die methodologischen Prämissen der multimodalen KA .....</b>	<b>46</b>
3.2.1 Datengrundlage – Technische Konservierungen natürlicher Interaktionen .....	47
3.2.2 Datenaufbereitung – Transkription audiovisueller Aufzeichnungen.....	49
3.2.3 Datenauswertung – die analytischen Maxime der multimodalen KA .....	53
3.2.4 Geschichtenerzählungen aus Perspektive der multimodalen KA.....	59
<b>3.3 Technologiegebrauch aus konversationsanalytischer Perspektive.....</b>	<b>64</b>
3.3.1 Die Ursprünge der konversationsanalytischen Untersuchung des Technologiegebrauchs .....	65
3.3.2 Gibsons wahrnehmungspsychologische Affordanztheorie .....	67
3.3.3 Die konversationsanalytische Adaption des Affordanzkonzepts .....	69
<b>4 Methodisches Vorgehen .....</b>	<b>74</b>
<b>4.1 Datenerhebung: Aufzeichnungen des ‚natürlichen‘ Smartphonegebrauchs in alltäglichen     Gesprächen .....</b>	<b>74</b>
<b>4.2 Reflektion der ‚Natürlichkeit‘ der aufgezeichneten Interaktionen.....</b>	<b>77</b>
<b>4.3 Datenkorpus und Kollektionsbildung .....</b>	<b>84</b>
<b>4.4 Datenauswertung: die Analyse von Kollektionen und Einzelfällen .....</b>	<b>88</b>

<b>5 Das Vorlesen von Textnachrichten in smartphone-assistierten</b>	
<b>Geschichtenerzählungen .....</b>	<b>92</b>
<b>5.1 Das Vorlesen von Textnachrichten in methodisch vorbereiteten Geschichtenerzählungen</b>	<b>99</b>
5.1.1 Das Vorlesen von Textnachrichten als konstitutives Element smartphone-assistierter	
Geschichtenerzählungen.....	100
5.1.2 Varianz der Teilnehmerahmen: Das simultane (Mit-)Lesen von Textnachrichten .....	108
5.1.3 Fazit: Praktiken des Vorlesens von Textnachrichten in methodisch vorbereiteten	
Geschichtenerzählungen.....	118
<b>5.2 Das Vorlesen von Textnachrichten in ‚unvorbereiteten‘ Geschichtenerzählungen und das</b>	
<b>Auftreten von interaktionalen Problemen .....</b>	<b>124</b>
5.2.1 Die ‚unvorbereite‘ smartphone-assistierter Geschichtenerzählungen und die Bearbeitung von	
Folgeproblemen .....	125
5.2.2 Der visuelle Zugriff zum Smartphonedisplay als Ressource für die Bearbeitung von interaktionalen	
Problemen.....	138
5.2.3 Fazit: ‚Unvorbereitete‘ Geschichtenerzählungen und die Bearbeitung interaktionaler Probleme ....	151
<b>5.3 „Stories in the making“: durch eingehende (Text-)Nachrichten veranlasste</b>	
<b>Geschichtenerzählungen.....</b>	<b>155</b>
5.3.1 Die Thematisierung von Textnachrichten an sequenziell passenden Positionen .....	158
5.3.2 „Eilmeldungen“ – disjunkt platzierte Thematisierungen eingehender Textnachrichten .....	172
5.3.3 Fazit: In der Entstehung befindliche, smartphone-assistierte Geschichtenerzählungen.....	184
<b>6 Die Integration digitaler Fotografien in smartphone-assistierten</b>	
<b>Geschichtenerzählungen .....</b>	<b>187</b>
<b>6.1 Bildgestützte, smartphone-assistierte Geschichtenerzählungen .....</b>	<b>194</b>
6.1.1 Die Eröffnung bildgestützter Geschichtenerzählungen durch informative Zeigesequenzen .....	196
6.1.2 Die Eröffnung bildgestützter Geschichtenerzählungen durch evokative Zeigesequenzen.....	207
6.1.3 Fazit: Bildgestützte, smartphone-assistierte Geschichtenerzählungen .....	217
<b>6.2 Bildgestützte, smartphone-assistierte Erwiderungen der ErzählerIn auf spaßhafte</b>	
<b>Provokationen.....</b>	<b>222</b>
6.2.1 An der Spaßmodalität orientierte, bildgestützte Erwiderungen.....	224
6.2.2 An der sachlichen Kritik orientierte, bildgestützte Erwiderungen .....	233
6.2.3 Fazit: Bildgestützte Erwiderungen auf spaßhafte Provokationen.....	248
<b>7 Fazit: Smartphone-assistierte Geschichtenerzählungen als multimodale Praktik .....</b>	<b>251</b>
<b>7.1 Zusammenfassung der zentralen Ergebnisse .....</b>	<b>251</b>
<b>7.2 Veränderungen des alltäglichen Erzählens durch die Integration mobiler Endgeräte.....</b>	<b>263</b>
<b>7.3 Resümee .....</b>	<b>269</b>
<b>Literaturverzeichnis .....</b>	<b>272</b>

# 1 Einleitung: Smartphone-assistierte Geschichtenerzählungen

## 1.1 Smartphonegebrauch in sozialen Situationen als Untersuchungsgegenstand

Das Smartphone ist für viele NutzerInnen<sup>1</sup> zu einer „digitalen Nahkörpertechnologie“ (Kaerlein, 2018) geworden – einem ständigen Begleiter, der in allen Lebenslagen mitgeführt wird. Schon lange sind die handlichen Endgeräte nicht mehr nur mobile Telefone, die Fernkommunikation mit Abwesenden über Telefonie und SMS ermöglichen. Es handelt sich vielmehr um kleine Computer, die unterschiedlichste zeit- und ortsunabhängige Formen der interpersonalen und massenmedialen Kommunikation ermöglichen: NutzerInnen verwenden verschiedenste Applikationen für Video- und Sprachtelefonie, für (vorwiegend) schriftsprachliche Fernkommunikation mittels Instant-Messaging, SMS und E-Mail, oder die vielfältigen, multi-medialen Kommunikationsangebote sozialer Netzwerke wie Facebook, Instagram oder TikTok. Darüber hinaus vereinen Smartphones mittlerweile eine Vielzahl nicht-kommunikativer Funktionen, die von Taschenlampe und Wecker über leistungsstarke Digitalkameras bis hin zu Streaming-Diensten für Musik oder Filme reichen.

Dabei ist es nicht nur das Smartphone selbst, das innerhalb der letzten 20 Jahre mit rasanter Geschwindigkeit weiterentwickelt wurde. Genauso haben sich die technische Infrastruktur des Internets, die allgemeine Netzabdeckung und die Verfügbarkeit mobiler Daten verändert. Während Anfang des Jahrtausends noch jede SMS einzeln abgerechnet wurde, so verfügen viele NutzerInnen mittlerweile über Flatrate-Tarife für ihre Smartphones, die unbeschränktes Telefonieren, Chatten und Surfen ermöglichen. Darüber hinaus können die mobilen Endgeräte mithilfe weniger Handgriffe mit in der Nähe befindlichen WLAN-Netzen verbunden werden. Auf diese Weise sind Smartphones zu Geräten geworden, mit denen es tatsächlich möglich ist „permanently online, permanently connected“ (Vorderer et al., 2017) zu sein, ohne dass finanzielle Kosten und Netzverbindung zu limitierenden Faktoren für die Nutzung werden.

Auf Grundlage dieser Entwicklungen ist der Gebrauch von Smartphones zu einem definierenden Merkmal unseres Lebens geworden, das viele Bereiche des Alltags durchdringt: So verbringen wir Wartezeiten am Bahngleis damit, über unser Handy Podcasts zu hören, während wir per WhatsApp mit Freunden Nachrichten austauschen. Wir wählen das Restaurant für ein Essen mit Freunden mithilfe von Google Maps über die Anzahl an 5-Sterne-Bewertungen aus

---

<sup>1</sup> Die Verwendung des ‚Binnen-I‘ in dieser Arbeit soll kenntlich machen, dass mit einem Begriff Menschen weiblicher, männlicher und diverser Geschlechtsidentitäten zugleich bezeichnet werden. Zur besseren Lesbarkeit wird im Singular immer der weibliche Artikel (bspw. „die NutzerIn“) verwendet und auf die zusätzliche Erwähnung des männlichen Artikels verzichtet. Wenn ich im empirischen Teil auf die TeilnehmerInnen dieser Studie Bezug nehme, verwende ich in Abhängigkeit der Geschlechtsidentität der konkreten Personen das Femininum oder Maskulinum („die Nutzerin“ bzw. „der Nutzer“).

und nutzen die gleiche Applikation, um den kürzesten Weg dorthin zu finden. Rechnungen bezahlen wir über unsere Banking-App, während wir zuhause auf der Couch eine Serie auf Netflix schauen. Und Urlaubsfotos werden noch vor Ort auf Instagram gepostet, um die eigenen Erlebnisse direkt mit persönlich bekannten und unbekanntem Followern zu teilen. Die Liste alltäglicher Aktivitäten, in die das Smartphone aufgrund der Vielzahl seiner Funktionen für viele NutzerInnen fest integriert ist, ließe sich fast beliebig fortsetzen. Dies bedeutet nicht nur, dass das Smartphone unsere Art, Bankgeschäfte zu erledigen oder uns zu orientieren, verändert hat, sondern auch, dass unterschiedliche Formen der durch das Smartphone vermittelten Kommunikation essenziell für die Pflege unserer sozialen Beziehungen geworden sind (bspw. Greschke et al., 2017; Ling & Yttri, 2002).

Aufgrund der vielfältigen Handlungsmöglichkeiten moderner Smartphones, ihrer zunehmenden Integration in alltägliche Aktivitäten sowie die damit einhergehende Ausbildung sozialer Erwartungen über wechselseitige Erreichbarkeit kommt es immer häufiger dazu, dass Smartphones von ihren BesitzerInnen auch im Kontext alltäglicher Interaktionen zwischen Anwesenden gebraucht werden: Uns allen sind vermutlich Situationen vertraut, in denen unsere GesprächspartnerIn kurz ein Anruf auf ihrem Handy annimmt, wir nebenbei eine WhatsApp-Nachricht beantworten oder schnell ein Erinnerungsfoto gemacht wird, bevor das gemeinsame Abendessen begonnen wird. Trotz dieser vermeintlichen Allgegenwart des Phänomens herrscht jedoch nach wie vor ein Mangel an empirisch fundiertem Wissen darüber, wie genau SmartphoneinhaberInnen<sup>2</sup> ihre mobilen Endgeräte in alltäglichen Situationen des sozialen Miteinanders verwenden und wie sich dies auf Face-to-Face-Interaktionen (im Folgenden „F2F-Interaktionen“) auswirkt (vgl. McGregor, 2020, S. 1). Gerade dieses empirische Wissen ist jedoch erforderlich, wenn die zentrale medien- und kommunikationswissenschaftliche Frage danach, wie die Handlungsangebote neuester Kommunikationstechnologien „die *Formen* der alltäglichen Kommunikation *modifizieren* oder gar grundlegend *verändern*“ (Keppler, 2013, S. 92, Herv. i. O.) beantwortet werden soll.

Dabei ist es nicht so, dass die gesellschaftlichen und sozialen Auswirkungen der zunehmenden Omnipräsenz mobiler Kommunikationstechnologien in unserem Alltag keinerlei Beachtung fänden: Sie werden sowohl im journalistischen als auch im wissenschaftlichen Diskurs angeregt diskutiert. Dabei geht es zum einen um die Auswirkungen auf das Individuum und zum anderen

---

<sup>2</sup> Mit den Begriffen „SmartphoneinhaberIn(nen)“ oder „GeräteinhaberIn(nen)“ werden im Folgenden die Personen bezeichnet, die das mobile Endgerät physisch manipulieren. Ich erkenne an, dass diese Bezeichnung potenziell missverständlich ist: Sie wird aus Gründen der besseren Lesbarkeit gebraucht, um die Verwendung von Relativsätzen, bspw. „InteraktionsteilnehmerIn, die das Smartphone bedient“, zu vermeiden. Auch wenn die Person, die ein Smartphone bedient, meist die BesitzerIn dieses Geräts ist, stellt dies keine notwendige Bedingung für die Verwendung des Begriffs in diesem Text dar.

um Veränderungen des gesellschaftlichen Miteinanders. Zumeist werden in den Diskussionen dabei die ‚problematischen‘ Dimensionen und Gefahrenpotentiale dieses Phänomens fokussiert. So informieren nationale und internationale Medienberichte über die negativen Konsequenzen der intensiven Nutzung mobiler Endgeräte für die psychische Gesundheit, das Konzentrationsvermögen oder die Schlafqualität individueller NutzerInnen<sup>3</sup>. Darüber hinaus wird auf teilweise reißerische Art auch vor den Gefahren der Selfie-Fotografie gewarnt, wenn berichtet wird, dass die fahrlässige Unvorsichtigkeit der Fotografierenden mittlerweile mehr Tote als Haiangriffe fordert<sup>4</sup>. An anderen Stellen werden die Risiken adressiert, die von sogenannten ‚Smombies‘ (Smartphone-Zombies) im Straßenverkehr ausgehen – HandynutzerInnen, die so sehr in die Nutzung ihres Telefons versunken sind, dass sie Gefahren in ihrer unmittelbaren Umwelt nicht mehr wahrnehmen<sup>5</sup>.

Neben den Risiken für die Gesundheit individueller NutzerInnen werden auch die Auswirkungen des Smartphonegebrauchs auf alltägliche Gespräche und soziale Beziehungen kritisch diskutiert. Insbesondere werden die LeserInnen dabei über die negativen Effekte des ‚Phubbings‘<sup>6</sup> aufgeklärt – diese Wortneuschöpfung aus ‚phone‘ und ‚snubbing‘ bezeichnet die Angewohnheit, sich mit dem eigenen Smartphone zu beschäftigen und dadurch anwesende GesprächspartnerInnen zu ignorieren. Es wird berichtet, dass sich diese zunehmend verbreitete Praktik negativ auf die Qualität unserer alltäglichen Unterhaltungen mit Anwesenden, der Verbundenheit mit unseren FreundInnen oder die Zufriedenheit in unseren romantischen Beziehungen<sup>7</sup> auswirkt. Diese alarmierenden Medienberichte, die in den kulturellen Praktiken des Smartphonegebrauchs eine Gefährdung unserer zwischenmenschlichen Beziehungen und alltäglichen Unterhaltungen vermuten, beruhen dabei oftmals auf den Ergebnissen medienpsychologischer

---

<sup>3</sup> Fiedler, M., Friebe, R., & Warnecke, T. (2019, Juni 23). Wie verändert uns das Smartphone? *Der Tagesspiegel*. <https://www.tagesspiegel.de/gesellschaft/panorama/staendig-online-wie-veraendert-uns-das-smartphone-/24483766.html>

<sup>4</sup> Kaiser, N. (2019, Januar 8). Wenn die Sucht nach Selfies tödlich endet. *WELT*. <https://www.welt.de/reise/Fern/article186675730/Gefahrliche-Posen-Wenn-die-Sucht-nach-Selfies-toedlich-endet.html>

<sup>5</sup> Lindner, M. (2018, Februar 12). Gefahr im Straßenverkehr: Hilfe, die Smombies kommen. *Augsburger Allgemeine*. <https://www.augsburger-allgemeine.de/augsburg/Gefahr-im-Strassenverkehr-Hilfe-die-Smombies-kommen-id44175991.html>

<sup>6</sup> *Phubbing* ist ein englisches Kofferwort aus ‚phone‘ (Telefon) und ‚snubbing‘, dem Gerundium von ‚to snub‘ (jmd. brüskieren). Es handelt sich um eine Wortneuschöpfung einer Werbeagentur, die das Wort für die Marketingkampagne eines australischen Wörterbuchverlags erfand, um einen angeblichen Trend im Umgang mit mobilen Telefonen zu beschreiben. Der Begriff fand auf diese Weise Eingang in den journalistischen und wissenschaftlichen Diskurs.

<sup>7</sup> Arnu, T. (2015, Juni 8). Der Feind in unserem Bett. *Süddeutsche Zeitung*. <https://www.sueddeutsche.de/leben/familie-partnerschaft-der-feind-im-bett-1.2504081>; Wagener, J. (2018). Warum deine Beziehung wahrscheinlich schlechter ist, wenn du viel am Handy hängst. *ze.tt*. Abgerufen 8. Januar 2021, von <https://ze.tt/wer-staendig-am-smartphone-haengt-ist-ungluecklicher-in-beziehungen/>

Arbeiten, die das Phänomen wissenschaftlich untersuchen (bspw. Chotpitayasunondh & Douglas, 2016, 2018; Roberts & David, 2016).

Innerhalb dieses medienpsychologischen Forschungsbereichs werden die Auswirkungen des Mediengebrauchs auf individuelle NutzerInnen und unterschiedlichste Sozialbeziehungen untersucht. Die „Phubbing“-Studien liefern Evidenz für den Schaden, den die routinierte Nutzung von Smartphones an unseren alltäglichen F2F-Interaktionen und sozialen Beziehungen anrichten kann: Die Verwendung mobiler Endgeräte im Kontext kopräsenter Interaktionen mit FreundInnen und Familienmitgliedern hat negative Auswirkungen auf die empfundene Verbundenheit der InteraktionspartnerInnen miteinander und die Gesprächsqualität (Chotpitayasunondh & Douglas, 2016, 2018; Przybylski & Weinstein, 2013). Die kulturelle Praxis der ständigen Smartphonennutzung führt nicht nur dazu, dass Menschen gemeinsame Zeit mit FreundInnen und Familie weniger genießen (Dwyer et al., 2018), sie wirkt sich auch negativ auf die Zufriedenheit in romantischen Paarbeziehungen aus (Roberts & David, 2016; Wang et al., 2017).

In der medienpsychologischen Verhaltensforschung wird ein überwiegend negatives Bild der Auswirkungen der Nutzung mobiler Endgeräte im Beisein relevanter Anderer gezeichnet<sup>8</sup>. Dass die Studien zu einem derartigen Urteil kommen, liegt nicht zuletzt an ihren Forschungsdesigns und Fragestellungen: Es werden Hypothesen zur Auswirkung des Mediengebrauchs auf bspw. das individuelle Wohlbefinden, Stress, Selbstkontrolle oder Normerfüllung aufgestellt und diese anschließend mithilfe von Surveys oder Experimenten überprüft. Damit beruhen die Ergebnisse auf retrospektiven Selbstauskünften der StudienteilnehmerInnen oder auf der Beobachtung von Verhalten unter experimentellen Bedingungen, die von der jeweiligen ForscherIn erschaffen wurden. Wird im Resultat eine negative Wirkung des Smartphonegebrauchs festgestellt, erscheinen die StudienteilnehmerInnen oftmals als NutzerInnen, die aufgrund ihrer ‚schlechten‘ Gewohnheiten und trotz ‚besserer‘ Absichten immer wieder in den Bann ihrer Smartphones gezogen werden (Aagaard, 2020).

Auf dieser Grundlage fehlt den kulturpessimistischen Einschätzungen des medialen und medienpsychologischen Diskurses somit eine hinreichende Erklärung dafür, warum sich die TeilnehmerInnen trotz aller (vermeintlichen) negativen Auswirkungen ständig erneut dafür entscheiden, ihr Smartphone in der Gegenwart von FreundInnen, Familienmitgliedern oder ArbeitskollegInnen zu nutzen und diese Interaktionen – jedem Alarmismus zum Trotz – im Alltag

---

<sup>8</sup> Diese Stimmen spiegeln somit kulturpessimistische Einschätzungen wider, die in vergleichbarer Form auch mit der Einführung von Handys – der Vorgängertechnologie moderner Smartphones – während des Jahrtausendwechsels geäußert wurden: Bereits zu diesem Zeitpunkt wurden Stimmen laut, die in der zunehmenden Verbreitung dieser neuen Kommunikationstechnologie eine essenzielle Gefahr für die Regeln unseres kulturellen Miteinanders sahen (bspw. Geser, 2005).

geordnet und unproblematisch verlaufen. Wissenschaftliche Studien, die untersuchen, wie mobile Endgeräte in alltäglichen Situationen und im Beisein anderer Anwesenden von den TeilnehmerInnen tatsächlich eingesetzt werden, finden sich jedoch im Feld der empirischen „Handyforschung“ (Eisentraut, 2016, S. 17). Anders als in der Medienpsychologie werden hier vor allem qualitative Studien durchgeführt, die von einer grundsätzlich ergebnisoffenen Haltung gekennzeichnet sind. Sie tragen nicht a priori zu überprüfende Annahmen über die Auswirkungen des Smartphonegebrauchs auf alltägliche Unterhaltungen an die Daten heran. Stattdessen wird mithilfe ethnografischer Beobachtungen, qualitativer Interviews und Audio- oder Videoaufnahmen untersucht, *wie* die mobilen Endgeräte im Alltag genutzt und diese Nutzungsweisen von den TeilnehmerInnen selbst wahrgenommen werden.

Dieser Forschungsbereich nimmt seine Ursprünge während der Jahrtausendwende in Reaktion auf die zunehmende Verbreitung von Handys. Zu Beginn werden zunächst vor allem Beobachtungsstudien durchgeführt, die in der Tradition der Soziologie Erving Goffmans untersuchen, welche Veränderungen der öffentlichen Interaktionsordnung mit der zunehmenden Nutzung von Mobiltelefonen an öffentlichen Orten einhergehen (bspw. Katz & Aakhus, 2002; Ling, 2005). In den letzten zehn Jahren entstanden mit der zunehmenden Verfügbarkeit leistungsstarker, mobiler Videokameras darüber hinaus eine zunehmende Anzahl von Studien, die aus Perspektive einer ethnomethodologisch geprägten, multimodalen Konversationsanalyse (im Folgenden ‚KA‘) den Gebrauch mobiler Kommunikationstechnologien in F2F-Interaktionen analysieren. Basierend auf audiovisuellen Aufzeichnungen alltäglicher Interaktionen kann hier im Detail rekonstruiert werden, wie die GesprächsteilnehmerInnen mit der Verwendung von Smartphones umgehen – in dezidiert multimodalen Analysen werden nicht nur die sprachlichen Äußerungen der TeilnehmerInnen untersucht, sondern das komplexe Zusammenspiel von Sprache, Mimik, Gestik und der Materialität der mobilen Endgeräte analysiert. So zeigen diese Studien, ob und wie sich die TeilnehmerInnen an Situationen des tatsächlichen Mediengebrauchs als etwas orientieren, das die Zeit mit anderen Anwesenden negativ beeinflusst, oder ob die mobilen Kommunikationstechnologien als Ressource für gemeinsame Aktivitäten eingesetzt werden und so das Gesprächsgeschehen beleben (vgl. Brown et al., 2015, S. 516; Keppler, 2013, S. 101).

Bis dato fokussieren diese Studien vor allem, wie die unterschiedlichen technischen Möglichkeiten der mobilen Endgeräte im Kontext von F2F-Interaktionen gebraucht werden. Im Ergebnis identifizieren sie diejenigen Verfahrensweisen, die von den TeilnehmerInnen eingesetzt werden, um den Gebrauch mobiler Endgeräte – im Zusammenspiel mit sprachlichen Äußerungen, Gesten und anderen verkörperten Handlungen – als sinnhaft verständliche Handlungen

hervorzubringen (vgl. McGregor, 2020, S. 5). Dabei rekonstruieren die Studien nicht nur, wie die TeilnehmerInnen die individuelle Nutzung des Smartphones zur Telekommunikation mit den (teils konkurrierenden) Anforderungen des Gesprächs zwischen den Anwesenden koordinieren (bspw. für SMS-Kommunikation während Gruppengesprächen: DiDomenico & Boase, 2013). Darüber hinaus werden auch neuartige und emergente Praktiken der kollaborativen Nutzung rekonstruiert, die eingesetzt werden, um die Handlungsmöglichkeiten des Smartphones gemeinschaftlich zu nutzen und sie so in Aktivitäten der F2F-Interaktion zu integrieren. Auf diese Weise konnte bspw. gezeigt werden, wie smartphone-gestützte Online-Suchen als gemeinsame Aufgabe der InteraktionsteilnehmerInnen hervorgebracht (Brown et al., 2015) werden, Zeigesequenzen digitaler Fotografien durch multimodale Adjazenzpaare organisiert (Oloff, 2019a) oder Textnachrichten sprachlich in die Interaktion integriert werden (Brown et al., 2018).

Anders als die eingangs beschriebenen kulturpessimistischen Einschätzungen argumentieren derartige konversationsanalytische Arbeiten, dass alltägliche Unterhaltungen in Kopräsenz nicht etwas sind, das durch den individuellen Gebrauch mobiler Endgeräte einseitig ‚beeinflusst‘ wird. Direkte mündliche Gespräche stellen vielmehr ein zentrales gesellschaftliches Forum für die Aneignung neuer und neuester (Kommunikations-)Technologien und Medien dar. In ihnen werden „intersubjektive Einstellungen überliefert und modifiziert, erprobt und erzeugt, auf die die Beteiligten in ihrem individuellen wie gemeinsamen Handeln zurückgreifen können“ (Keppler, 2013, S. 90). Diese Aushandlungsprozesse finden zum einen statt, indem die GesprächsteilnehmerInnen gemeinsam über neue Medien(inhalte) sprechen (Gehrau, 2019; Gehrau & Goertz, 2010). Zum anderen stellen F2F-Interaktionen zunehmend Kontexte des alltäglichen Mediengebrauchs dar, in denen die TeilnehmerInnen (un-)angemessene Formen der Smartphonennutzung erproben, aushandeln und etablieren. In diesem Prozess des alltäglichen Gebrauchs findet eine Aneignung der technischen und kommunikativen Möglichkeiten der mobilen Endgeräte statt. Dabei werden nicht nur die von den EntwicklerInnen intendierten Nutzungsweisen des *individuellen* Gebrauchs – bspw. zur mobilen Telefonie – etabliert, sondern auch neuartige und unvorhergesehene Praktiken entwickelt, mit denen das Smartphone in gemeinsame Aktivitäten der F2F-Interaktion integriert und *kollaborativ* verwendet wird (vgl. McGregor, 2020, S. 2).

Anhand der fruchtbaren Ergebnisse dieses Forschungsbereichs wird deutlich, dass mit der ethnomethodologischen Konversationsanalyse ein Forschungsansatz vorliegt, der spezifisch dafür geeignet ist, zu untersuchen, wie die TeilnehmerInnen an F2F-Interaktionen diese durch ihre (sprachlichen) Handlungen als sinnhaft ‚geordnet‘ hervorbringen (Bergmann, 1988a; Sacks,

1995). Das Ziel konversationsanalytischer Untersuchungen ist es, diejenigen Methoden zu identifizieren, mittels derer die TeilnehmerInnen „in und mit ihren Äußerungen und Handlungen die charakteristischen Strukturmerkmale und die ‚gelebte Geordnetheit‘ . . . des interaktiven Geschehens, in das sie verwickelt sind, hervorbringen“ (Bergmann, 2000, S. 528). Anhand des analytischen Interesses an den TeilnehmerInnen-Methoden zeigt sich dabei, dass die KA auf den theoretischen Grundannahmen der von Harold Garfinkel (1967) gegründeten Ethnomethodologie (im Folgenden „EM“) beruht.

In der EM wird soziale Ordnung als lokale Hervorbringungsleistung verstanden, die in und durch die wechselseitig aufeinander bezogenen Handlungen der TeilnehmerInnen (re-)produziert wird. Ein zentrales, methodologisches Argument der EM ist es dabei, dass soziale Handlungen von den Mitgliedern eines Settings mittels Ethnomethoden – verstanden als kulturell mehr oder minder verbreitete Verfahrensweisen – „accountable“, d. h. erkennbar, verstehbar, beschreibbar, berichtbar und erklärbar (vgl. Garfinkel, 1967, S. vii)<sup>9</sup> gemacht werden. Ziel der EM ist eine Identifikation eben dieser Ethnomethoden, mit denen die TeilnehmerInnen ihre Handlungen auf geordnete Weise produzieren, sodass ihr Sinn intersubjektiv verständlich wird. In wechselseitiger Bezugnahme aufeinander zeigen sich die TeilnehmerInnen im Vollzug von Interaktionen durch ihre sprachlichen Äußerungen und verkörperten Handlungen beständig an, wie sie das Verhalten der Anderen verstehen und wie sich selbst verstanden wissen wollen. Auf diese Weise wird der Sinn sozialer Handlungen interaktionsöffentlich hervorgebracht, sodass sich der Prozess der lokalen Hervorbringung sinnhafter Ordnung durch eine genaue Beobachtung des Verhaltens der TeilnehmerInnen rekonstruieren lässt.

Die KA macht sich diese Eigenschaft sozialer Handlungen nun bereits seit knapp 50 Jahren zunutze, um die Organisationsprinzipien sprachlicher Interaktionen basierend auf der Analyse technischer Aufzeichnungen möglichst natürlicher Interaktionen zu identifizieren. Dabei wurde bereits in der Entstehungsphase der KA anerkannt, dass soziale Interaktionen von den TeilnehmerInnen nicht nur durch den Einsatz sprachlicher Ressourcen hervorgebracht werden, sondern auf dem Zusammenspiel unterschiedlicher, wechselseitig miteinander verknüpfter Modalitäten – wie Sprache, Mimik, Gestik und anderen verkörperte Handlungen – beruhen (vgl. Stivers &

---

<sup>9</sup> Die in der Ethnomethodologie und KA verwendeten Begriffe sind dadurch geprägt, dass die ersten grundlegenden theoretischen Werke und empirischen Studien in englischer Sprache veröffentlicht wurden. Insbesondere die Wortschöpfungen Garfinkels lassen sich teilweise nur unzureichend übersetzen – nicht ohne Grund erschien erst 2020 die erste deutsche Übersetzung der *Studies in Ethnomethodology* (Garfinkel, 1967). Ich verwende in dieser Arbeit teilweise die englischen Fachtermini, die Eingang in die deutsche Fachsprache gefunden haben, z. B. *accountable*. Diese werden bei der ersten Erwähnung in Anführungszeichen eingeführt und danach kursiv gesetzt. Teilweise verwende ich auch die Übersetzungen, die sich in deutschsprachigen Publikationen eingebürgert haben. Wenn deren Bezeichnungen nicht eindeutig sind, werden sie als Übersetzungen durch Verweise auf den Originalausdruck und entsprechende Literaturangaben kenntlich gemacht.

Sidnell, 2005, S. 1). Mit der zunehmenden Verbreitung leistungsstarker und mobiler Videokameras entwickelte sich eine multimodale KA (Deppermann, 2013; Mondada, 2013a), die systematisch dieses Zusammenspiel unterschiedlicher Ressourcen in der interaktiven Hervorbringung sozialer Ordnung untersucht.

Die vorliegende Arbeit begann als Dissertationsprojekt mit dem Vorhaben, sich durch die detaillierte Untersuchung alltäglicher Situationen des Smartphonegebrauchs in Settings physischer Kopräsenz mit anderen Anwesenden in diesen Forschungsbereich der multimodalen KA einzuschreiben. Basierend auf Audio- und Videoaufzeichnungen zwangloser Unterhaltungen zwischen FreundInnen und Bekannten sollte das Wissen über den Umgang mit und die Auswirkungen der tatsächlichen Nutzung mobiler Kommunikationstechnologien in F2F-Interaktionen weiter vertieft werden. Aufgrund des ergebnisoffenen Zugangs und im Verlauf des iterativ-zyklischen Forschungsprozesses wurde meine Aufmerksamkeit dabei jedoch nicht auf die Nutzung einer konkreten Handlungsmöglichkeit des Smartphones (bspw. der Möglichkeit zur Online-Suche), sondern zunehmend auf eine spezifische kommunikative Aktivität alltäglicher Unterhaltungen gelenkt: Das spontane Erzählen von Geschichten, das als Methode zur gemeinsame Interpretation unserer alltäglichen Erlebnisse einen zentralen Stellenwert im kommunikativen Haushalt unserer Gesellschaft besitzt. Immer wieder lässt sich in meinem Datenkorpus beobachten, dass die TeilnehmerInnen sich gegenseitig Geschichten erzählen, in die mithilfe des Smartphones Medieninhalte integriert werden – sei es, indem gespeicherte Textnachrichten vorgelesen oder digitale Fotografien gezeigt werden. Der Gebrauch des Smartphones erfolgt hierbei nicht in Konkurrenz zur Geschichtenerzählung, die integrierten Medieninhalte stellen für die ErzählerInnen viel mehr eine konstitutive Ressource dar, um erzählwürdige Ereignisse ihres Lebens auf intersubjektiv nachvollziehbare Weise narrativ zu rekonstruieren. Ausgehend von dieser Beobachtung wurde die zentrale Fragestellung dieser Arbeit entwickelt: *Wie werden die Handlungsmöglichkeiten von Smartphones von den TeilnehmerInnen methodisch als Ressource für die interaktive Hervorbringung von Geschichtenerzählungen eingesetzt?* Damit verknüpft ist auch die Frage danach, *welche Funktionen die Integration der über die mobilen Endgeräte abrufbaren Medieninhalte für die narrative Rekonstruktionen übernehmen.* Die Analyse von derartigen „smartphone-assistierten Geschichtenerzählungen“ ist dabei in zweifacher Hinsicht geeignet, die medien- und kommunikationswissenschaftliche Frage danach, wie die Handlungsangebote neuester Kommunikationstechnologien „die *Formen* der alltäglichen Kommunikation *modifizieren* oder gar grundlegend *verändern*“ (Keppler, 2013, S. 92, Herv. i. O.) empirisch zu beantworten. Diese zwei Argumente sollen im Folgenden ausgeführt werden.

1. In der vorliegenden Arbeit werden keine beliebigen Veränderungen der Formen alltäglicher Kommunikation durch den Gebrauch mobiler Kommunikationstechnologien untersucht. Mit der Analyse der Smartphoneintegration in Geschichtenerzählungen wird viel mehr eine Aktivität alltäglicher Gespräche fokussiert, die einen zentralen Stellenwert im kommunikativen Haushalt unserer Gesellschaft besitzt: Das Erzählen von Geschichten stellt eine etablierte Lösung für das Problem der Flüchtigkeit sozialer Situationen dar (vgl. Luckmann & Bergmann, 1995, S. 293) und übernimmt eine essenzielle Rolle in der gemeinsamen Interpretation und Ausdeutung unseres sozialen Erlebens im Alltag.

Das Erzählen von Geschichten und Alltag stehen dabei in einem wechselseitigen Verhältnis zueinander: Zum einen „werden im alltäglichen Erzählen Erlebnisse, Begebenheiten, Inhalte der als Alltag erfahrenen Lebenswelt zum Gegenstand der narrativen Rekonstruktion gemacht“ (König & Oloff, 2018a, S. 278). Geschichtenerzählungen haben also „Alltag“ zum Thema, sie stellen ein zentrales Verfahren dar, mit dem wir Andere über wichtige, problematische oder amüsante Ereignisse unseres alltäglichen Lebens in Kenntnis setzen und gemeinsam ihre Bedeutung verhandeln. Zum anderen findet Erzählen oftmals eingebettet in alltägliche Situationen der TeilnehmerInnen statt: Es handelt sich um ungeplante Aktivitäten, die aus der Situation heraus und durch die interaktiven Beiträge aller Anwesenden hervorgebracht werden. Erzählen stellt also eine alltägliche Aktivität dar und „konstituiert somit selbst wiederum Alltag“ (König & Oloff, 2018a, S. 279). Die vorliegende Arbeit untersucht, wie Smartphones und über sie abrufbare Medieninhalte – im Zusammenspiel mit sprachlichen Äußerungen, Gesten und anderen verkörperten Handlungen – als Ressourcen für die Hervorbringung von Geschichtenerzählungen eingesetzt werden. Auf diese Weise wird zum einen eine kommunikative Aktivität analysiert, die einen zentralen Stellenwert für die Thematisierung und Ausdeutung der zunehmend medialen Erfahrungen unseres Alltags besitzt. Zum anderen wird so empirisch rekonstruiert, wie der Gebrauch der mobilen Kommunikationstechnologien im Kontext von Geschichtenerzählungen zur Konstitution eines Alltags beiträgt, der immer stärker von Medien durchdrungen ist.

2. Wenn die Frage nach den Veränderungen kommunikativer Aktivitäten durch die Nutzung mobiler Kommunikationstechnologien empirisch beantwortet werden soll, müssen die Organisationsprinzipien dieser Aktivität in ihrer ‚nicht-mediatisierten‘ Form bekannt sein: Bei Geschichtenerzählungen in alltäglichen Gesprächen handelt es sich um ein konversationsanalytisch gut erforschtes Phänomen. Eine Erzählung wird verstanden als „in Form einer Diskurs-einheit realisierte verbale Rekonstruktion eines Ablaufs realer oder fiktiver Handlungen oder

Ereignisse, die im Verhältnis zum Zeitpunkt des Erzählens zurückliegen“ (Gülich & Hausendorf, 2000, S. 373).

In unzähligen Untersuchungen wurde gezeigt, dass Geschichtenerzählungen durch die Beiträge *aller* InteraktionsteilnehmerInnen produziert werden, indem sie gemeinsam verschiedene Aufgaben der Interaktionsorganisation bearbeiten. Diese lassen sich wie folgt umreißen: Alltägliche Gespräche sind zunächst von einer prinzipiellen Konkurrenz um das Rederecht gekennzeichnet (Sacks et al., 1974). Da narrative Rekonstruktionen durch mehrere sprachliche Äußerungen – ein sogenanntes „big package“ (Sacks, 1995, Vol. II, S. 354) – realisiert werden, gilt es für die potenzielle ErzählerIn somit zunächst, ihren InteraktionspartnerInnen die Verfügbarkeit einer Erzählung anzuzeigen und sich von diesen ein „ticket“ (vgl. Sacks, 1971, S. 312) für ihre Produktion einzuholen. Diese Projektion von Erzählungen findet in spezifisch designten Prä-Sequenzen statt, die auch der Bearbeitung weiterer Aufgaben dienen: In ihnen wird das Verhältnis der Erzählung zum vorangegangenen Gespräch (bspw. als thematisch verbunden oder unverbunden) *accountable* gemacht (Jefferson, 1978, S. 220), die Erzählwürdigkeit des Ereignisses dargestellt sowie die Haltung der ErzählerIn zu den zu rekonstruierenden Ereignissen vorbereitet und den RezipientInnen somit ein bestimmtes Verständnis des Folgenden nahegelegt (Sacks, 1974, S. 341). In der Kernphase der Erzählung wird hauptsächlich adressatenspezifische Rekonstruktions- und Bewertungsarbeit geleistet (Kotthoff, 2017, S. 57). Durch diese wird eine Ereignisabfolge (meist) in chronologischer Abfolge dargestellt und die TeilnehmerInnen dieser Ereignisse als Figuren der Erzählwelt animiert (vgl. Couper-Kuhlen, 2007, S. 81). Abschließend muss die ErzählerIn für ihre ZuhörerInnen *accountable* machen, dass die Erzählung ihr Ende erreicht hat. Diesen kommt anschließend die Aufgabe zu, durch die Produktion thematisch relevanter Erwidierungen anzuzeigen, dass sie die Bedeutung der Geschichte verstanden haben (Jefferson, 1978).

Vor dem Hintergrund dieser umfangreichen Erkenntnisse über die sequenzielle Organisation von Geschichtenerzählungen und die methodischen Verfahrensweisen, die zu ihrer Hervorbringung eingesetzt werden, können die Veränderungen dieser Aktivität, die mit der Integration des Smartphones einhergehen, eindeutig empirisch bestimmt werden. Darüber hinaus ermöglicht es der vorhandene Forschungsstand, das Funktionspotential der integrierten Medieninhalte zu identifizieren, d.h. zu rekonstruieren, zur Lösung welches Problem der Interaktionsorganisation sie eingesetzt werden.

## 1.2 Aufbau der Arbeit

Die Frage danach, wie und mit welcher Funktion Medieninhalte über das Smartphone in alltägliche Geschichtenerzählungen integriert werden, erfolgt dabei in folgenden Schritten: In den Kap. 2 bis 4 werde ich einen analytischen Rahmen für die anschließende Ergebnisdarstellung schaffen. In Kap. 2.1. gilt es zunächst, das Verhältnis dieser Arbeit zu mediensoziologischen und mediatisierungstheoretischen Perspektiven auf den Gebrauch mobiler Kommunikationstechnologien in alltäglichen Interaktionen auszuloten. Dabei werde ich aufzeigen, warum beide Forschungsbereiche nicht unmittelbar für diese Arbeit anschlussfähig sind. Anschließend werde ich in Kap. 2.2. den empirischen Forschungsstand zur Nutzung mobiler Endgeräte im Kontext kopräsender Situationen darstellen – dabei wird die Arbeit im Feld der empirischen „Handyforschung“ (Eisentraut, 2016, S. 17) verortet. In groben Zügen wird die Entwicklung dieses Forschungsfeldes nachgezeichnet: Um die Jahrtausendwende wurden zunächst zahlreiche, von der Soziologie Goffmans geprägte, ethnografische Studien durchgeführt, die die Auswirkungen des Handygebrauchs auf die öffentliche Interaktionsordnung und das Führen mobiler Telefonate in sozialen Situationen untersuchten. Seit einiger Zeit werden nun konversationsanalytische Untersuchungen des individuellen und gemeinsamen Smartphonegebrauchs in F2F-Interaktionen basierend auf Audio- und Videoaufzeichnungen alltäglicher Unterhaltungen durchgeführt. Der Fokus der Darstellung liegt darauf, die vorhandenen Erkenntnisse zum Gebrauch von Smartphones als Ressource für die Produktion sozialer Handlungen in F2F-Interaktionen zusammenzufassen und dabei das Desiderat zu identifizieren, das mit der Untersuchung *smartphone-assistierter Geschichtenerzählungen* bearbeitet wird: Anders als vorhandene Studien, untersucht die vorliegende Arbeit die Integration *unterschiedlicher* Nutzungsmöglichkeiten des Smartphones in *eine spezifische* kommunikative Aktivität und nicht den Umgang mit einer spezifischen Nutzungsmöglichkeit im allgemeinen Kontext alltäglicher F2F-Interaktionen.

In Kap. 3 wird die multimodale KA als Forschungsansatz vorgestellt. Sie vereint die theoretischen Grundannahmen der EM mit spezifisch für die Erforschung von F2F-Interaktionen entwickelten methodologischen Prämissen und analytischen Maximen. Aus diesem Grund dient dieses Kapitel sowohl dazu, das dieser Arbeit zugrundeliegende, theoretische Grundverständnis sozialer Ordnung zu explizieren, als auch die Forschungslogik der KA zusammenzufassen. Nach einer kurzen Übersicht über den Forschungsstand zu Geschichtenerzählungen wird darüber hinaus die konversationsanalytische Adaption des ursprünglich wahrnehmungspsychologischen Konzepts der „Affordanzen“ (Gibson, 1977) vorgestellt. Der Begriff bezeichnet die materiell verfestigten Bedingungen der (Handlungs-)Möglichkeiten von technologischen

Artefakten. Durch seine konversationsanalytische Adaption wird ein analytischer Rahmen für die empirische Untersuchung der Konsequenzen des Technologiegebrauchs auf die sequenzielle und zeitliche Entwicklung von Geschichtenerzählungen geschaffen.

In Kap. 4 werde ich mein eigenes methodisches Vorgehen vorstellen und offenlegen, auf welche Weise Aufzeichnungen alltäglicher Interaktionen angefertigt, die Daten aufbereitet und ausgewertet wurden. Spezifisch wird dabei auf die Herausforderungen der Kollektionsbildung in der multimodalen Analyse *smartphone-assistierter Geschichtenerzählungen* eingegangen.

Kap. 5 und 6 stellen das Kernstück dieser Arbeit dar und präsentieren die Ergebnisse der multimodalen Analyse *smartphone-assistierter Geschichtenerzählungen*. Der Analyseteil gliedert sich in zwei Teile, in denen die Integration unterschiedlicher Medieninhalte – zunächst von Textnachrichten und anschließend von digitalen Fotografien – fokussiert wird: In Kap. 5 werden Geschichtenerzählungen diskutiert, die konstitutiv durch Praktiken des Vorlesens von Textnachrichten hervorgebracht werden. In den Kap. 5.1. und 5.2. steht dabei das Smartphone als „Archiv sozialer Originale“ (vgl. Keppler, 2013, S. 99) im Fokus, das es den TeilnehmerInnen jederzeit ermöglicht, gespeicherte Nachrichtenverläufe aufzurufen und in all ihren Details in die Interaktion einzuspielen. In Kap. 5.1. werden methodisch vorbereitete, *smartphone-assistierte Geschichtenerzählungen* untersucht, die durch die Hervorbringung einer Prä-Sequenz eindeutig vom vorangegangenen „turn-by-turn talk“ (Jefferson, 1978, S. 220) abgegrenzt werden und anschließend durch das Vorlesen von Textnachrichten ausgestaltet werden. In diesem Kapitel wird zum einen rekonstruiert, welche spezifischen Aufgaben der Interaktionsorganisation in der Prä-Sequenz *smartphone-assistierter Geschichtenerzählungen* bearbeitet werden müssen. Zum anderen werden die konstitutiven multimodalen Praktiken identifiziert, mit denen diese Geschichtenerzählungen als solche *accountable* gemacht werden – die Analysen anschließender Kapitel bauen wiederum auf einem Verständnis dieser konstitutiven Praktiken auf. In Kap. 5.2. werden ‚unvorbereitete‘ Geschichtenerzählungen analysiert, die durch einen regulären Redebeitrag initiiert und „erst in der Folge zu einer Erzählung ausgearbeitet“ (König & Oloff, 2018a, S. 303) werden. Diese Variante zeigt sich anfällig für das Auftreten von Verstehensproblemen und Missverständnissen in der sequenziellen Abfolge. Die Analysen dieses Kapitels verdeutlichen somit zum einen, welche Aufgaben der Interaktionsorganisation für die *Accountability smartphone-assistierter Geschichtenerzählungen* bearbeitet werden müssen. Zum anderen zeigen sie, wie die TeilnehmerInnen das Smartphone als Ressource in der Reparatur von interaktionalen Problemen verwenden.

Im finalen Kap. 5.3. dieses Teils werden Geschichtenerzählungen untersucht, die durch den Empfang einer Textnachricht ausgelöst werden, wobei nun die Möglichkeit des Smartphones

zur orts- und zeitunabhängigen Telekommunikation eine zentrale Rolle spielt. In und durch diese spezifischen Geschichtenerzählungen werden noch nicht abgeschlossene Ereignisse rekonstruiert, die eine Relevanz für zukünftige Handlungsabfolgen der ErzählerIn besitzen.

Anschließend wird in Kap. 6 die smartphone-gestützte Integration digitaler Fotografien in Geschichtenerzählungen diskutiert. Dazu wird zunächst das dieser Arbeit zugrundeliegende Verständnis dieser visuellen Medieninhalte expliziert, bevor in Kap. 6.1. bildgestützte, *smartphone-assistierte Geschichtenerzählungen* untersucht werden, in denen die Integration einer oder mehrerer digitaler Fotografien eine konstitutive Rolle für die narrative Rekonstruktion vergangener Ereignisse übernimmt. Durch die Integration fotografischer Repräsentationen vergangener Eindrücke oder Handlungen werden ‚durch und durch‘ multimodale Erzählungen hervorgebracht, in denen eine Ereignisabfolge durch das Zusammenspiel von sprachlichen Äußerungen und visuellen Medieninhalten rekonstruiert wird. Die konstitutive Multimodalität dieser Variante *smartphone-assistierter Geschichtenerzählungen* fordert dabei auf besondere Weise das ‚hergebrachte‘ konversationsanalytische Verständnis von Erzählungen als verbalen Rekonstruktionen heraus.

Während in allen vorangehenden Kapiteln die Integration von Medieninhalten im Hauptteil von Geschichtenerzählungen untersucht wird, fokussiert Kap. 6.2. abschließend Erzählungen, in denen Fotografien erst in Erwiderung auf spezifische Äußerungen der RezipientInnen integriert werden. In diesem Kapitel wird diskutiert, wie von ErzählerInnen *smartphone-assistierte, bildgestützte Erwiderungen* auf „Teasings“ (Drew, 1987) – spaßhafte Provokationen – der RezipientInnen eingesetzt werden. Bei spaßhaften Provokationen handelt es sich um Äußerungen, mit denen ein Element der narrativen Rekonstruktion von den RezipientInnen als „abweichend“ – d. h. nicht ihren Normalitätserwartungen entsprechend – interpretiert wird. Derartige Provokationen machen es anschließend erforderlich, dass die ErzählerIn sich zu dieser Interpretation positioniert. Anhand der Analyse bildgestützter, *smartphone-assistierter Erwiderungen* in *Teasing*-Sequenzen lassen sich zum einen die grundsätzlichen Herausforderungen der Integration von Smartphones in als *turn-by-turn-talk* organisierten Sequenzen identifizieren. Zum anderen ist sie geeignet, zu rekonstruieren, inwiefern das Zeigen visueller Medieninhalte die RezipientInnen zur Teilnahme an einer eröffneten Sequenz (Alignierung) und inhaltlichen Gleichgerichtetheit mit dem Sprecher (Affiliation) leiten kann (Raclaw et al., 2016, S. 377).

Im abschließenden Kap. 7 werden zunächst die Ergebnisse dieser Arbeit zusammengefasst. Anschließend werden die zentralen Veränderungen der Formen des alltäglichen Erzählens durch den Gebrauch mobiler Kommunikationstechnologien reflektiert und ein Resümee formuliert.

## 2 Theoretischer Hintergrund und Forschungsstand

Die vorliegende Arbeit untersucht aus konversationsanalytischer Perspektive, wie und mit welcher Funktion Smartphones und die auf ihnen befindlichen Medieninhalte in Geschichtenerzählungen in alltäglichen Gesprächen integriert werden. Auf diese Weise werden die Veränderungen alltäglicher, sozialer Situationen durch den Gebrauch mobiler Kommunikationstechnologien anhand der exemplarischen Analyse der Smartphone-Nutzung in einer konkreten kommunikativen Aktivität – des alltäglichen Erzählens – untersucht. Mit der Einnahme einer konversationsanalytischen Perspektive verschreibt sich diese Arbeit zum einen – wie ich ausführlich in Kap. 3 darstellen werde – bestimmten theoretischen Grundannahmen: zuvorderst der Auffassung, soziale Wirklichkeit als „Vollzugswirklichkeit“ zu verstehen, „d. h. als eine Wirklichkeit, die lokal (also: vor Ort, im Ablauf des Handelns), endogen (also: in und aus der Handlungssituation), audiovisuell (also: durch Hören und Sprechen, durch Wahrnehmen und Agieren) in der Interaktion der Beteiligten erzeugt wird“ (Bergmann, 1981, S. 12). Zum anderen liegt ihr ein spezifischer, eng gefasster Medienbegriff zugrunde: Medien werden grundsätzlich als Technologien verstanden, „die sich dadurch auszeichnen, dass sie die Möglichkeiten der menschlichen Kommunikation erweitern. Sie treten zwischen zwei Akteure und ermöglichen so die Ausdehnung von Kommunikation über die Grenzen der unmittelbaren Interaktion hinaus“ (Bergmann, 2011, S. 13). Ausgehend von dieser Definition versteht diese Arbeit die Nutzung mobiler Endgeräte in F2F-Interaktionen als Mediengebrauch – sie untersucht jedoch eine spezifische Nutzungsweise, in der diese Technologien nicht zur Kommunikation mit Abwesenden, sondern als Erweiterung der Möglichkeiten der Kommunikation mit kopräsenten Anderen gebraucht werden. Der Einnahme dieser spezifischen Perspektive liegt die Auffassung zugrunde, dass alltägliche Gespräche in Kopräsenz ein zentrales Forum der Medienaneignung darstellen und die sozialen Wirkungen neuer Kommunikationstechnologien nur zu verstehen sind, wenn die Formen ihres Gebrauchs in alltäglichen Handlungszusammenhängen empirisch untersucht werden (Keppler, 2013, S. 90; vgl. a. McGregor, 2020, S. 2).

Ziel dieses Kapitels ist es, die vorliegende Arbeit im Kontext vorhandener theoretischer Ansätze und der Ergebnisse existierender empirischer Studien zu verorten. Dazu werde ich zunächst das Verhältnis dieser Arbeit zur Mediensoziologie (Kap. 2.1.1) sowie zum medien- und kommunikationswissenschaftlichen Mediatisationsansatz (Kap. 2.1.2) diskutieren. Anschließend wird sie im Feld der qualitativen Mobilkommunikationsforschung verortet (Kap. 2.2), in dem die Folgen der zunehmenden Allgegenwart mobiler Informations- und Kommunikationstechnologien auf den gesellschaftlichen Alltag empirisch untersucht werden.

## 2.1 Theoretische Ansätze der Erforschung mobiler Kommunikationstechnologien

### 2.1.1 Mediensoziologische Zugänge zu mobilen Kommunikationstechnologien

Bei der Konversationsanalyse handelt es sich ursprünglich um einen soziologischen Forschungsansatz (Bergmann, 2000, S. 525), der für die systematische Analyse sprachlicher Interaktion entwickelt wurde (Sacks, 1984). Neben alltäglichen Gesprächen in Körperpräsenz wurden mit Telefongesprächen (bspw. Schegloff, 1979, 2002) schon früh eine spezifische Form medial-vermittelter Kommunikation und später unterschiedlichste Kontexte des Technologiegebrauchs untersucht (bspw. Suchman, 1985; Heath & Luff, 2000). Aus diesem Grund erscheint eine Anbindung dieser Arbeit an eine *Soziologie der Medien* zunächst naheliegend.

Hinsichtlich der Entwicklung einer *Mediensoziologie* als eigenständige soziologische Disziplin ist jedoch – mindestens im deutschsprachigen Raum – festzustellen, dass es sich um einen stark fragmentierten Forschungsbereich handelt, in dem eine Vielzahl unterschiedlicher Theorien zur Anwendung kommt<sup>10</sup> und kaum dezidiert soziologische Medientheorien<sup>11</sup> existieren. Eine Ausnahme stellt der in der Systemtheorie verwendete Medienbegriff dar, der etwa Geld, Macht, Liebe, Kunst oder Wahrheit als symbolisch generalisierte Kommunikationsmedien (Luhmann, 1997, S. 316 ff.) versteht. Ausgehend von einem solch weit gefassten Medienverständnis könnten letztlich auch Sinnesorgane, Sprache oder Architektur (vgl. bspw. Ziemann, 2014, S. 18 ff.) als Medien begriffen werden: Eine derart verstandene *Mediensoziologie* verfügt zum einen nicht über klar umrissene Phänomenbereiche, die ihren Gegenstand bildeten, sondern „hätte die ganze Welt zu ihrem Thema“ (Bergmann, 2011, S. 14). Zum anderen verliert eine derartige Konzeption von Medien jeglichen Bezug zum alltäglichen Verständnis des Begriffs – und damit zu den primären Sinnstrukturen<sup>12</sup> der Alltagswelt der TeilnehmerInnen, die sich eine qualitative *Mediensoziologie* deutend erschließen möchte. Aus diesen Gründen wird dieses weite Begriffsverständnis für eine empirisch orientierte Medienforschung, wie sie hier verfolgt wird, als nicht anschlussfähig erachtet. Stattdessen wird die oben vorgestellte, enger gefasste Begriffsbedeutung verwendet und somit der Fokus auf Medien als Kommunikationstechnologien gelegt.

---

<sup>10</sup> Diese Theorienpluralität spiegelt sich auch im Aufbau mediensoziologischer Sammelbände wider, in denen theoretische Zugänge des Felds bspw. system- und poststrukturalistische Ansätze, Kritische Theorie, Cultural Studies, Handlungstheorien, praxeologische Perspektiven und die Akteur-Netzwerk-Theorie vorgestellt werden (vgl. Hoffmann & Winter, 2018).

<sup>11</sup> Neben der im Folgenden dargestellten, systemtheoretischen Konzeption liegen vereinzelt weitere Entwürfe eines mediensoziologischen Forschungsprogramms und eines eigenständigen soziologischen Medienbegriffs vor (vgl. Thiedeke, 2012; Ziemann, 2014). Aber auch diese erscheinen – aufgrund der starken Einflüsse einer systemtheoretischen Perspektive auf ihr Verständnis von Medien – nicht mit der hier verfolgten Perspektive kompatibel.

<sup>12</sup> Nach Schütz' (1971) Auffassung sind sozialwissenschaftliche Begriffe, Konzepte und Theorien ‚Konstrukte zweiter Ordnung‘, die den primären Sinnstrukturen der AkteurInnen als ‚Konstrukten erster Ordnung‘ aufrufen müssen. Diese Argumentation bildet eine grundsätzliche, methodologische Prämisse der qualitativen Sozialforschung (Bergmann, 2011, S. 23).

Wirft man einen Blick in unterschiedliche Herausgeberbände und Monografien, die die *Soziologie der Medien* bzw. Medien- und Kommunikationssoziologie aus ihren unterschiedlichen Perspektiven vorstellen (bspw. Neumann-Braun & Müller-Doohm, 2000; Jäckel, 2005; Hoffmann & Winter, 2018; Scholtz, 2020), so fällt zudem auf, dass die soziologische Aufmerksamkeit vor allen Dingen den klassischen Massenmedien (Fernsehen, Radio und Print-Journalismus) und weniger Medien der interpersonalen Kommunikation gilt. Auch wenn dies zum Teil dem Erscheinungsdatum der Werke geschuldet sein mag, so ist diese einseitige Fokussierung durchaus immer noch kennzeichnend für die *Mediensoziologie*, die sich aus diesem Grund dem Vorwurf ausgesetzt sieht, eigentlich eine „Massenmediensoziologie“ (Thiedeke, 2012, S. 13) zu sein.

Darüber hinaus lassen sich an der soziologischen Auseinandersetzung mit (Kommunikations-)Medien die häufig vertretenen, entweder einseitig technik- oder sozialdeterministischen Perspektiven kritisieren (vgl. Eisentraut, 2016, S. 9). So fragt bspw. Thiedeke (2012) kritisch, „[o]b die soziale Bedeutung von Medien soziologisch umfassend untersucht ist, wenn man nur danach fragt, was Medien mit Rezipienten oder Rezipienten mit Medien ‚machen‘“ (2012, S. 21). Beide Perspektive nehmen eine Engführung vor, die empirisch – und aus Perspektive der KA – nicht haltbar ist: Weder determinieren Medien aufgrund ihrer materiellen Eigenschaften oder Inhalte das Verhalten und die Einstellungen ihrer NutzerInnen; noch sind Medien durch die sozialen Aushandlungsprozesse der NutzerInnen beliebig form- oder interpretierbar (vgl. Hutchby, 2001a, S. 14–15).

Deswegen nimmt diese Arbeit eine ‚dritte‘ Position ein, aus der davon ausgegangen wird, dass in allen Situationen des Mediengebrauchs die „interne Logik, die Konfiguration ihrer spezifischen Hardware und Software . . . auf eine soziale Logik, das heißt auf unterschiedliche Konstellationen der Ausnutzung der von den jeweiligen Medien eröffneten Möglichkeiten“ (Keppler, 2013, S. 85, Herv. i. O.) trifft. Erst in diesem Aufeinandertreffen der materiell verfestigten Handlungsmöglichkeiten der Kommunikationstechnologien und den Organisationsprinzipien der jeweiligen sozialen Situationen entfalten Medien ihre sozialen Wirkungen und werden (interaktions-)öffentlich wirksam. Daraus folgt auch, dass theoretisch und empirisch sowohl die ‚Logik der Medien‘ als auch die ‚Logik des Sozialen‘ – hier die Organisationsprinzipien alltäglicher Gespräche – berücksichtigt werden müssen, um rekonstruieren zu können, wie Medien gebraucht werden und welche Auswirkungen ihre Nutzung auf das soziale Miteinander hat. Dies erfolgt in der vorliegenden Arbeit durch die Adaption einer *affordanz*theoretischen Perspektive, die in Kap. 3.3 ausführlich erläutert wird.

### 2.1.2 Der medien- und kommunikationswissenschaftliche Mediatisierungsansatz

Während die *Mediensoziologie* vor allem massenmediale Kommunikation fokussiert, findet eine theoretische Konzeption und empirische Untersuchung interpersonaler, medial-vermittelter Kommunikation im Alltag der NutzerInnen vor allem im benachbarten Fach der Medien- und Kommunikationswissenschaft statt. Hier wurde im Zuge der Auseinandersetzung mit der stetig zunehmenden Relevanz unterschiedlichster Formen medialer Kommunikation in unserem gesellschaftlichen Alltag mit dem *Mediatisierungsansatz* (bspw. Krotz, 1995) eine Meta-Theorie entwickelt, die den Anspruch verfolgt, „eine theoretische Fundierung bereitzustellen, welche hilft, die derzeitigen Prozesse des sozialen und medialen Wandels adäquat zu beschreiben, aber auch zu erklären“ (Hepp & Hartmann, 2010, S. 35). Dieser Ansatz fand im deutschsprachigen Raum insbesondere durch die Publikationen von Friedrich Krotz Verbreitung (1995, 2007, 2008; Krotz et al., 2014, 2017), er wird aber auch international in unterschiedlichen Ausprägungen diskutiert (bspw. Thompson, 1995; Schulz, 2004; Hjarvard, 2008; Lundby, 2009). Die verschiedenen Ausprägungen sollen hier nicht in ihren Einzelheiten<sup>13</sup> vorgestellt werden – es soll lediglich betont werden, dass ihnen gemeinsam ist, *Mediatisierung* als fortlaufende Verbreitung technischer Kommunikationsmedien in zeitlicher, räumlicher und sozialer Hinsicht (Krotz, 2007, S. 37 ff.) und als Metaprozess sozialen Wandels zu verstehen, der gemeinsam und verknüpft mit anderen Prozessen ‚der Moderne‘ wie Individualisierung und Globalisierung existiert (vgl. Krotz, 2007, S. 12). „Metaprozess“ meint dabei,

dass Mediatisierung weder räumlich noch zeitlich noch in seinen sozialen und kulturellen Folgen begrenzt ist und dass auch die Konsequenzen dieser Entwicklung nicht als getrennt zu untersuchende Folge verstanden werden können, sondern einen konstitutiver [*sic!*] Teil von Mediatisierung ausmachen. (Krotz, 2007, S. 12; vgl. a. Hepp & Hartmann, 2010, S. 12)

Kommunikativem Handeln wird dabei als Form des sozialen Handelns ein zentraler Stellenwert im Prozess der Mediatisierung zugemessen. Entsprechend wird davon ausgegangen, dass Medien nicht in erster Linie über ihre Inhalte auf Individuen und Gesellschaft wirken, sondern dadurch, wie sie menschliche Kommunikation strukturell beeinflussen. Die Folgen der Mediatisierung sind nicht auf technische Innovation reduzierbar, sondern bestehen in der Aneignung

---

<sup>13</sup> Es existieren Ansätze, die Mediatisierung als „long-lasting process, whereby social and cultural institutions and modes of interaction are changed as a consequence of the growth of the media’s influence“ (Hjarvard, 2008, S. 114) verstehen. Damit vertreten sie ein stärker institutionalistisches Verständnis und gehen davon aus, dass sich soziale und kulturelle Institutionen durch eine Anpassung an die Logiken ‚neuer‘ Medien verändern. Von diesen lassen sich Ausprägungen differenzieren, die aus einer sozialkonstruktivistischen Perspektive fragen, welche Rolle die spezifischen Eigenschaften von Medien in der kommunikativen Konstruktion von Wirklichkeit (Couldry & Hepp, 2016) übernehmen. Die vorliegende Arbeit grenzt sich insbesondere von den sozialkonstruktivistischen Varianten des Mediatisierungsansatzes ab, deren theoretische Grundannahmen stärkere Überschneidungen zur ethnomethodologisch-konversationsanalytischen Position (vgl. Kap. 3.1) aufweisen.

und Integration der Medien in den Alltag ihrer NutzerInnen. Diese Integration wiederum generiert veränderte Formen von Alltag, sozialen Beziehungen, Kultur und Gesellschaft (Krotz, 2007, S. 11–13).

Innerhalb des theoretischen Rahmens des *Mediatisierungsansatzes* entstanden eine Vielzahl empirischer Studien, die auch interpersonale medien-vermittelte Kommunikationsformen untersuchen, z. B. ihre Rolle in transnationalen Eltern-Kind-Beziehungen (Greschke et al., 2017) oder für die Vergemeinschaftung junger Menschen im Internet (Hepp et al., 2014). Trotz der Vielzahl existierender Fallstudien bleibt der *Mediatisierungsansatz* aufgrund seines umfassenden Erklärungsanspruchs allgemein – unklar ist, wie und wo Mediatisierung ihre Konkretisierung erfährt (vgl. Hartmann, 2010, S. 36). Der Konzeption von Krotz folgend wird sie in der Lebenswelt, d. h. im Alltag der AkteurInnen verortet – denn die alltägliche Kommunikation wird als Basis für den Prozess der Mediatisierung verstanden (Krotz, 2007, S. 17–18).

Innerhalb der Mediatisierungsforschung selbst wird insbesondere der Domestizierungsansatz als Lösung für die bemängelte Allgemeinheit und als Programm für die empirische Erforschung des Mediatisierungsprozesses vorgeschlagen (Hartmann, 2010, 2013). Domestizierung bezeichnet grundsätzlich den Prozess der Eingliederung (neuer) Medien-Technologien in vorhandene Routinen und Orte des Alltags. Während letztere zunächst in der Sphäre des Haushalts verortet wurden, werden inzwischen, insbesondere durch die Verbreitung mobiler Kommunikationstechnologien, auch andere Umfelder untersucht (Höflich, 2011, S. 17). Zentral für Domestizierungsprozesse sind „die Aushandlungen, die bei solchen Integrations- (und Interpretations-) Prozessen nötig sind – auf der individuellen und vor allem der sozialen Ebene; im Hinblick auf den Alltag, aber auch im Hinblick auf Werte, Geschichten u. a. mehr“ (Hartmann, 2010, S. 37). Dabei wird Domestizierung als Prozess auf der Mikro-Ebene verortet und als grundlegende Voraussetzung für die auf der Meta-Ebene stattfindende Mediatisierung verstanden (vgl. Hartmann, 2010, S. 41). Mit der Domestizierungsperspektive wird somit eine konkrete Verortung des Wie und Wo von Mediatisierungsprozessen geschaffen, die durch ethnografische Studien (Hartmann, 2013, S. 53–70) empirisch erforscht werden (können).

Das grundsätzliche Interesse des *Mediatisierungsansatzes* an der Entwicklung von Medien und den damit verbundenen Folgen für menschliche Kommunikation (Krotz, 2007, S. 12) scheint zunächst mit der in dieser Arbeit verfolgten Perspektive kompatibel. Der *Mediatisierungsansatz* kann hier jedoch lediglich als allgemeiner Rahmen dienen, vor dessen Hintergrund die Nutzung mobiler Endgeräte in alltäglichen Gesprächen als Aspekt eines allgemeinen gesellschaftlichen Entwicklungsprozesses erscheint: Denn auch wenn die Bedeutung der kommunikativen Aneignung und Integration von Medien in den Alltag ihrer NutzerInnen betont wird,

wird weder theoretisch noch empirisch konkretisiert, wie und wo Prozesse der Mediatisierung (empirisch) zu verorten sind. Auch der Ansatz, Domestizierungsprozesse als Basis der Mediatisierung auf der Mikro-Ebene empirisch zu erforschen, ist für diese Arbeit nicht anschlussfähig. Die Stärke des Domestizierungsansatzes besteht darin, in einer Längsschnittperspektive aufzuzeigen, wie (neue) Medien und Kommunikationstechnologien einen Platz in sozialen Settings und im kommunikativen Alltag der NutzerInnen einnehmen (oder verlieren). „Was in diesen Studien jedoch nicht eigens zum Gegenstand wird, sind die kommunikativen Prozesse selbst, in Form der beim Mediengebrauch stattfindenden oder durch ihn induzierten Gespräche“ (Keppler, 2018, S. 78). Der Domestizierungsansatz stellt weder theoretisches noch methodologisches Instrumentarium<sup>14</sup> bereit, das geeignet wäre, mit der Eigenlogik konkreter Situationen des Mediengebrauchs umzugehen und die interaktiven Prozesse der lokalen Hervorbringung sinnhafter Ordnung zu rekonstruieren. Wenn jedoch analysiert werden soll, welchen Beitrag mobile Kommunikationstechnologien im Vollzug der Hervorbringung sozialer Wirklichkeit leisten, dann muss das „Augenmerk auf die Formen ihres alltäglichen Gebrauchs und das heißt: ihrer alltäglichen Aneignung“ (Keppler, 2013, S. 90) gelegt werden.

Wie ich in Kap. 3 ausführlich darstellen werde, bietet die ethnomethodologische KA sowohl ein theoretisches als auch methodologisches Fundament, um zu analysieren, wie und mit welcher Funktion Smartphones in alltägliche F2F-Interaktionen methodisch integriert werden. Durch eine systematische, konversationsanalytische Untersuchung alltäglicher Gesprächssituationen lässt sich „zu einer empirischen Bestimmung dessen . . . gelangen, wie technische Kommunikationsmedien . . . in die alltäglichen Formen des diskursiven Geschehens systematisch integriert werden und zu welchen Veränderungen des interaktiven Verhaltens dies führt“ (Keppler, 2013, S. 94).

---

<sup>14</sup> Ein Blick in Hartmann (2013) zeigt bspw., dass als Methode der Domestizierungsforschung lediglich ethnografisches Forsuchen mithilfe von Interviews, Mediennutzungstagebüchern und visuellen Materialien vorgestellt wird (vgl. Hartmann, 2013, S. 53–70) und keinerlei interaktionstheoretische Ansätze diskutiert werden (bspw. symbolischer Interaktionismus, Pragmatismus, oder ethnomethodologische KA), die ein grundsätzliches, theoretisches Fundament für die Analyse sozialer Interaktion böten.

## 2.2 Forschungsstand – Der Gebrauch mobiler Kommunikationstechnologien in sozialen Situationen

In den vorangegangenen Ausführungen wurden theoretische und programmatische Perspektiven der Mediensoziologie sowie der Medien- und Kommunikationswissenschaft skizziert und eine (kritische) Verortung der vorliegenden Arbeit vorgenommen. Im Folgenden möchte ich den Fokus auf die empirische Untersuchung der Nutzung mobiler Kommunikationstechnologien in alltäglichen, sozialen Situationen legen, um das konkrete Desiderat zu identifizieren, das mit der vorliegenden Untersuchung *smartphone-assistierter Geschichtenerzählungen* bearbeitet wird.

Wie bereits erwähnt, findet die empirische Untersuchung interpersonalen, medien-vermittelter Kommunikation vorwiegend im Bereich der (internationalen) Medien- und Kommunikationswissenschaft statt. Dort entwickelte sich in den 1990er-Jahren mit der zunehmenden Verbreitung mobiler Kommunikationstechnologien – zunächst ausgehend von ForscherInnen aus skandinavischen Ländern – eine Art interdisziplinäre „Handyforschung“ (Eisentraut, 2016, S. 11); ein fruchtbares Forschungsfeld, das in den letzten zwei Jahrzehnten maßgeblich von ForscherInnen wie Richard Ling (bspw. 1997, 2004, 2008) oder Joachim Höflich (bspw. 1996, 2011, 2016) vorangetrieben wurde. Im Rahmen zahlreicher Monographien (bspw. Goggin, 2006; Kopomaa, 2000; Ling, 2004, 2008; Weber, 2008) und Sammelbände (bspw. Arndt et al., 2010; Glotz et al., 2005; Goggin, 2013; Höflich et al., 2010; Höflich & Hartmann, 2006; Katz & Aakhus, 2002; Nyiri, 2005; Wimmer & Hartmann, 2014) sowie unzähliger Aufsätze – insbesondere im Journal *Mobile Media & Communication* – werden die sozialen Veränderungen thematisiert, die mit der zunehmenden Einbettung mobiler Kommunikationstechnologien in unseren Alltag einhergehen.

Eine lückenlose Darstellung dieses umfangreichen und heterogenen Forschungsstands ist hier weder möglich, noch wäre sie zielführend, denn analytisch relevant sind für die vorliegende Arbeit diejenigen Studien, die aus soziologischer Perspektive den Zusammenhang zwischen dem Gebrauch mobiler Kommunikationstechnologien und alltäglichen Interaktionen untersuchen. Mit dieser Entscheidung werden Untersuchungen angrenzender Fächer ausgeschlossen, die sich zwar durchaus mit dem Gebrauch mobiler Endgeräte im Alltag auseinandersetzen, aber in ihrer Forschung von grundsätzlich anderen Prämissen geleitet werden. Dies betrifft bspw. die Medienpsychologie, die die motivationalen Voraussetzungen der Mediennutzung sowie die kognitiven, emotionalen und verhaltensmäßigen Wirkungen von Medien untersucht; die Medienpädagogik mit ihrer (normativen) Fokussierung auf Medienbildung und -erziehung sowie die

Usability-Forschung mit ihrem anwendungsbezogenen Interesse an der Benutzerfreundlichkeit von Informations- und Kommunikationstechnologien.

Die Veränderungen von Interaktionen durch den Gebrauch mobiler Kommunikationstechnologien betreffen wiederum drei verschiedene Phänomenbereiche, für die sich jeweils eigenständige Untersuchungsfelder der Mobilkommunikationsforschung herausgebildet haben: 1.) Die durch Handys und Smartphones ermöglichten Formen der Telekommunikation, 2.) der Gebrauch mobiler Endgeräte im öffentlichen Raum, und 3.) der Gebrauch mobiler Endgeräte in F2F-Interaktionen. Die sozial- und kulturwissenschaftlichen Untersuchungen des ersten Forschungsbereichs nehmen ihren Ausgangspunkt bei einer Grundfunktion mobiler Kommunikationstechnologien – der Ermöglichung ortsunabhängiger Telekommunikation mit körperlich Abwesenden. Die empirischen Analysen der durch mobile Endgeräte ermöglichten Formen der medial-vermittelten Interaktion sind für die vorliegende Arbeit nur mittelbar relevant. Deshalb werde ich nur eine skizzenhafte Übersicht über die untersuchten Phänomene geben, bevor ich mich anschließend den anderen beiden Forschungsbereichen ausführlicher zuwende.

Die im Bereich der Telekommunikationsforschung untersuchten Phänomene sind so vielfältig wie die über die mobilen Endgeräte ermöglichten Kommunikationsformen. Sie reichen von mobilen Telefonaten (bspw. Laurier, 2001; Weilenmann, 2003), über textbasierte, asynchrone Formen der Kommunikation, wie SMS (bspw. König, 2015; Schmidt & Androutsopoulos, 2004) oder WhatsApp (bspw. König, 2019; Wyss & Hug, 2016) bis hin zur Kommunikation über digitale Fotografien (bspw. Koskinen, 2004; Koskinen et al., 2002; Meyer & Meier zu Verl, 2017; Oksman, 2005) oder Sprachnachrichten (Hector, 2017; König & Hector, 2017). Gemeinsam ist den Studien dieses Forschungsbereichs, dass sie die Formen der handyvermittelten Telekommunikation als Phänomene ‚in eigenem Recht‘ (vgl. Goffman, 1964, S. 134) untersuchen und diejenigen Praktiken identifizieren, mit denen die TeilnehmerInnen ihre kommunikativen Zwecke in medial-vermittelten, alltäglichen Interaktionen verfolgen.

Darüber hinaus zeigen derartige Studien auf, wie handyvermittelte Formen der Telekommunikation die Pflege familiärer oder freundschaftlicher Beziehungen der NutzerInnen über räumliche Distanzen hinweg ermöglichen (bspw. Döbler, 2014; Greschke, 2014; Ling, 2008; Rivière, 2006, S. 126–177). Mit der Einbettung der Mobilkommunikation in die sozialen Beziehungen der NutzerInnen entwickeln sich sowohl verstärkte Kontrollmöglichkeiten (bspw. Burkart, 2000a, S. 210, 224; Eisentraut, 2016, S. 100–111), aber vor allem wechselseitige Erwartungen über die ständige Erreichbarkeit der HandynutzerInnen (bspw. Burkart, 2000b, S. 225). Auch wenn diese Untersuchungen damit keine unmittelbare Relevanz für diese Arbeit besitzen, so erlauben die Erkenntnisse über Praktiken der handyvermittelten Telekommunikation und

Erwartungen an die GeräteinhaberInnen jedoch mittelbar ein besseres Verständnis des beobachtbaren Smartphonegebrauchs in F2F-Interaktionen. Auf diese Weise tragen Studien aus diesem Forschungsbereich zu einer theoretischen Sensibilisierung für den Untersuchungsgegenstand bei und werden meine Analyse vereinzelt im Hintergrund bereichern.

### **2.2.1 Der Gebrauch von Handys im öffentlichen Raum.**

Eben diese Ausbildung wechselseitiger Erwartungen über die ständige Erreichbarkeit der NutzerInnen mobiler Endgeräte hängt eng mit den im zweiten Untersuchungsbereich der Mobilkommunikationsforschung analysierten Phänomenen zusammen: Die zunehmende Sicht- und Hörbarkeit von mobilen Telefongesprächen in der Öffentlichkeit sorgte insbesondere in der ersten Phase der gesellschaftlichen Verbreitung von Mobiltelefonen für ein reges wissenschaftliches Interesse. Es wurden zahlreichen Studien durchgeführt, in denen die neuen räumlichen und zeitlichen Kontexte der handygestützten Telekommunikation fokussiert wurden. Der Gebrauch von Mobiltelefonen wurde in dieser Phase – sowohl von den ForscherInnen als auch den StudienteilnehmerInnen – als mit den üblichen sozialen Normen und Verhaltensweisen in der Öffentlichkeit konfligierend beschrieben (vgl. Oloff, 2019b, Z. 91–94)<sup>15</sup>. In ethnografischen, zumeist auf Beobachtungen und Interviews basierenden Studien wurde entsprechend der Frage nach beobachtbaren Veränderungen der öffentlichen Interaktionsordnung<sup>16</sup> durch die Integration von Mobiltelefonen in den kommunikativen Alltag ihrer NutzerInnen nachgegangen (bspw. Burkart, 2000a; Ling, 2002, 2005).

Veränderungen der öffentlichen Interaktionsordnung werden zum einen hinsichtlich der Aufmerksamkeit konstatiert, die sich einander fremde Personen in der Öffentlichkeit schenken. Durch die Möglichkeiten und Anforderungen mobiler Kommunikationstechnologien entsteht zunehmend eine „challenge of the absent presence“ (Gergen, 2002): An die Stelle einer höflichen Nichtbeachtung („civil inattention“, Goffman, 1963, S. 83 ff.) tritt aufgrund der ständigen medialen Verbindung zu Abwesenden zunehmend eine versunkene Unaufmerksamkeit und vollständige Nichtbeachtung Anderer, bspw. von Mitfahrenden in öffentlichen Verkehrsmitteln (Höflich et al., 2014, S. 34–35; Ito & Okabe, 2005). Diese mit der Handynutzung einhergehenden Aufmerksamkeitsverluste werden meist als (unintendierte) Effekte einer dauerhaften Nutzung der mobilen Kommunikationstechnologien verstanden. Wie neuere Untersuchungen

---

<sup>15</sup> Zum Zeitpunkt der Fertigstellung dieser Arbeit ist Oloff (2019b) ausschließlich als Diskussionspapier im Journal für Medienlinguistik zugänglich, für das statt Seitenzahlen nur Zeilenangaben verfügbar sind.

<sup>16</sup> Ursprünglich von Erving Goffman (1974) für die US-amerikanische Gesellschaft der zweiten Hälfte des 20. Jahrhunderts beschrieben wurde –

zeigen, können mobile Medien jedoch auch strategisch zur Minimierung der eigenen Adressierbarkeit in der Öffentlichkeit verwendet werden, um so eine Involvierung in Interaktionen mit Anderen zu verhindern (Ayaß, 2014; Eisentraut, 2016, S. 210).

Ein zweites zentrales Moment dieser Forschung geht von dem akustischen Eindringen des Handygebrauchs in den öffentlichen Raum aus, das durch das Klingeln der Geräte und hörbare Führen von Telefonanten entsteht. Aufgrund dieser „intrusive nature“ (Ling, 2004, S. 123–143) von Mobiltelefonen kommt es zu einer Privatisierung der Öffentlichkeit (Fortunati, 2002; Harper, 2001; Höflich, 2006): Telefongespräche, die vormals im Privaten stattfanden, werden nun an öffentlichen Orten geführt. Sie drängen sich als „acts of unreciprocated communication“ (Katz, 2011, S. 44) ‚unüberhörbar‘ in das akustische Wahrnehmungsfeld Anderer und ‚nötigen‘ unbeteiligte Dritte, private Gespräche fremder Personen mitzuhören. Auf Seiten der Telefonierenden führt dies zu einer potenziell unerwünschten Preisgabe und für die Mithörenden zu einem – ebenso unerwünschten – Erhalt von Informationen (Ling, 2008, S. 93–95). Dieser Umstand macht eine Aushandlung von Regeln und Normen des angemessenen Umgangs mit Telefonanten (Ling, 1997; Nardi & O’Day, 1999), sowie das Erlernen neuer Praktiken des Telefonierens vor Publikum (Burkart, 2007, S. 23; Murtagh, 2002) – bspw. durch das Aufsuchen eines hörgeschützten Raumes – erforderlich.

Die beobachtbare Varianz im Umgang mit öffentlichen Telefongesprächen ist dabei sowohl mit kulturellen und situativen Unterschieden, als auch den Einstellungen der TeilnehmerInnen verbunden: Mobiltelefonate werden grundsätzlich eher in transitorischen Räumen, bspw. in öffentlichen Verkehrsmitteln (Lasen, 2006; Schlote & Linke, 2010, 2014), als an Orten geführt, an denen hochgradig ritualisierte Formen der Interaktion stattfinden, bspw. Restaurants (Geser, 2004, S. 26; Höflich et al., 2014; Ling, 1997). Inwiefern Mithörende diese mobilen Telefonate als akzeptiert ansehen, ist oftmals davon abhängig, ob sie im Freien oder in geschlossenen Räumen geführt werden (Ling, 1997, 2002; Taylor, 2005) bzw. wie formal Interaktionen an den betreffenden Orten – bspw. Restaurants, Kinos oder Schulen (Ling & Yttri, 2002; Taylor, 2005, S. 160–163) – organisiert sind. Die Akzeptanz mobiler Telefonate hängt jedoch nicht nur mit dem sozialen und räumlichen Kontext zusammen: Im Laufe der Zeit entwickeln sich durchaus auch kulturelle Differenzen hinsichtlich anerkannter Formen des Mobiltelefongebrauchs (Katz & Aakhus, 2002; Lasen, 2005; Rivière & Licoppe, 2005): Während bspw. in Japan mobile Telefongespräche in öffentlichen Verkehrsmitteln als unhöflich empfunden werden, wird in Italien oder Spanien wesentlich ‚unverkrampfter‘ damit umgegangen (vgl. Höflich, 2016, S. 163).

Die Analysen des Einzugs mobiler Kommunikationstechnologien in den öffentlichen Raum konstatieren als Ergebnis also beobachtbare Veränderungen im Verhältnis des Einzelnen zu anwesenden Fremden: Diese lassen sich als gegensätzliches Moment der Absentierung der NutzerInnen mobiler Kommunikationstechnologien bei gleichzeitiger Privatisierung der Öffentlichkeit durch ihre Telefonate beschreiben. Beide Dimensionen des problematischen Charakters des öffentlichen Mobiltelefongebrauchs entstehen dadurch, dass die NutzerInnen der mobilen Endgeräte mehrere soziale Situationen gleichzeitig ‚besetzen‘ (Palen et al., 2001, S. 110; s. u. a. Lasen, 2006; Palen et al., 2000): den physikalischen Raum der NutzerIn und anderer Anwesender sowie den virtuellen Raum des Telefongesprächs. Die Studien dieses Forschungsbereich sensibilisieren zum einen für die Relevanz des sozialen und räumlichen Kontexts für die Beurteilung des Gebrauchs mobiler Endgeräte. Zum anderen lenken sie den Blick auf die Etablierung und Aushandlung sprachlicher und verkörperter Praktiken des Umgangs mit Telekommunikationssituationen in Anwesenheit Anderer. Sie untersuchen jedoch im Vergleich zur vorliegenden Arbeit „nicht-zentrierte“ und nicht „zentrierte“ Interaktionen (Goffman, 2018). Der Frage danach, wie über das Mobiltelefon Kommunizierende im Kontext von F2F-Interaktionen mit eben dieser Herausforderung umgehen, dass sie zeitgleich in die soziale Situation vor Ort und die medial-vermittelte Kommunikation involviert sind, rückt nun in den Fokus der Studien des dritten Forschungsbereichs.

### **2.2.2. Der individuelle Gebrauch mobiler Endgeräte in F2F-Interaktionen**

Aufgrund der Praktiken einer ständigen Erreichbarkeit über das Mobiltelefon ist verstärkt „die Möglichkeit der Interferenz zweier kommunikativer Situationen gegeben.“ (Burkart, 2000a, S. 218; s. a. Höflich, 2016, S. 168–169). Denn das Schreiben von Textnachrichten oder mobile Telefonate können in räumlich-zeitlicher Überschneidung mit F2F-Interaktionen stattfinden. Mit den Veränderungen, die mit eben diesem Einzug von Mobiltelefonen in zentrierte Interaktionen zwischen körperlich kopräsenten Akteuren einhergehen, beschäftigt sich ein dritter Forschungsbereich der Mobilkommunikationsforschung. Ebenso wie die Forschung zur Handynutzung im öffentlichen Raum ist dieser Bereich in den ersten Jahren zunächst maßgeblich von der Soziologie Goffmans geprägt. Dieser etablierte die Untersuchung sozialer Situationen als Phänomene ‚in eigenem Recht‘ (Goffman, 1964, S. 134), deren spezifische Dynamik aufgrund der Möglichkeit der wechselseitigen Wahrnehmung aller Anwesenden entsteht. Insbesondere in zentrierten Interaktionen („encounters“, Goffman, 1964, S. 134–135), bei denen zwischen den Interaktionspartnern ein gemeinsamer visueller und kognitiver Fokus gebildet wird (ebd., 1964, S. 134), handeln Menschen im Bewusstsein darüber, von den jeweils Anderen im eigenen

Verhalten wahrgenommen zu werden. In derartigen Situationen versuchen sie die Eindrücke, die sie vermitteln, so zu gestalten, dass ein möglichst positives Bild ihrer Person entsteht, gleichzeitig wird auch den InteraktionspartnerInnen eine positive Selbstdarstellung und die Wahrung des eigenen Gesichts ermöglicht (Goffman, 1959, S. 132–151, 1967, S. 5–46). Dieses Eindrucksmanagement und die Vermeidung gesichtsbedrohender Verhaltensweisen erfordern von allen Beteiligten ein hohes Maß an Aufmerksamkeit sowie die feine Abstimmung kommunikativer Aktivitäten. Aus dieser Perspektive wird die Überschneidung kommunikativer Situationen, die bspw. durch einen eingehenden Anruf während eines Gesprächs zwischen Freunden entsteht, insbesondere als Herausforderung verstanden, die rituelle Aspekte von Interaktion (Goffman, 1967) und damit Fragen des Eindrucksmanagements und Faceworks (ebd. 1967, S. 5–47) betrifft. Mit der Überschneidung von zwei kommunikativen Situationen entstehen somit „dual front interactions“ (Humphreys, 2005, S. 819), in denen die TeilnehmerInnen gleichzeitig Eindrucksmanagement an den ‚Fronten‘ von zwei Interaktionen (der medial vermittelten und F2F-Interaktion<sup>17</sup>) betreiben müssen.

Der Umgang mit den Herausforderungen derartiger Situationen und die Herausbildung sozialer Normen – einer Etikette des Umgangs mit mobiler Kommunikation (Nardi & O’Day, 1999) in Situationen des körperlichen Beisammenseins – ist Gegenstand zahlreicher Studien (bspw. Caron & Caronia, 2007, S. 217 ff.; Ling, 2004, S. 123–143, 2008, S. 93–116). In ihnen werden unterschiedlichste Praktiken des Eindrucksmanagements identifiziert, die das ‚gesichtsbedrohende‘ Potenzial der Interferenz multipler Kommunikationssituationen entschärfen können: Bspw. wird während des mobilen Telefonats der kopräsenten InteraktionspartnerIn über andere Kommunikationskanäle (gestisch, mimisch) signalisiert, dass es der Telefonierenden unangenehm ist, dieses Gespräch ausgerechnet jetzt führen zu müssen (Humphreys, 2005, S. 820). Darüber hinaus existieren zahlreiche Methoden, mit denen die Anwesenden versuchen, einen ‚privaten Raum‘ (Humphreys, 2005, S. 818) für das Telefonat schaffen: Diese reichen von der Auflösung der wechselseitigen Blickverschränkung, über demonstratives Wegschauen und die körperliche Abwendung (Lasen, 2005, S. 94; Ling, 2004, S. 135–136; Murtagh, 2002, S. 85–86; Plant, 2001, S. 53) bis hin zur Herstellung räumlicher Distanz durch die Telefonierenden (Geser, 2004; Lasen, 2005, S. 94) oder die Aufnahme alternativer Beschäftigungen durch die Anwesenden (Humphreys, 2005, S. 817; Plant, 2001, S. 34). Mobile Telefonate in fokussierten Interaktionen können von den TeilnehmerInnen jedoch auch als Verletzungen wechselseitiger

---

<sup>17</sup> Goffman betont selbst, dass Eindrucksmanagement und Facework grundsätzlich sowohl in kopräsenten als auch medial vermittelten Interaktionen stattfinden, auch wenn sie sich in ihrer Form unterscheiden (mögen) (ebd., 1967, S. 33). Er hat jedoch nie selbst Analysen medial vermittelter Interaktionen vorgenommen.

Höflichkeitserwartungen verstanden werden. Dies zeigen die Anwesenden den Telefonierenden non-verbal über abschätzige Blicke oder sprachlich durch deutlich hörbare kritische Kommentare (Ito & Okabe, 2005; Lasen, 2005, S. 78; Ling, 1997; Plant, 2001, S. 32–34). Insgesamt zeigen diese Studien, dass mit der Ausbreitung mobiler Kommunikationstechnologien auch ein Prozess der Entwicklung von Praktiken des Umgangs mit neuen Situationen und eine Aushandlung normativer Erwartungen über das Verhalten in F2F-Interaktionen einhergeht.

Aufgrund der stetig wachsenden Zahl der NutzerInnen durchdringt der Gebrauch mobiler Endgeräte mittlerweile die unterschiedlichsten Situationen sozialer Praxis. Damit erweitern sich die Möglichkeiten der systematischen Beobachtung dieses Phänomens (Oloff, 2019b, Z. 202–209) und Fragestellungen, die an das Phänomen herangetragen werden können. Insbesondere aus konversationsanalytischer Perspektive werden seit einigen Jahren basierend auf Audio- und Videoaufzeichnungen natürlicher Situationen Analysen der handygestützten Telekommunikation durchgeführt. In diesen wird jedoch nicht a priori davon ausgegangen, dass der Handygebrauch für die Anwesenden – aufgrund ritueller Aspekte der Organisation sozialer Situationen – problematisch ist. Stattdessen werden sowohl die Koordination des Gebrauchs mobiler Endgeräte mit den Anforderungen von F2F-Interaktionen als praktische Hervorbringungsleistung als auch die situierte Aushandlung damit verbundener moralischer Aspekte (Robles et al., 2018, S. 151) untersucht.

Die Herausforderung, mehrere, potenziell miteinander konkurrierende Aktivitäten zu koordinieren, ist zunächst nicht spezifisch für den Mobiletelefongebrauch in F2F-Interaktionen, sondern ein alltägliches Phänomen, das in unterschiedlichsten Settings auftritt, bspw. wenn Menschen gemeinsam Essen und sich unterhalten (Mondada, 2009; König & Oloff, 2018a). Mittlerweile hat sich ein eigener, konversationsanalytisch geprägter Forschungsbereich (Haddington et al., 2014) herausgebildet, in dem derartige Multiaktivitätssettings (C. Goodwin, 1984, S. 225) untersucht werden. Analytisch wird Multiaktivität dabei als praktisches Problem aufgefasst, an dem sich die TeilnehmerInnen beobachtbar orientieren und es durch das kompetente zeitliche Arrangement von Ressourcen und Handlungen bearbeiten. Dabei lassen sich insbesondere drei zeitliche Ordnungen von Multiaktivität identifizieren:

in the *parallel* order, two or more courses of action are simultaneously achieved without any hitches; in the *embedded* order, both courses of action are organised together, occasioning mutual adjustments and slowing down or accelerating the pace of one or the other course of action within a mutual re-synchronisation in relation to the parsing of the action, with a notable attention towards unit completion. . . . In the *exclusive* order, one or more courses of action are abandoned in favour of a single and exclusive other course of action. (Mondada, 2014a, S. 69–70, Herv. i. O.)

Ein weiteres zentrales Ergebnis dieser Studien ist, dass Multiaktivität von den TeilnehmerInnen in konvergierender oder divergierender Weise gemanagt werden kann: „they can mutually align with the production of relevant next actions . . . and display shared attention towards common objects; they can dis-align as well, in not attending to the other’s talk or in not paying attention“ (Mondada, 2012a, S. 252). Unabhängig davon, ob sich die Teilnehmer nun gemeinsam (konvergierend) oder voneinander losgelöst (divergierend) an unterschiedlichen Aktivitäten ausrichten, findet das Multiaktivitätsmanagement immer hochgradig sensitiv gegenüber den sequenziellen Strukturen der sprachlichen Interaktion statt (ebd., 2012a, S. 227).

In Untersuchungen der handyvermittelten Telekommunikation in F2F-Interaktionen werden aus konversationsanalytischer Perspektive in diesem Sinne divergente Nutzungsweisen, „i.e., when the mobile device use is disconnected from the ongoing conversation“ (Oloff, 2019b, Z. 235–236), fokussiert – also Formen des individuellen Gebrauchs, in denen kein konvergenter Fokus der Anwesenden auf die Handynutzung als gemeinsame Aktivität hervorgebracht wird<sup>18</sup>. In den wenigen konversationsanalytischen Studien divergenter Nutzungsweisen im Kontext von F2F-Interaktionen wird empirisch vor allem schriftsprachliche handyvermittelte Telekommunikation untersucht (das Versenden und Empfangen von SMS oder anderen Kurznachrichten, bspw. WhatsApp). Anhand detaillierter Analysen des sprachlichen und verkörperten Verhaltens der InteraktionsteilnehmerInnen zeigen sie, wie die Smartphone NutzerInnen ihren Wechsel in der Aufmerksamkeit zwischen F2F-Interaktion und „Texting“ (DiDomenico et al., 2018; DiDomenico & Boase, 2013) managen. Die GeräteinhaberInnen zeigen durch ihr verkörpertes Verhalten an, inwiefern sie sich an der sequenziellen Organisation des Gesprächs unter den Anwesenden sowie an den mit SprecherInnen- oder ZuhörerInnen-Rollen verbundenen normativen Erwartungen (Zimmerman, 1998) orientieren. Auf diese Weise wird für ihre InteraktionspartnerInnen erkennbar, ob sie die Gerätenutzung als ‚Haupt- oder Nebentätigkeit‘

---

<sup>18</sup> Im Folgenden werden ausschließlich Studien vorgestellt, in denen der Mobiltelefon- oder Smartphonegebrauch im Kontext alltäglicher Unterhaltungen untersucht wird. Es existieren weitere Studien – insbesondere aus dem Bereich der „mobility studies“ (Urry, 2000) –, die aus Perspektive der KA die Koordination des situierten Gebrauchs mobiler Endgeräte mit der sequenziellen Organisation von Mobilitätspraktiken untersuchen: Analysiert wird dabei vor allem das Management der simultanen Mobiltelefonnutzung mit den praktischen Anforderungen des Autofahrens (Esbjörnsson et al., 2007; Haddington & Rauniomaa, 2011; Laurier, 2004) oder Fahrten in öffentlichen Verkehrsmitteln (Figeac & Chaulet, 2018; Licoppe & Figeac, 2015, 2018). Außerdem wird die smartphone-gestützte Navigation über Applikationen wie ‚Google Maps‘ im Kontext von Stadtausflügen oder Wanderungen analysiert (Brown et al., 2014, S. 125–127, 2013, S. 1037–1039; Laurier & McGregor, 2016; McGregor, 2020, S. 121–141). Alle diese Studien untersuchen basierend auf Videoaufzeichnungen natürlicher Situationen des Mediengebrauchs „the relationships between device actions, interactions, physical actions, movement and conversation“ (Brown et al., 2013, S. 1039) und zeigen auf, wie die TeilnehmerInnen diese Multiaktivitätssettings kompetent und in Abhängigkeit der zeitlichen und sequenziellen Verläufe der verschiedenen Aktivitäten miteinander koordinieren.

behandeln. Auch diese Differenzierung geht auf analytische Konzepte Erving Goffmans (1963) zurück:

A main involvement is one that absorbs the major part of an individual's attention and interest, visibly forming the principal current determinant of his actions. A side involvement is an activity that an individual can carry on in an abstracted fashion without threatening or confusing simultaneous maintenance of main involvement. (ebd., 1963, S. 43)

Die Begriffe des *main involvement* und *side involvement* werden bspw. von DiDomenico und Boase (2013; s. a. DiDomenico et al., 2018) konversationsanalytisch fruchtbar gemacht, um empirisch zu zeigen, dass es nicht per se an der Art der Tätigkeit liegt, ob diese eine Haupt- oder Nebentätigkeit darstellt, sondern dass die TeilnehmerInnen diese erst durch ihr Verhalten als solche hervorbringen. Die AutorInnen zeigen, dass sich der Status der Handynutzung dynamisch im Interaktionsgeschehen ändern kann und die TeilnehmerInnen dabei den sequenziellen Verlauf sowie die Anforderungen der lokalen Interaktion berücksichtigen<sup>19</sup>.

Grundsätzlich zeigen derartige Studien, dass die individuelle Nutzung mobiler Endgeräte im Kontext von F2F-Interaktionen nicht per se als problematisch behandelt wird. Die TeilnehmerInnen beurteilen und managen den Gerätegebrauch viel mehr in Bezug auf unterschiedliche situative Relevanzen, wie den verwendeten Medienkanal (Telefonate oder Textnachrichten), die thematische und sequenzielle Passung zum anhaltenden Gespräch sowie die Konstellation und Anzahl der InteraktionspartnerInnen (Oloff, 2019b, Z. 792–794). Ob und inwiefern der Gebrauch mobiler Endgeräte problematisch wird, „is related to the management of recurrent practical issues in social interaction, such as availability and sequence organization“ (Oloff, 2019b, Z. 825–828). Problempotenzial besitzt der individuelle Mediengebrauch in F2F-Interaktion also insofern, dass er das Management und die Bewältigung von praktischen Aufgaben der Interaktionsorganisation, bspw. des Sprecherwechsels, erschweren kann.

Dies bedeutet jedoch nicht, dass sich in den Details der aufgezeichneten Interaktionen empirisch keine Hinweise auf potenzielle oder tatsächliche Problematiken identifizieren lassen, die

---

<sup>19</sup> Das Verhältnis zwischen *main* und *side involvement* in Interaktion ist nie definitiv festgelegt, sondern wird dynamisch und durchgehend zwischen den InteraktionsteilnehmerInnen ausgehandelt (Mondada, 2014a, S. 46). Dies zeigt sich insbesondere daran, dass die Studien, die den individuellen Gebrauch mobiler Endgeräte untersuchen, in ihren Analysen darauf verweisen, dass die TeilnehmerInnen über etablierte Methoden verfügen, mit denen die handygestützte Telekommunikation als ‚gemeinsame‘ Hauptaktivität der F2F-Interaktionen hervorgebracht wird: u. a. werden empfangene Text- oder Chatnachrichten zum gemeinsamen Thema der sprachlichen Interaktion gemacht (DiDomenico et al., 2018, S. 687–691; DiDomenico & Boase, 2013, S. 126–130; McGregor, 2020, S. 100–104; Relieu, 2009, S. 218–220), Anwesende als ‚BeraterInnen‘ für die Produktion handyvermittelter Antworten involviert (Brown et al., 2018, S. 18; s. a. McGregor, 2020, S. 108–113) oder als dritte InteraktionsteilnehmerInnen in Videotelefonate integriert (Isaacs et al., 2012, S. 619). Durch derartige Praktiken wird die telekommunikative Handynutzung nicht nur zur *main activity* der GeräteinhaberInnen, sondern als konvergente bzw. gemeinsame Handynutzung zu einer Aktivität, die (vorübergehend) im Zentrum der Aufmerksamkeit von allen anwesenden InteraktionspartnerInnen steht.

mit dem individuellen Smartphonegebrauch in F2F-Interaktionen einhergehen: Zunächst verwenden die TeilnehmerInnen durchaus sprachliche Praktiken wie Entschuldigungen, Rechtfertigungen und Erklärungen vor der Initiierung der Gerätenutzung und zeigen so an, dass diese als potenziell erklärungsbedürftige Angelegenheit behandelt wird (Oloff, 2019b, Z. 425–431; Porcheron et al., 2016a, S. 1654). Auch in den Details des Gesprächsverlaufs zeigt sich, dass (insbesondere) der (andauernde) Smartphonegebrauch zu Problemen führen kann, denn es lassen sich „interruptions, recapitulations of the conversation for members re-joining, displays of attentiveness despite ostensible inattentiveness, and prompts of absent-minded members“ (Porcheron et al., 2016a, S. 1657) beobachten. In einigen Fällen zeigt sich sogar, dass Versuche Anwesender, die Aufmerksamkeit der HandynutzerInnen zu rekrutieren, mehrfach wiederholt werden müssen oder ganz scheitern (Mantere & Raudaskoski, 2017, S. 149). Das Smartphone kann also zum „sticky media device“ (Mantere & Raudaskoski, 2017; Raudaskoski et al., 2019) werden, das die Aufmerksamkeit ihrer NutzerInnen absorbiert und von dessen Display sie kaum ihren Blick lösen können.

Konversationsanalytische Untersuchungen der individuellen Handynutzung unterscheiden sich darin, ob sie stärker die Kompetenz der TeilnehmerInnen herausstellen, die potenziell miteinander konkurrierenden Anforderungen von Handygebrauch und F2F-Interaktion methodisch zu koordinieren, oder ob auf die problematischen Konsequenzen der Interferenz der multiplen Kommunikationssituationen verwiesen wird. Gemeinsam ist ihnen jedoch, dass sie den gleichen zentralen Aspekt des Gebrauchs mobiler Endgeräte als praktisch zu bearbeitende Herausforderung der Interaktionsorganisation identifizieren: Mit Begriffen wie „bystander inaccessibility“ (Mantere, 2020; Raudaskoski et al., 2019), „bystander ignorance“ (Mantere et al., 2018) oder „opacity“ (Oloff, 2019b, Z. 811) wird darauf verwiesen, dass aufgrund der materiellen Eigenschaften des Geräts (Größe und Displayplatzierung), der Steuerung über Bediengesten sowie der Multifunktionalität der Geräte für andere Anwesende – ohne visuellen Zugriff auf das Gerätedisplay – nicht ersichtlich ist, zu welchem Zweck das Gerät genutzt wird (Mantere et al., 2018, S. 106). Daraus entstehen spezifische Herausforderungen, den intersubjektiven Sinn des Gerätegebrauchs und eine geteilte Situationsdefinition unter den Anwesenden zu etablieren. Im Ergebnis zeigt sich, dass die TeilnehmerInnen bereits über etablierte (Ethno-)Methoden verfügen, diese Herausforderung zu bewältigen, aber auch, dass hierin das Problempotenzial des individuellen Gebrauchs mobiler Endgeräte liegt.

Im Gegensatz zu den Studien, die aus Goffmans Perspektive das Problem der individuellen Mobiltelefonnutzung darin erkennen, dass es geltenden Höflichkeitsnormen widerspricht, bieten konversationsanalytische Studien somit eine alternative Erklärung an: Ob die

Mobiltelefonnutzung problematisch wird, beurteilen und managen die TeilnehmerInnen hinsichtlich unterschiedlicher situativer Relevanzen (Medienkanal, sequenzielle Passung, Teilnehmerkonstellation) und in Bezug auf die praktischen Aufgaben der Organisationsprinzipien sozialer Interaktion. Als problematisch wird die individuelle Handynutzung vor allem dann behandelt, wenn sie mit der normativen Anforderung konfligiert, dass die TeilnehmerInnen mit ihren Handlungen zur *Accountability*, d. h. zur Beobachtbarkeit-und-Berichtbarkeit (vgl. Garfinkel, 2020, S. 35) der Aktivitäten in einem sozialen Setting beitragen und sicherstellen, dass ihre Handlungen für andere Anwesende intersubjektiv verständlich sind.

### **2.2.3. Konvergente Handy- und Smartphonennutzung in F2F-Interaktionen**

Während Formen des individuellen Gebrauchs mobiler Endgeräte in einer Vielzahl ethnografischer Studien untersucht werden, existieren nur vereinzelte Untersuchungen, die kollaborative Nutzungsformen wie das ‚Teilen‘ von auf dem Telefon gespeicherten Medieninhalten in F2F-Interaktionen aus dieser Perspektive analysieren (Taylor, 2005; Taylor & Harper, 2003, S. 156; Weilenmann & Larsson, 2000, 2002). Bereits Anfang dieses Jahrtausends konnten bspw. Weilenmann und Larsson (2002) aufzeigen, dass Jugendliche gemeinsam telefonieren, gespeicherte Textnachrichten mit Anwesenden teilen oder kollaborativ Erwiderungen verfassen. Entsprechend forderten die Autorinnen eine systematische Erforschung dieses vernachlässigten Aspekts der gemeinsamen Handynutzung (Weilenmann & Larsson, 2002, S. 93). Abgesehen von vereinzelten jüngeren Studien<sup>20</sup>, bspw. der Untersuchung von Holfelder und Ritter (2013), die rekonstruieren, wie durch die gemeinsame kommunikative Aneignung von Handyvideos gefilmte Situationen von singulären zu kollektiv bedeutsamen Erfahrungen in der Peergroup werden (ebd. 2013, S. 21 f.), besteht dieser Mangel jedoch nach wie vor fort. So monieren auch Raclaw et al. (2016) über zehn Jahre später immer noch, dass „[t]he everyday use of mobile technology in interaction is still relatively unresearched, and much of what exists has not attended to the use of these technologies alongside the production of social action in face-to-face conversation“ (ebd., 2016, S. 378).

In den letzten zehn Jahren lässt sich jedoch eine deutliche Zunahme konversationsanalytischer Untersuchungen konvergenter bzw. kollaborativer Formen der Smartphonennutzung in

---

<sup>20</sup> Darüber hinaus existieren einige ethnomethodologisch fundierte Beobachtungsstudien zur konvergenten Smartphone-Nutzung in mobilen Settings, die von Weilenmann und ihren MitarbeiterInnen selbst durchgeführt wurden. Dabei werden Praktiken der lokalen Produktion von Social-Media-Content während Zoo- (Weilenmann & Hillman, 2019) bzw. Museumsbesuchen (Hillman et al., 2016; Weilenmann et al., 2013) rekonstruiert. Diese Studien untersuchen zwar kollaborative Praktiken des Smartphonegebrauchs – diese sind jedoch nicht in ‚klassische‘ F2F-Interaktionen bzw. gemeinsame Unterhaltungen, sondern den gemeinsamen Besuch öffentlicher Bildungseinrichtungen eingebettet und werden deshalb hier nur am Rande erwähnt.

alltäglichen Interaktionen beobachten. Im Folgenden werden die Erkenntnisse dieses Forschungsbereichs zunächst übersichtsartig dargestellt – den analytischen Kapiteln werden später ausführlichere Diskussionen derjenigen Praktiken des Smartphonegebrauchs vorangestellt, die unmittelbar für das Verständnis der untersuchten Phänomene relevant sind. Zunächst werde ich jedoch aufzeigen, dass sich der aktuelle Forschungsstand hauptsächlich dem Einsatz unterschiedlicher Handlungsmöglichkeiten des Smartphones widmet und dabei die kontextunabhängigen Herausforderungen und methodischen Lösungen des konvergenten Mediengebrauch in F2F-Interaktionen herausarbeitet. Im Gegensatz zur vorliegenden Arbeit rückt dabei die Frage, in welche gemeinsamen (sprachlichen) Aktivitäten der Anwesenden die Smartphonennutzung eingebettet ist, eher in den Hintergrund.

Ganz grundsätzlich zeigen die Untersuchungen des kollaborativen Handygebrauchs auf, dass Smartphones aufgrund ihrer vielfältigen technischen Möglichkeiten von den TeilnehmerInnen als Ressource für die thematische Entwicklung von alltäglichen Gesprächen eingesetzt werden können (Brown et al., 2018, S. 7–9; Keppler, 2013, 2019; Porcheron et al., 2016a). Die Nutzung mobiler Endgeräte wird hierbei nicht nur zum Thema, über das von den Anwesenden gesprochen wird; die Geräte werden dazu eingesetzt, die unterschiedlichsten Medieninhalte in das Gespräch zu integrieren (vgl. Keppler, 2019, S. 178): Dies können bspw. über Online-Suchen recherchierte Informationen, archivierte Textnachrichten oder Fotografien sein.

Die smartphone-gestützte Integration von Medieninhalten in die sprachliche Interaktion wird zum einen dadurch ermöglicht, dass die meisten modernen Endgeräte einen ortsunabhängigen Zugriff auf das World Wide Web und damit auf eine schier unbegrenzte Menge an Informationen und Medieninhalten erlauben. Die Initiierung einer internetgestützten Suchanfrage wird dabei meist durch die gesprächsweise Hervorbringung eines „searchable object“ (Brown et al., 2015) veranlasst. Als ‚suchbare Objekte‘ werden Informationen und Medieninhalte behandelt, die im ‚Hier und Jetzt‘ der lokalen Interaktion gesucht, gefunden und gemeinsam rezipiert werden können, bspw. die Antwort auf eine Frage oder ein kurzer Videoclip<sup>21</sup>. Durch die sprachliche Interaktion der TeilnehmerInnen wird ein sequenzieller Kontext geschaffen, in dem die Initiierung einer Suchanfrage als verständliche Handlung relativ ‚natürlich‘ erfolgen kann (Brown et al., 2015, S. 511).

---

<sup>21</sup> Vereinzelt existieren auch Untersuchungen von Gesprächen, in denen die TeilnehmerInnen keine gemeinsame Sprache oder eine Lingua franca sprechen. In diesen Fällen können die Möglichkeiten smartphone-gestützter Online-Suchen genutzt werden, um Ausdrücke oder Äußerungen zu übersetzen. Auf diese Weise stellt das Smartphone Ressourcen bereit, die für Selbst-Reparaturen bei der Produktion von Äußerungen in einer Fremdsprache (Greer, 2016) oder für die Übersetzung von Äußerungen in die Sprache der RezipientIn genutzt werden (İkizoğlu, 2019). Es lässt sich argumentieren, dass das Smartphone in derartigen Kontexten nicht nur als Ressource für die thematische Entwicklung des Gesprächs fungiert, sondern intersubjektiv verständliche Kommunikation überhaupt erst ermöglicht.

Der Beginn des Suchprozesses wird dabei meist von den SmartphoneinhaberInnen mit sprachlichen Markierungen (bspw. „warte, ich such’s kurz“) angekündigt: „Damit zeigt der Smartphonebesitzer . . . an, dass er seine Aufmerksamkeit für eine gewisse Zeit seinem Gegenüber entzieht und eine andere Tätigkeit nun ins Zentrum der Aufmerksamkeit rückt“ (Keppler, 2019, S. 179). Während des Suchprozesses stehen die SmartphoneinhaberInnen nun vor der Herausforderung, die Partizipationsmöglichkeiten der anderen Anwesenden auf eine Weise zu managen, dass diese nicht von der Bedienung des mobilen Endgeräts ausgeschlossen sind (Brown et al., 2015, S. 511). Indem die SmartphoneinhaberInnen den Anwesenden Fragen stellen und ihre Antworten bei der Formulierung von Suchbegriffen oder Auswahl von -ergebnissen berücksichtigen, wird der Suchprozess als kollaborativ zu bearbeitende Aufgabe hervorgebracht (Brown et al., 2015, S. 512–513; McGregor, 2020, S. 77–93; Suderland, 2019, S. 102–111). Derartige *kollaborative Suchen* zeichnen sich durch eine ‚engmaschige Geordnetheit‘ (vgl. „close ordering“, Schegloff & Sacks, 1973, S. 297) aus, denn sie werden von den SmartphoneinhaberInnen durch die Produktion erster Paarteile von Adjazenzpaaren (bspw. Fragen oder Vorschlägen) strukturiert, die im unmittelbaren Anschluss bestimmte Erwidern konditionell relevant machen. Alternativ produzieren die SmartphoneinhaberInnen *suchbegleitende Kommentierungen*<sup>22</sup>, die den Anwesenden zwar keine Möglichkeiten zur Teilnahme am Suchprozess, aber einen „rudimentären Zugang zu den Wahrnehmungen der Geschehnisse auf dem Smartphonedisplay“ (Suderland, 2019, S. 116; s. a. Brown et al., 2015, S. 513) verschaffen. Im Unterschied zur *kollaborativen Suche* weisen derartige Phasen der *suchbegleitenden Kommentierung* kaum dialogische Strukturen auf, sie machen für andere Anwesende lediglich ersichtlich, dass der Suchprozess noch nicht abgeschlossen ist (Suderland, 2019, S. 116).

In beiden Varianten zeigt sich eine beobachtbare Orientierung der TeilnehmerInnen an der potenziellen *bystander inaccessibility* des Smartphonegebrauchs und der Einsatz von Praktiken, die dazu dienen, diesen *accountable* zu machen. Dass dieses Problem maßgeblich mit der Bedienung des mobilen Endgeräts über stumme Bediengesten zusammenhängt, wird besonders im Vergleich zu Analysen von Online-Suchanfragen deutlich, die mithilfe von Sprachbefehlen über intelligente persönliche Assistenten (IPA) wie Siri, Google Home oder Alexa durchgeführt werden (Porcheron et al., 2016b, 2018): Die Sprachbefehle übernehmen dabei eine doppelte

---

<sup>22</sup> Der Begriff der *suchbegleitenden Kommentierung* wird in Anlehnung an das Konzept des „online commentary“ verwendet: Dieses stammt ursprünglich aus der Analyse von Arzt-Patienten-Interaktionen und bezeichnet eine Gesprächspraktik, „that describes or evaluates what the physician is seeing, feeling or hearing during the physical examination of the patient“ (Heritage & Stivers, 1999, S. 1501). Analog zu dieser Definition verweist der Begriff der *suchbegleitenden Kommentierung* darauf, dass die sprachlichen Praktiken dazu dienen, den Verlauf der physischen Manipulation des Smartphones (und die Geschehnisse auf dem oftmals nicht einseharen Gerätedisplay) für andere Anwesende durch Beschreibungen und Einschätzungen sinnhaft verständlich zu machen.

Funktion (Schegloff, 2007a, S. 169), denn sie initiieren die Suchanfrage über den IPA, während sie gleichzeitig als hörbare Äußerungen für andere Anwesende *accountable* machen, welchem Zweck die Geräte-Nutzung dient<sup>23</sup>. Die Sprachsteuerung eines IPA veranlasst die InteraktionsteilnehmerInnen zunächst zur interaktiven Produktion von Stille (Porcheron et al., 2016b, S. 212). Anschließend stumme (d. h. lediglich auf dem Display sichtbare) ‚Reaktionen‘ des IPA werden wiederum mittels sprachlicher Praktiken der SmartphoneinhaberIn – bspw. Erklärungen, rhetorische Erwiderungen oder Wiederholungen der Sprachbefehle (Porcheron et al., 2016b, S. 215) – für Anwesende verständlich gemacht. Auf diese Weise zeigt sich eindrücklich, welche Relevanz die Art und Weise der Gerätebedienung für die Frage der *Accountability* der Gerätenutzung im Kontext von F2F-Interaktionen besitzt.

Neben der Möglichkeit der Integration von Medieninhalten über Online-Suchen können Handys und Smartphones auch als *Archive sozialer Originale* (vgl. Keppler, 2013, S. 99) fungieren, über die sich gespeicherte Episoden computer-vermittelter Kommunikation in all ihren Details in die Unterhaltungen einspielen lassen: Über die mobilen Endgeräte werden zahlreiche Applikationen zur schriftsprachlichen und multimedialen Kommunikation (bspw. Emailprogramme, SMS, oder IM-Applikationen wie WhatsApp etc.) genutzt, in denen vergangene Interaktionen gespeichert werden und sich jederzeit wieder aufrufen lassen. Darüber hinaus verfügen moderne Smartphones über integrierte Kameras und einen lokalen Festplattenspeicher, in dem digitale Fotografien und Videos archiviert und darüber ebenfalls in die F2F-Interaktion eingespielt werden können (Keppler, 2019; Oloff, 2019a; Raclaw et al., 2016). Auch derartige, gespeicherte Medieninhalte müssen – bevor sie in das Gespräch integriert werden können – von den SmartphoneinhaberInnen zunächst gesucht und gefunden werden. Diese Suchprozesse können ebenfalls durch den Einsatz der gleichen sprachlichen Praktiken *accountable* gemacht werden. Aufgrund der privaten Natur der Medieninhalte erfolgt die Suche jedoch meist nicht kollaborativ (Suderland, 2019, S. 116).

Ist die Suche nach den Informationen oder Medieninhalten abgeschlossen, können anschließend sogenannte „mobile-supported sharing activities“ (Raclaw et al., 2016, S. 362) eröffnet werden, in denen die Medieninhalte in die F2F-Interaktion integriert werden. Die Art und Weise, wie diese von den SmartphoneinhaberInnen sprachlich in die Interaktion eingeführt

---

<sup>23</sup> Brown et al. (2013) weisen darauf hin, dass auch stumme Bediengesten eine vergleichbare Doppelfunktion übernehmen können: „Using the device, when it is visible to someone else, produces actions that are ‚oriented-to phenomena‘. That is, those who see the screen can be expected by others to have seen the action and to understand what that action means“ (ebd, 2013, S. 1036). Damit stumme Bediengesten nicht nur der Manipulation des mobilen Endgeräts dienen, sondern auch als sinnhafte Handlung für Anwesende erkennbar sind, ist für diese ein visueller Zugang zum Gerätedisplay erforderlich. Ohne diesen wird durch die Bediengesten für andere InteraktionsteilnehmerInnen zwar erkennbar, dass das Smartphone bedient wird; ohne weitere sprachliche Kommentierungen wird jedoch nicht *accountable*, welchem Zweck die Bedienung gilt.

werden, ist dabei zum einen von der Medialität des Inhalts abhängig, also davon, ob es sich um schriftsprachliche, statische visuelle (bspw. Bilder oder Fotografien) oder dynamische (audio-)visuelle (bspw. GIFS oder Videos) Medien handelt. Zum anderen wird die Prä-Sequenz der Integration in Abhängigkeit davon ausgestaltet, ob mit ihr eine präferierte oder dispräferierte Handlung (bspw. ein Witz oder eine Beschwerde) realisiert werden soll (Raclaw et al., 2016, S. 364).

Werden über das Smartphone schriftsprachliche Medieninhalte in das Gespräch integriert, bieten insbesondere die *Archive sozialer Originale* von Instant-Messaging-Applikationen eine vielfältig einsetzbare Ressource: „the ability to go back over archived text and images allows for different sorts of reporting on those conversations – gisting, storytelling, verbatim reading, and the like, each of them designed with respect to local arrangements” (Brown et al., 2018, S. 23). Die unterschiedlichen Formen der sprachlichen Bezugnahme auf die archivierte Kommunikation werden jeweils auf den lokalen Kontext der F2F-Interaktion und die spezifischen RezipientInnen angepasst. Insbesondere Textnachrichten werden von den SmartphoneinhaberInnen oftmals durch sprachliche Äußerungen im Modus der direkten Redewiedergabe zitiert (Brown et al., 2018, S. 19–21; Keppler, 2019, S. 186; Raclaw et al., 2016, S. 377). Dies bedeutet jedoch nicht, dass es sich um ‚neutrale‘ Wiedergaben des Wortlauts handelt. Die SprecherInnen vermitteln vielmehr durch den Einsatz sprachlicher und vokaler Ressourcen implizit und explizit ihre Haltung zu den wiedergegebenen Nachrichten, um die RezipientInnen so zu einer angemessenen Erwiderung zu leiten (Brown et al., 2018, S. 20).

Den RezipientInnen wird im Zuge der Integration schriftsprachlicher Medieninhalte immer wieder auch visueller Zugriff zum Gerätedisplay ermöglicht (Brown et al., 2015, S. 514): Dieser wird bspw. genutzt, um zu verhindern, dass potenziell Mithörende durch die lautsprachliche Realisierung von den Inhalten der Nachrichten erfahren (Keppler, 2019, S. 186); aber auch um den RezipientInnen Zugang zu visuellen Medieninhalten zu ermöglichen, die gemeinsam mit den Textnachrichten verschickt wurden (Brown et al., 2018, S. 14). Während der visuelle Zugriff zum Gerätedisplay bei der Integration (vorwiegend) textsprachlicher Medieninhalte nicht notwendigerweise, sondern hinsichtlich der spezifischen lokalen und kontextuellen Relevanzen ermöglicht wird, ist er ein konstitutiver Bestandteil der Integration visueller Medieninhalte.

Der visuelle Zugriff zu bspw. digitalen Fotografien wird für die RezipientInnen durch die Hervorbringung „initialer Zeigesequenzen“ (Oloff, 2019a) ermöglicht: Diese werden durch Adjazenzpaare organisiert, von denen Teile auch verkörperter Natur sein können, und durch die interaktiv die Aufmerksamkeits- und Displayfokussierung sichergestellt, sowie Betrachtungsanweisungen durch die Zeigenden formuliert werden (Oloff, 2019a, S. 214–215). Die

RezipientInnen des Zeigens antworten meist zunächst „im Sinne der projizierten Positionierung, ohne sich jedoch auf den eigentlichen Inhalt des Bildes zu beziehen“ (Oloff, 2019a, S. 214). Diese erste vage Evaluation dient vor allem der Alignierung mit der Zeigeaktivität und der Affiliation mit der der SmartphoneinhaberIn. Konkretere Evaluationen erfolgen erst nach eingehender Betrachtung des Medieninhalts und werden von den InteraktionsteilnehmerInnen oft gemeinsam bearbeitet.

Das Teilen visueller Medieninhalte wird auf diese Weise als gemeinschaftliche Aktivität hervorgebracht, die sich aus der physischen Handhabung des mobilen Endgeräts, der visuellen Fokussierung des Displays und der auf die angezeigten Bilder bezogenen sprachlichen Interaktion zusammensetzt (vgl. Aaltonen et al., 2014, S. 129)<sup>24</sup>. Die Zugänglichkeit des visuellen Medieninhalts auf dem Gerätedisplay fügt der Interaktion eine zusätzliche semiotische Ressource hinzu, die die sprachliche Interaktion unterstützt. Wenn sich die digitalen Bildinhalte im geteilten Wahrnehmungsraum der TeilnehmerInnen befinden, ändert sich die sequenzielle Organisation der F2F-Interaktion insofern, als dass die Turnkonstruktion nicht mehr nur über sprachliche Äußerungen, sondern über multimodale semiotische Einheiten stattfinden kann, d. h. dass „incomplete utterances that might not be sufficient in themselves may become pragmatically sufficient in connection with the photographs“ (Aaltonen et al., 2014, S. 139).

Neben den Untersuchungen von Oloff (2019a) und Aaltonen et al. (2014), die in ihren Analysen das mobile Endgerät systematisch als „physisches, berühr- und manipulierbares Objekt“ (Oloff, 2019a, S. 193) berücksichtigen, existieren weitere Studien, die vor allem die Funktionen der smartphone-gestützten Bildintegration analysieren: Fotografien werden nicht nur eingesetzt, „um bestimmte Zustände, Ereignisse oder Erlebnisse im Gespräch zu belegen“ (Keppler, 2019, S. 184); je nachdem, wie die visuellen Medieninhalte sprachlich in die Interaktion eingeführt werden, können sie selbst als physische Repräsentation einer vergangenen Erfahrung zur ‚Geschichte‘ dieses Ereignisses werden (Isaacs et al., 2012, S. 622). In jedem Fall fügt die Integration visueller Medieninhalte den Gesprächen eine zusätzliche Interpretationsebene hinzu, die eindeutige interaktionale Konsequenzen für die TeilnehmerInnen hat: „they provide expanded epistemic access to an assessable referent that is ultimately used to guide recipients toward alignment and affiliation“ (Raclaw et al., 2016, S. 377). Welche konkrete Bedeutung ein Medieninhalt für die jeweilige Interaktion übernimmt, muss dabei immer im Zuge einer

---

<sup>24</sup> Die AutorInnen untersuchen die Integration von Fotografien durch den Gebrauch von Digitalkameras in atypischen Interaktionen. Aufgrund zentraler Gemeinsamkeiten zwischen Digitalkameras und Smartphones als tragbaren Geräten mit Displays, auf denen gespeicherte digitale Fotografien angezeigt werden können, sind die Ergebnisse jedoch für die Analysen der smartphone-gestützten Integration in F2F-Interaktionen relevant (s. ausführlich Kap. 6).

gemeinsamen kommunikativen Aneignung etabliert werden: Denn was sich für die GesprächsteilnehmerInnen in einem visuellen Medieninhalt zeigt, hängt stets davon ab, „was sie in ihm sehen, d. h. wie sie es sich – sei es im Konsens, sei es im Dissens – im Gang der Konversation aneignen“ (Keppler, 2019, S. 185).

#### **2.2.4 Zusammenfassung: konvergenter Smartphonegebrauch in F2F-Interaktionen**

In der Gesamtschau zeigen die Untersuchungen der kollaborativen Handynutzung in F2F-Interaktion, dass über die mobilen Endgeräte abrufbare Medieninhalte – bspw. Informationen, Textnachrichten oder Fotografien – im Zusammenspiel mit sprachlichen und verkörperten Handlungen als Ressource in der Produktion sozialer Handlungen eingesetzt werden können. In und durch *mobile-supported sharing activities* werden nicht nur zusätzliche thematische Ressourcen für das Gespräch der Anwesenden bereitgestellt: Durch die handygestützte Integration von Medieninhalten erhalten die RezipientInnen einen (im Vergleich zur sprachlichen Beschreibung gesteigerten) epistemischen Zugang zu ansonsten ‚abwesenden‘ Referenten (Raclaw et al., 2016, S. 377), der ihnen deren Einschätzung oder Bewertung überhaupt erst ermöglicht (vgl. Pomerantz, 1984, S. 58–59). Die geteilten Medieninhalte sind mit einer erhöhten Glaubwürdigkeit (Keppler, 2013, S. 99, 2019, S. 215) ausgestattet, sodass die RezipientInnen mit ihnen zur Affiliation mit der Haltung der SmartphoneinhaberInnen geleitet werden können (Raclaw et al., 2016, S. 377).

Die Phasen der Smartphoneintegration sind dabei grundsätzlich von einem asymmetrischen Verhältnis zwischen den InteraktionsteilnehmerInnen geprägt, das auch als „driver-passenger-relationship“ (Brown et al., 2015, S. 511) bezeichnet wird: Aufgrund der materiellen Konfiguration der mobilen Endgeräte – die auf die Bedienung durch eine einzelne NutzerIn ausgelegt sind – entsteht ein Setting, in dem die GeräteinhaberIn das Smartphone in den Händen hält und physisch handhabt, während andere Anwesende (meist) höchstens visuellen Zugang zum Display erhalten. Die SmartphoneinhaberIn steht vor der praktischen Aufgabe, die für die Integration des mobilen Endgeräts notwendigen physischen Manipulationen durchzuführen (Suderland, 2019, S. 108). Dies führt dazu, dass sie mit exklusiven kommunikativen Rechten ausgestattet wird und mit ihren Gesprächsbeiträgen die Interaktion strukturiert (Keppler, 2013, S. 95). Die TeilnehmerInnen orientieren sich an diesem asymmetrischen Verhältnis in der Regel nicht als ‚problematisch‘. Voraussetzung dafür ist jedoch, dass es den TeilnehmerInnen durch den Einsatz beobachtbarer und berichtbarer Praktiken gelingt, eine „local order of mobile phone use“ (McGregor, 2020, S. 13) hervorzubringen und aufrechtzuerhalten, die Teilnahme(-möglichkeiten) am Gebrauch mobiler Endgeräte zu managen und die Gerätenutzung

zeitlich und sequenziell mit der sprachlichen Interaktion zu koordinieren (vgl. Brown et al., 2015, S. 516). Ist dies erfolgreich, so übernehmen die über die mobilen Endgeräte eingespielten Medieninhalte in alltäglichen Gesprächen die „Funktionen von Gesprächsförderern und nicht von Gesprächsverhinderern“ (Keppler, 2013, S. 101).

Durch detaillierte Analysen konkreter Interaktionssituationen können konversationsanalytische Arbeiten aufzeigen, wie die TeilnehmerInnen als „Akteure ihres alltäglichen Lebens“ (Keppler, 2013, S. 87) in und durch ihre praktischen Handlungen die intersubjektive Sinnhaftigkeit der Nutzung mobiler Kommunikationstechnologien im Vollzug der F2F-Interaktionen hervorbringen und diese als Ressource in der Produktion erkennbarer sozialer Handlungen einsetzen (McGregor, 2020, S. 148). Auf diese Weise trägt die KA zu einem besseren Verständnis davon bei, wie soziale Ordnung hervorgebracht und aufrechterhalten wird, während die TeilnehmerInnen die vielfältigen Aktivitäten alltäglicher Unterhaltungen mit dem Gebrauch mobiler Endgeräte koordinieren.

Auch wenn die Anzahl konversationsanalytischer Untersuchungen der gemeinsamen Nutzung mobiler Endgeräte in den letzten Jahren stetig gestiegen ist, befindet sich dieses Forschungsfeld noch in den Anfängen. Dies zeigt sich vor allem an den analytischen Foki der Arbeiten: Sie alle untersuchen den Einsatz bestimmter Handlungsmöglichkeiten des Smartphones, bspw. Online-Suchen oder das Teilen archivierter Textnachrichten. Anhand der exemplarischen Analyse einzelner Datenstücke werden typischen Formen der konvergenten Nutzung mobiler Endgeräte rekonstruiert, ihre sequenziellen Platzierungen ausgelotet und diejenigen Praktiken identifiziert, mit denen die Smartphoneintegration als Teil sinnhafter Handlungen *accountable* gemacht wird. Identifiziert werden auf diese Weise vor allem die kontextunabhängigen Herausforderungen der Interaktionsorganisation, die mit dem Gebrauch des Smartphones verbunden sind, und ihre methodischen Lösungen sowie das Funktionspotenzial von *mobile-supported sharing activities*. Die konversationsanalytisch zentrale Frage danach, in welche gemeinsamen (sprachlichen) Aktivitäten der Anwesenden die Smartphoneintegration eingebettet ist, rückt dabei jedoch meist in den Hintergrund. Auf diese Weise entsteht zwar ein zunehmend umfassendes Bild davon, welche Handlungsmöglichkeiten des Smartphones überhaupt in F2F-Interaktionen genutzt und in welchen Varianten diese realisiert werden. Es existieren jedoch bis dato kaum systematische Analysen der kontextsensitiven Anpassung der Smartphoneintegration an die je relevante gemeinsame Gesprächsaktivität (s. jedoch: Brown et al., 2018; Raclaw et al., 2016).

Ausgehend von dem hier skizzierten Forschungsstand macht es sich die vorliegende Arbeit zur Aufgabe, zu untersuchen, wie und mit welcher Funktion Smartphones als Ressource in der

Hervorbringung einer spezifischen kommunikativen Aktivität – dem alltäglichen Geschichtenerzählen – eingesetzt werden. Erzählen ist eine rekonstruktive Gattung der Alltagskommunikation, die eine etablierte Lösung des kommunikativen Problems der Flüchtigkeit sozialer Ereignisse (Luckmann & Bergmann, 1995, S. 293) und der gemeinsamen Interpretation sozialen Erlebens dient (König & Oloff, 2018b, S. 208). Erzählen und Alltag stehen dabei in einem wechselseitigen Verhältnis zueinander: Zum einen „werden im alltäglichen Erzählen Erlebnisse, Begebenheiten, Inhalte der als Alltag erfahrenen Lebenswelt zum Gegenstand der narrativen Rekonstruktion gemacht“ (König & Oloff, 2018a, S. 278) – d. h. TeilnehmerInnen erzählen ihren InteraktionspartnerInnen von ihrem Alltag. Zum anderen findet Erzählen oftmals eingebettet in die alltäglichen Situationen der TeilnehmerInnen statt: Es handelt sich um ungeplante Aktivitäten, die aus der Situation heraus und durch die interaktiven Beiträge aller Anwesenden hervorgebracht werden: „Erzählen konstituiert somit selbst wiederum Alltag“ (König & Oloff, 2018a, S. 279). Wie zu zeigen sein wird, werden Smartphones, die über sie stattfindende Kommunikation sowie über sie abrufbare Medieninhalte heutzutage immer wieder zum Thema alltäglicher Erzählungen und konstituieren – indem sie als Ressource in der Hervorbringung der Erzählung eingesetzt werden – einen zunehmend von Medien durchdrungenen Alltag selbst mit. Indem die alltägliche Aktivität des Geschichtenerzählens fokussiert wird, beschränkt sich diese Arbeit also nicht auf die Frage, welche Handlungsmöglichkeiten mobiler Endgeräte in F2F-Interaktionen genutzt werden. Vielmehr wird spezifisch untersucht, welchen Beitrag diese unterschiedlichen Handlungsmöglichkeiten als Ressource in der Hervorbringung von Geschichtenerzählungen und für die Bearbeitung der für diese Aktivität charakteristischen Aufgaben leisten können – aber auch welche neuen Herausforderungen damit einhergehen. Aus Perspektive der multimodalen KA wird das Smartphone dabei als „physisches, berühr- und manipulierbares Objekt“ (Oloff, 2019a, S. 193) berücksichtigt, sodass rekonstruiert werden kann, wie der Gebrauch der mobilen Endgeräte die multimodale Praxis des Geschichtenerzählens modifiziert bzw. grundlegend verändert (vgl. Keppler, 2013, S. 92).

Im nächsten Kapitel wird der Forschungsansatz der multimodalen KA mit seinen theoretischen Grundannahmen und methodologischen Prämissen vorgestellt, sowie die Adaption einer *af-fordanz*theoretischen Perspektive expliziert, durch die der Mediengebrauch in Interaktion analytisch greifbar gemacht werden soll.

### **3 Multimodale Konversationsanalyse als Forschungsansatz zur Untersuchung des Smartphonegebrauchs in F2F-Interaktionen**

In dieser Arbeit wird der Gebrauch mobiler Endgeräte als Ressource für die interaktive Hervorbringung von Geschichtenerzählungen in alltäglichen F2F-Interaktionen aus Perspektive der multimodalen Konversationsanalyse untersucht. Bei der multimodalen KA handelt es sich um eine Weiterentwicklung der KA, die in den 1960er und 70er Jahren von Harvey Sacks, Emanuel Schegloff und Gail Jefferson als Forschungsprogramm für die Untersuchung sprachlicher Interaktion erarbeitet wurde. In Anknüpfung an die von Harold Garfinkel (1967) gegründete Ethnomethodologie wird Interaktion aus Perspektive der KA als fortwährender Prozess der Hervorbringung sinnhafter sozialer Ordnung verstanden. Ihr Ziel ist es:

diejenigen Verfahren empirisch zu bestimmen, mittels derer die Teilnehmer an einem Gespräch im Vollzug ihrer (sprachlichen) Handlungen die Geordnetheit der (sprachlichen) Interaktion herstellen, das Verhalten ihrer Handlungspartner auf die in ihm zum Ausdruck kommende Geordnetheit hin analysieren und die Resultate dieser Analysen wiederum in ihren Äußerungen manifest werden lassen. (Bergmann, 1988a, S. 4)

Von Beginn an war es für die KA kennzeichnend, als Datengrundlage technische Aufzeichnungen von Gesprächen zu verwenden, die in alltäglich-praktischen Interaktionszusammenhängen stattfinden. Die gerade in der Anfangszeit verbreitete Anfertigung von Audioaufzeichnungen war dabei nicht von einem vorgängigen Interesse an Sprache motiviert, sondern stellte den Versuch dar, die methodische Hervorbringung sinnhaft geordneter Interaktionen mit den technisch verfügbaren Mitteln der Zeit in ihren Details aufzuzeichnen und rekonstruieren zu können (vgl. Sacks, 1984, S. 26). Mithilfe von Audioaufzeichnungen können alltägliche Interaktionen als flüchtige Phänomene technisch konserviert, wiederholt analytisch zugänglich und Interpretationen am Datenmaterial intersubjektiv überprüfbar gemacht werden (vgl. Bergmann, 1988a, S. 3).

Dass soziale Interaktionen von den TeilnehmerInnen dabei nicht nur durch den Einsatz sprachlicher Ressourcen hervorgebracht werden, sondern auf dem Zusammenspiel unterschiedlicher, wechselseitig miteinander verknüpfter Modalitäten – wie Sprache, Mimik, Gestik und anderen verkörperte Handlungen – beruhen (vgl. Stivers & Sidnell, 2005, S. 1), wurde schon in der Entstehungsphase der KA anerkannt. Bereits früh finden sich Studien, die durch die Analyse von Videoaufnahmen multimodale Aspekte von F2F-Interaktionen untersuchen: Beispielhaft lassen sich hier die Arbeiten von Charles und Marjorie Goodwin anführen, in denen systematisch die Relevanz der Blick-Organisation und der wechselseitigen visuellen Beobachtung für die Organisation der Teilnahme, des Sprecherwechsels und der Turnkonstruktion in alltäglichen Gesprächen analysiert wurden (C. Goodwin, 1979, 1980, 1981; M. H. Goodwin, 1980).

Seit Beginn dieses Jahrhunderts steigt die Anzahl videobasierter Studien, die die Koordination der verschiedenen, von den TeilnehmerInnen zur Interaktionsorganisation genutzten Ressourcen untersuchen, stetig an und die multimodale Konversations- bzw. Interaktionsanalyse<sup>25</sup> hat sich als systematisch ausgearbeitetes Forschungsprogramm etabliert. Diese Entwicklung wurde maßgeblich durch die zunehmende Verfügbarkeit portabler Videokameras mit großen Speicherkapazitäten begünstigt. Grundsätzlich ist die systematische Untersuchung der Multimodalität ein konsequenter Schritt für die KA, die sich als Disziplin versteht, „which aims at a comprehensive understanding of human interaction and which sets as its goal to uncover the practices by which social interaction is produced“ (Deppermann, 2013, S. 2). Innerhalb dieses Forschungsprogramms wird der Begriff „Multimodalität“<sup>26</sup> dabei verwendet,

to refer to the various resources mobilized by participants for organizing their action – such as gesture, gaze, facial expressions, body postures, body movements, and also prosody, lexis and grammar. The plurality of ‘modalities’ referred to in this term treats multimodality as constitutive and primary. This encourages a view of modalities as constitutively intertwined, and language as integrated within this plurality as one among other resources, without any a priori hierarchy. (Mondada, 2016a, S. 338)

Mit Multimodalität werden also Prosodie, Lexik, und Grammatik als Aspekte des natürlichen Sprachgebrauchs, sowie Gestik, Mimik, Körperhaltungen und -bewegungen – also verkörperte Handlungen – als Ressourcen der Interaktionsorganisation gefasst. Multimodalität wird dabei als konstitutive Eigenschaft von Interaktionen verstanden und a priori keine analytische Hierarchisierung der unterschiedlichen Ressourcen vorgenommen. Vielmehr wird davon ausgegangen, dass diese im Vollzug der Interaktion wechselseitig miteinander verbunden werden

---

<sup>25</sup> In der Literatur wird teilweise zwischen „multimodaler Konversationsanalyse“ und „multimodaler Interaktionsanalyse“ unterschieden (bspw. König & Oloff, 2018b, S. 221; Schmitt & Knöbl, 2013, S. 444–445). Beide Ansätze haben das Ziel, die Relevanz der verschiedenen verfügbaren und eingesetzten Ressourcen basierend auf der empirischen Untersuchung der praktischen Aktivitäten und beobachtbaren Orientierungen der TeilnehmerInnen zu rekonstruieren. Wenn diese Differenzierung getroffen wird, bezieht sie sich meist darauf, ob „visuelle Ausdrucksressourcen bei der multimodalen Erweiterung verbal definierter Konzepte“ (Schmitt & Knöbl, 2013, S. 444, eig. Herv.) berücksichtigt werden (multimodale KA), oder analytisch mit einem grundsätzlichen Interesse an interaktiven Ordnungsstrukturen vorgegangen wird, ohne ein primäres Interesse für verbale Strukturen zu verfolgen (multimodale Interaktionsanalyse). Ich werde in dieser Arbeit ausschließlich den Begriff der multimodalen KA verwenden, da weder in der ursprünglichen Konzeption der KA (Sacks, 1984, S. 26) noch in der multimodalen KA (bspw. Mondada, 2016b, S. 338) ein theoretisches Primat von Sprache angelegt ist.

<sup>26</sup> „Multimodalität“ und „multimodale Interaktion“ sind Begriffe, die in anderen Forschungskontexten entwickelt und mit verschiedenen Bedeutungen für die Erforschung unterschiedlicher Phänomenbereiche verwendet werden: In verwandten Disziplinen wird der Begriff zur Bezeichnung verschiedener (Kommunikations-)Kanäle, Medien sowie materieller Repräsentationen und Zeichen verwendet, die unterschiedliche semiotische Effekte ermöglichen – bspw. Texte, statische und bewegte Bilder oder multimediale Botschaften (vgl. Kress & Leeuwen, 2001; Mondada, 2014b, S. 138). Die unterschiedliche Begriffsverwendung in den verschiedenen Disziplinen birgt die grundsätzliche Gefahr einer Uneindeutigkeit der Verwendung von „Multimodalität“ als Label für konversationsanalytische Studien, da unter dem gleichen Begriff semiotische, diskurs- und medienwissenschaftliche Untersuchungen verfasst werden, die wenig oder nichts mit der konversationsanalytischen Konzeption des Begriffs gemein haben (vgl. Deppermann, 2013, S. 2).

und Sprache als eine unter vielen Ressourcen im Prozess der Hervorbringung sinnhaft strukturierter Interaktionen zum Einsatz kommt (vgl. Mondada, 2018, S. 86).

Ziel dieses Kapitels ist die Darstellung der multimodalen KA als Forschungsansatz, in dem die theoretischen Grundannahmen der Ethnomethodologie sowie spezifisch für die Erforschung von F2F-Interaktionen entwickelte methodologische Prämissen und analytische Maxime ineinandergreifen. Vor dem Hintergrund dieser Darstellung wird die anschließende Offenlegung des konkreten methodischen Vorgehens dieser Arbeit nachvollziehbar. Darüber hinaus ermöglicht die ausführliche Diskussion des Forschungsansatzes eine Explikation einiger grundsätzlicher Eigenschaften und Organisationsprinzipien alltäglicher F2F-Interaktionen, auf denen die spätere empirische Analyse des Smartphonegebrauchs als Ressource für die Hervorbringung von Geschichtenerzählungen aufbaut. Im Folgenden werde ich zunächst die sozialtheoretischen Grundannahmen der Ethnomethodologie vorstellen (Kap. 3.1), die das theoretische Fundament für das Forschungsprogramm der multimodalen KA bieten (Kap. 3.2). Nach einer kurzen Darstellung konversationsanalytischer Erkenntnisse über die Organisationsprinzipien von Geschichtenerzählungen in alltäglichen Gesprächen werde ich abschließend diskutieren, wie das ursprünglich wahrnehmungspsychologische Konzept der *Affordanzen* in der KA adaptiert wird (Kap. 3.3), um den Gebrauch mobiler Kommunikationstechnologien in Interaktion empirisch untersuchen zu können.

### **3.1 Die ethnomethodologischen Grundlagen der multimodalen KA**

Die (multimodale) KA ist als Forschungsprogramm für die Erforschung (sprachlicher) Interaktion aus der von Harold Garfinkel in den 1950er und 60er Jahren entwickelten EM (1967) hervorgegangen<sup>27</sup>. Die Verfahrensgrundsätze der KA lassen sich entsprechend nur vor dem

---

<sup>27</sup> Die Entwicklung der KA wurde konzeptionell außerdem von der Soziologie Erving Goffmans (1967, 1974, 2003) beeinflusst. Das komplizierte Verhältnis zwischen KA und Goffmans Perspektive soll jedoch im Folgenden nicht selbst diskutiert werden und wird deshalb hier kurz umrissen: Goffmans Einflüsse bestehen zum einen darin, dass sein Werk maßgeblich dazu beitrug, F2F-Interaktionen als eigenständigen, soziologischen Forschungsbereich analytisch zu erschließen (vgl. Schegloff, 1988, S. 89–90). Zum anderen beeinflusste sein Verständnis der Interaktionsordnung als ‚Realität sui generis‘ die Entwicklung der methodologischen Prämissen der KA. Für Goffman stellt die sinnhafte Ordnung sozialer Situationen einen gesonderten Erfahrungsbereich der gesellschaftlichen Wirklichkeiten dar (vgl. Raab, 2014, S. 11). Soziale Situationen besitzen eine eigene Logik, die nicht über den Rückgriff auf Makrophänomene oder die Eigenschaften der TeilnehmerInnen erklärt werden kann, sondern im Vollzug der Situation interaktiv von den TeilnehmerInnen hervorgebracht wird (vgl. Rawls, 1987, S. 145). Wie noch zu zeigen sein wird, ähnelt diese Konzeption Garfinkels Auffassung von der lokalen Produktion sozialer Ordnung im Interaktionsvollzug. Programmatisch formulierte Goffman: „I assume that the proper study of interaction is not the individual and his psychology, but rather the syntactical relations among the acts of different persons mutually present to one another“ (ebd. 1967, S. 2). Die Fokussierung auf die ‚syntaktischen Relationen zwischen Handlungen verschiedener TeilnehmerInnen an sozialen Interaktionen‘ entspricht durchaus dem Vorhaben der KA, die Methoden der TeilnehmerInnen zu rekonstruieren, mit denen diese in wechselseitiger Bezugnahme aufeinander im Interaktionsvollzug sinnhafte soziale Ordnung hervorbringen. Es existieren jedoch auch einige grundlegende Differenzen den Perspektiven von Goffman und der KA auf soziale Interaktion: Zum einen

Hintergrund der ihnen zugrundeliegenden sozial- und handlungstheoretischen Grundannahmen der EM verstehen. Ethnomethodologische Studien

seek to treat practical activities, practical circumstances, and practical sociological reasoning as topics of empirical study, and by paying to the most commonplace activities of daily life the attention usually accorded extraordinary events, seek to learn about them as phenomena in their own right. (Garfinkel, 1967, S. 1)

Sie untersuchen also alltägliche, praktische Aktivitäten und die in und durch sie stattfindenden Prozesse der Sinnkonstitution als Gegenstände ‚in eigenem Recht‘. Mit dem Ziel, eine – zu der Zeit ihrer Entwicklung neuartige – Antwort auf die Möglichkeiten und Bedingungen sozialer Ordnung zu finden, werden dabei diejenigen *Methoden* untersucht, „welche die Alltagshandelnden in fragloser und unanalysierter Weise verwenden, um eine gemeinsame soziale Welt zu erzeugen“ (Eberle, 2009, S. 100). Aus Perspektive der EM wird soziale Wirklichkeit also in und durch soziales Handeln konstituiert, d. h., dass wir das, „was wir im alltäglichen Handeln als vorgegebene soziale Tatsachen, als objektive Sachverhalte, als unabhängig von unserem Zutun existierende Realitäten wahrnehmen und behandeln, erst in unseren Handlungen und Wahrnehmungen als solche produzieren“ (Bergmann, 1981, S. 11). Diese „Vollzugswirklichkeit“ (Bergmann, 1981, S. 12) existiert nicht unabhängig von den sozialen Handlungen, mit denen sie erzeugt wird, und besteht nur solange, „wie sie durch Handlungen und Interaktionen fortlaufend hervorgebracht wird“ (Eberle, 2009, S. 101).

Da alle Gesellschaftsmitglieder durch ihre wechselseitig aufeinander bezogenen Handlungen und konzertierten Aktivitäten am Prozess der Wirklichkeitserzeugung teilhaben, kann dieser nicht in subjektiv beliebiger Manier verlaufen. Vielmehr weist er formale Strukturen auf, die sich als kulturell (mehr oder minder) verbreitete Methoden der Handelnden beschreiben lassen, auf die sich diese in ihrem Handeln beziehen und durch die sie soziale Begebenheiten erzeugen. „Ethnomethodologie bezeichnet daher die von den Mitgliedern einer Gesellschaft im Handlungsvollzug praktizierte Methodologie, über welche die – von den Handelnden als vorgegeben erfahren und selbstverständlich hingenommene – gesellschaftliche Wirklichkeit und soziale Ordnung erst produziert werden“ (Bergmann, 1981, S. 11–12). Darin begründet sich, dass für die EM Aussagen über soziale Wirklichkeit immer Aussagen darüber sind, wie in und durch das alltägliche Handeln und die praktischen Erklärungen der Gesellschaftsmitglieder eine intersubjektiv geteilte und sinnhaft geordnete Wirklichkeit hervorgebracht wird. Zur zentralen

---

arbeitet Goffman vorwiegend mit ethnografischen Beobachtungen, während die KA auf technische Aufzeichnungen von Interaktionen besteht. Darüber hinaus wird Goffman seinem eigenen programmatischen Anspruch nicht durchgehend gerecht, sondern greift für seine Erklärungen sozialer Phänomene immer wieder auf Individuen und ihre Psyche zurück (s. dazu Goffman selbst 1967; und aus Perspektive der KA: Schegloff, 1988, S. 94). Die Gemeinsamkeiten und Differenzen werden ausführlich in Schegloff (1988) diskutiert. In den Arbeiten von Goffman selbst zeigen sie sich vor allem in den Aufsätzen in *Forms of Talk* (1981).

Aufgabe wird es deshalb, zu untersuchen, *wie* durch die Handlungen der Gesellschaftsmitglieder „die alltägliche Gewissheit von einer real existierenden sozialen Wirklichkeit intersubjektiv hergestellt“ (Bergmann, 1988b, S. 51) wird.

Die Möglichkeiten der Erzeugung intersubjektivem Sinns im Interaktionsvollzug basieren dabei auf einem spezifischen Verständnis sozialer Handlungen als „praktischen Handlungen“, die den „unausweichlichen Bedingungen der Realisierung von Handlungen in alltäglichen Situationen und Kontexten“ (Bergmann, 1988b, S. 27) unterliegen. Dies bedeutet, dass die AkteurInnen in praktischen Handlungskontexten bei jeder Handlungsrealisierung eine Auswahl aus verschiedenen, möglichen Alternativen treffen. Da die AkteurInnen in ihrem Alltag diesem Entscheidungszwang konkreter Handlungssituationen fortwährend ausgesetzt sind, lautet – in den Worten Garfinkels (1967) – „the practical question *par excellence*: ‚What to do next?‘“ (1967, S. 12, Herv. i. O.). Die Beantwortung dieser Frage stellt für die AkteurInnen jedoch keinen Gegenstand theoretischer Reflektion dar – da sie mit der Handlungsrealisierung versuchen, praktische Ziele zu erreichen, muss die Handlungsselektion an die Prozessualität und Zeitlichkeit der jeweiligen Handlungskontexte angepasst werden, wodurch der „praktische Entscheidungsdruck des alltäglichen Handelns“ (Bergmann, 1988b, S. 29) entsteht. Für den Umgang mit dieser praktischen Herausforderung der Entscheidungsfindung besitzen menschliche AkteurInnen als (kompetente) Mitglieder der jeweiligen Handlungssettings („members“, Garfinkel & Sacks, 1986, S. 163) folgende Lösung: Indem die TeilnehmerInnen ihre Handlungsentscheidungen basierend auf praktischen Theorien über Handlungs- und Sinnzusammenhänge treffen, wird es ihnen möglich, unter dem Handlungsdruck alltäglicher Aktivitätskontexte Selektionen zu treffen und praktische Handlungen durchzuführen.

Dass diese Handlungen zur Hervorbringung der sinnhaften Geordnetheit alltäglicher Aktivitätskontexte beitragen, wird aus Perspektive der EM insbesondere über zwei Eigenschaften praktischer Handlungen erklärt: ihrer Indexikalität und der Reflexivität praktischer Erklärungen. Ausgehend von diesen zentralen Konzepten nimmt die EM an, dass die alltäglichen Handlungsweisen, durch die praktische Aktivitätskontexte hervorgebracht werden, identisch sind mit den Verfahrensweisen, die diese Aktivitäten auf sichtbar-sinnhafte und damit intersubjektiv verständliche Weise ordnen (Garfinkel, 1967, S. vii). Diese Auffassung impliziert einen reflexiven Zusammenhang zwischen dem *sinnhaften* und dem *sinnkonstituierenden* Charakter sozialer Handlungen.

Um diese Konzeption zu verstehen, muss zunächst expliziert werden, was es bedeutet, dass die EM von einer grundsätzlichen Indexikalität – d. h. „Kontextgebundenheit“ (Bergmann, 1988b, S. 34) – jedweder Äußerung und Handlung ausgeht. Die Sinnkonstitution jeder Handlung oder

Äußerung ist maßgeblich davon abhängig, an welchem Ort, zu welchem Zeitpunkt, unter welchen Bedingungen, sowie von welcher Person sie realisiert und an wen sie adressiert wird etc. Sprachliche Verfahren, an denen die Kontextgebundenheit sozialen Sinns besonders augenscheinlich wird, sind situationsabhängige Referenzmittel, bspw. Possessivpronomina wie „meins“, „deins“ oder „unsers“. Die Verwendung derartiger Referenzformen fordert die RezipientIn heraus, „auf den pragmatischen Kontext der Redesituation zurückzugreifen, um die Referenzobjekte lokalisieren, den Sinngehalt einer Äußerung feststellen“ (Bergmann, 1988b, S. 34) zu können. Diese Kontextabhängigkeit derartiger sprachlicher Mittel wurde zum Zeitpunkt der Entwicklung der EM bereits in der Sprachwissenschaft unter dem Begriff der Deixis diskutiert. Die Neuerung der ethnomethodologischen Perspektive lag nun darin, dass es Garfinkel um die Anerkennung der grundsätzlichen Indexikalität aller menschlichen Äußerungen und Handlungen ging, die immer auf den Kontext ihrer Produktion bezogen sind und deren Sinn nur durch Bezugnahme auf ihren Realisierungskontext zu erschließen ist (Garfinkel, 1967, S. 4–11; Garfinkel & Sacks, 1986, S. 169–170).

Für alle sprachlichen Äußerungen gilt damit, dass sie ihre sinnhafte Bedeutung durch eine Kombination aus Inhalt und Realisierungskontext erlangen. Sprachliche Äußerungen besitzen einen gewissen Inhalt und sind aufgrund des *sinnkonstituierenden* Potenzials von Sprache – als sozial objektiviertes Zeichensystem, in dem typische Erfahrungsschemata sedimentiert sind (vgl. Schütz & Luckmann, 2017, S. 319) – in der Lage, spezifische Handlungen zu realisieren. Für die *Sinnkonstitution* muss jedoch immer die Relevanz des Kontexts berücksichtigt werden: Selbst die Bedeutung der formelhaftesten Äußerung wie bspw. „Okay“ ist immer abhängig vom Kontext ihrer Realisierung (Heritage, 2011, S. 159–160). Gleiches gilt ebenso für verkörperte Handlungen wie Gesten oder mimische Reaktionen. Die sinnhafte Bedeutung jeder Äußerung und Handlung ist also unauflöslich mit ihrem jeweiligen Realisierungskontext verknüpft. Garfinkel und Sacks sprechen in diesem Zusammenhang auch davon, dass „the properties of indexicals . . . remain obstinately unavoidable and irremediable“ (1986, S. 170). Aufgrund dieser ‚unheilbaren Indexikalität‘ von Äußerungen ist ihr Sinn „in sozial organisierten Handlungszusammenhängen strukturell ungewiss“ (Bergmann, 1988b, S. 39). Diese Vagheit jeglicher Äußerungen stellt für die AkteurInnen in alltäglichen Verständnisprozessen jedoch kein Problem dar. Sie ist viel mehr als konstitutiv für den situationsangemessenen natürlichen Sprachgebrauch anzusehen, denn sie „ist eine konstitutive Bedingung für Sinngewissheit, also für sinnhaftes Erleben und Handeln“ (Bergmann, 1988b, S. 40). Wenn AkteurInnen in kommunikativen Handlungskontexten eine Äußerung realisieren, gehen sie zunächst ‚naiv‘ davon aus, dass ihre InteraktionspartnerInnen dieser den gleichen Sinn zuschreiben würden, wenn sie an ihrer

Stelle wären. Sie agieren unter der Annahme einer Konsensunterstellung („pretence of agreement“, Garfinkel, 1959, S. 62) und – bis zum Beweis des Gegenteils – im Vertrauen darauf, dass ihr Gegenüber den Sinn einer Äußerung aufgrund ihrer kontextuellen Einbettung verstehen und sich selbst der Sinn von nicht unmittelbar verständlichen Kommunikationsbeiträgen im Interaktionsvollzug etablieren wird (vgl. Abels, 2013, S. 90).

Die Lösung für den vermeintlichen Widerspruch, dass die ‚unheilbare‘ Unbestimmtheit des Sinns jedweder Äußerung oder Handlung gleichzeitig die Bedingung der Möglichkeit wechselseitigen Verstehens und sinnhaften Handelns darstellt, besteht in einem spezifischen ‚Mechanismus von Handlungsrealisationen, der von Garfinkel (1967) folgendermaßen beschrieben wird: „The activities whereby members produce and manage settings of organized everyday are identical with members’ procedures for making those settings ‚account-able‘. The ‚reflexive‘ . . . character of accounting practices and accounts makes up the crux of that recommendation” (ebd., 1967, S. 1). Die These ist also, dass die Methoden, durch deren Realisierung im Interaktionsvollzug gesellschaftliche Begebenheiten ihren Wirklichkeitscharakter erhalten, identisch mit denjenigen Verfahren sind, durch die diese Begebenheiten als sinnhaft geordnet und verständlich – als *accountable* – erscheinen. Wenn *accountable* in deutschsprachigen Publikationen überhaupt übersetzt wird, dann zumeist mit „beobachtbar und berichtbar“ (Garfinkel, 2020, S. 35). Diese Wortkombination wird eingesetzt, um zu erfassen, dass Begriff der *Accountability* im Original „sowohl die sinnhaft-verstehende Aneignung eines Geschehens . . . als auch dessen sprachliche Bezeichnung und Weiterverarbeitung“ (Bergmann, 1988b, S. 45) als konstitutive Ordnungsleistungen des sinnhaften Erkennens, Beschreibens und Erklärens umfasst. Die TeilnehmerInnen an praktischen Handlungskontexten sind ständig damit befasst, die Vorgänge dieser Situation als bestimmte Vorgänge zu erkennen, zu beschreiben und sich mit anderen darüber zu verständigen. Diese *Beobachtbarkeit-und-Berichtbarkeit* wird im Vollzug konkreter Interaktionen durch den Einsatz praktischer Erklärungen hervorgebracht, durch die „the member makes familiar, commonplace activities of everyday life recognizable as familiar, commonplace activities” (Garfinkel, 1967, S. 9, Herv. i. O.). Sinngebungsprozesse werden also nicht als individuelle Bewusstseinsvorgänge, sondern als (interaktions-)öffentliche Prozesse des ‚Sehens-und-Sprechens‘ konzeptualisiert. Die Nicht-Berücksichtigung individueller Bewusstseinsvorgänge<sup>28</sup> in der Frage nach Möglichkeiten sinnhafter sozialer Ordnung begründet sich darin, dass Prozesse des Verstehens und Erklärens nur dann intersubjektiv bedeutungsvoll

---

<sup>28</sup> Die anti-kognitivistische Haltung der EM wird in folgendem Zitat besonders deutlich: „I shall exercise a theorist's preference and say that meaningful events are entirely and exclusively events in person's behavioral environment . . . Hence there is no reason to look under the skull since nothing of interest is to be found there but brains.“ (Garfinkel, 1963, S. 190)

und für die Wirklichkeitskonstitution als „ongoing accomplishment“ (Garfinkel, 1967, S. 1) folgenreich werden können, wenn sie als ‚öffentlich‘ beobachtbare Prozesse stattfinden.

Weil die EM davon ausgeht, dass die beständig ablaufenden Prozesse des Verstehens und Erklärens beobachtbare Phänomene darstellen, schreibt sie sozialen Handlungskontexten einen selbst-explikativen Charakter zu: Da die sinnhaften Handlungen eines Aktivitätskontext und ihre sinnkonstituierenden *accounts* – ihre praktischen Erklärungen – mit sich identisch sind, organisieren sich soziale Handlungskontexte beständig selbst als praktisch erklärbare und sinnhaft geordnete Kontexte mit rationalen Eigenschaften (vgl. Garfinkel, 1967, S. 33). Die TeilnehmerInnen an alltäglichen Handlungskontexten nutzen die konstitutive *Reflexivität* praktischer Erklärungen auf selbstverständliche und unhinterfragte Weise, um die Geordnetheit und Rationalität dieser Settings beobachtbar und berichtbar zu machen. *Accounts* sind dabei nicht als Erklärungen im eigentlichen Sinne zu verstehen, sondern als indexikale Formulierung, die „bloß Markierungen bieten, zwischen denen die durch Konsensunterstellung, Alltagswissen und Interpolation einholbaren ‚Erklärungen‘ hängen“ (Bergmann, 1988b, S. 50). Sie sorgen dafür, dass die Handlungskontexte als etwas potenziell Erklärbares hervorgebracht werden und somit als sinnhaft verständlich behandelt werden können.

Ausgehend von den theoretischen Grundannahmen einer *Vollzugswirklichkeit*, des *indexikalen Charakters* sozialer Handlungen, der *Reflexivität praktischer Erklärungen* und der *Selbst-Explikativität* praktischer Handlungsvollzüge wurden in der (multimodalen) KA methodologische Prämissen für die Untersuchung von F2F-Interaktionen entwickelt. Dies ermöglichen es, empirisch die *formalen* Organisationsprinzipien und *Methoden* zu identifizieren, mittels derer die „Teilnehmer an einem Gespräch in und mit ihren Äußerungen und Handlungen die charakteristischen Strukturmerkmale und die ‚gelebte Geordnetheit‘ . . . des interaktiven Geschehens, in das sie verwickelt sind, hervorbringen“ (Bergmann, 2000, S. 528). Die methodologischen Prämissen der (multimodalen) KA werden im Folgenden vorgestellt.

### **3.2 Die methodologischen Prämissen der multimodalen KA**

Der konversationsanalytische Forschungsprozess ist grundsätzlich davon geprägt, dass aufgrund des ethnomethodologischen Erbes ablehnt wird, ein methodisches Regelwerk auszuformulieren (Bergmann, 1988a, S. 5). Die KA richtet sich damit gegen die verbreitete wissenschaftliche Praxis der Kanonisierung methodischer Vorgehensweisen, die unabhängig von Fragestellung und Gegenstandskonstitution in jeder Studie auf die exakt gleiche Weise befolgt werden könnten. In der KA unterliegen die für eine Fragestellung angemessenen Analyseverfahren stattdessen der methodologischen Prämisse eines „unique adequacy requirement“

(Garfinkel & Wieder, 1992, S. 182): Die Entdeckungsschritte, die zur Identifikation und Rekonstruktion eines Phänomens führen, sind selbst Bestandteil des Phänomens und werden ebenso wie die Analyseergebnisse als Ziel der Untersuchung verstanden. Die ‚passenden‘ Erhebungs- und Analyseverfahren werden erst im Verlauf des Forschungsprozess ‚entdeckt‘ und müssen jeweils den spezifischen Eigenschaften des Phänomens angemessen sein (Bergmann, 2000, S. 530). Trotz dieser Forderung nach einer strikt empiriegeleiteten Entwicklung des methodischen Vorgehens verlaufen konversationsanalytische Studien nicht willkürlich. Sie werden von einer spezifischen „analytic mentality“ (Schenkein, 1978) geleitet und folgen einigen grundsätzlichen methodologischen Prämissen, die im Forschungsprozess auf den jeweiligen Untersuchungsgegenstand zugeschnitten oder angemessen ergänzt werden (müssen).

Im Folgenden werde ich diese methodologischen Prämissen vorstellen<sup>29</sup>. Das Vorgehen der in dieser Arbeit verwendeten multimodalen KA entspricht grundsätzlich den ursprünglich entwickelten Verfahrensgrundsätzen der KA – aufgrund der Audiovisualität des Datenmaterials und der systematischen Berücksichtigung der Multimodalität von Interaktionen müssen jedoch einige dieser Prinzipien gegenstandsangemessen modifiziert werden. Ich werde diese Modifikationen in der folgenden Darstellung an den entsprechenden Stellen explizit kenntlich machen.

### **3.2.1 Datengrundlage – Technische Konservierungen natürlicher Interaktionen**

Der Anspruch der KA ist es, „to discover the natural living order of social activities as they are endogenously organized in ordinary life, without the exogenous intervention of researchers imposing topics and tasks or displacing the context of action“ (Mondada, 2012b, S. 34). Um die ‚gelebte Geordnetheit‘ alltäglicher Aktivitäten rekonstruieren zu können, werden in konversationsanalytischer Studien ausschließlich „registrierende Konservierungen“ (Bergmann, 1985) – d.h. Audio- oder Videoaufnahmen – von möglichst ‚natürlichen‘ Interaktionen als Datengrundlage verwendet, d. h. soziale Situationen aufgezeichnet, die so auch ohne das Zutun der ForscherIn stattgefunden hätten (Schu, 2001, S. 1015).

---

<sup>29</sup> Ausführliche Darstellungen der KA finden sich beispielsweise in ten Have (2007), Silverman (1998), Psathas (1995) und im deutschsprachigen Raum vor allem in den Studienbriefen von Bergmann (1988a) oder bei Gülich & Mondada (2008). Umfassend lässt sich die Entwicklung der KA – ähnlich wie die der EM – nur in Abgrenzung von den zur Zeit ihrer Entwicklung vorherrschenden Strömungen in den Sozialwissenschaften verstehen. Sowohl Garfinkel als auch Sacks zeigten sich hochgradig skeptisch gegenüber den gängigen soziologischen Verfahren, die entweder in der Tradition Durkheims eine quantitative Ermittlung durchschnittlicher Ausprägungen der Eigenschaften sozialer Prozesse anstrebten oder in der Tradition Max Webers Idealtypen sozialer Zusammenhänge entwickelten. Aus Perspektive der EM und KA führen beide Verfahren dieser „constructive analysis“ (Garfinkel & Sacks, 1986, S. 162) notwendigerweise dazu, dass die durch ihre spezifischen Details konstituierten Eigenschaften der untersuchten sozialen Ereignisse verloren gehen, sodass soziologische Konzepte und Verallgemeinerungen in einer unbestimmten Beziehung zu den Ereignissen stehen, aus denen sie entwickelt wurden (Sacks, 1963; s. a., Heritage, 1984a, S. 234–235).

Diese methodologische Restriktion ist eine unmittelbare Konsequenz aus der theoretischen Prämisse, dass „every feature of an activity’s sense, facticity, objectivity, accountability, communability is to be treated as a contingent accomplishment of organized common practices“ (Garfinkel, 1967, S. 33). Wenn die Faktizität sozialer Wirklichkeit als lokal, in und durch die Handlungsvollzüge der TeilnehmerInnen im Vollzug konkreter Interaktionen hervorgebracht verstanden wird, muss eben diese Ereignishaftigkeit sozialer Geschehnisse durch eine technische *Registrierung* festgehalten und in all ihren Details für die Analyse verfügbar gemacht werden. Für sprachliche Interaktion bedeutet dies, dass (mindestens) Audioaufzeichnungen der hörbaren Interaktionsbeiträge aller TeilnehmerInnen angefertigt werden. Für multimodale Konversationsanalysen gilt es, die Videokamera(s) so zu positionieren, dass der gesamte Interaktionsraum und damit alle für die Interaktionsorganisation relevanten sprachlichen und verkörperten Beiträge der TeilnehmerInnen aufgezeichnet werden (Mondada, 2006, S. 55). Die Forderung, soziale Aktivitäten in ihrem natürlichen Kontext zu untersuchen, begründet sich in der Annahme, dass sich nicht im Vorhinein bestimmen lässt, welche Aspekte des Kontexts für die TeilnehmerInnen beobachtbare Handlungsrelevanz besitzen und diese erst in der Untersuchung der endogenen Organisationsprinzipien entdeckt werden müssen.

Durch technische Aufzeichnungen wird dem Umstand Rechnung getragen, dass es sich bei Interaktionen um flüchtige Phänomene (Bergmann, 1985, S. 317) handelt, und sich erst durch ihre technische Fixierung „die im praktischen Handeln reflexiv eingebetteten Prozesse der sinnhaften Strukturierung . . . erkennen und beschreiben“ (vgl. Bergmann, 1988a, S. 10–11) lassen. Einmal festgehalten, bieten die Aufzeichnungen die Möglichkeit, die Interaktionen beliebig oft abzuspielen, Analysen hinsichtlich ihrer Gegenstandsangemessenheit mit dem ursprünglichen Datum zu vergleichen (Sacks, 1984, S. 26) und sich dem gleichen Material mit neuen Fragestellungen zu widmen (Pomerantz & Fehr, 1997, S. 170; Schu, 2001, S. 1015).

Andere Datentypen, die durch Verfahren der „rekonstruierenden Konservierung“ (Bergmann, 1985, S. 305) generiert werden, bspw. Interviews oder in Feldnotizen festgehaltene Beobachtungen der ForscherIn, werden aus konversationsanalytischer Perspektive abgelehnt. Die Begründung dafür ist, dass *rekonstruierende Konservierungen* durch konstruktive Prozesse geschaffen werden, „in which the specific details of naturally situated interactional conduct are irretrievably lost and are replaced by idealizations about how interaction works“ (Heritage, 1984a, S. 236): Gemeint ist damit, dass

1. jede nachträgliche Rekonstruktion eines flüchtigen sozialen Ereignisses – bspw. die in einem Interview elizitierte Stegreiferzählung einer TeilnehmerIn – zwangsläufig eine Deutung dieses

Geschehens darstellt, die „gegenüber dem primären Sinnzusammenhang des sich vollziehenden sozialen Geschehens einen sekundären Sinnzusammenhang“ (Bergmann, 1988a, S. 12) bildet.

2. *Rekonstruierende* Konservierungen nicht dem Verlauf des originalen Geschehens folgen, sondern je eigenen Darstellungsprinzipien unterliegen. Dies führt dazu, dass der Detaillierungsgrad und die chronologische Reihenfolge des Originalgeschehens nach den Relevanzen der jeweiligen Gattung neu geordnet und kondensiert werden (Bergmann, 1985, S. 305–306).

3. Für jedes Ereignis eine theoretisch unendliche Anzahl rekonstruktiver Versionen existiert, die eine *indexikale* Struktur besitzen, da sie immer abhängig von den jeweils adressierten RezipientInnen und den mit ihnen verfolgten Zwecken ausgestaltet werden. *Registrierende Konservierungen* erlauben hingegen eine Aufzeichnung der primären Sinnzusammenhänge des Originalgeschehens, die dem Verlauf dieses Ereignisses folgt und mit dessen Beendigung abgeschlossen ist.

Registrierende Aufzeichnungen möglichst natürlicher Interaktionen sind entsprechend Voraussetzung dafür<sup>30</sup>, die ethnomethodologische Frage nach dem „Wie“ der lokalen und methodischen Hervorbringung sinnhafter sozialer Ordnung (vgl. Bergmann, 1988b, S. 54; Abels, 2013, S. 104) in F2F-Interaktionen empirisch beantworten zu können.

### **3.2.2 Datenaufbereitung – Transkription audiovisueller Aufzeichnungen**

Registrierende Aufzeichnungen stellen also die Datengrundlage für konversationsanalytische Untersuchungen dar – als Audio- oder Videoaufnahmen entziehen sie sich jedoch noch einem einfachen analytischen Zugriff. Die technisch aufgezeichneten Details der Interaktion müssen deshalb von der AnalytikerIn zunächst transkribiert, d.h. in eine schriftliche Form überführt werden (vgl. ten Have, 2007, S. 6). Im konversationsanalytischen Forschungsprozess stellen die Aufzeichnungen die Daten dar, während die Transkripte als ihre schriftlichen

---

<sup>30</sup> Mit einer vergleichbaren Begründung werden auch experimentelle Forschungsdesigns in der KA größtenteils abgelehnt. Diese Ablehnung beruht zum einen darauf, dass das Verhalten der TeilnehmerInnen in experimentellen Forschungsdesign gezielt beeinflusst wird und es das Ziel der KA ist, die ‚natürliche‘ Geordnetheit sozialer Interaktionen zu rekonstruieren (Mondada, 2012b, S. 34). Zum anderen begründet sich der Skeptizismus der KA in den „at present non-calculable effects of imposed artifice on the conduct of interaction“ (Schegloff, 1996c, S. 28). Aus Perspektive der KA lässt sich die Frage danach, welche Verhaltensweisen der TeilnehmerInnen als Effekte des experimentellen Forschungsdesign interpretiert werden müssen (Heritage, 1984a, S. 238; Schegloff, 1991, S. 56), nur beantworten, wenn die Eigenschaften der gelebten Geordnetheit des interessierenden Phänomens bekannt sind. Diese erforderliche Kenntnis ist jedoch in den meisten Experimentalstudien nicht gegeben, sodass sich die ökologische Validität der Ergebnisse nicht oder nur schwer beurteilen lässt (Kendrick, 2017, S. 7). In den letzten 10–15 Jahren werden jedoch zunehmend konversationsanalytisch fundierte Experimentalstudien durchgeführt, die Erkenntnisse ‚klassischer‘ Konversationsanalysen in quasi-experimentellen Settings überprüfen und durch neue Methoden der Datenaufzeichnung vertiefen (bspw. Holler & Kendrick, 2015; Hömke et al., 2017; Kendrick & Holler, 2017; Stivers et al., 2009). Für eine ausführliche Diskussion der methodologischen Bedingungen der Vereinbarkeit von KA und Experimentalforschung s. Kendrick (2017) sowie Ruiter und Albert (2017).

Repräsentationen behandelt werden (Hutchby & Wooffitt, 1998, S. 74; S. a. Mondada, 2007). Sie dienen einerseits als praktisches Hilfsmittel, um sich auf die Daten beziehen zu können, zum anderen werden sie in Publikationen als Belege für die empirische Verankerung der Analysen verwendet und fungieren somit als zentrale „objects of knowledge“ (C. Goodwin, 2015b, S. 387) in den Praktiken der konversationsanalytischen Wissensproduktion.

Damit Transkripte diese Funktionen übernehmen können, wurden spezifische Transkriptionskonventionen ausgearbeitet, die dem Anspruch folgen, das soziale Geschehen in seinem zeitlichen Verlauf und in all seinen Details zu bewahren. Dies soll eine Rekonstruktion derjenigen Teilnehmermethoden ermöglichen, mit denen die AkteurInnen Interaktionen in ihrem zeitlichen Verlauf als sinnhaft geordnet hervorbringen. Aufgrund der in den ersten Jahrzehnten der KA überwiegenden Arbeit mit Audioaufzeichnungen wurden dabei zunächst Transkriptionskonventionen ausgearbeitet, die beanspruchen, „nicht nur den Inhalt des Gesprächs bewahren, d. h. *was* gesagt wurde, sondern auch die Details dessen, *wie* etwas gesagt wurde, also die besondere Realisierungsform“ (Bergmann, 1988a, S. 19, Herv. i. O.). Das von Gail Jefferson entwickelte Transkriptionssystem (1985, 1996) wurde über lange Zeit universell in KA-Studien verwendet. Im Laufe der Zeit wurden davon ausgehend weitere Konventionen entwickelt, wobei sich hier vor allem das gesprächsanalytische Transkriptionssystem (GAT, Selting et al., 1998, bzw. GAT2, 2009) im deutschsprachigen Raum durchsetzen konnte.

Schon bei der Transkription von Tonaufzeichnungen geht die KA mit ihren Konventionen dabei einen Kompromiss aus Genauigkeit und Lesbarkeit ein: Zum einen wird versucht, alle Details gesprochener Sprache festzuhalten. Entsprechend werden die Transkripte nicht von den typischen Merkmalen des natürlichen Sprachgebrauchs wie Versprechern, Abbrüchen, Stockungen, Überlappungen, dialektalen Färbungen, paraverbalen Äußerungen und Überlappungen bereinigt, sondern diese festgehalten und so der wiederholten Analyse zugänglich gemacht. Auf diese Weise entstehen im Transkriptionsprozess von links nach rechts und von oben nach unten lesbare, schriftliche Repräsentationen des Interaktionsvollzugs, die für sich beanspruchen, die spezifische Realisierungsform der sprachlichen Interaktion darzustellen. Gleichzeitig müssen diese Transkripte jedoch auch für die größtenteils nicht sprachwissenschaftlich ausgebildeten RezipientInnen lesbar sein: Um dies zu gewährleisten, wird auf eine phonetische Transkription verzichtet und stattdessen das lateinische Alphabet verwendet, das um zahlreiche spezielle Zeichen ergänzt wird, mit denen sich die Realisierungsformen sprachlicher Äußerungen – bspw. Überlappungen oder Wortabbrüche – konventionalisiert repräsentieren lassen (vgl. Ayaß, 2015, S. 508). Trotz aller Bemühungen um Lesbarkeit sind diese Transkripte jedoch für LeserInnen,

die nicht mit den Annotationskonventionen vertraut sind, nur schwer zu verstehen (vgl. Bucholtz, 2000, S. 1461).

Diese Ambivalenz von Genauigkeit und Lesbarkeit verschärft sich für die Transkription von audiovisuellen Daten noch einmal. Denn während die Herausforderung bei der Transkription von Audioaufzeichnungen darin besteht, „die auf Tonträgern verfügbar gewordene gesprochene Sprache mit ihren linguistischen Spezifika so getreu wie möglich ins Medium der Schrift zu übersetzen“ (Stukenbrock, 2009, S. 147), gilt es bei Videoaufnahmen nun, „einen Strom ineinander übergehender, sowohl sequenziell als auch simultan vielfach ineinander verschlungener audio-visueller Phänomene auf den unterschiedlichsten Ausdrucksebenen schriftlich zu fixieren. (ebd., 2009, S. 147). Während für die Transkription von Audiodaten fest etablierte Konventionen existieren, ist die Entwicklung einer empirisch fundierten Transkriptionskonvention multimodaler Interaktion, „which defines which phenomena need to be transcribed and in which granularity“ (Deppermann, 2013, S. 3), bis dato noch nicht abgeschlossen. Es existieren verschiedene alternative Vorschläge, bspw. Entwürfe zu Transkription von Gesten (Kendon, 1990), von Blicken (C. Goodwin, 1981) oder beidem (Heath, 1986; Streeck, 1993). In den letzten Jahren finden jedoch die von Lorenza Mondada entwickelten und auf den GAT2-Prinzipien aufbauenden Konventionen (Mondada, 2007, 2016a, 2016b) zunehmend Verbreitung in der wissenschaftlichen Gemeinschaft.

Unabhängig von der konkreten Wahl des Annotationssystems besteht die Herausforderung im Prozess der Transkription darin, alle Phänomene der multimodalen Interaktion in eine schriftliche Form zu überführen und damit linear, in klar voneinander abgegrenzten Einheiten darzustellen (Ayaß, 2015, S. 512). Dabei muss eine Darstellungsweise gefunden werden, die in der Lage ist, die Zeitlichkeit multimodaler Interaktion, „that combines multiple successive and simultaneous lines of conduct“ (Mondada, 2018, S. 88) zu erfassen. Das für sprachliche Interaktion kennzeichnende „one party talks at one time“-Prinzip (Sacks et al., 1974, S. 699) gilt nicht für den Einsatz visueller Ressourcen, die zeitgleich zur Äußerung einer SprecherIn von dieser selbst oder anderen Anwesenden realisiert werden können. Die Verschränkung der verschiedenen Zeitlichkeiten der unterschiedlichen eingesetzten Ressourcen stellt eine konstitutive Eigenschaft multimodaler Interaktion dar. Sie erzeugt Herausforderungen für die schriftliche Repräsentation im Transkript, die zwar durch die Integration von Standbildern bearbeitet (Ayaß, 2015; Stukenbrock, 2009), jedoch nicht vollständig gelöst werden kann.

Dies liegt nicht zuletzt daran, dass die Reichhaltigkeit audiovisueller Daten jeglichen Versuch einer lückenlosen, multimodalen Transkription zu einem fiktiven Vorhaben macht: „Die Vorstellung, dass das Transkript sozusagen ‚alle‘ wesentlichen Informationen enthalte . . . ist nicht

mehr aufrecht zu erhalten“ (Stukenbrock, 2009, S. 147), denn ein Transkript, das alle sprachlichen Äußerungen und jegliches beobachtbares körperliches Verhalten schriftlich repräsentieren würde, könnte dem Anspruch an Lesbarkeit (endgültig) nicht mehr gerecht werden: Multimodales Transkribieren ist dementsprechend ein hochgradig selektiver Prozess<sup>31</sup>: Die Entscheidungen darüber, welche Elemente der Interaktion im Transkript abgebildet werden, werden in Abhängigkeit davon getroffen, welche Aspekte des multimodalen Verhaltens von den TeilnehmerInnen auf beobachtbare Weise als relevant für die Interaktionsorganisation behandelt werden (vgl. Mondada, 2018, S. 88). Eine Beurteilung dieser Relevanz kann dabei nur ausgehend von einer vorherigen Analyse der Interaktionsepisode erfolgen. Aufgrund der vielfachen analytischen und konstruktiven Entscheidungen der ForscherIn (vgl. Davidson, 2009, S. 38), die bei der Erstellung multimodaler Transkripte getroffen werden, hängen „Gegenstand und Zweck der Transkription, Analyse und Darstellung audio-visueller Daten . . . viel enger zusammen, als dies beim Verbaltranskript der Fall ist“ (Stukenbrock, 2009, S. 168). Stärker noch als Verbaltranskripte müssen multimodale Transkripte deshalb als Ergebnis der detaillierten Analyse einer Interaktionsepisode verstanden werden. Entsprechend ermöglichen sie in einem engen Sinne keine intersubjektive Nachvollziehbarkeit der Analysen (vgl. Deppermann, 2013, S. 3). Sie fungieren viel mehr als Wissensobjekte, die es der LeserIn ermöglichen sollen, nachzuvollziehen, welche Aspekte der multimodalen Interaktion analytisch relevant sind.

Bereits diese kurze Diskussion der Herausforderungen des Transkriptionsprozesses verweist darauf, dass die Auffassung von Transkripten als schriftlichen Repräsentationen der Daten nicht darüber hinwegtäuschen darf, dass Transkribieren einen konstruktiven Prozess der Übersetzung von Ton (und Bild) in Text darstellt (Duranti, 2007; ten Have, 2007): Zum einen werden konversationsanalytische Transkriptionskonventionen von den theoretischen Grundannahmen (vgl. Ochs, 1979, S. 44) über den Vollzugscharakter sozialer Wirklichkeit durchdrungen und wurden entwickelt, um diesen im Analyseprozess als methodische Hervorbringungsleistung der TeilnehmerInnen rekonstruieren zu können. Nur durch die Anfertigung derart detaillierter

---

<sup>31</sup> Die Selektivität des Transkriptionsprozesses wird bei der Transformation audiovisueller Daten in Schrift besonders deutlich. Das gleiche Argument trifft jedoch – in geringerem Ausmaß – auch auf die Arbeit mit Audiodaten zu: Betrachtet man konversationsanalytische Studien, in denen Audioaufzeichnungen mithilfe der weitverbreiteten Konventionen von GAT2-Basistranskripten (Selting et al., 2009, S. 373) schriftlich repräsentiert werden, so fällt bspw. auf, dass die Prosodie gesprochener Sprache meist nur insofern berücksichtigt wird, als dass Tonhöhenverschiebungen am Ende von Redezügen transkribiert werden. Damit wird impliziert, dass der Einsatz prosodischer Ressourcen an anderen turn-internen Positionen eine geringere Relevanz für die Interaktionsorganisation besitzt, als andere verbale Ressourcen, die umfassender im Transkript berücksichtigt werden. Dass derartige Entscheidungen als konstruktive Leistungen der ForscherIn in der Konstitution des Datenmaterials reflektiert werden müssen (Ochs, 1979), wurde über lange Zeit in der konversationsanalytischen Literatur wenig reflektiert, findet jedoch seit einigen Jahren vor allem aufgrund der ‚Offensichtlichkeit‘ derartiger Prozesse im Umgang mit Videodaten vermehrt statt.

Transkripte können die Details der ‚gelebten Geordnetheit‘ von Interaktionsphänomenen wie bspw. Weinen (Hepburn, 2004) oder Lachen (Jefferson, 1985) überhaupt erst entdeckt und beschrieben werden (vgl. Davidson, 2009, S. 37). Da niemals alle Details multimodaler Interaktionen schriftlich repräsentiert werden können, ist Transkribieren darüber hinaus notwendigerweise immer ein selektiver Prozesse, der sowohl die theoretischen Ziele und Gegenstandsdefinitionen der KA (vgl. Hutchby & Wooffitt, 1998, S. 76), als auch die Entscheidungen der ForscherIn reflektiert, die diese vor dem Hintergrund ihres spezifischen Erkenntnisinteresses<sup>32</sup> trifft (vgl. Davidson, 2009, S. 38). Derartige konstruktivistische Überlegungen sprechen gegen eine ‚naive‘ Behandlung von Transkripten als ‚neutralen‘ schriftlichen Repräsentationen der untersuchten Interaktionen. Sie erkennen an, dass (multimodale) Transkripte das Resultat theoriegeleiteter Konventionen und situierter Praktiken (Mondada, 2007, S. 810) darstellen, in ihrer konkreten Gestalt durch vielschichtige Entscheidungen der ForscherIn geprägt sind und als Ergebnisse des Analyseprozesses für die LeserInnen nachvollziehbar machen, welche Aspekte des multimodalen Verhaltens der TeilnehmerInnen für die Hervorbringung sinnhaft geordneter Interaktion relevant sind. Von welchen Verfahrensgrundsätzen dieser Analyseprozess gekennzeichnet ist, werde ich im Folgenden darstellen.

### **3.3.3 Datenauswertung – die analytischen Maxime der multimodalen KA**

Liegen mit Videoaufnahmen technische Aufzeichnungen möglichst natürlicher Interaktionssituationen und Transkripte als praktische Hilfsmittel vor, kann die ‚eigentliche‘ analytische Arbeit beginnen. Der Analyseprozess ist dabei grundsätzlich von einer Geordnetheitsunterstellung gekennzeichnet (vgl. Schegloff & Sacks, 1973, S. 290)<sup>33</sup>. Diese besagt, dass kein Element der

---

<sup>32</sup> In diesem Zusammenhang wird seit einiger Zeit auch der veränderliche Status von Transkripten in der methodologischen Literatur reflektiert: Der Prozess der Datenaufbereitung wird in der KA selbst als konstitutiver Bestandteil der analytischen Arbeit verstanden (Hutchby & Wooffitt, 1998) und in der Regel von der ForscherIn selbst übernommen (Ayaß, 2015, S. 508). Im Zuge des Transkriptionsprozesses gewinnt diese immer detaillierteres Wissen über ihren Datenkorpus und generiert analyseleitende Ideen. Mit den Veränderungen des Wissens der ForscherIn über ihren Gegenstand geht einher, dass sich Transkripte einzelner Interaktionsepisoden in diesem Prozess verändern können (Mondada, 2007) und andere Interaktionsphänomene ins Zentrum des Interesses rücken (Davidson, 2009, S. 42), die wiederum mit spezifischen Anforderungen an das Transkript verbunden sind. Diese Veränderungen sind dabei nicht als kumulative Schritte hin zu einem zunehmend ‚besseren‘ Transkript zu verstehen, denn in Abhängigkeit des analytischen Interesses kann der Detaillierungsgrad sowohl erhöht als auch verringert werden. Darüber hinaus können rezipientenspezifische Überlegungen oder Anforderungen der HerausgeberInnen schriftlicher Publikationen die konkrete Ausgestaltung der Transkripte beeinflussen. Entsprechend muss ein Transkript als „evolving flexible object“ (Mondada, 2007, S. 810) verstanden werden, das keinesfalls in einem statischen Abbildungsverhältnis zum Originalgeschehen steht. Transkripte erhalten ihre ‚finale Form‘ immer für alle praktischen Zwecke einer Präsentation oder Publikation, in der ausgewählte Interaktionsdetails für eine rezipientenspezifische ausgearbeitete Analyse hervorgehoben werden (vgl. Mondada, 2007, S. 819).

<sup>33</sup> Die Ordnungsprämisse wurde von Sacks bereits in seinen frühen „Lectures on Conversation“ (1995) und in kritischer Auseinandersetzung zu alternativen soziologischen Erklärungsansätzen formuliert: Diesen warf er vor, bereits a priori Entscheidungen zu treffen, welche Phänomene als „good problems“ (Sacks, 1984, S. 22) eine hinreichende Relevanz für ihre sozialwissenschaftliche Erforschung besitzen (bspw. soziale Ungleichheit,

Interaktion von Vorneherein als mögliches Untersuchungsobjekt ausgeschlossen, sondern „vielmehr zunächst einmal als Bestandteil einer sich im Handeln der Beteiligten reproduzierenden Ordnung betrachtet“ (Bergmann, 1988a, S. 27) wird. Diese Grundannahme verpflichtet die ForscherIn zu einer radikalen Offenheit gegenüber der Entdeckung von Prozessen der Hervorbringung sozialer Ordnung selbst in den unscheinbarsten Interaktionsphänomenen, ohne a priori Entscheidungen darüber zu treffen, welche Details der Interaktion für die TeilnehmerInnen relevant sind.

Die für den Analyseprozess kennzeichnende *analytische Mentalität* (Schenkein, 1978) der KA verlangt von der ForscherIn dabei sowohl Sensibilität für die kontextsensitiven Bedeutungsnuancen einzelner Interaktionsepisoden als auch Abstraktionsfähigkeit, um wiederkehrende und bis zum einem gewissen Grad kontextunabhängige Ordnungsmuster identifizieren zu können (vgl. Bergmann, 1988a, S. 7). Im Prozess der Identifikation der beobachtbaren, ordnungsgenerierenden Verfahren der TeilnehmerInnen stellen die praktischen Handlungskompetenzen und das Alltagswissen der ForscherIn eine wichtige Ressource dar: Die im alltäglichen Handeln der TeilnehmerInnen eingebetteten Methoden machen die interaktive Kompetenz von Gesellschaftsmitgliedern aus. Untersucht eine ForscherIn nun ihr vertraute Phänomene, verfügt sie ebenfalls über diese Kompetenzen und kann sie als Ressource nutzen, um die sinnhafte Strukturierung der Interaktionsepisoden zu erkennen. Dieser – immer stattfindende – Rückgriff auf das eigene (TeilnehmerInnen-)Wissen muss jedoch stets reflektiert werden, um zu explizieren, „welche (ethno-)analytischen Mittel und Techniken“ (Bergmann, 1988a, S. 39) der ForscherIn dazu verholfen haben<sup>34</sup>, eine Gesprächspraktik als solche zu erkennen. KonversationsanalytikerInnen bedienen sich dabei mit der Heuristik des „doing“ (Sacks, 1985)<sup>35</sup> einer systematischen Befremdungstechnik, die diesen Prozess unterstützen soll: Jede, noch so vertraute Aktivität wird als „doing“ – als etwas, das ‚getan‘ werden muss – aufgefasst, um die analytische Aufmerksamkeit auf diejenigen formalen (TeilnehmerInnen-)Methoden zu lenken, deren Realisierung eine Aktivität wie bspw. einen Streit, ein Witz oder eine Erzählung überhaupt erst als solche erkennbar macht (vgl. Churchill, 1971, S. 183).

---

Klassenkonflikte, deviantes Verhalten etc.). Im Gegensatz dazu vertritt die KA nun die Haltung, dass „we may alternatively take it that there is order at all points“ (Sacks, 1984, S. 22) und macht es sich zur Aufgabe, die ordnungsgenerierenden Handlungsprinzipien zu rekonstruieren.

<sup>34</sup> Das wohl bekannteste Beispiel einer derartigen Explikation ethno-analytischer Techniken der Sinnerschließung liefert Harvey Sacks in seinem Aufsatz „On the analyzability of stories by children“ (1972), in dem er aufzeigt, mittels welcher Interpretations- und Kategorisierungsverfahren die zwei Sätze „The baby cried. The mommy picked it up“ als Rekonstruktion eines Ereignisses ‚gehört‘ werden, in dem eine Mutter *ihr* weinendes Kind auf den Arm nimmt.

<sup>35</sup> Harvey Sacks (1985) untersucht in diesem paradigmatischen Aufsatz, welche praktischen Leistungen erbracht und welche formalen Methoden von den Mitgliedern eines Settings für die erkennbare Hervorbringung von „doing being ordinary“ – gewöhnlichem Verhalten – in sozialen Kontexten eingesetzt werden..

Mit einem derartig befremdenden Blick auf alltägliche und oftmals hochgradig vertraute Interaktionsphänomene beginnen konversationsanalytische Studien ihre Analysen meist mit einer ‚unmotivierten‘ Untersuchung des Datenmaterials, die nicht von vorher definierten Fragestellungen oder Annahmen, sondern dem ‚Bemerken‘ von vorher unscheinbaren Eigenschaften des sprachlichen und verkörperten Verhalten der TeilnehmerInnen (vgl. Schegloff, 1996a, S. 172) geleitet wird. Ausgehend von „einfachen Beobachtungen“ (Bergmann, 1988a, S. 30) von Interaktionsphänomenen wird im weiteren Verlauf der Analyse eruiert, ob es sich dabei um ein „Ordnungselement“ (Bergmann, 1988a, S. 32) – vorläufig verstanden als wiederholt auftretendes Phänomen im Korpus – handelt, das sich in unterschiedlichen Kontexten und Realisierungsformen in verschiedenen Fällen des Datenmaterials identifizieren lässt. Aus diesen Fällen werden Kollektionen gebildet, die die eigentliche Datenbasis für konversationsanalytische Studien bilden.

In der Arbeit mit Videoaufnahmen entstehen bei der Kollektionsbildung aufgrund der Reichhaltigkeit des audiovisuellen Datenmaterials einige Herausforderungen (vgl. Mondada, 2008): Die AnalytikerIn muss eine große Bandbreite hör- und sichtbarer Eigenschaften des Datenmaterials gleichzeitig berücksichtigen und Phänomene, die an der linguistischen Gesprächsoberfläche Gemeinsamkeiten aufweisen, können durchaus von zentralen Differenzen gekennzeichnet sein, wenn man den gleichzeitigen Einsatz verkörperter, räumlicher und materieller Ressourcen berücksichtigt (Deppermann, 2013, S. 3–4). Bei der Arbeit mit Videodaten multiplizieren sich entsprechend die Fragen danach, welche Eigenschaften multimodaler, sozialer Handlungen simultane Relevanz besitzen, und fordern die ForscherIn heraus, diejenigen Aspekte der verschiedenen Modalitäten zu identifizieren, die für die Handlungsrealisierung essenziell, optional oder irrelevant sind (Deppermann, 2013, S. 4). Bei der Kollektionsbildung muss also der Einsatz unterschiedlicher Ressourcen berücksichtigt werden, um zusammengehörende Fälle zu identifizieren, die als methodische Hervorbringung eines Ordnungselements (Bergmann, 1988a, S. 32) klassifiziert werden können<sup>36</sup>.

Anschließend muss der Nachweis entwickelt werden, dass es sich bei diesem Ordnungselement nicht lediglich um eine wiederholt auftretende Regelmäßigkeit, sondern um „interaktiv erzeugte Geordnetheit“ (Bergmann, 1988a, S. 34) handelt: D.h. es muss rekonstruiert werden,

---

<sup>36</sup> Diese Herausforderungen führen dazu, dass sich systematische Verfahrensvorschläge für die Kollektionsbildung multimodaler Konversationsanalysen erst in der Entwicklung befinden (Kamunen, 2019, S. 423) und bis dato viele Studien – ungeachtet dessen, ob sie eine begrenzte Anzahl multimodaler Details (bspw. die Kombination sprachlicher Äußerungen mit einer bestimmten Geste) oder komplexe *multimodale Gestalten* untersuchen (Mondada, 2018, S. 87) – auf detaillierte Einzelfallanalysen abstellen. Eine explizite methodologische und theoretische Argumentation, warum auch derartige Studien zu Ergebnissen gelangen können, die über den jeweiligen Einzelfall hinausweisen, findet sich bei Schegloff (1987a). Ich werde diese Argumentation in der Beschreibung meines eigenen analytischen Vorgehens in Kap. 4.4 ausführlicher erläutern.

mittels welcher Methoden die TeilnehmerInnen diese Geordnetheit auf beobachtbare und berichtbare Weise hervorbringen und welches ‚Problem‘ der Interaktionsorganisation dabei von ihnen bearbeitet wird. Der Begriff des Problems wird in der KA dabei „über seine Alltagsbedeutung hinausreichend, für alle Aufgaben, Funktionen, Zwecke und Ziele, an denen die Interaktanten die Gestaltung ihrer Gesprächsbeiträge“ (Deppermann, 2008, S. 81) und verkörperten Handlungen ausrichten, verwendet. Die zentrale Herausforderung besteht dabei oftmals darin, dass sich in den Details des Verhaltens der TeilnehmerInnen lediglich die methodischen Lösungen potenzieller Probleme, nicht jedoch das Auftreten des Problems selbst beobachten lässt (vgl. Schegloff & Sacks, 1973, S. 290). Für die ForscherIn stellt sich deshalb die Aufgabe, dieses Einzelfallanalysen und den strategischen Kollektionsaufbau zu identifizieren, sodass es gelingen kann, „den methodischen Apparat zu beschreiben, der dafür sorgt, dass die strukturellen Probleme der Gesprächsorganisation für die Interagierenden ‚unproblematische Probleme‘ bleiben“ (Bergmann, 1988a, S. 37).

Allen konversationsanalytischen Studien ist grundsätzlich gemein, dass sie im Ergebnis diejenigen Verfahren und Regeln rekonstruieren, die von TeilnehmerInnen eingesetzt werden, um soziale Ordnung und die sinnhafte Strukturierung von Interaktion hervorzubringen. Sie unterscheiden sich jedoch darin, wie sehr sie auf die Identifikation der generischen, kontextunabhängigen Eigenschaften der Organisationsprinzipien oder die kontextsensitive Partikularisierung der formalen (TeilnehmerInnen-)Methode abstellen. Dies stellt jedoch keinesfalls einen Widerspruch dar, denn die KA erachtet – wie bereits diskutiert – die Kontextgebundenheit jedweden menschlichen Handelns als konstitutive Eigenschaft von Interaktion. Gerade weil die verschiedensten im Rahmen von F2F-Interaktionen realisierbaren Aktivitäten an die unterschiedlichsten Umstände und InteraktionspartnerInnen angepasst werden können, geht die KA davon aus, dass ein formaler Apparat existiert, der selbst kontextunabhängig funktioniert. Kontextunabhängigkeit und Kontextsensitivität stehen also in einem komplementären Verhältnis: Die kontextunabhängigen Regeln (bspw. des Sprecherwechsels) der ‚Interaktionsmaschinerie‘ (vgl. Sacks, 1984, S. 29) operieren kontextsensitiv, indem sie auf spezifische Weise und an distinkten sequenziellen Positionen für den jeweiligen Kontext der konkreten Interaktionssituation adaptiert und partikularisiert werden (vgl. Sacks et al., 1974, S. 699; Bergmann, 1981, S. 30).

Mit dem Ziel der Identifikation der kontextunabhängigen Eigenschaften der Organisationsprinzipien von Interaktion ist auch der spezifische Verallgemeinerungsanspruch der KA verbunden: Ihre Ergebnisse beanspruchen, „ein ‚*generatives Prinzip*‘ [zu] enthalten, das nicht nur die untersuchten Fälle erklären, sondern auch weitere ähnliche Fälle produzieren könnte“ (Herv. i. O.

Bergmann, 1988a, S. 42). Dabei muss immer der empirische Nachweis erbracht werden, dass sich die TeilnehmerInnen in ihrem Verhalten beobachtbar an den postulierten Regelmäßigkeiten als handlungsleitend orientieren (vgl. Schegloff, 1991, S. 52ff.). Für die Validierung ihrer Ergebnisse nutzt die KA dabei insbesondere zwei analytische Verfahrensweisen: Die „next-turn proof procedure“ (vgl. Sacks et al., 1974, S. 728–729) sowie die Untersuchung von „abweichenden Fällen“ (Bergmann, 1988a, S. 46).

Mit abweichenden Fällen sind Interaktionsepisoden gemeint, an deren Verlauf sich die TeilnehmerInnen als Abweichung orientieren, bspw. indem sie die auftretenden Unregelmäßigkeit reparieren, sie als abweichendes Geschehen markieren oder von Anwesenden praktische Erklärungen für die Nichteinhaltung eines erwarteten Ordnungsprinzips einfordern (vgl. Bergmann, 1988a, S. 48). Derartige Fälle können eine starke Evidenz für das postulierte Organisationsprinzip liefern, denn sie stellen empirische Beweise dafür dar, dass dieses Ordnungsprinzip für die TeilnehmerInnen handlungsleitend ist. Diese Validierungsmethode konversationsanalytischer Ergebnisse ist jedoch auf das Auftreten<sup>37</sup> entsprechender *abweichender Fälle* angewiesen. Bei der Verwendung der *next-turn proof procedure* macht sich die KA dagegen ein fundamentales Organisationsprinzip sprachlicher Interaktion zunutze, um die eigenen Analysen im Datenmaterial zu verankern. Eines der zentralen Erkenntnisse der KA ist es, aufzuzeigen, dass die sequenzielle Organisation von Interaktionen für die TeilnehmerInnen eine „architecture for intersubjectivity“ (Heritage, 1984a, S. 254; s. a. Schegloff, 1992) bietet: In Interaktionen zeigen die RezipientInnen in ihren Nachfolgeäußerungen ihr Verständnis der vorangegangenen Äußerungen an. Im Interaktionsverlauf produzieren die TeilnehmerInnen somit fortwährend füreinander Interpretationen des Verhaltens ihrer InteraktionspartnerInnen und etablieren so die intersubjektiven Bedeutungen ihrer Äußerungen und Handlungen. Da diese Interpretationen – in den beobachtbaren Handlungen der TeilnehmerInnen – interaktions-öffentlich verfügbar sind, können sie der KA als Evidenzquelle dienen (Sacks, 1995, S. 286–299; Schegloff, 1995, S. lxiii): „The display of those understandings in the talk of subsequent turns affords both a resource for the analysis of prior turns and a proof procedure for professional analyses of prior turns – resources intrinsic to the data themselves“ (Sacks et al., 1974, S. 729). Für die Analyse sprachlicher Interaktionen folgt daraus ein strikt sequenzanalytisches Vorgehen, das durch das Prinzip der lokalen Kohärenz (vgl. Sacks, 1987, S. 54–56) die Indexikalität sprachlicher

---

<sup>37</sup> ForscherInnen gehen teilweise soweit zu fordern, dass alle Kollektionen konversationsanalytischer Arbeiten abweichende Fälle enthalten sollten, um den handlungsleitenden Charakter eines postulierten Ordnungsprinzips empirisch validieren zu können (vgl. Robinson, 2007, S. 71). Ob sich abweichende Fälle identifizieren lassen, ist jedoch hochgradig von der Art des Phänomens, der Frequenz seines Auftretens und Robustheit der untersuchten Praktik abhängig (Sidnell, 2012, S. 99) und kann deshalb nicht gegenstandsunabhängig als Gütekriterium formuliert werden.

Äußerungen systematisch berücksichtigt: Die unmittelbar vorangehende Äußerung bildet den Bezugsrahmen für eine aktuelle Handlung und die TeilnehmerInnen orientieren sich beobachtbar an der wechselseitigen Erwartung, dass dieser sequenzielle Kontext zur Erschließung des Sinns von Äußerungen herangezogen werden kann<sup>38</sup>. Dies zeigt sich u. a. daran, dass spezifische Methoden wie Deplatzierungsmarkierungen (Schegloff & Sacks, 1973, S. 319) eingesetzt werden, um anzuzeigen, wenn eine nächste Handlung nicht auf die unmittelbar vorangegangene reagiert.

In der Arbeit mit Videoaufnahmen muss die Anwendung der *next-turn proof procedure* nun modifiziert werden, denn aus Perspektive der multimodalen Konversationsanalyse wird Sprache als eine – und nicht notwendigerweise relevanteste – Ressource verstanden, die von den TeilnehmerInnen zur methodischen Hervorbringung sinnhaft strukturierter Interaktionen eingesetzt wird (Mondada, 2018, S. 86). Daraus folgt für die Analyse zunächst, dass es auch multimodale Erwidern geben kann, die ausschließlich durch verkörpertes Verhalten realisiert werden (Deppermann, 2013, S. 4): Dies betrifft bspw. uns allen aus dem Alltag vertraute Phänomene wie ein Kopfnicken oder -schütteln als Antwort auf eine Frage. Konsequenterweise muss analytisch eine „next-action proof procedure“ (Mondada, 2018, S. 361) zur Anwendung kommen, die es der ForscherIn erlaubt, zu rekonstruieren, wie die TeilnehmerInnen ihre Erwidern durch den Einsatz sprachlicher, verkörperter oder materieller Ressourcen produzieren. Die sequenzanalytische Forschungslogik muss jedoch noch grundsätzlicher reflektiert werden, wenn systematisch die Multimodalität von Interaktion erforscht werden soll. Wie bereits erwähnt, ist der Einsatz mimischer, gestischer und anderer verkörperter Ausdrucksressourcen nicht an die lineare Sequenzialität sprachlicher Interaktion gebunden, sondern kann zeitgleich zu verbalen Äußerungen realisiert werden. Dies führt dazu, dass Handlungen bereits vor der Produktion einer sprachlichen Äußerung begonnen und Erwidern sehr früh, sogar während Produktion der vorangehenden Handlung initiiert werden können. In multimodalen Interaktionen können Handlungen als komplexe „multimodale Gestalten“ (Mondada, 2014b, S. 140) realisiert werden, in denen sich die Zeitlichkeiten unterschiedlicher Ressourcen auf komplexe Weise miteinander verschränken (Mondada, 2016a, S. 347)

---

<sup>38</sup> Dass der sequenzielle Kontext von den TeilnehmerInnen als zentrale Ressourcen für die lokale und prozessuale Hervorbringung sinnhafter Interaktion eingesetzt wird, wurde in unzähligen konversationsanalytischen Studien aufgezeigt. Die KA schließt nicht grundsätzlich aus, dass auch andere Kontexte relevant für das Sinnverständnis sprachlicher Äußerungen in Interaktion sein können, operiert jedoch mit der methodischen Restriktion, dass sich in Äußerungen der TeilnehmerInnen eine beobachtbare, handlungsrelevante Orientierung an einem Kontext zeigen muss, um diesen plausibel als Erklärung für das untersuchte Phänomen heranziehen zu können (Schegloff, 1997, S. 166–167) und dieser nicht von außen durch die AnalytikerIn an das Phänomen herangetragen werden darf.

Auch wenn multimodale Interaktionen also nicht auf die gleiche Weise wie sprachliche Interaktion linear organisiert, sondern durch den simultanen Einsatz verschiedener Ressourcen gekennzeichnet sind, bleibt Sequenzialität als zentrales Organisationsprinzip von Interaktion und methodologische Prämisse der KA bestehen:

In order to make sense of and identify the relevance of multimodal cues, co-participants as well as analysts must rely on the way they are assembled in an orderly way, and their composition as well as their position. Sequentiality remains the fundamental principle on which the organization of human action is based. (Mondada, 2016a, S. 361)

Die Hervorbringung sinnhafter Ordnung im zeitlichen und sequenziellen Verlauf multimodaler Interaktion lässt sich entsprechend – durch die gegenstandsangemessen angepassten – Verfahrensgrundsätze der KA untersuchen. Ziel ist es dabei, zu rekonstruieren, wie die TeilnehmerInnen die unterschiedlichen, ihnen zur Verfügung stehenden Ressourcen im Zusammenspiel miteinander einsetzen, um methodisch Aufgaben der Interaktionsorganisation zu bearbeiten und diejenigen Aktivitäten, in die sie involviert sind, auf beobachtbare und berichtbare Weise hervorzubringen.

### **3.2.4 Untersuchungsgegenstand – Geschichtenerzählungen aus Perspektive der multimodalen KA**

Mithilfe dieser Verfahrensgrundsätze haben konversationsanalytische Studien seit den 1970er Jahren einen stetig wachsenden Korpus an Erkenntnissen über die Organisationsprinzipien von Gesprächsphänomenen unterschiedlichster Größenordnung hervorgebracht. Gemeinsam tragen diese Arbeiten zur Rekonstruktion des methodischen Apparats bei (vgl. Sacks, 1984, S. 29), der von den TeilnehmerInnen routiniert eingesetzt wird, um in Interaktion sinnhafte soziale Ordnung hervorzubringen. Die multimodale KA berücksichtigt nun systematisch die grundsätzliche Multimodalität von F2F-Interaktionen (Stivers & Sidnell, 2005, S. 2) und erweitert die Ergebnisse konversationsanalytischer Untersuchungen sprachlicher Interaktion mithilfe des durch Videodaten ermöglichten Zugriffs: Auf diese Weise konnte gezeigt werden, dass verkörperte Handlungen und die körperlichen Arrangements der TeilnehmerInnen eine entscheidende Rolle für die Interaktionsorganisation – wie Praktiken der Turnkonstruktion, des Sprecherwechsels und der Sequenzorganisation (bspw. Mondada, 2013b) – spielen. Selbst eine übersichtsartige Darstellung der unterschiedlichen Forschungsbereiche der KA wäre aus Platzgründen in dieser Arbeit nicht zu leisten und auch analytisch wenig zielführend<sup>39</sup>. Ich möchte

---

<sup>39</sup> Systematische Darstellungen der wichtigsten Untersuchungsbereiche der KA finden sich bspw. bei Bergmann (1988c) oder in Sidnell und Stivers (2012).

stattdessen einige zentrale konversationsanalytische Erkenntnisse über das Geschichtenerzählen in Interaktion vorstellen, das als kommunikative Aktivität im Zentrum meiner Analysen steht. Beginnend mit Studien, die Geschichtenerzählungen als verbale Praktiken untersuchen, werde ich anschließend aufzeigen, welche Rolle der Einsatz körperlich-visueller Ressourcen für die Hervorbringung von Erzählungen in zeitlich-räumlicher Kopräsenz spielen. Auf diese Weise soll exemplarisch der Mehrwert einer multimodalen Perspektive illustriert, sowie die späteren Analysen vorbereitet werden.

Geschichtenerzählungen besitzen als rekonstruktive Gattung der Alltagskommunikation einen zentralen Stellenwert in unserem kommunikativen Haushalt, denn sie bieten eine fest etablierte Lösung für das kommunikative Problem der Flüchtigkeit sozialer Situationen (Luckmann & Bergmann, 1995, S. 293): In und durch Geschichtenerzählungen vermittelt die ErzählerIn den ZuhörerInnen kommunikativ ihre Erlebnisse (Quasthoff, 1980, S. 11) und der soziale Sinn dieser Erfahrungen wird gemeinsam verhandelt. Die Erzählung einer Geschichte wird konversationsanalytisch verstanden als eine „in Form einer Diskurseinheit realisierte verbale Rekonstruktion eines Ablaufs realer oder fiktiver Handlungen oder Ereignisse, die im Verhältnis zum Zeitpunkt des Erzählens zurückliegen oder zumindest ... als zurückliegend dargestellt werden“ (Gülich & Hausendorf, 2000, S. 373). In Interaktionen kann durch die Erzählung einer Geschichte dabei eine Vielzahl sozialer Handlungen realisiert werden – mit ihnen können Beschwerden zum Ausdruck gebracht (Drew, 1998; Günthner, 1999a), Verantwortung zugeschrieben (Mandelbaum, 1993), Witze erzählt (Sacks, 1974, 1978), oder Belege für vorangegangene Behauptungen erbracht werden (Mandelbaum, 2013, S. 493).

Konversationsanalytische Studien betonen die Prozesshaftigkeit interaktionalen Erzählens, das durch die Beiträge aller TeilnehmerInnen hervorgebracht wird. Geschichtenerzählungen werden durch „big packages“ (Sacks, 1995, Vol. II, S. 354) – d.h. durch mehrere sprachliche Äußerungen, mit denen ein vergangener Ereignisverlauf rekonstruiert wird – realisiert. Da alltägliche Gespräche von einer grundsätzlichen Konkurrenz um das Rederecht geprägt sind, müssen für Hervorbringung von Geschichtenerzählungen einige spezifische Aufgaben der Interaktionsorganisation bearbeitet werden: Die potenzielle ErzählerIn muss anderen Anwesenden zunächst die Verfügbarkeit einer Erzählung anzeigen und sich ein *ticket* (vgl. Sacks, 1971, S. 312) für die Produktion eines „multi-unit-turns“ (C. Goodwin & Heritage, 1990, S. 299) einholen.<sup>40</sup> Diese Aushandlung findet in spezifisch designten Prä-Sequenzen statt, die auch der

---

<sup>40</sup> Das System des Sprecherwechsels alltäglicher Gespräche ermöglicht grundsätzlich eine Übergabe des Rede-rechts am erkennbaren Ende jeder Turnkonstruktionseinheit („turn-constructive unit“, Sacks et al., 1974, S. 703). Dieser Mechanismus wird durch die Prä-Sequenz vorübergehend außer Kraft gesetzt, sodass die ErzählerIn die Möglichkeit erhält, die Geschichte zu entwickeln (Sacks, 1971).

Bearbeitung weiterer Aufgaben dienen: In ihnen wird das Verhältnis der Erzählung zum vorangegangenen Gespräch (bspw. als thematisch verbunden oder unverbunden) *accountable* gemacht (Jefferson, 1978, S. 220), die Erzählwürdigkeit des Ereignisses dargestellt, sowie die Haltung der Erzählerin zu den zu rekonstruierenden Ereignissen vorbereitet und den RezipientInnen somit ein bestimmtes Verständnis des Folgenden nahegelegt (Sacks, 1974, S. 341). Durch die Hervorbringung derartiger Prä-Sequenzen wird die Geschichtenerzählung eindeutig vom umgebenden *turn-by-turn-talk* abgegrenzt.

Findet in der Prä-Sequenz eine Alignierung – verstanden als formale Ausrichtung an der projizierten Aktivität (Kotthoff, 2017, S. 12; Mandelbaum, 2013, S. 500) – durch die RezipientInnen statt, indem sie die Erzählung ratifizieren und sich als ZuhörerInnen positionieren, wird in der Kernphase hauptsächlich Rekonstruktions- und Bewertungsarbeit geleistet (Kotthoff, 2017, S. 57) und die Erzählung adressatenzogen ausgestaltet (Quasthoff, 2001, S. 1301). Dabei verwendet die ErzählerIn<sup>41</sup> sprachliche Methoden, mit denen die chronologische Abfolge und kausalen Verknüpfungen der rekonstruierten Ereignisse dargestellt werden. Dies geschieht u.a. dadurch, dass sie Methoden der indirekten und direkten Redewiedergabe (im Folgenden „IRS“ für „indirect reported speech“ und „DRS“ für „direct reported speech“, vgl. Holt & Clift, 2007, S. 3) einsetzt, um vergangene sprachliche Handlungen zusammenzufassen oder wörtlich wiederzugeben. Durch die Verwendung unterschiedlicher Rede- und Sprechstile werden die verschiedenen Figuren der Erzählwelt animiert und für die ZuhörerInnen unterscheidbar inszeniert. Insbesondere durch den Einsatz prosodischer Ressourcen zeigt die ErzählerIn dabei implizit ihre Bewertung der wiedergegebenen Äußerungen an (Günthner, 1997).

Auch wenn die Verantwortung für die Ausgestaltung der Geschichte hauptsächlich bei der ErzählerIn liegt (Kotthoff, 2017, S. 21), spielt das Verhalten der ZuhörerInnen eine maßgebliche Rolle für ihren Verlauf: „[t]he ‚audience‘ is in fact the co-author [...] with recipient turns playing a crucial role in shaping and even constituting the ongoing course of the storytelling.“ (Mandelbaum, 2013, S. 501). Die RezipientInnen zeigen ihre andauernde Alignierung vor allem durch die Produktion von Rezeptionssignalen an, mit denen sie die SprecherIn anhalten, die Erzählung fortzuführen, ohne dabei selbst die Übernahme des Rederechts zu beanspruchen (C. Goodwin, 1986; Jefferson, 1984; Schegloff, 1982). Darüber hinaus können die RezipientInnen ihre (Dis)Affiliation – verstanden als inhaltliche (Un-)Gleichgerichtetheit in der Haltung zum Erzählten (Kotthoff, 2017, S. 12; Niemelä, 2011, S. 18) – darstellen, bspw. indem sie erzählbegleitende Bewertungen realisieren (C. Goodwin, 1986). Die Untersuchungen des

---

<sup>41</sup> Geschichten müssen nicht von einer einzelnen ErzählerIn produziert werden, sondern können auch kollaborativ durch mehrere TeilnehmerInnen erzählt werden (bspw. Lerner, 1992).

RezipientInnenverhaltens zeigen also zum einen, dass dieses eine zentrale Rolle für die Entwicklung der Geschichtenerzählung spielt. Zum anderen verdeutlichen sie, wie im Verlauf der Interaktion veränderbare Formen der Zuhörerschaft hervorgebracht werden (C. Goodwin, 1986), indem die ZuhörerInnen ihrer (Dis-)Alignierung mit der Erzählung und (Dis-)Affiliation mit der Erzählhaltung der SprecherIn anzeigen.

Auch die Beendigung von Geschichtenerzählungen stellt sich als interaktiv komplexes Geschehen dar: Die ErzählerInnen müssen *accountable* machen, dass die Erzählung ihr Ende erreicht hat, während ihre ZuhörerInnen wiederum anzeigen müssen, dass sie die Bedeutung der Geschichte verstanden haben, indem sie thematisch relevante Erwidern, bspw. Bewertungen, produzieren (Jefferson, 1978). An diesen Erwidern zeigt sich auch, ob die RezipientInnen eine mit der ErzählerIn inhaltlich gleichgerichtete Haltung zu den rekonstruierten Ereignissen einnehmen und diese als relevant für das vorangegangene Gespräch anerkennen. Eine etablierte Methode, mit der die RezipientInnen ihr Verständnis der Geschichte anzuzeigen, ist dabei die Initiierung einer zweiten, thematisch verbundenen Erzählung (Sacks, 1995, Vol. I, S. 771), sodass ganze Serien von Geschichten (Ryave, 1978) aufeinanderfolgen können. Die kollaborative Rückkehr zum *turn-by-turn-talk* erfolgt, nachdem es genügend Hinweise darauf gegeben hat, dass die ZuhörerInnen die Geschichten verstanden und gewürdigt haben (De Fina & Georgakopoulou, 2012, S. 89).

Der bis hierhin skizzierte Forschungsstand fokussiert ausschließlich verbale und vokale Verfahren des Erzählens. Bei Geschichtenerzählungen in F2F-Interaktion handelt es sich jedoch um eine multimodale Praktik, „zu deren Realisierung sich die Erzählenden verschiedener kinestischer, räumlicher oder materieller Ressourcen bedienen“ (König & Oloff, 2018b, S. 209). Mittlerweise existieren einige Studien<sup>42</sup>, die die Koordination multimodaler Ressourcen in der Hervorbringung von Geschichtenerzählungen in alltäglichen Interaktionen analysieren.

Der durch die Videoaufnahmen ermöglichte Zugriff auf das sichtbare Verhalten der TeilnehmerInnen ermöglicht zum einen eine differenziertere Betrachtung davon, wie RezipientInnen sowohl ihre (Dis-)Alignierung mit der im Vollzug befindlichen Geschichtenerzählung sowie ihre (Dis-)Affiliation mit der Haltung der ErzählerIn anzeigen. Durch ihre körperliche Positionierung zueinander und ihre Blickführung bringen die TeilnehmerInnen im Interaktionsvollzug die für die Erzählaktivität erforderlichen Teilnehmerahmen („participation framework“, C.

---

<sup>42</sup> Darüber hinaus finden sich auch Interaktionsanalysen, in denen narrative Aktivitäten als Datenmaterial verwendet werden, um multimodale Phänomene zu untersuchen. Diese Studien fokussieren zwar nicht spezifisch die Charakteristika der narrativen Sequenzen, untersuchen jedoch bspw. Gesten als verkörperte Verweise auf den und im Raum – bspw. als Orts- oder Richtungsangaben (Haviland, 2000), die multimodale Herstellung von Lächeln (Ford & Fox, 2010) oder Handlungsanzeigen durch turn-einleitendes Lächeln (Kaukomaa et al., 2013).

Goodwin & M. H. Goodwin, 1992, C. Goodwin, 2007a)<sup>43</sup> hervor. Durchgehend orientieren sich die TeilnehmerInnen dabei durch visuelles „Monitoring“ an verkörperten Anzeigen von Zuhörerschaft und Aufmerksamkeit (M. H. Goodwin, 1980, S. 309; Heath, 1982). Dabei konnten sowohl alignierende Verhaltensweisen, wie verkörperten Rezeptionssignale (bspw. Nicken), mit denen das Publikum die ErzählerIn zur Fortsetzung der Geschichte animiert, als auch non-verbales Störverhalten, wie bspw. die Eröffnung von Nebensequenzen durch die Blickaufnahme mit anderen (nicht-erzählenden) Anwesenden (M. H. Goodwin, 1997) identifiziert werden. Auch ihre (Dis-)Affiliation mit der ErzählerIn zeigen die RezipientInnen durch ihr verkörpertes Verhalten an: Erzählbegleitendes Kopfnicken oder Gesichtsausdrücke können an passenden sequenziellen Positionen als verkörperte Anzeigen einer zustimmenden oder ablehnenden Haltung („stance displays“) fungieren (Kupetz, 2014; Selting, 2010, 2012). Am Ende einer Erzählung wird Nicken jedoch als zu minimal und deshalb als disaffiliative Erwiderung behandelt, die die ErzählerIn oftmals zur Erweiterung der Geschichte veranlasst (Stivers, 2008).

Der Einsatz körperlich-visueller Ressourcen der ErzählerIn wurde vor allem für bestimmte Abschnitte narrativer Sequenzen untersucht: So konnte bspw. gezeigt werden, wie potenzielle ErzählerInnen körperlich-visuelle und sprachliche Ressourcen einsetzen, um selbstinitiierte Erzählungen in Kontexten zu platzieren, in denen das Rederecht umkämpft ist (Morek, 2018), oder wie Blicke genutzt werden, um in kollaborativen Erzählungen, das Rederecht zwischen zwei ErzählerInnen zu verteilen (Zima, 2017). Für die direkte Redewiedergabe in der Kernphase des Erzählens wurde gezeigt, wie die ErzählerInnen neben DRS und prosodischen Modulationen auch Blickwechsel und Gesten einsetzen, um verschiedene Figuren der Erzählwelt zu animieren (Ehmer, 2011; Koike, 2001) und die wiedergegebenen Äußerungen voneinander abzugrenzen (Stec et al., 2015, 2016). Verkörperte Ressourcen kommen jedoch nicht nur in Kombination mit sprachlichen Äußerungen zum Einsatz, es können auch ganze Elemente der Geschichte – bspw. Handlungen der Figuren der Erzählwelt – durch körperliche Darstellungen realisiert werden (Sidnell, 2006). Derartige multimodale Darstellungen erhöhen die

---

<sup>43</sup> Der Begriff des „participation framework“ wurde ursprünglich von Goffman (1979a, S. 11) entwickelt und bezeichnet dort das Verhältnis der SprecherInnen zu ihren Äußerungen. Das Konzept wurde von Charles Goodwin in seinen multimodalen Interaktionsanalysen übernommen und weiterentwickelt: „participation frameworks“ werden bei Goodwin als durch die wechselseitige verkörperte Ausrichtung der TeilnehmerInnen zueinander konstituiert verstanden. Sie bilden einen dynamischen Rahmen, indem der soziale Sinn der indexikalischen Äußerungen und verkörperten Handlungen verständlich wird (C. Goodwin, 2007a, S. 57): „Such frameworks encompass at least two different types of phenomena: 1) displays through which participants make visible their current engagement in the activity, and 2) structures that provide for the relevance of particular types of displays at specific moments in time.“ (C. Goodwin & Goodwin, 1992, S. 96). In deutschsprachigen Publikationen wird „participation framework“ sowohl als „Teilnehmer-“ (bspw. König & Oloff, 2018b, S. 218) als auch als „Teilnehmerahmen“ (bspw. Gülich & Mondada, 2008, S. 121) übersetzt. Ich verwende im Folgenden ausschließlich den Begriff des Teilnehmerahmens.

Aufmerksamkeit und Beteiligung der RezipientInnen (Schmitt & Deppermann, 2009), sodass sich insbesondere die Höhepunkte von Erzählungen – die im unmittelbaren Anschluss Erwidern der RezipientInnen relevant machen – durch eine verstärkte Kombination verbaler, vokaler und visueller Ressourcen auszeichnen (Niemelä, 2011, S. 38; Selting, 2017).

Auch wenn sich die Erforschung der Multimodalität des alltäglichen Erzählens noch in ihren Anfängen befindet, wird bereits anhand dieser skizzenhaften Darstellung deutlich, dass „Erzählen in zeitlich-räumlicher Ko-Präsenz . . . eine grundlegend körperlich und räumlich situierte Aktivität [ist, Anm. D. S.], die entsprechend in einem multimodal ausgerichteten Zugang erschlossen werden sollte.“ (König & Oloff, 2018b, S. 221). Erzählen stellt eine multimodale Praktik dar, die von allen Anwesenden kollaborativ im Interaktionsvollzug durch den Einsatz verbaler, prosodischer und verkörperter Ressourcen im räumlichen Setting ko-konstruiert wird. Mithilfe von Videoaufnahmen lässt sich dabei auch rekonstruieren, ob diese Erzählungen in Multiaktivitätssettings stattfinden (Haddington et al., 2014) und die Zeitlichkeit des Erzählens mit simultan stattfindenden Tätigkeiten verknüpft ist (König & Oloff, 2018a, S. 303). Ziel dieser Arbeit ist es, eben dieses komplexe Zusammenspiel einer Bandbreite multimodaler Ressourcen für die sinnhafte Hervorbringung *smartphone-assistierter Geschichtenerzählungen* zu analysieren. Da es sich bei diesem Untersuchungsgegenstand um einen spezifischen Fall des Gebrauchs mobiler Endgeräte in Interaktion handelt und die multimodale Konversationsanalyse in ihren Verfahrensgrundsätzen keine Konzeption für die empirische Analyse von Technologiegebrauch enthält, werde ich im letzten Teil dieses Kapitels die konversationsanalytische Adaption der *affordanztheoretischen* Perspektive vorstellen.

### **3.3 Technologiegebrauch aus konversationsanalytischer Perspektive**

Mit *smartphone-assistierten Geschichtenerzählungen* untersucht die vorliegende Arbeit nicht nur die grundsätzliche Multimodalität alltäglichen Erzählens, sondern auch den Gebrauch mobiler Kommunikationstechnologien im Kontext dieser spezifischen kommunikativen Aktivität. In der multimodalen KA ist jedoch keine theoretische Konzeption für die Analyse technologischer Artefakte verankert. Um ihren Gebrauch als Ressource für die lokale Produktion sinnhafter sozialer Ordnung in Interaktionen analytisch beschreibbar zu machen, hat sich in den letzten 20 Jahren in der KA eine Adaption der ursprünglich von James Gibson (1977) entwickelten *affordanztheoretischen* Perspektive durchgesetzt (bspw. Hutchby, 2001a, 2001b; Arminen et al., 2016). Ziel dieses Kapitels ist es, zunächst die Anfänge der konversationsanalytischen Untersuchungen technologischer Artefakte nachzuzeichnen, bevor anschließend der Begriff der *Affordanzen* eingeführt und in seiner konversationsanalytischen Adaption vorgestellt wird.

### 3.3.1 Die Ursprünge der konversationsanalytischen Untersuchung des Technologiegebrauchs

Betrachtet man die Anfänge der KA, erscheint es zunächst einigermaßen erstaunlich, dass eine theoretische Konzeption fehlt, mit der die Materialität technologischer Objekte in ihrer Relevanz für die Hervorbringung sozialer Ordnung für die empirische Forschung beschreibbar gemacht werden könnte. Überraschend ist dies insofern, als dass bereits seit der Gründung der KA mit Telefongesprächen (bspw. Schegloff, 1968, S. 1080 ff., 1979) zumindest eine spezifische Form medial-vermittelter Kommunikation untersucht wird. Schon in einer frühen Vorlesung von Sacks findet sich ein Hinweis darauf, dass eine systematische Berücksichtigung der spezifischen Handlungsmöglichkeiten von Kommunikationstechnologien in der Analyse von Interaktionen notwendig und gewinnbringend wäre:

. . . an examination in which the modes of interaction were considered and the telephone's distinctive features were located, . . . It may well then be that institutions could get examined [by members] for their unique possibilities, and when their unique possibilities are found, they're employed. We might then look to other developed institutions for which we have some handle on their history, to see the sorts of work done on them which have nothing particularly to do with 'what they're supposed to do' or anything like that, and which then get elaborated on to make them both formally analysed and comfortable institutions. (Sacks, 1995, Vol. II, S. 162)

Sacks fordert hier die empirische Untersuchung davon, wie sich InteraktionsteilnehmerInnen an den spezifischen (Handlungs-)Möglichkeiten des Telefons orientieren und sich diese für ihre praktischen Zwecke aneignen. Dabei klingt auch die Möglichkeit an, dass die Praktiken des Gebrauchs, die sich in der Aneignung entwickeln, nicht den intendierten Nutzungsweisen der EntwicklerInnen entsprechen (müssen). Obwohl damit das Potenzial der Berücksichtigung der einzigartigen Möglichkeiten von Kommunikationstechnologien früh erkannt wurde, entwickelte sich zunächst kein systematisches Interesse der KA an der Frage nach den Konsequenzen einer medialen Vermittlung von Kommunikation für die Interaktionsorganisation. Vielmehr wurden Telefongespräche als „tool for the analysis of talk-in-interaction“ (Schegloff, 2002, S. 293) angesehen, da die TeilnehmerInnen in ihnen sprachliche Interaktion ohne die Möglichkeiten des visuellen Feedbacks oder gestisch-mimischer Verhaltensweisen organisieren müssen. Dies ermöglichte es KonversationsanalytikerInnen zu untersuchen, wie die Hervorbringung sinnhafter sozialer Ordnungen in Unterhaltungen ausschließlich durch verbale und vokale Ressourcen der TeilnehmerInnen gelingen kann (vgl. Schegloff, 1979, S. 24).

Erst mit der zunehmenden Verbreitung von Informations- und Kommunikationstechnologien und der steigenden Relevanz der Frage nach den gesellschaftlichen Auswirkungen dieser

Entwicklung entstand ein explizites Interesse der KA an mediatisierten Interaktionen (vgl. Arminen et al., 2016, S. 294): Unter diesem Überbegriff werden Studien zu insbesondere drei Phänomenbereichen durchgeführt: Medial- bzw. computervermittelte Kommunikation (bspw. Telefongespräche, Online-Foren- oder WhatsApp-Kommunikation), Mensch-Maschine-Interaktion (bspw. Interaktion zwischen menschlichen NutzerInnen und technischen Geräten, persönlichen Assistenzsystemen oder Robotern) sowie der Gebrauch technologischer Objekte in zentrierten Interaktionen (bspw. die Nutzung von Mobiltelefonen in Alltagsgesprächen).

Eine wegweisende Studie, die maßgeblich zur Entwicklung dieses konversationsanalytischen Forschungsbereichs beitrug, war Lucy Suchmans *Plans and Situated Action* (1985): In ihrer Untersuchung des Gebrauchs von Kopiergeräten in institutionellen Kontexten erprobte sie empirisch, ob sich die konversationsanalytischen Techniken der Analyse menschlicher Interaktion auch auf die Interaktion zwischen Menschen und Maschinen übertragen ließen (vgl. Hutchby, 2001a, S. 8). Diese Vorgehensweise wurde damit begründet, dass die Kopiergeräte eine grafische Benutzeroberfläche besitzen, auf der die NutzerIn mithilfe von Bediengesten verschiedene Funktionen aktivieren kann. Die Interaktion zwischen Kopiergerät und NutzerIn besitzt in diesem Sinne einen dialogischen Charakter, der jedoch maßgeblich durch die technisch-materiellen Eigenschaften der Technologien geprägt ist. Suchman konnte zeigen, dass auch Mensch-Maschine-Interaktionen – so wie alle kommunikativen Handlungen – situiert, d. h. abhängig von spezifischen, lokalen Umständen sind. Mensch und Maschine unterscheiden sich jedoch maßgeblich darin, inwiefern sie diese Umstände in ihren Interaktionsbeiträgen berücksichtigen können:

The user of a machine deploys a full complement of observational, linguistic, and inferential skills in order to construe – and misconstrue – the machine's behavior according to the behavior's particular circumstances. At the same time, the machine is mapping a restricted set of effects, left by the user's actions, onto a prescribed template of possible meanings. As a consequence of the asymmetrical access of user and machine to the situation of action, the ordinary collaborative resources of human interaction are unavailable. (Suchman, 1985, S. 124)

Die in diesem Zitat anklingende Perspektive ist bis heute kennzeichnend für konversationsanalytische Studien, die Mensch-Maschine-Interaktionen untersuchen: Der Gebrauch von Technologien wird in und durch situierte Praktiken realisiert, für die auf Seite der NutzerInnen sprachliche, inferentielle und beobachtende Kompetenzen erforderlich sind. Gleichzeitig reagieren technologische Artefakte nur auf eine begrenzte Anzahl vordefinierter Effekte, die durch die Bedienung der NutzerIn ausgelöst werden. Die technologischen Artefakte ermöglichen – aufgrund ihrer materiellen Eigenschaften – also spezifische Handlungsweisen, während sie gleichzeitig andere einschränken oder unmöglich machen.

Dieses ethnomethodologisch-konversationsanalytische Verständnis des Verhältnisses von Technologiegebrauch und Interaktion lässt sich verkürzt als ‚dritte Position‘ zwischen Technik- und Sozialdeterminismus bezeichnen: Weder wird davon ausgegangen, dass Technologien aufgrund ihrer Eigenschaften die Muster menschlichen Verhaltens mechanistisch verändern (Technikdeterminismus), noch dass Technologien durch menschliche Aktivitäten in ihrer Bedeutung und Wirkung beliebig form- und veränderbar wären (Sozialdeterminismus). Hutchby (2001a) expliziert diese Perspektive folgendermaßen:

Technologies do not impose themselves on society, mechanistically altering the pattern of human relations and social structures. Neither does human agency encounter technologies as blank slates, as infinitely malleable forms. Technologies for communication possess materiality not only in the physical sense but in the sense of their very conditions of possibility. Technologies do not make humans; but humans make what they do of technologies in the interface between the organized practices of human conversation, and the technology's array of communicative affordances. (2001a, S. 206)

Es wird also davon ausgegangen, dass Technologien ihre sozialen Wirkungen (immer nur) in konkreten Situationen ihres Gebrauchs und durch die Wechselwirkungen zwischen den Organisationsprinzipien sozialer Interaktionen und ihren kommunikativen *Affordanzen* entfalten. Mit dem Begriff der *Affordanzen* – in der deutschsprachigen Literatur auch mit „Angebotscharakter“ (Zillien, 2008, S. 163) übersetzt – wird auf die Materialität von Technologien (und anderen Objekten) nicht lediglich im Sinne ihrer physikalischen Eigenschaften, sondern hinsichtlich der in ihnen materiell verfestigten Bedingungen ihrer (Handlungs-)Möglichkeiten verwiesen. Der Begriff geht auf den Wahrnehmungspsychologen James Gibson (1977) zurück und wurde von der KA auf spezifische Weise adaptiert. Um zu konkretisieren, in welcher Bedeutung er in der konversationsanalytischen Erforschung des Technologiegebrauchs verwendet wird, werde ich zunächst Gibsons ursprüngliche Begriffskonzeption vorstellen und anschließend die konversationsanalytischen Modifikationen dieser theoretischen Perspektive nachzeichnen.

### 3.3.2 Gibsons wahrnehmungspsychologische Affordanztheorie

Während sich der englische Begriff „to afford“ als „bieten“, „erlauben“ oder „ermöglichen“ übersetzen lässt<sup>44</sup>, handelt es sich bei dem Begriff der „affordances“ um eine Neuschöpfung Gibsons, deren Bedeutung von ihm folgendermaßen beschrieben wird:

---

<sup>44</sup> Bspw. <https://www.dict.cc/?s=afford>, zuletzt abgerufen am 20.11.2020.

Unter den Angeboten (affordances) der Umwelt soll das verstanden werden, was sie den Lebewesen anbietet (offers), was sie zur Verfügung stellt (provides) oder gewährt (furnishes), sei es zum Guten oder zum Bösen . . . Ich meine damit etwas, das sich auf die Umwelt und das Lebewesen gleichermaßen bezieht und zwar auf eine Art, die kein gebräuchliches englisches Wort auszudrücken vermag. Zum Ausdruck bringen soll es die Komplementarität von Lebewesen und Umwelt. (Gibson, 1982, S. 137)

Mit *Affordanzen* soll also der Angebotscharakter der Umwelt für Lebewesen bezeichnet werden. Das Konzept ist dabei grundsätzlich relational zu verstehen: *Affordanzen* sind gleichzeitig Eigenschaft des Objekts und des Subjekts. Sie lassen sich in gewisser Hinsicht als objektive Eigenschaften der materiellen Umwelt bezeichnen, da der Angebotscharakter eines Objekts sich nicht in Abhängigkeit „von den Bedürfnissen, der Aufmerksamkeit oder Einschätzung eines Beobachters“ (Zillien, 2008, S. 164) verändert. Gemeint ist damit, dass solange eine materielle Umwelt für ein bestimmtes Lebewesen die Möglichkeit für eine spezifische Aktivität bereitstellt, davon gesprochen werden kann, dass diese *Affordanz* existiert (vgl. Bærentsen & Trettvik, 2002, S. 52f.; Hutchby, 2001b, S. 448). Im Unterschied zu den physikalischen Eigenschaften eines Gegenstands (bspw. seiner Größe oder seines Gewichts) lassen sich *Affordanzen* jedoch nicht als unabhängige Größe messen, sondern sind immer nur in Relation zu einem wahrnehmenden Lebewesen zu verstehen (vgl. Gibson, 1986, S. 127). Der Begriff verweist somit immer gleichzeitig auf die Umwelt und auf ein wahrnehmendes Lebewesen: In dieser Hinsicht besitzt er also auch eine subjektive Seite, denn welche Möglichkeiten eine Umwelt einem Lebewesen bietet, erschließt sich diesem jeweils immer nur vor dem Hintergrund seiner spezifischen körperlichen Ausstattung und aktuellen Handlungsmöglichkeiten<sup>45</sup>. *Affordanzen* besitzen jedoch nicht nur eine funktionale Seite: In dem Maße wie sie für Lebewesen mit bestimmten Kapazitäten Handlungsmöglichkeiten bieten, schränken sie andere Handlungsweisen ein oder verunmöglichen sie vollständig (vgl. Hutchby, 2001b, S. 448). Während Gibson vor allem die *Affordanzen* der Umwelt und natürlicher Objekte fokussierte, existieren viele weitere Arten von Handlungsangeboten, bspw. von menschengemachten Artefakten.

Diese bereits ursprünglich angelegte „Wechselseitigkeit von Artefakt und Wahrnehmendem“ (Zillien, 2008, S. 165) stellt einen der wichtigsten Anknüpfungspunkte für die spätere

---

<sup>45</sup> Sharrock und Coulter (1998) wenden hier kritisch ein, dass für die Begriffsbestimmung diese Verwendung von ‚objektiv‘ und ‚subjektiv‘ nicht erforderlich wäre. Der gleiche Sinn würde vermittelt, wenn Gibson herausstellen würde, „that an affordance is a relational and dispositional property of an object, such that, for example, something has the property of 'affording eating' only relative to the needs and digestive proclivities of a particular kind of animal or person, and thence the terminology of 'objective' and 'subjective' can be dispensed with entirely“ (ebd., 1998, S. 150).

konversationsanalytische (und breitere mediensoziologische<sup>46</sup>) Adaption des *Affordanz*konzepts dar<sup>47</sup>: Auf diese Weise hebt Gibson hervor, dass Menschen ihre Umwelt nicht vornehmlich hinsichtlich ihrer physikalischen Eigenschaften, sondern der von ihnen angebotenen Handlungsmöglichkeiten wahrnehmen. Die Welt – und die in ihr auffindbaren Dinge, Objekte und Artefakte – werden für Menschen zu einem bedeutungsvollen Ort, da sie in Abhängigkeit davon wahrgenommen wird, welche Handlungsziele die AkteurInnen verfolgen und welche Handlungsmöglichkeiten sie aufgrund ihrer körperlichen Fähigkeiten wahrnehmen (vgl. Bærentsen & Trettvik, 2002, S. 52). Der Entwurf eines derartigen ontologischen Relationismus, der die unauflösbare Wechselwirkung zwischen Umwelt und Lebewesen und die Relevanz der Wahrnehmung von Handlungsangeboten betont, öffnet die Gibson'sche *Affordanz*theorie für verschiedene soziologische Adaptionen, in deren Zentrum oftmals die Frage nach dem Verhältnis von Technologien und Gesellschaft bzw. Sozialität steht. Als wahrnehmungspsychologische Theorie enthält sie jedoch auch zahlreiche Elemente, die für ihre konversationsanalytische Adaption einer Neurahmung bedürfen. Diese mit den theoretischen Grundannahmen der KA unvereinbaren Dimensionen werden im Folgenden identifiziert und aufgezeigt, welche Modifikationen in konversationsanalytische Adaption erfolgen.

### 3.3.3 Die konversationsanalytische Adaption des *Affordanz*konzepts

Für eine multimodale KA, die verstehen will, in welchem Verhältnis der Angebotscharakter technologischer Artefakte und die Organisationsprinzipien sinnhaft strukturierter Interaktionen zueinanderstehen, ermöglicht die Adaption des *Affordanz*konzepts die Entwicklung eines theoretischen Bezugsrahmens, der für die empirische Analyse der Nutzung und Wirkung neuer Informations- und Kommunikationstechnologien geeignet ist. Anders als frühere Mediumstheorien (bspw. Innis, 2008; McLuhan, 1968) verliert dieser Bezugsrahmen dabei nicht die materiellen Grundlagen aus dem Blick, die den Möglichkeiten für die Realisierung praktischer Handlungen in alltäglichen Aktivitäten zugrunde liegen (vgl. Hutchby, 2001b, S. 450; Zillien, 2008, S. 163).

---

<sup>46</sup> S. dazu ausführlich Zillien (2008).

<sup>47</sup> Ausgehend von der Annahme dieses „ontologischen Relationismus“ – einer grundlegenden Wechselwirkung zwischen Umwelt und Lebewesen – bietet Gibsons *Affordanz*theorie somit einen systematischen Entwurf für die Überwindung wissenschaftshistorisch gewachsener Dualismen von Subjekt und Objekt (Gibson, 1982, S. 43) sowie Natur und Kultur (Gibson, 1982, S. 140; s. a., Costall, 1995, S. 477). Auch wenn Gibson die Frage nach der Wahrnehmung von Handlungsmöglichkeiten der materiellen Umwelt fokussiert, sind theoretische Überlegungen über die Konsequenzen der Handlungsrealisierung in seinen Schriften angelegt (bspw. Gibson, 1982, S. 43). Dies macht seinen Theorieentwurf für unterschiedliche konstruktivistische Ansätze anschlussfähig, die ebenfalls mit diesen hergebrachten Dualismen brechen.

Zunächst besteht für die Verwendung des *Affordanz*begriffs in der KA eine grundsätzliche Problematik darin, dass sich in unterschiedlichen Forschungsrichtungen uneinheitliche Verwendungen des Begriffs etabliert haben. Neben der von Gibson (1977, 1982) entwickelten *affordanz*theoretischen Wahrnehmungspsychologie hat der Begriff vor allem im Bereich der Usability-Forschung weite Verbreitung erfahren. Das Konzept wurde hier von Norman (1988) eingeführt und von anderen AutorInnen weiterentwickelt (bspw. Gaver, 1991). In diesem Forschungsbereich wird der Begriff der *Affordanzen* mit der Bedeutung als „Gebrauchseigenschaft“ (Norman, 1989, S. 21) verwendet. Er dient als theoretisches Konzept in Usability-Studien, die untersuchen, inwiefern die von den Entwicklern intendierten Gebrauchseigenschaften von Technologien aufgrund des Objektdesigns für die NutzerInnen visuell wahrnehmbar sind. Diese Definition von *Affordanzen* als „near magical properties making things easy for the ‚user‘ of artifacts“ (Bærentsen & Trettvik, 2002, S. 52) fand weite Verbreitung.<sup>48</sup>

Darüber hinaus werden *affordanz*theoretische Adaptionen auch in mediensoziologischen Ansätzen (vgl. bspw. Orlikowski, 2000; Vyas et al., 2006; Zillien, 2008) diskutiert. Die mediensoziologische Kritik an der ursprünglich wahrnehmungspsychologischen Konzeption des *Affordanz*begriffs überschneidet sich in vielerlei Hinsicht mit der der KA. Die mediensoziologischen Adaptionen unterscheiden sich jedoch insofern, als dass sie konstruktivistische Positionen vertreten, die zwar die historische und gesellschaftliche Gewachsenheit ‚sozialer Tatsachen‘ anerkennt (vgl. bspw. Zillien, 2008, S. 174–177), im Gegensatz zur KA jedoch die lokalen (Re-)Produktionsprozesse von sozialer Wirklichkeit in Interaktion nicht hinreichend berücksichtigen.

Die konversationsanalytische Adaption zeichnet sich insbesondere dadurch aus, dass die folgenden Dimensionen des ursprünglichen *Affordanz*konzepts kritisch reflektiert und vor dem Hintergrund der eingangs beschriebenen theoretischen und methodologischen Grundannahmen der KA modifiziert werden:

1. In ihrer ursprünglichen Konzeption fokussiert die Gibson'sche *Affordanz*theorie die Wahrnehmung der Handlungsmöglichkeiten von Objekten durch (menschliche) Lebewesen. Aus Perspektive der explizit anti-kognitivistischen KA kann die Frage nach den

---

<sup>48</sup> Wie einflussreich diese Begriffsverwendung der Usability-Forschung ist, lässt sich beispielhaft an der kritischen Diskussion der *Affordanz*theorie des Soziologen Rich Ling (2004) illustrieren. Ausgehend von der anekdotischen Beobachtung einer familiären Situation der Computernutzung, in der Lings Tochter eine Instruktion für die Bewegung des Mauszeigers auf dem Bildschirm logisch fehlinterpretiert (und die Maus vom Mauspad hochhebt), kommt Ling zu dem Fazit, dass *affordanz*theoretisch fälschlicherweise angenommen würde, dass Objekte ihre möglichen Funktionen für jeden ausstrahlen (vgl. 2004, S. 25). Dieses Verständnis der *affordanz*theoretischen Perspektive entspricht jedoch eher der Norman'schen Adaption als „Gebrauchseigenschaft“ als der Originalkonzeption von Gibson, in der die Wahrnehmung von *Affordanzen* untrennbar mit den Handlungsmöglichkeiten und -kapazitäten der AkteurIn verknüpft ist.

Organisationsprinzipien des sinnhaftem Technologiegebrauchs (Hutchby, 2001a, S. 206) jedoch nicht über individuelle Wahrnehmungsprozesse, sondern nur über eine Analyse der beobachtbaren Praktiken des tatsächlichen Technologiegebrauchs erfolgen. Damit verbunden ist die Grundannahme, dass technologische Objekte ihre soziale Wirkung nur dadurch entfalten, dass „[i]hre *interne* Logik, die Konfiguration ihrer spezifischen Hardware und Software, . . . in allen Kontexten des Mediengebrauchs auf eine *soziale* Logik, das heißt auf unterschiedliche Konstellationen der Ausnutzung der von den jeweiligen Medien eröffneten Möglichkeiten“ (Keppler, 2013, S. 85, Herv. i. O.) trifft. Konsequenterweise geht die KA empirisch der Frage nach, ob und wie sich die TeilnehmerInnen in ihrem Verhalten beobachtbar an dem Angebotscharakter technologischer Objekte orientieren (Arminen et al., 2016, S. 302) und inwiefern diese die Realisierung sozialer Handlungen im Kontext von Interaktionen ermöglichen, begünstigen, erschweren oder unmöglich machen.

2. Eine weitere Differenz besteht zum Verständnis sozialer Wirklichkeit, das in der ursprünglichen *Affordanz*theorie angelegt ist: In Gibsons Theorie wird von einer bereits gegebenen Welt ausgegangen (vgl. Shotter, 1983, S. 20), die von Subjekten hinsichtlich ihres Angebotscharakters wahrgenommen wird. Ausgehend von der Annahme über den Vollzugscharakter sozialer Wirklichkeit können die Handlungsangebote von Objekten aus konversationsanalytischer Perspektive jedoch nicht als invariant existierende Eigenschaften der materiellen Umwelt verstanden werden. *Affordanzen* werden vielmehr als in und durch die Handlungen der TeilnehmerInnen hervorgebracht aufgefasst (Hutchby, 2001b, S. 448): Die *Affordanzen* eines technologischen Artefakts manifestieren sich also erst dadurch, dass sich TeilnehmerInnen in ihren Handlungen beobachtbar an ihnen orientieren.

3. Als weiterer Kritikpunkt kann angeführt werden, dass in der ursprünglichen Konzeption der *Affordanz*theorie die Relevanz des sozialen und kulturellen Kontexts konkreter Situationen des Objektgebrauchs nicht hinreichend berücksichtigt wird (Zillien, 2008, S. 170). Grundsätzlich geht Gibson davon aus, dass Lebewesen den Angebotscharakter von Objekten unmittelbar wahrnehmen (Gibson, 1986, S. 174). Auch wenn von ihm anerkannt wird, dass die Wahrnehmung von Handlungsangeboten erlernt ist (vgl. Guski, 1996, S. 48) und durch Erfahrung geschult wird (vgl. Gibson, 1986, S. 235f.), wird nicht systematisch berücksichtigt, dass menschliche AkteurInnen ihre Umwelt ausgehend von ihrer Sozialisierung, (Aus-)Bildung und praktischen Erfahrung wahrnehmen. Ausgehend von den theoretischen Grundannahmen der EM, wird aus konversationsanalytischer Perspektive jedoch davon ausgegangen<sup>49</sup>, dass *Affordanzen*

---

<sup>49</sup> Diese Kritik ist nicht nur theoretisch begründet, denn wie in Kap. 2.2.1. gezeigt wurde, können die Ergebnisse mediensoziologischer Untersuchungen zur Nutzung spezifischer Handlungsangebote von Mobiltelefonen als

von AkteurInnen nicht ‚direkt‘ wahrgenommen werden können, sondern erkannt werden, weil sie sie bereits in der Vergangenheit genutzt haben oder sie uns beigebracht wurden (vgl. Sharrock & Coulter, 1998, S. 155). Darüber hinaus ist das Erkennen von *Affordanzen* mit konventionellen Regeln des Objektgebrauchs verbunden: Neben den materiell oder technisch möglichen Handlungsweisen wird der Angebotscharakter von – insbesondere menschengemachten – Objekten durch soziale Konventionen ihrer Nutzung geprägt (Hutchby, 2001b, S. 448–449). Damit wird die ursprüngliche *Affordanz*konzeption um eine moralische Dimension ergänzt: Die *Affordanz* eines Objekts mag eine Handlung prinzipiell ermöglichen; wenn diese in einem interaktionalen Setting von den TeilnehmerInnen durch ihr beobachtbares Verhalten als relevante nächste Handlung projiziert wird, können Anwesende moralisch dafür verantwortlich (gemacht) werden, diese *Affordanz* zu nutzen (vgl. Arminen et al., 2016, S. 295).

Ausgehend von dieser kritischen Reflektion der Problemdimensionen der ursprünglichen Konzeption von *Affordanzen* und der Darstellung ihrer Anpassungen an das ethnomethodologisch-konversationsanalytische Forschungsprogramm ergibt sich eine spezifische Perspektive auf Technologiegebrauch-in-Interaktion, die insbesondere von Ian Hutchby (2001a, 2001b, 2003) ausgearbeitet wurde. In seiner konversationsanalytischen Adaption verweist der Begriff der *Affordanzen* auf die spezifische Materialität von (technologischen) Objekten und bezeichnet die in ihnen materiell verfestigten Bedingungen ihrer (Handlungs-)Möglichkeiten<sup>50</sup>. *Affordanzen* stellen grundsätzlich relationale Eigenschaften dieser Objekte dar, sie werden jedoch nicht als invariante, sondern als emergente Eigenschaften verstanden, die ausgehend vom ethnomethodologischen Verständnis von sozialer Wirklichkeit als *Vollzugswirklichkeit* (Bergmann, 1981, S. 12) immer erst im tatsächlichen Aufeinandertreffen von Objekten und TeilnehmerInnen und durch ihre praktischen Handlungen lokal hervorgebracht werden (Hutchby, 2001a, S. 129, 2001b, S. 448). Die beobachtbare Orientierung an den *Affordanzen* technologischer Objekte findet immer im Kontext konkreter Interaktionssituationen statt, in denen die TeilnehmerInnen vor dem Hintergrund konventioneller Nutzungsweisen der Objekte ihre praktischen Zwecke

---

empirische Belege für die Relevanz sozialer und kultureller Kontexte angeführt werden. Diese Studien fördern das ursprünglich als kulturunabhängig konzeptualisierte *Affordanz*verständnis empirisch heraus.

<sup>50</sup> An dieser Stelle sei darauf hingewiesen, dass sich konversationsanalytische Adaptionen des *Affordanz*konzepts entwickelt haben, die den Begriff auch auf die erkennbaren strukturellen Muster sprachlicher Interaktion beziehen: So werden unmittelbar erkennbare Projektionen oder wahrnehmbare Konsequenzen der Realisierung einer bestimmten sprachlichen Äußerungen ebenfalls als *Affordanz* dieser Äußerung betrachtet (vgl. Forrester, 1999, S. 88). Ich werde den *Affordanz*begriff in dieser Arbeit ausschließlich dazu verwenden, um den emergenten Angebotscharakter materieller und technologischer Objekte zu bezeichnen. Dies begründet sich zum einen damit, dass die KA bspw. mit dem Begriff der „sequential implicativeness“ (Schegloff & Sacks, 1973, S. 296) bereits über ein gut entwickeltes technisches Vokabular für die Beschreibung der „*Affordanzen*“ sprachlicher und verkörperter Handlungen-in-Interaktion verfügt. Zum anderen ist es eben gerade nicht die Materialität sprachlicher Äußerungen, sondern die Eigenlogik von Interaktionen und die interaktiven Aushandlungsprozesse der TeilnehmerInnen, durch die der Angebotscharakter von Äußerungen hervorgebracht wird.

verfolgen. Der Angebotscharakter von Technologien wird somit nicht als ‚Kontext-für-Handlungen‘ verstanden, der sich aufgrund der invarianten Eigenschaften der Objekte bestimmen und seine Effekte vorhersagen ließe. Stattdessen wird aus Perspektive der KA empirisch untersucht, welche Konsequenzen sich aus der beobachtbaren Orientierung der TeilnehmerInnen an den Bedingungen der (Handlungs-)Möglichkeiten des Objekts für den Verlauf und die sinnhafte Organisation von Interaktion ergeben (vgl. Arminen et al., 2016, S. 305–306).

Die multimodale KA stellt in dieser Arbeit also die grundsätzlichen methodologischen und analytischen Verfahrensweisen bereit, um empirisch zu untersuchen, wie die TeilnehmerInnen an alltäglichen F2F-Interaktionen methodisch unterschiedliche – sprachliche, verkörperte, räumliche und materielle – Ressourcen für die Hervorbringung *smartphone-assistierter Geschichtenerzählungen* einsetzen. Mithilfe der Adaption der hier skizzierten *affordanztheoretischen* Perspektive lässt sich rekonstruieren, wie sich die TeilnehmerInnen dabei beobachtbar an den materiell verfestigten Bedingungen der (Handlungs-)Möglichkeiten von Smartphones orientieren. Das *Affordanz*konzept lenkt den analytischen Blick darauf, die Folgen der Mediatisierung alltäglicher Unterhaltungen – hier verstanden als Smartphonegebrauch in F2F-Interaktionen – als Konsequenzen für den sequenziellen Verlauf und die zeitliche Entwicklung der Interaktion zu untersuchen. Im Fokus wird dabei stehen, zu rekonstruieren, welche Handlungspotenziale der mobilen Endgeräte in der kollaborativen Hervorbringung *smartphone-assistierter Geschichtenerzählungen* als besonders auffällig, begünstigt oder eingeschränkt erscheinen (vgl. Arminen et al., 2016, S. 306). Es muss also expliziert werden, welche durch Smartphones bereitgestellten Handlungsmöglichkeiten eine konstitutive Rolle für *smartphone-assistierte Erzählungen* übernehmen, welche methodischen Verfahrensweisen erst durch die *Affordanzen* der Geräte ermöglicht werden, aber auch, welche Handlungsweisen durch die spezifische Materialität der mobilen Endgeräte erschwert werden. Dies erfolgt basierend auf der detaillierten, empirischen Analyse von Videoaufzeichnungen konkreter Situationen des Smartphonegebrauchs – die dazu eingesetzten Techniken der Datenerhebung, Kollektionsbildung und das analytische Vorgehen werden im folgenden Kapitel vorgestellt.

## 4 Methodisches Vorgehen

### 4.1 Datenerhebung: Aufzeichnungen des ‚natürlichen‘ Smartphonegebrauchs in alltäglichen Gesprächen

Im vorangegangenen Kapitel wurden die theoretischen Grundlagen, methodologischen Prämissen und analytischen Maxime der multimodalen KA vorgestellt. Im Folgenden werde ich nun darlegen, wie ich diese Verfahrensgrundsätze im konkreten methodischen Vorgehen umgesetzt habe. Das Ziel der Datenerhebung war es, alltägliche Unterhaltungen aufzuzeichnen, um empirisch Formen des spontan auftretenden Smartphonegebrauchs in F2F-Interaktionen untersuchen zu können. Die Datenerhebung gliederte sich dabei in zwei Phasen: von 2016 bis 2017 wurden Audio- und von 2018 bis 2019 Videoaufnahmen angefertigt. In beiden Phasen war die Datenerhebung von der konversationsanalytischen Anforderung geprägt, möglichst natürliche Interaktionen aufzuzeichnen, für die ein echter Sprechanlass der TeilnehmerInnen vorlag und die auch ohne das Zutun der ForscherIn stattgefunden hätten (Bergmann, 2000, S. 531).

In der ersten Phase der Datenerhebung wurden zwanglose Unterhaltungen zwischen miteinander befreundeten oder verwandten TeilnehmerInnen aufgezeichnet. Die meisten der Gespräche fanden in privaten Räumlichkeiten statt, nur vereinzelt wurden Aufnahmen in teilöffentlichen Räumen (bspw. Cafés) gemacht. In dieser Phase entstanden insgesamt 57 Aufnahmen mit einem Gesamtumfang von ca. 26 Stunden<sup>51</sup>. Alle Aufnahmen wurden mithilfe der Aufzeichnungsfunktion eines Smartphones angefertigt. Dies geschah unter der Annahme, dass sichtbar platzierte mobile Endgeräte heutzutage oftmals ‚natürlicher Bestandteil‘ alltäglicher Gespräche sind und ihre Verwendung als Aufzeichnungsgeräte möglichst wenig Aufmerksamkeit der Beforschten auf sich zieht. Um den TeilnehmerInnen maximale Verhaltensspielräume zu ermöglichen (Schu, 2001, S. 1015) und die ‚natürlicherweise‘ auftretenden Formen des Smartphonegebrauchs beobachten zu können, wurde den Beforschten vor Beginn der Aufzeichnung lediglich mitgeteilt, dass es sich um ein medien- und kommunikationswissenschaftliches Forschungsprojekt zu „Alltagskommunikation heute“ handelt. Das eigentliche Forschungsinteresse wurde erst nach Beendigung der Aufzeichnung in einer Nachbesprechung offengelegt: In dieser erhielten die TeilnehmerInnen detaillierte Informationen über das Forschungsvorhaben und den Umgang mit den Daten, sodass am Ende schriftliche Einverständniserklärungen

---

<sup>51</sup> Darüber hinaus erhielt ich dankenswerter Weise Zugriff zum Datenkorpus des 2012–2014 an der Universität Mannheim durchgeführten DFG-Projekts: „Alltagskommunikation unter mediatisierten Bedingungen“. Dadurch standen mir noch einmal Audioaufnahmen im Gesamtumfang von ca. 42 Stunden zur Verfügung.

ausgefüllt werden konnten, mit denen die Beforschten ihre Erlaubnis zur Verwendung der Aufnahmen in dieser Arbeit gaben.

Da abgesehen von der gezielten Aufzeichnung alltäglicher Unterhaltungen keinerlei spezifische Elizitationsformen (Schu, 2001, S. 1019) eingesetzt wurden, um das Auftreten des interessierenden Phänomens zu beeinflussen, enthielten etliche Aufnahmen keine oder kaum Fälle des Smartphonegebrauchs in Interaktion. Entsprechend musste ein umfangreicher Datenkorpus generiert werden, um wiederkehrende Arten des Gebrauchs mobiler Endgeräte in den Daten identifizieren und analysieren zu können. In Tonaufzeichnungen lassen sich dabei nur diejenigen Interaktionsepisoden identifizieren, in denen sich die TeilnehmerInnen sprachlich explizit auf den Gebrauch der mobilen Endgeräte beziehen. Durch die Analyse dieser Passagen konnten erste Erkenntnisse über die sprachlichen Praktiken des Umgangs mit Smartphones in alltäglichen F2F-Interaktionen generiert werden. Gleichzeitig ergaben sich in der Arbeit mit den Audioaufzeichnungen jedoch auch zahlreiche Fragen nach der Relevanz verkörperter Praktiken des Smartphonegebrauchs, die sich mithilfe dieses Datentyps nicht beantworten lassen.

Um der konstitutiven Multimodalität und Materialität der Praktiken des Gebrauchs mobiler Endgeräte gegenstandsadäquat zu begegnen, wurden in der zweiten Erhebungsphase von 2018 bis 2019 Videoaufnahmen angefertigt. Dabei wurden sowohl TeilnehmerInnen in Cafés rekrutiert, als auch Möglichkeiten zur Aufzeichnung in privaten Wohnungen wahrgenommen. Insgesamt konnten so noch einmal 25 Aufnahmen mit einem Gesamtumfang von ca. 19 Stunden generiert werden. Diese Videodaten stellen den Kernkorpus dieser Arbeit dar – die Analysen der Audiodaten ergänzen die spätere Ergebnisdarstellung lediglich und geben vor allem Aufschluss über die verwendeten sprachlichen Praktiken während des Smartphonegebrauchs in F2F-Interaktionen.

In Vorbereitung für die Videoaufnahmen in Cafés wurden zunächst die BetreiberInnen kontaktiert, um eine offizielle Dreherlaubnis einzuholen. Anschließend folgte eine Phase der Bekanntmachung mit den räumlichen Begebenheiten vor Ort, um die Aufnahmen vorzubereiten (Heath et al., 2010, S. 43). Auch für die Aufzeichnungen im privaten Raum wurden vergleichbare Vorbereitungen getroffen, u. a. um einen geeigneten Platz für die Kamera auszuwählen und eine Aufnahmeleitung im Umgang mit der Kamera zu instruieren. Auch in dieser Phase wurde versucht, die alltäglichen Unterhaltungen vollständig – d. h. von ihrer Eröffnung bis zu ihrer Beendigung durch die TeilnehmerInnen – aufzuzeichnen (Mondada, 2006, S. 55). Dies gestaltete sich jedoch insbesondere bei den Aufnahmen in gastronomischen Betrieben als schwierig, denn hier wurden Gäste spontan angesprochen und als TeilnehmerInnen rekrutiert. Diese Strategie führte unweigerlich dazu, dass ich Gruppen von FreundInnen ansprach, die gemeinsam an der

Lokalität angekommen waren und deren Gespräche bereits begonnenen hatten, sodass die Aufzeichnung der Gesprächseröffnung nicht möglich war. In den Aufnahmen sind außerdem nicht alle Gesprächsbeendigungen festgehalten, denn es wurde vorher mit den TeilnehmerInnen vereinbart, dass sie die Aufnahme jederzeit abbrechen können. Einige wenige TeilnehmerInnen machten von dieser Option Gebrauch, insbesondere mit der Begründung, sich über intimere Themen unterhalten zu wollen.

Alle Aufnahmen wurden mit einer einzelnen, auf einem Stativ installierten Kamera aufgezeichnet. Es wurde aus einer fixierten Kameraperspektive gefilmt, so dass alle InteraktionsteilnehmerInnen im Bildausschnitt zu sehen sind (Heath et al., 2010, S. 40). Ziel war es, einen Kamelwinkel zu finden, aus der der gesamte Interaktionsraum und damit alle für die Interaktionsorganisation relevanten sprachlichen und verkörperten Beiträge der TeilnehmerInnen aufgezeichnet werden konnten (Mondada, 2006, S. 55). In einigen Räumlichkeiten mussten dabei jedoch praktische Kompromisse eingegangen werden: In privaten Wohnungen war der Platz teilweise so beengt, dass lediglich die Oberkörper der TeilnehmerInnen fokussiert werden konnten; in den Cafés musste darauf geachtet werden, dass möglichst wenig andere Gäste im Bild waren und das Kamerastativ das Personal nicht bei der Arbeit behinderte. Sobald die Aufzeichnungen gestartet wurden, verließ ich die Kamera, um möglichst wenig zusätzliche Aufmerksamkeit auf die Aufnahmesituation zu ziehen (Laurier & Philo, 2006, S. 186). Bei Aufnahmen in privaten Settings wurden die TeilnehmerInnen vorher instruiert, wie sie die Aufzeichnung selbstständig beenden können; bei Aufnahmen in Cafés nahm ich einige Tische entfernt und außerhalb der Hörweite Platz. Mit den TeilnehmerInnen wurde vereinbart, dass sie mich jederzeit per Handzeichen rufen können, um die Aufnahme zu beenden.

Mit Aufnahmen, die mit einer statisch positionierten Kamera angefertigt werden, geht die Einschränkung einher, nur in den seltensten Fällen das Display der mobilen Endgeräte aufzeichnen zu können. Diese Schwierigkeit resultiert zum einen aus der „micro-mobility“ (Luff & Heath, 1998) der Kommunikationsgeräte – sie werden aus Handtaschen geholt, auf Tischen abgelegt, alleine und gemeinsam genutzt oder an andere TeilnehmerInnen weitergegeben –, sodass je nach Verwendung jeweils ein anderer Kamerawinkel für die Aufzeichnung des Displays vonnöten wäre (vgl. Relieu, 2009, S. 217). Zum anderen sind die Displays oftmals verspiegelt, sodass die angezeigten Medieninhalte in der Aufnahme nicht erkennbar sind, selbst wenn sie zur Kamera gedreht sind (Brown et al., 2013, S. 1033). Diese Einschränkungen in der Beobachtbarkeit des Displays wurden in einem forschungspraktischen Kompromiss in Kauf genommen: Die Möglichkeit, möglichst natürliche Gespräche und den spontan stattfindenden Smartphonegebrauch filmen zu können, ohne die TeilnehmerInnen im Vorhinein über dieses

Forschungsinteresse aufzuklären, wurde gegenüber aufwendigeren – und damit reichhaltigeren – Aufzeichnungsformen priorisiert.<sup>52</sup>

Nach Beendigung der Aufzeichnung erhielten die TeilnehmerInnen Zeit, die schriftliche Einverständniserklärung zu lesen und wurden anhand von beispielhaften Transkripten darüber informiert, in welcher Form die Daten in schriftlichen Publikationen repräsentiert werden. Sie konnten in einer mehrstufigen Einverständniserklärung entscheiden, ob und welchen Publikationen die Videoaufnahmen gezeigt werden dürfen. Allen TeilnehmerInnen wurde bei Interesse die Videoaufnahme digital übermittelt. In den ausführlichen Nachbesprechungen wurde darüber hinaus mit den Beforschten über das tatsächliche Erkenntnisinteresse und ihre Wahrnehmung der Aufnahmesituation gesprochen. Daraus ergaben sich zum einen oftmals angeregte Gespräche über Haltung der Beforschten zur zunehmenden Allgegenwart mobiler Kommunikationsgeräte. Zum anderen führte die nachträgliche Offenlegung des Forschungsinteresses oftmals dazu, dass die TeilnehmerInnen ihre eigene Smartphonennutzung während der Aufzeichnung reflektierten: Eindrücklich zeigte sich dabei die Schwierigkeit, diese aus der Erinnerung zu rekonstruieren (vgl. Vanden Abeele et al., 2019, S. 35–36). Einige der TeilnehmerInnen berichteten mir ‚entschuldigend‘, dass sie ihre Handys während der Aufnahme gar nicht genutzt hätten – die anschließende Betrachtung des Videos zeigte jedoch oftmals, dass die mobilen Endgeräte sowohl individuell als auch gemeinsam verwendet wurden. Diese Beobachtung unterstreicht noch einmal (anekdotisch) die Bedeutung technischer Aufzeichnungen für eine empirische Untersuchung der Nutzung mobiler Endgeräte in alltäglichen Gesprächssituationen, die als routinierte Tätigkeit nicht ohne Weiteres in Selbstauskünften rekonstruiert werden kann.

## **4.2 Reflektion der ‚Natürlichkeit‘ der aufgezeichneten Interaktionen**

Durch die Videoaufzeichnungen wurden registrierende Konservierungen der alltäglichen Interaktionen der TeilnehmerInnen angefertigt und diese somit in stabile, dem Anspruch nach ‚naturalistische‘ Daten für die anschließenden Analysen transformiert. Alle Aufzeichnungen wurden zeitnah nach ihrer Anfertigung einer initialen Betrachtung unterzogen. Dies diente zum einen dazu, sich eine Übersicht über den Gesprächsverlauf und die Aktivitäten der

---

<sup>52</sup> Es existieren Studien (Brown et al., 2013, 2015, 2018), die mit hohem technischen Aufwand – d. h. durch die Aufnahme mit mehreren Kameras oder die Installation von Screen-Capture-Software – auch die Displayaktivitäten aufzeichnen. Auf diese Weise wird die Reichhaltigkeit der Daten erhöht, es lassen sich jedoch nur schwer Interaktionen aufzeichnen, die so auch ohne das Zutun der ForscherIn stattgefunden hätten. Insbesondere wenn Screen-Capture-Software auf den Endgeräten der NutzerInnen installiert werden muss, ist es nicht mehr möglich, die TeilnehmerInnen erst im Nachhinein darüber aufzuklären, dass die Nutzung mobiler Endgeräte im Zentrum des Erkenntnisinteresses steht.

TeilnehmerInnen zu verschaffen. Zum anderen galt es, kritisch zu reflektieren, inwiefern die entstandenen Aufnahmen dem Natürlichkeitsanspruch der KA genügen. Denn auch wenn die technische Aufzeichnung natürlich auftretender Interaktionen seit Jahrzehnten eine fest etablierte Praxis der KA darstellt, wird dieser Datenerhebungsmethode oftmals mit einer gewissen Skepsis begegnet: Die Kritik gründet dabei zumeist in der von Labov (1971) formulierten Idee des ‚Beobachter-Paradoxes‘, die von ihm folgendermaßen formuliert wird:

das Ziel der sprachwissenschaftlichen Erforschung der Gemeinschaft muss sein, herauszufinden, wie Menschen sprechen, wenn sie nicht systematisch beobachtet werden; wir können die notwendigen Daten jedoch nur durch systematische Beobachtung erhalten. (Labov, 1971, S. 135)

Da aus forschungsethischen Gründen eine informierte Einverständnis zur Aufnahme unerlässlich ist, wird durch die technische Aufzeichnung eine beliebige soziale Situation in eine öffentliche Situation transformiert, was für die TeilnehmerInnen mit dem Anspruch verbunden ist, „sich in einer öffentlichen Situation nicht-öffentlich zu verhalten“ (Schu, 2001, S. 97). Kritisch wird dabei oftmals angenommen, dass ein Bewusstsein über die Aufnahme die entstehenden Interaktionen zwangsläufig ‚kontaminieren‘ müsse und die Präsenz der Aufzeichnungstechnologie die TeilnehmerInnen dazu anhalte, sich in einer idealisierten Weise darzustellen (vgl. Tuncer, 2016, S. 1). Aus dieser Perspektive erscheinen die angenommenen Verzerrungen des TeilnehmerInnenverhaltens als unüberwindbares Hindernis, das es unmöglich macht, Situationen zu analysieren, die sich genauso auch ohne das Zutun der ForscherIn abgespielt hätten.

Aus konversationsanalytischer Perspektive lässt sich dieser Skepsis in zweierlei Weise begegnen: Erstens gilt das ethnomethodologisch-konversationsanalytische Interesse grundsätzlich der Untersuchung derjenigen Teilnehmermethoden, die im routinierten Vollzug alltäglicher Aktivitäten zur erkennbaren Hervorbringung eben dieser eingesetzt werden (vgl. Bergmann, 2000, S. 525). Unabhängig davon, ob eine Interaktion aufgezeichnet wird oder nicht, bedeutet dies für die TeilnehmerInnen, dass sie intersubjektiv verständliche (Ethno-)Methoden zur Realisierung sozialer Handlungen einsetzen müssen und dabei nur auf die (sprachlichen und verkörperten) Ressourcen zurückgreifen können, in deren Verwendung sie kompetent sind. Wenn die Beforschten also keine Vorgaben hinsichtlich ihres Verhaltens erhalten, werden Aktivitäten aufgezeichnet, die insofern ‚natürlich‘ sind, als dass sie für die Mitglieder eines Aktivitätssettings *accountable* sind.

Zweitens wird mit dieser Argumentation jedoch nicht impliziert, dass die Aufzeichnung keinerlei Auswirkungen auf das beobachtbare Verhalten der TeilnehmerInnen haben kann. Aus Perspektive der KA wird nicht a priori davon ausgegangen, dass Aufnahmetechnologien zwangsläufig eine Omnirelevanz für die TeilnehmerInnen besitzen, nur weil sie omnipräsent

sind (vgl. Speer & Hutchby, 2003, S. 334). Stattdessen wird die Beurteilung der Auswirkungen der Aufnahmesituation auf den Interaktionsverlauf zu einer empirischen Frage gemacht und reflektiert, ob und inwiefern sich die TeilnehmerInnen beobachtbar an der Interaktion als ‚aufgezeichneter Interaktion‘ orientieren. Konsistent mit den methodologischen Prämissen der KA muss eine Orientierung der TeilnehmerInnen an der Aufnahmesituation für andere Anwesende *accountable* gemacht werden, sodass diese auch für die AnalytikerIn (interaktions-)öffentlich in der Aufzeichnung verfügbar ist. Entsprechend lässt sich durch eine genaue Betrachtung der Aufnahmen empirisch rekonstruieren, an welchen Stellen eine Orientierung der TeilnehmerInnen an der Aufnahmesituation in den Vordergrund tritt.

Unter diesem Blickwinkel wurden alle Aufnahmen des Datenkorpus daraufhin untersucht, inwiefern die TeilnehmerInnen in ein beobachtbares „doing-being-observed“ (Hazel, 2016) involviert waren und sich an der Präsenz der Videokamera orientierten. Während einige Publikationen derartige Fragen zum alleinigen Gegenstand ihrer Untersuchung machen (bspw. Hazel, 2016; Laurier & Philo, 2006; Speer & Hutchby, 2003; Tuncer, 2016), möchte ich im Folgenden lediglich anhand einiger Datenausschnitte veranschaulichen, wie die TeilnehmerInnen damit umgehen, dass ihre Interaktionen durch die Aufzeichnung zum Untersuchungsgegenstand der Arbeit gemacht werden und abschließend reflektieren, wie sich diese Beobachtungen eines *doing-being-observed* zum Natürlichkeitsanspruch der KA verhalten. Besonders herausgreifen möchte ich dabei Situationen, in denen die TeilnehmerInnen ihr Verhalten beobachtbar an die Aufzeichnung anpassen und in denen die Kamera von ihnen als Ressource für die Produktion praktischer Handlungen genutzt wird.

Die Durchsicht des Datenkorpus zeigt immer wieder kurze Momente, in denen die TeilnehmerInnen direkt in die Kamera schauen, und sich somit durch ihr verkörpertes Verhalten an der hintergründigen Präsenz der Aufnahmetechnologie orientieren (Laurier & Philo, 2006, S. 188) – ohne dass dies jedoch zum Thema oder Fokus der lokalen Interaktion zu machen. Vereinzelt finden sich auch Momente, in denen „participants *adjust* to the camera in the organization of the course of their affairs and for their practical categorization as being filmable or not“ (Mondada, 2006, S. 61, Herv. i. O.). Eine derartige Anpassung des eigenen Verhaltens aufgrund der Aufnahmesituation findet bspw. in folgendem Ausschnitt einer Audioaufnahme statt, indem die zwei FreundInnen Peter (P) und Lara (L) gerade gemeinsam ein Foto machen:

### Transkript 1: 050217\_„Hat sie nich gesagt“<sup>53</sup>

125 P: mehr LÄcheln.  
126 L: zeig ma deine ZÄ:hne:; hihi  
127 <<lachend>sei ein guter SKLA[:ve.]  
128 P: [hohohoho  
129 P: [((lachen))]  
130 L: [((lachen))] hihihi  
131 P: hat sie NICH ge:sagt  
132 °hh  
133 stimme EINS sagt hm;  
134 un stimme zwei ruft zur VERnunft.  
135 (--)  
136 L: wie BEwusst dir das ist dass das aufgenommen wird;  
137 das ist SAU <<:-)>lu:stich>

P's Aufforderung, auf dem Foto mehr zu Lächeln, wird von L als Anlass für einen derben Witz (Z. 126–127) verwendet, indem sie das Lächeln für die Kamera mit dem Überprüfen der Zähne eines Sklaven vergleicht. Im weiteren Verlauf der Interaktion wird dieser Witz von P als heikle Angelegenheit im Kontext einer aufgezeichneten Interaktion kategorisiert. Indem P in Z. 133–134 den unmittelbar vorangegangenen Austausch der InteraktionspartnerInnen mit spezifischen Zitationsmarkern („stimme eins“, „stimme zwei“) rekonstruiert, erinnert er seine Interaktionspartnerin daran, dass das Gespräch aufgezeichnet wird und von einem unbestimmten Publikum gehört werden kann. Dadurch, dass P den Witz nicht selbst wiederholt (Z. 133) und sich als Stimme der „Vernunft“ positioniert (Z. 134), zeigt er seine Haltung zu dem Witz an und markiert diesen als unangemessen. L's Äußerungen in Z. 136–137 orientieren sich wiederum daran, welchen Umgang mit dem Witz sie von P unter ‚natürlichen‘ Bedingungen erwartet hätte und wie sehr sie sein Verhalten als beeinflusst von der Aufnahme interpretiert. Die immer expliziter werdende Orientierung an der Aufnahmesituation wird von den TeilnehmerInnen in diesem Fall als Ressource genutzt, sich gegenseitig damit zu konfrontieren, welches Bild sie von sich vermitteln und welche impliziten Erwartungen hinsichtlich der Interaktionsrollen sie an ihr Gegenüber richten (vgl. Tuncer, 2016, S. 18).

Neben Fällen, in denen InteraktionsteilnehmerInnen retrospektiv die vorangegangenen Äußerungen einer SprecherInnen als unangemessen für eine aufgezeichnete Interaktion kategorisieren (Mondada, 2006, S. 60), findet das *doing-being-observed* in anderen Situationen seinen Ausdruck in einer prospektiven Orientierung daran, eine projizierte Handlung aufgrund der

---

<sup>53</sup> Für die illustrativen Beispiele in diesem Kapitel wurden lediglich GAT2-Basistranskripte der sprachlichen Interaktion angefertigt.

Öffentlichkeit der Aufzeichnung nicht zu realisieren. Ein Beispiel dafür findet sich in folgendem Ausschnitt aus einem Mehrpersonengespräch, in dem parallel mehrere Unterhaltungen stattfinden. Transkribiert ist lediglich die Interaktion zwischen den zwei Interaktionspartnern Amir (A) und Markus (M) ganz rechts im Bild:

**Transkript 2: 190513\_„das sind die privaten Sachen“**

035 A: ey ich muss dir von SAMStagabend erzählen;=  
036 was ne kaTASTrophe.  
037 M: was?  
038 A: wo ich bei nem GEburtstag war-  
039 <<p>> aber ich erzähl hier NIX;>  
040 und HIER schon mal gar [nich  
041 M: [hhhihi  
042 (--)  
043 A: <<f>so schon gar nichts> <<ff>so #schOn GAR  
nIchts.>

#1



044 das sind die PRiVaten sachen;  
045 M: WELcher geburtstag?  
046 A: ja erzähl ich dir SPÄter;

In einer Prä-Sequenz kündigt A zunächst die Verfügbarkeit einer Erzählung an (Z. 037), woraufhin sich M in Z. 038 durch seine Frage als potenzieller Zuhörer positioniert. A entwickelt diese Erzählung jedoch anschließend nicht und begründet dies mit einer doppelten Unangemessenheit des Kontexts der Gesprächssituation: In Z. 039 scheint er auf die Gruppengröße und -zusammensetzung, in Z. 040 anschließend auf die Aufnahmesituation zu verweisen. Diese Orientierung an der Öffentlichkeit der Aufzeichnung wird in Z. 043 explizit, als A direkt in die Kamera schaut, mit deutlich erhöhter Lautstärke formuliert: „<<f>so schon gar nichts> <<ff>so schOn GAR nIchts.>“ und anschließend den privaten Charakter der Geschehnisse als Begründung anführt (Z. 044). An diesem Beispiel zeigt sich, dass eine projizierte Handlung

aufgrund der Aufnahmesituation von den Teilnehmern nicht realisiert wird. Durch Kommentare, die direkt an die Kamera gerichtet werden, wird der Inhalt der Geschichte als zu privat für eine Aufzeichnung kategorisiert (Mondada, 2006, S. 61).

Im weiteren Verlauf des Gesprächs (nicht im Transkript enthalten) zeigt sich, dass die TeilnehmerInnen alltäglicher Gespräche durchaus über kreative Wege verfügen, mit der Öffentlichkeit der Aufnahmesituation umzugehen: Die zwei Teilnehmer stehen kurz darauf auf und unterhalten sich außerhalb der Sicht- und Hörweite der Kamera weiter. Die problematisierte Aufnahmesituation bietet auf diese Weise Gelegenheit, spezifische Identitäts- und Beziehungsarbeit zu leisten (Gordon, 2012, S. 314), bspw. ‚geheime‘ Gespräche zu führen, die interaktiv ein Vertrauensverhältnis zwischen den Interaktionspartnern sowie Grenzen zu den anderen Anwesenden und der abwesenden ForscherIn herstellen. Derartige Praktiken des Umgangs mit aufgezeichneten Interaktionen führen jedoch auch dazu, dass bestimmte Aktivitäten für die AnalytikerIn nicht beobachtbar werden.

Neben retrospektiv und prospektiv orientierten Anpassungen des Verhaltens an die Aufnahmesituation wurde die auf dem Stativ installierte Kamera von den TeilnehmerInnen auch selbst als Ressource für die Produktion sozialer Handlungen genutzt. Dies möchte ich beispielhaft an einigen Screenshots illustrieren, in denen TeilnehmerInnen versehentlich ins Bild der Kamera gerieten oder sich gezielt in dieses bewegten, um einen visuellen ‚Witz‘ zu produzieren. Derartige Situationen entwickelten sich bspw., wenn TeilnehmerInnen aufstanden und zu ihrem Sitzplatz zurückkehrten, wenn weitere Personen sich zu den (bereits über die Aufnahme) informierten Beforschten hinzugesellten oder wenn das Personal der Cafés während seiner Arbeit ‚durch das Bild‘ lief.

In den folgenden Screenshots sind einige dieser humoristischen Interaktionen mit der Kamera festgehalten.



**Abb. 1: Beispiele für humoristische Interaktionen mit der Kamera**

Indem die TeilnehmerInnen der Kameralinse mit ihrem Gesicht oder den Händen ‚zu nahe‘ kommen, manipulieren sie die Optik der Aufnahme und stellen ihr Erscheinen im Bild der eingestellten Aufnahmetiefe gegenüber, die auf den Interaktionsraum der Beforschten ausgerichtet ist. Die TeilnehmerInnen nutzen ihre Kompetenz, zu erkennen, welchen Aufnahmebereich die Kamera fokussiert, um einen Witz zu produzieren: Dieser besteht zum einen aus einem Spiel

mit der ‚erwarteten‘ Distanz, die Personen zu einer Kamera einhalten, welche beobachtbar auf einen bestimmten Bereich ausgerichtet ist, und orientiert sich so der Erkennbarkeit des Witzes bei der nachträglichen Betrachtung der Aufzeichnung. Zum anderen stehen derartige Witze in Relation zum Einverständnis der TeilnehmerInnen, sich bei ihrem gewöhnlichen Verhalten in alltäglichen Situationen aufzeichnen zu lassen. Für einen flüchtigen Moment wird die Kamera durch die demonstrativen, verkörperten Handlungen thematisiert und ihre Anwesenheit für andere Anwesende relevant gemacht, die diese performativen Witze zumeist mit Lachen und amüsierten Kommentaren goutieren. Die TeilnehmerInnen machen sich so gemeinsam über die ‚Aufzeichnungs-Bürokratie‘ der Aufnahmesituation lustig (vgl. Laurier & Philo, 2006, S. 185–186) und bringen die Kamera durch ihr Verhalten als ‚Fremdkörper‘ im Kontext ihrer alltäglichen Interaktionen hervor.

Auch wenn derartige Situationen und Praktiken des *doing-being-observed* in den Daten selten auftreten, ist nicht von der Hand zu weisen, dass durch die Einführung von Aufnahmetechnologien in die natürlich Ökologie der Interaktionen spezifische Sequenzen veranlasst werden, die ohne die Aufzeichnung in dieser Form nicht aufgetreten wären (Potter, 2004, S. 612). Diese, an der Aufnahmesituation orientierten Sequenzen repräsentieren jedoch keine ‚unnatürlichen‘ Datenstücke. Es lässt sich vielmehr argumentieren, dass sie ‚natürliche‘ Verhaltensweisen im Kontext von Situationen darstellen, die technisch aufgezeichnet werden: Es ist ein konstitutives Merkmal alltäglicher Unterhaltungen, dass sie eine starke lokale Sensibilität besitzen, die es den TeilnehmerInnen erlaubt, Objekte und Ereignisse in der situativen Umgebung der Interaktion zu thematisieren: „It is therefore the most ‚natural‘ thing for interactants who know that their conversation is recorded, to comment on the recording itself“ (Bergmann, 1990, S. 219). Die beobachtbaren Orientierungen an der Aufnahmesituation stellen insofern selbst ‚natürliche Interaktionen‘ dar, als in ihnen (Ethno-)Methoden des Umgangs mit der Öffentlichkeit von aufgezeichneten Interaktionssituationen zum Einsatz kommen.

In derartigen, hier beispielhaft illustrierten Episoden orientieren sich die TeilnehmerInnen also an einer ‚präsenten Präsenz‘ („present presence“, Laurier & Philo, 2006, S. 189) der Kamera, wodurch eindeutig erkennbar wird, wie die Öffentlichkeit der Aufnahmesituation das TeilnehmerInnenverhalten beeinflussen kann. Das Auftreten von Sequenzen, in denen die Orientierung an der Aufnahmesituation in den Vordergrund tritt, macht jedoch nicht den gesamten Datenkorpus unbrauchbar: Derartige Episoden lassen sich eindeutig von denjenigen Aktivitäten unterscheiden, die ‚natürlicherweise‘ Bestandteil alltäglicher Interaktionen zwischen den Anwesenden sind (bspw. durch beobachtbare Differenzen in der Blickführung, der Adressierung der abwesenden ForscherIn oder verkörperte Orientierungen am Bildausschnitt). Diese Passagen

wurden – unabhängig davon, ob in ihnen mobile Endgeräte gebraucht werden oder nicht – nicht bei der Kollektionsbildung berücksichtigt, sondern als Artefakte der Aufnahmesituation behandelt. Die als ‚naturalistisch‘ zu behandelnden und damit potenziell analyserelevanten Daten wurden somit durch die konstitutive Differenz zwischen der beobachtbaren An- oder Abwesenheit der Aufzeichnungstechnologie in den Videodaten produziert (vgl. Laurier & Philo, 2006, S. 191).

Es ist grundsätzlich anzuerkennen, dass die Präsenz der Aufnahmetechnologie die Frequenz des Auftretens bestimmter Phänomene beeinträchtigen kann. Das Bewusstsein der TeilnehmerInnen über die Aufnahmesituation kann dazu führen, dass bestimmte Aktivitäten – wie etwa das Erzählen von Geschichten, deren Thema als zu privat kategorisiert wird – der Erforschung durch Videoaufnahmen verborgen bleiben und die ForscherIn niemals ein vollständiges Bild aller Aktivitäten erhält, die in dem untersuchten Setting auftreten können (vgl. Drew, 1989). Dies war jedoch auch nicht das Ziel dieser Arbeit: Die Videoaufnahmen wurden angefertigt, um „naturally organized ordinary activities“ (Lynch, 2002) – die emischen Organisationsprinzipien alltäglicher Aktivitäten – in diesem Fall des *smartphone-assistierten Geschichtenerzählens* in alltäglichen Unterhaltungen – aus Perspektive der multimodalen KA untersuchen zu können. Das Verhalten der TeilnehmerInnen muss dabei für ihre InteraktionspartnerInnen zur *Accountability* der jeweils im Vollzug befindlichen Aktivitäten beitragen, sodass „[t]he practices through which cooperation and intersubjectivity are produced and maintained, and their organization, do not substantially vary because of cameras“ (Tuncer, 2016, S. 18).

### **4.3 Datenkorpus und Kollektionsbildung**

Neben dem Ausschluss aller Gesprächsepisoden, in denen sich die TeilnehmerInnen beobachtbar an der Präsenz der Kamera orientieren, wurden bei der Betrachtung der Videoaufnahmen schriftliche Inventare des Gesprächsablaufs angelegt, die in Anlehnung an die Vorschläge von Deppermann (2008, S. 34) formatiert wurden: Um eine Übersicht über den in der Entwicklung befindlichen Datenkorpus zu erlangen, wurde die grobe thematische Entwicklung der Gespräche notiert und exakte Zeitmarken für die unterschiedlichsten Formen der Nutzung mobiler Endgeräte festgehalten. Am Ende der zweiten Erhebungsphase war ein Korpus von 25 Aufzeichnungen entstanden, in dem insgesamt 213 Episoden der Smartphonennutzung identifiziert werden konnten. Diese lassen sich noch einmal dahingehend differenzieren, ob es sich um Formen der individuellen Handynutzung, „i.e., when the mobile device use disconnected from the

ongoing conversation“ (Oloff, 2019b, Z. 235–236), oder um kollaborative Formen handelt, in denen das Smartphone für gemeinsame Aktivitäten der Anwesenden genutzt wurde.

<b>Datenkorpus (25 Aufnahmen = 19h)</b>	<b>Individuelle Smart- phonennutzung</b>	<b>Kollaborative Smartphonennut- zung</b>	<b>Smartphonennut- zung gesamt</b>
Zweipersonengespräche (15 Aufnahmen = 8h 20min)	42	37	79
Mehrpersonengespräche (10 Aufnahmen = 10h 40min)	91	43	134
<b>Gesamtanzahl Episoden der SP-Nut- zung</b>	<b>133</b>	<b>80</b>	<b>213</b>

Abb. 2: Tabellarische Übersicht der Smartphonennutzung im Datenkorpus<sup>54</sup>

Anhand dieser Übersicht lässt sich zunächst erkennen, dass es in einem Großteil der Aufnahmen sowohl zu individuellen als auch kollaborativen Formen der Handynutzung kam.<sup>55</sup> Die individuellen Nutzungsformen umfassen Aktivitäten wie das Telefonieren mit Abwesenden sowie das Lesen und Beantworten von Textnachrichten oder Emails – in den allermeisten Fällen handelt es sich jedoch um „checking behaviors“ (Oulasvirta et al., 2012, S. 105): um kurze Fokussierungen des Smartphones, mit denen die TeilnehmerInnen überprüfen, ob auf dem Display Neuigkeiten, bspw. Textnachrichten, Emails oder Social-Media-Updates, angezeigt werden. Auffällig ist, dass Formen der individuellen Smartphonennutzung in Mehrpersonengesprächen verhältnismäßig häufiger auftreten: Dies hängt nicht nur damit zusammen, dass gleichzeitig mehrere AkteurInnen anwesend sind, die in der Regel jeweils über ein eigenes Smartphone verfügen. Wie DiDomenico et al. (2018, S. 681) zeigen konnten, bieten Mehrpersonengespräche den TeilnehmerInnen größere Freiräume, ihr Engagement an der F2F-Interaktion zu reduzieren und sich vorübergehend ihrem mobilen Endgerät zuzuwenden, da immer auch andere Anwesende mit sprachlichen Äußerungen adressiert werden oder sich selbst als nächste

<sup>54</sup> Diese quantitative Darstellung suggeriert eine Eindeutigkeit, die der Komplexität des Datenmaterials eigentlich nicht angemessen ist: Für die Auszählungen wurde der Verlauf von der Initiierung bis zur Beendigung (den Beginn der Bedienung bis zum Weglegen) des Smartphonegebrauchs als *eine* Episode der Nutzung des mobilen Endgeräts verstanden. Konversationsanalytisch lassen sich dabei zum einen Episoden identifizieren, in denen ein Wechsel zwischen individuellen und kollektiven Nutzungsformen stattfindet. Zum anderen spräche einiges dafür, bspw. Phasen der Smartphonennutzung, in der zunächst einer ChatpartnerIn geantwortet und anschließend die Nachricht(en) einer anderen KommunikationspartnerIn vorgelesen wird, als *zwei* Episoden des Smartphonegebrauchs zu interpretieren. Eine derartige, sequenzanalytische Identifikation aller im Korpus auftretenden Episoden des Smartphonegebrauchs war jedoch nicht Ziel der Arbeit. Diese tabellarische Darstellung soll lediglich dazu dienen, einen Eindruck davon zu vermitteln, wie häufig bzw. selten das Phänomen auftritt.

<sup>55</sup> Dabei muss berücksichtigt werden, dass in drei Aufnahmen gar keine mobilen Endgeräte verwendet wurden.

SprecherInnen auswählen können. Insbesondere wenn TeilnehmerInnen aus ihrer Rolle als ZuhörerInnen den individuellen Smartphonegebrauch initiieren, wird dieser (meist) nicht als rechtfertigungsbedürftiges Verhalten behandelt, das zum Thema der Interaktion werden müsste (Oloff, 2019b, Z. 450).

Für diese Arbeit von besonderem Interesse sind diejenigen Gesprächsepisoden, in denen das Smartphone als Ressource für gemeinsame Aktivitäten der Anwesenden gebraucht wird. Aus diesen Fällen wurde die zu analysierende Kollektion gebildet: Ausgehend von der einfachen Beobachtung, dass die TeilnehmerInnen ihre Smartphones als Ressource in der Hervorbringung von Geschichtenerzählungen einsetzen, wurde dieses wiederkehrende Phänomen als Ordnungselement identifiziert (vgl. Bergmann, 1988a, S. 30–32). Alle Fälle, in denen die multimodalen Praktiken des Smartphonegebrauchs eine konstitutive Rolle für die Hervorbringung einer Geschichtenerzählung übernehmen, bilden die für diese Arbeit zentrale Kollektion (vgl. Sidnell, 2012, S. 88–92). Mit Geschichtenerzählungen werden dabei größere Einheiten im Gespräch fokussiert, die durch eine Vielzahl von Einzelhandlungen und Sequenzpaaren hervorgebracht werden. Derartige *big packages* sind für die TeilnehmerInnen als Geschichtenerzählungen gut zu erkennen, in ihrer multimodalen Ausgestaltung jedoch schwer analytisch präzise zu erfassen (vgl. König & Oloff, 2018b, S. 225). Die Kollektionsbildung stellte deshalb aus verschiedenen Gründen eine praktisch zu bewältigende Herausforderung dar. Diese Herausforderungen sollen im Folgenden diskutiert werden, um anschließend das analytische Vorgehen sowie die Form der Ergebnisdarstellung nachvollziehbar zu machen.

Wie bereits diskutiert, erhöht sich durch die Arbeit mit Videodaten grundsätzlich die Komplexität der untersuchten Interaktionen in beträchtlichem Maße und macht die Berücksichtigung der verschiedenen Zeitlichkeiten des Einsatzes verkörperter, materieller und räumlicher Ressourcen erforderlich. Für die Kollektionsbildung muss entsprechend der Einsatz unterschiedlicher Ressourcen berücksichtigt werden, um zusammengehörende Fälle zu identifizieren, die als methodische Hervorbringung eines Ordnungselements (Bergmann, 1988a, S. 32) klassifiziert werden können. Für die Untersuchung von Geschichtenerzählungen als multimodaler Praktik bedeutete dies, sich „von einem (angenommenen) Ideal des identischen Musters trennen“ (König & Oloff, 2018b, S. 225) zu müssen und stattdessen zu versuchen, prototypische Handlungsmuster zu identifizieren und zu beschreiben (Mondada, 2016a). Im Rahmen dieser Arbeit wurden deshalb bspw. Erzählungen, in denen sich die TeilnehmerInnen in unterschiedlicher Sitzposition zueinander befinden – die Interaktion damit also keine sichtbare Gleichheit hinsichtlich ihres Teilnahmerahmens (C. Goodwin, 2007a) aufweist – als Fälle einer gemeinsamen Kollektion behandelt, wenn sie durch andere, vergleichbare *multimodale Gestalten*

konstituiert wurden, bspw. durch das Vorlesen von Textnachrichten als Methode der Ausgestaltung der Erzählung.

Die Gegenstandskonstitution erforderte darüber hinaus eine weitere Differenzierung der untersuchten Fälle: Geschichtenerzählungen sind (vereinfacht) durch eine dreigliedrige Struktur aus Erzählankündigung und Rahmensetzung, der Kernaktivität der Erzählung und Abschluss gekennzeichnet. In den Phasen der Prä-, Kern- und Postsequenz gilt es für die TeilnehmerInnen jeweils spezifische Aufgaben der Interaktionsorganisation zu bearbeiten (bspw. Quasthoff, 2001). Entsprechend wurde innerhalb der Kernkollektion unterschieden, in welcher Phase der Erzählung der Smartphonegebrauch stattfindet und der Bearbeitung welcher Aufgabe er dient. Auf diese Weise entstanden Sub-Kollektionen, in denen bspw. Fälle versammelt wurden, in denen in der Prä-Sequenz digitale Fotografien über das Smartphone integriert wurden, um Figuren der narrativen Rekonstruktion visuell einzuführen; in denen Medieninhalte eingespielt wurden, um in der Kernphase die Erzählung auszugestalten, oder das Smartphone erst in der Post-Sequenz in Reaktion auf die Erwiderungen der TeilnehmerInnen eingesetzt wurden.

Die dreigliedrige Grundstruktur stellt ein generisches Merkmal von Geschichtenerzählungen dar. Erzählungen werden jedoch auf unterschiedliche Weise kontextsensitiv ausgestaltet, um eine Bandbreite unterschiedlicher Handlungen zu realisieren. Abhängig davon, ob mit der Geschichte ein Witz erzählt (bspw. Sacks, 1978), ein Problem dargestellt (bspw. Jefferson, 1988), eine Beschwerde formuliert (bspw. Günthner, 1999a) oder das eigene Verhalten erklärt (bspw. Mandelbaum, 1993) wird, verändern sich die spezifischen, zu bearbeitenden Aufgaben und die zur Realisierung eingesetzten (TeilnehmerInnen-)Methoden. Entsprechend musste berücksichtigt werden, welchem interaktionalen Zweck die untersuchten Fälle von Geschichtenerzählungen dienten, und welchen spezifischen Beitrag die Integration des Smartphones für diesen Typ der Erzählung leistete.

Während die vorangegangenen Aspekte eine Differenzierung der Kollektion aufgrund von Eigenschaften der Praktiken des Geschichtenerzählens erforderlich machten, wird eine weitere zentrale Differenz durch die Medialität der über das Smartphone integrierten Inhalte konstituiert. Die Analyse der Fälle zeigt, dass die Frage, ob es sich um schriftsprachliche oder visuelle Medieninhalte handelte, einen Unterschied für die Art und Weise der Integration, der Rolle des visuellen Zugriffs zum Gerätedisplay sowie das Funktionspotenzial der Medieninhalte als Ressource für die Hervorbringung der Erzählung macht. So wie bspw. Textnachrichten über multimodale Praktiken des Vorlesens in einer Erzählung integriert werden können, wird die Integration einer digitalen Fotografie über Praktiken des Zeigens realisiert.

Die grundsätzliche Multimodalität *smartphone-assistierter Geschichtenerzählungen*, ihre dreigliedrige Struktur, ihr Einsatz für die Realisierung verschiedenster sozialer Handlungen sowie die unterschiedliche Medialität der integrierten Medieninhalte sorgten also – noch bevor die Details spezifischer Praktiken in den Blick genommen werden konnten – dafür, dass die Kernkollektion in zahlreiche Sub-Kollektionen organisiert werden musste. Dies machte eine spezifische Anpassung des analytischen Vorgehens und der Ergebnisdarstellung erforderlich, die im Folgenden dargestellt wird.

#### **4.4 Datenauswertung: die Analyse von Kollektionen und Einzelfällen**

Um die Fragen nach den Methoden des Einsatzes von Smartphones als Ressource in der Hervorbringung von Geschichtenerzählungen sowie der Funktion der integrierten Medieninhalte für die Erzählungen beantworten zu können, wurden zu Beginn ausgewählte Fälle als individuelle Interaktionsepisoden einer detaillierten Sequenzanalyse (Deppermann, 2008, S. 53–78) unterzogen. Dafür wurden zunächst Videoclips der entsprechenden Gesprächsepisoden sowie GAT2-Basistranskripte (Selting et al., 2009, S. 369–377) der sprachlichen Interaktion angefertigt. Die analytische Arbeit erfolgte am Videomaterial, die Transkripte fungierten dabei als Hilfsmittel, die zumindest die sprachliche Interaktion in einer schriftlich fixierten Version repräsentierten (Hutchby & Wooffitt, 1998, S. 74). In den detaillierten Sequenzanalysen wurden die konstitutiven Ressourcen der Praktiken *smartphone-assistierter Geschichtenerzählungen* identifiziert und Annahmen über ihre praktischen Aufgaben – als Methoden zur Bearbeitung von Problemen der Interaktionsorganisation – formuliert. Zu diesem Zeitpunkt waren die Analysen vor allem auf eine möglichst vollständige Erschließung der Sinnhaftigkeit des Smartphonegebrauchs im Kontext der jeweiligen Einzelfälle fokussiert. Einzelne Exemplare *smartphone-assistierter Geschichtenerzählungen* wurden also als kontextsensitive Lösungen spezifischer Aufgaben in ihrem Fallkontext rekonstruiert. In anschließenden fallübergreifenden Analysen wurde versucht, die – bereits andeutungsweise beschriebene – Varianz *smartphone-assistierter Geschichtenerzählungen* auszuarbeiten. Ziel dieser Analysephase war es, die kontextunabhängigen Verfahrensweisen des Einsatzes von Smartphones als Ressource in der Hervorbringung der Geschichtenerzählungen zu rekonstruieren.

Durch die fallvergleichende Analyse trat besonders deutlich das reflexive Verhältnis zwischen den Organisationsprinzipien von Geschichtenerzählungen und den Praktiken des Smartphonegebrauchs hervor: Über das Smartphone einspielbare Medieninhalte können auf vielfältige Weise zur Bewältigung der interaktiv zu bearbeitenden Aufgaben in Geschichtenerzählungen eingesetzt werden. Damit das Smartphone zur sinnhaften Hervorbringung dieser spezifischen

Aktivität beitragen kann, sind jedoch wiederum spezifische Aufgaben – bspw. die Projektion des Smartphonegebrauchs, die Suche nach Medieninhalten oder die Organisation von Zeigesequenzen – zu bewältigen, die innerhalb der Grundstrukturen der Geschichtenerzählung eingebettet werden müssen und mittels spezifischer Praktiken bearbeitet werden.

Die fallübergreifenden Analysen trugen außerdem zu einer veränderten Gegenstandskonstitution bei: Zunächst wurden hauptsächlich Fälle von Geschichtenerzählungen untersucht, die insofern eine kanonische Form aufweisen, als dass sie durch methodisch hervorgebrachte Prä- und Post-Sequenzen eindeutig vom umgebenden *turn-by-turn-talk* abgegrenzt sind (Quasthoff, 2001, S. 1302). Durch die vergleichende Analyse konnten auch Gesprächsepisoden als *smartphone-assistierte Geschichtenerzählungen* identifiziert und in die Kollektion aufgenommen werden, die nicht eindeutig vom umgebenden *turn-by-turn-talk* abgegrenzt, sondern stärker fragmentiert (vgl. Kotthoff, 2018; Spreckels, 2008) sind oder sich auf nicht abgeschlossene, noch andauernde Ereignisse beziehen (vgl. Georgakopoulou, 2008; Günthner, 2012a). Im Verlauf der Analysen zeigte sich dabei, dass der Umfang der Kollektion durchaus geeignet ist, auch die stärker kontextunabhängigen Strukturen der Praktiken der Integration mobiler Endgeräte in Geschichtenerzählungen zu identifizieren. Aufgrund des begrenzten Umfangs der verschiedenen Subkollektionen gestaltete es sich jedoch schwieriger, die Spezifik der Smartphoneintegration in den unterschiedlichen Phasen von Geschichtenerzählungen sowie für die Ausgestaltung von Erzählungen in Abhängigkeit von der realisierten Handlung (bspw. Beschwerden oder Erklärungen) zu rekonstruieren.

Aus diesem Grund spielt für die Arbeit die methodologische Begründung der KA für die Verallgemeinerbarkeit der Ergebnisse von Einzelfallanalysen<sup>56</sup> eine wichtige Rolle: Die KA geht zunächst davon aus, dass „[t]here is constitutive order to singular occasions of interaction, and to the organization of action within them“ (Schegloff, 1987a, S. 102, Herv. i. O.). Der Anspruch des konversationsanalytischen Forschungsprogramms ist es, die ‚Maschinerie von Interaktionen‘ zu rekonstruieren, deren kontextunabhängige Ordnungsprinzipien in konkreten Interaktionsepisoden kontextsensitiv partikularisiert werden. Eine derartige Maschinerie muss also in der Lage sein, mit den Details einzelner Interaktionsepisoden umzugehen. Die Eigenschaften und Organisationsprinzipien dieser Maschinerie finden sich in konversationsanalytischen Untersuchungen unterschiedlichster Gesprächsphänomene: „In a sort of exercise, the resources of past work on a range of phenomena and organizational domains in talk-in-interaction are

---

<sup>56</sup> Dieser Verfahrensweise bedient sich bspw. Harvey Sacks in einer Vielzahl seiner „Lectures on Conversation“ (1995), aber auch Jefferson (1980, 2015), Goodwin (1984) und Schegloff (1984). Ausführlich wird die Logik von Einzelfallanalysen in Schegloff (1987a) diskutiert.

brought to bear on the analytic explication of a *single fragment* of talk” (Schegloff, 1987a, S. 101, Herv. i. O.). Für die Analyse einzelner Datenfragmente, die nicht systematisch in Kollektionen eingebunden werden können, stellen vorhandene Erkenntnisse über Interaktionsphänomene also eine zentrale Ressource dar, um ihre methodische Geordnetheit zu rekonstruieren. Da es sich bei Geschichtenerzählungen um ein seit der Gründung der KA in einer Vielzahl von Studien untersuchtes Gesprächsphänomen handelt, konnte ich mir für die vorliegende Arbeit diese Vorgehensweise zu Nutze machen, um die (Ethno-)Methoden der Smartphoneintegration in den unterschiedlichen Phasen und Formen von Geschichtenerzählungen zu explizieren. Mit diesem Ansatz wird außerdem eine konstante Rückbindung der eigenen Analysen zum konversationsanalytischen Forschungsstand sichergestellt.

Indem ich aufzeige, wie der Smartphonegebrauch im Interaktionsvollzug mit den Organisationsprinzipien von Geschichtenerzählungen koordiniert wird, wird das *affordanztheoretische* Argument, dass „humans make what they do of technologies in the interface between the organized practices of human conversation, and the technology's array of communicative affordances“ (Hutchby, 2001a, S. 206) anhand einer konkreten kommunikativen Aktivität empirisch umgesetzt: Wie zu zeigen sein wird, findet die Art und Weise der Nutzung der Handlungsmöglichkeiten von Smartphones eingebettet in die organisierten Praktiken von Geschichtenerzählungen in Interaktion statt. Die Frage, wie und mit welcher Funktion Smartphones in Geschichtenerzählungen integriert werden und welche Praktiken die so hervorgebrachten *smartphone-assistierten Geschichtenerzählungen* auszeichnen, wird in den folgenden analytischen Kapiteln beantwortet.

Die Ergebnisdarstellung erfolgt dabei grundsätzlich in Form von Einzelfallanalysen, mit der unterschiedliche Varianten *smartphone-assistierter Geschichtenerzählungen* anhand von exemplarischen Fällen rekonstruiert werden. Sowohl hinsichtlich der Analysen von Geschichtenerzählungen (König & Oloff, 2018b) als auch der Smartphoneintegration in eben diese kommunikative Aktivität fokussiert diese Arbeit Phänomene, die aus Perspektive der multimodalen KA bis dato noch wenig erforscht sind. In dem Versuch, die Systematik und Varianz *smartphone-assistierter Geschichtenerzählungen* zu rekonstruieren, wird diese Form der Einzelfallanalyse durchgehend beibehalten, um eine intersubjektive Nachvollziehbarkeit der Analysen dieser neuartigen Form narrativer Rekonstruktionen sicherzustellen.

Alle Transkripte wurden dabei nach den von Lorenza Mondada (2016b) entwickelten Transkriptionskonventionen für multimodale Interaktionen ausgearbeitet. Diese basieren auf dem GAT2-System (Selting et al., 2009) und werden um folgende Konventionen für die Annotation

verkörperter Handlungen ergänzt, um das zeitliche und sequenzielle Zusammenspiel der unterschiedlichen, von den TeilnehmerInnen eingesetzten Ressourcen schriftlich zu repräsentieren:

* *	Gestures and descriptions of embodied actions are delimited between
+ +	two identical symbols (one symbol per participant)
Δ Δ	and are synchronized with corresponding stretches of talk.
*--->	The action described continues across subsequent lines
---->*	until the same symbol is reached.
>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
....	Action's preparation.
----	Action's apex is reached and maintained.
,,,,	Action's retraction.
ric	Participant doing the embodied action is identified when (s)he is not the speaker.
fig	The exact moment at which a screen shot has been taken
#	is indicated with a specific symbol showing its position within the turn at talk.

**Abb. 3: Multimodale Transkriptionskonventionen nach Mondada (2016b, S. 1)**

Die schriftlichen Repräsentationen der multimodalen Interaktionen werden durch Standbilder<sup>57</sup> ergänzt, die prägnante Momente der Smartphoneintegration in die Erzählungen in ihrer holistischen, *multimodalen Gestalt* visuell darstellen (Stukenbrock, 2009, S. 154). Die Transkripte in den folgenden Kapiteln stellen somit das Ergebnis der Analysen des jeweiligen Falles dar (Mondada, 2018): Sie repräsentieren diejenigen sprachlichen, verkörperten, materiellen und räumlichen Ressourcen, an denen sich die TeilnehmerInnen beobachtbar als relevant für die interaktive Hervorbringung von *smartphone-assistierten Geschichtenerzählungen* orientieren.

---

<sup>57</sup> Ebenso wie die Transkripte wurden auch alle Standbilder anonymisiert: mittels einer Bildbearbeitungssoftware wurden sie so verfremdet, dass die InteraktionsteilnehmerInnen nicht erkennbar sind.

## 5 Das Vorlesen von Textnachrichten in smartphone-assistierten Geschichtenerzählungen

„Seit 1992 die erste SMS verschickt wurde, hat die handyvermittelte schriftliche Kommunikation immer mehr an Bedeutung in der Alltagsinteraktion gewonnen“ (König & Bahlo, 2014, S. 1) Mit der zunehmenden Verbreitung und der technologischen Weiterentwicklung der mobilen Endgeräte ist auch die schiere Menge versendeter Nachrichten exponentiell gewachsen:

There are currently an estimated 269 billion e-mails, 55 billion Whatsapp messages, over 9.5 billion Instagram photos and videos, and around 500 million tweets per day. Online interaction is a worldwide phenomenon, with users of Chinese social media site Weibo sending over 100 million messages each day and about 38 billion messages sent every day on WeChat, the Chinese instant messaging service. (Meredith, 2019, S. 241)

Der Großteil der für diese Kommunikation erforderlichen Software kann auf modernen Smartphones installiert werden und wird auf ihnen in den unterschiedlichsten Situationen im Alltag der NutzerInnen verwendet. Die technischen Möglichkeiten für diese asynchronen, medial-vermittelten Formen der Kommunikation werden oftmals durch Instant-Messaging-Applikationen (IM-Applikationen) bereitgestellt, mit denen Textnachrichten und andere multimediale Botschaften wie Bilder, Videos, GIFs und Dokumente versendet werden können. Durch ihre Nutzung entstehen auf den Smartphones der NutzerInnen *Archive sozialer Originale* (Keppler, 2013, S. 99), in denen diese Kommunikation in all ihren Details archiviert wird. Denn anders als es noch bei SMS üblich war, werden von Emailprogrammen und IM-Applikationen nicht mehr lediglich die empfangenen Nachrichten gespeichert, sondern diese nun gemeinsam mit versendeten Nachrichten als „Thread“ – in der Reihenfolge ihres zeitlichen Eingangs auf dem Server der jeweiligen Applikation – archiviert und dargestellt (vgl. bspw. Wyss & Hug, 2016, S. 241): Auf diese Weise bleiben die Kommunikationsbeiträge aller Beteiligten erhalten und die archivierte Interaktion lässt sich in ihrer Entstehung ‚Zug-um-Zug‘ nachvollziehen.

Die damit für die Smartphone-NutzerInnen verbundene Möglichkeit, die gespeicherten Kommunikationsverläufe jederzeit wieder aufrufen zu können, kann von ihnen auf vielfältige Art und Weise genutzt werden, um sich im Kontext von F2F-Interaktionen sprachlich auf dieses Archiv zu beziehen:

In its modern threaded chat design text messaging supports a resource to be drawn on in different ways in support of face-to-face communication. Moreover, the ability to go back over archived text and images allows for different sorts of reporting on those conversations—gisting, storytelling, verbatim reading, and the like, each of them designed with respect to local arrangements. (Brown et al., 2018, S. 23)

Aufgrund der Speicherung aller Kommunikationsbeiträge im Thread-Format stellen derartige Kommunikationsplattformen für die TeilnehmerInnen eine Ressource dar, auf die sich in F2F-

Interaktionen auf unterschiedliche Weise sprachlich beziehen können: Bspw. durch Zusammenfassungen, Geschichtenerzählungen, oder wörtliches Vorlesen, wobei alle diese sprachlichen Aktivitäten spezifisch auf den lokalen Kontext der F2F-Interaktion und die jeweiligen RezipientInnen angepasst werden.

In den folgenden Kapiteln (Kap. 5.1–5.3) werde ich eine spezifische Form des Einsatzes archiverter Kurznachrichten-Threads in alltäglichen Gesprächen untersuchen: Im Fokus werden *smartphone-assistierte Geschichtenerzählungen* stehen, die konstitutiv durch das Vorlesen archiverter Textnachrichten hervorgebracht werden. Über den gesamten Korpus hinweg lässt sich immer wieder beobachten, dass die GesprächsteilnehmerInnen nicht nur über diese Form medial-vermittelter Kommunikation sprechen, sondern das Smartphone als materielle Ressource einsetzen, um die archivierten Textnachrichten in situ aufzurufen, diese auf direkte – wörtlich zitierende – oder indirekte – zusammenfassende – Weise wiederzugeben und andere Anwesende (mit-)lesen zu lassen. In Kombination mit dem methodischen Einsatz sprachlicher und verkörperter Praktiken werden auf diese Weise *smartphone-assistierte Geschichtenerzählungen* hervorgebracht, die für andere Anwesende als solche verständlich sind und deren Bedeutung anschließend interaktiv etabliert wird. Allen für dieses Kapitel analysierten Fällen ist dabei gemeinsam, dass Praktiken der smartphone-assistierten Wiedergabe von Textnachrichten eine zentrale Ressource in der Hervorbringung der Erzählung darstellen. Analytisch werden derartige Erzählungen in den folgenden Kapiteln dahingehend unterschieden, wie sie von den SmartphoneinhaberInnen eröffnet und in das Gespräch eingeführt werden.

Im ersten analytischen Kapitel (5.1) werde ich smartphone-assistierte Erzählungen diskutieren, die von der aktuellen SprecherIn und SmartphoneinhaberIn methodisch vorbereitet und auf eine Weise in die sprachliche Interaktion eingeführt werden, dass das Verhältnis der projizierten Erzählung zum vorangehenden Gespräch *accountable* gemacht wird. Im anschließenden Kapitel (5.2) werde ich eine Variante des Phänomens vorstellen, in der die smartphone-assistierte Erzählung ‚unvorbereitet‘ – d. h. ohne die Produktion einer elaborierten Prä-Sequenz – eröffnet wird und rekonstruieren, welche Folgeprobleme durch die uneindeutige Verbindung der Geschichte zum vorangehenden Gespräch entstehen können. In beiden Varianten spielt dabei die *Affordanz* des Smartphones als *Archiv sozialer Originale* eine zentrale Rolle, da sie es den GeräteinhaberInnen erlaubt, mittels weniger Bediengesten gespeicherte Threads aufzurufen und in die Interaktion einzuspielen. Im abschließenden Kapitel (5.3.) dieses Teils der Arbeit werde ich eine Variante diskutieren, in der nun die *Affordanz* des Smartphones zur orts- und zeitunabhängigen Telekommunikation im Zentrum steht: Es werden smartphone-assistierte Erzählungen untersucht, die dadurch ausgelöst werden, dass die SmartphoneinhaberIn während der

Unterhaltung den Erhalt einer neuen (Text-)Nachricht wahrnimmt. Konstitutiv für derartige Erzählungen ist, dass diese Nachrichten durch die SmartphoneinhaberIn explizit sprachlich für die Interaktion zwischen den Anwesenden relevant gemacht (vgl. DiDomenico et al., 2018, S. 22) und als Ressource für die interaktive Hervorbringung einer Geschichtenerzählung eingesetzt werden. Wie zu zeigen sein wird, handelt es sich dabei um Erzählungen, die sich auf noch nicht abgeschlossene Ereignisse beziehen und im Vollzug der F2F-Interaktionen weiterentwickelt werden können (vgl. Georgakopoulou, 2015, S. 260).

Wie bereits erwähnt, ist allen in diesem Kapitel analysierten Varianten *smartphone-assistierter Geschichtenerzählungen* gemein, dass sie konstitutiv durch sprachliche Praktiken der Wiedergabe archivierter Textnachrichten hervorgebracht werden. Oftmals werden die Textnachrichten dabei von der ErzählerIn wörtlich zitiert. Dies führt dazu, dass derartige smartphone-assistierte Erzählungen an der linguistischen Gesprächsoberfläche große Ähnlichkeit zu den konversationsanalytisch gut erforschten Modi der direkten und indirekten Redewiedergabe aufweisen. Aufgrund der Verwandtschaft dieser Phänomene bieten die vorhandenen Erkenntnisse über die Praktiken der DRS und IRS, mit denen GesprächsteilnehmerInnen im Vollzug von F2F-Interaktionen ihre eigenen oder die Äußerungen anderer aus vergangenen (Gesprächs-)Situationen wiedergeben, einen analytischen Rahmen für dieses Kapitel. Um meine eigenen Analysen im Forschungsstand zu verorten, werde ich ihnen im Folgenden deshalb eine Übersicht über die zentralen konversationsanalytischen Ergebnisse der Erforschung von Praktiken der Redewiedergabe voranstellen sowie – in einem zweiten Schritt – vorhandene Erkenntnisse zur Wiedergabe von Textnachrichten diskutieren.

In konversationsanalytischen Studien, die Praktiken der Wiedergabe von vergangenen sprachlichen Äußerungen, untersuchen, wird grundsätzlich zwischen DRS und IRS (vgl. Holt & Clift, 2007, S. 3) unterschieden. DRS wird als solche erkennbar, wenn die sprachlichen Äußerungen die persönliche, räumliche und zeitliche Deixis aus Sicht der ursprünglichen SprecherIn wiedergegeben, um sowohl die Form als auch den Inhalt der originalen Äußerung zu rekonstruieren (vgl. Holt, 1996, S. 222; Holt & Clift, 2007, S. 5). Im Unterschied dazu werden in der IRS die wiedergegebenen Äußerungen zusammengefasst, ohne dass beansprucht wird, die ursprünglichen Äußerungen in den Details ihrer Realisierung zu rekonstruieren. Deiktische Referenzen, Pronomen, Anredeformen und die Zeitformen von Verben werden an den Kontext der aktuellen Gesprächssituation angepasst. Durch die kurzen, detailarmen Zusammenfassungen sprachlicher Handlungen wird durch IRS der Fokus auf die zu realisierende Handlung und nicht die Realisierungsform der vergangenen Äußerung gelegt (vgl. Holt, 2017, S. 185).

Konversationsanalytische Arbeiten rekonstruieren neben diesen lexiko-semantischen Merkmalen von IRS und DRS auch sprachliche Praktiken, die der Vorbereitung der Redewiedergabe dienen. Mithilfe von Zitationsmarkern, die meist in turn-initialer Position platziert werden, machen die SprecherInnen *accountable*, wessen Rede in einer Äußerung wiedergegeben wird. Die verbreitetsten Formen der Zitationsmarker bestehen aus der Kombination von Pronomen und *verbum dicendi* (bspw. „Sie sagte...“) (vgl. Bergmann, 1998, S. 117; Holt, 2017, S. 173). Es existieren jedoch auch alternative Formen wie „be like“ (Blyth et al., 1990) im Englischen oder „und ich/er/sie so“ (Golato, 2000) im Deutschen, die ohne Einsatz eines *verbum dicendi* anzeigen, wessen Rede in der folgenden Äußerungen wiedergegeben wird.

Darüber hinaus spielt auch der Einsatz von Prosodie und Stimmqualität eine konstitutive Rolle, die Redewiedergabe als solche kenntlich zu machen (Couper-Kuhlen, 1999; Günthner, 1997, 1999b, 2012b; Niemelä, 2011). So konnte bspw. gezeigt werden, dass „[d]ifferent prosodic and voice quality features are a central means for reporters to signal not only where reported speech starts and ends but also whose voice is being quoted“ (Günthner, 1999b, S. 687; s. a. Couper-Kuhlen, 1999, S. 12). Die prosodischen Ressourcen werden bei der Rekonstruktion vergangener Äußerungen methodisch dazu eingesetzt, sowohl *accountable* zu machen, wann die Redewiedergabe beginnt und endet, als auch um kenntlich zu machen, wessen Äußerungen wiedergegeben werden. Die SprecherInnen setzten Modulationen der Stimmqualität außerdem ein, um die Redewiedergabe mit ihren Evaluationen zu durchsetzen und die RezipientInnen zu einem angemessenen Verständnis der rekonstruierten Äußerungen im jeweiligen situativen Kontext der Interaktion zu bewegen (vgl. Günthner, 1997, S. 257).

Spätestens wenn berücksichtigt wird, dass prosodische Ressourcen von SprecherInnen dazu eingesetzt werden, ihre Haltung zu den wiedergegebenen Äußerungen zu kommunizieren, wird ersichtlich, dass insbesondere DRS einen Doppelcharakter zwischen „Authentisierung und Fiktionalisierung“ (Bergmann, 1998, S. 117) besitzt. Einerseits wird durch die Konstruktion als wörtliche Zitation ein Anspruch auf Authentizität der reproduzierten Äußerung erhoben (Bergmann, 1998, S. 118). Andererseits findet simultan eine Stilisierung der ‚originalen‘ Äußerung bezüglich der aktuellen kommunikativen Ziele der Beteiligten statt (vgl. Günthner, 1999b, S. 693). Diese hat oftmals das Ziel, „ein ‚bis zur Kenntlichkeit verzerrtes‘ Portrait des Zitierten zu entwerfen und so dessen ‚eigentlichen‘ Charakter zum Vorschein zu bringen“ (Bergmann, 1998, S. 121)<sup>58</sup>. Der Bedeutungsgehalt wörtlicher Zitate in alltäglichen

---

<sup>58</sup> Aus diesem Grund verzichten einige Autoren – bspw. Goffman (1981) und Tannen (1989) – vollkommen auf die Bezeichnung ‚Wiedergabe‘ und verwenden stattdessen ausschließlich den Begriff ‚animieren‘, „um zu verdeutlichen, dass es sich im Mündlichen immer um Stilisierungen von Redewiedergaben handelt, deren Authentizität kaum ausmachbar ist“ (Kotthoff, 2018, S. 237, FN 11). Ich werde jedoch sowohl von „Animation“ als auch

Unterhaltungen besteht also nicht einfach in einer neutralen, versachlichten Information, sondern er ergibt sich im spezifischen sequenziellen Kontext der jeweiligen Aktivität und ist durchwirkt von den kommunikativen Absichten der SprecherIn.

Das für IRS und DRS konstitutive Zusammenspiel sprachlicher Praktiken und prosodischer Ressourcen wird hauptsächlich im Kontext narrativer Aktivitäten untersucht<sup>59</sup>, wobei der Fokus auf der Komposition und sequenziellen Platzierung von Praktiken der Redewiedergabe im Prozess der interaktiven Hervorbringung von Geschichtenerzählungen in (überwiegend alltäglichen) Gesprächen liegt (vgl. Holt & Clift, 2007, S. 2). Im Kontext von Erzählungen ermöglichen es die Praktiken der Redewiedergabe, vergangene Interaktionen in ihrem zeitlichen und sequenziellen Verlauf sowie in den Details ihrer Realisierung zu rekonstruieren. Der Wechsel zwischen direkten und indirekten Formen der Redewiedergabe kann u. a. dazu beitragen, den Hintergrund einer Erzählung zunächst durch IRS zu entwickeln und so die Klimax der narrativen Rekonstruktion vorzubereiten, die im Modus der DRS als solche *accountable* gemacht wird (bspw. Holt, 2017, S. 174; Mayes, 1990).

Insbesondere die DRS wird von den ErzählerInnen für eine implizite Evaluation der berichteten Ereignisse eingesetzt und ermöglicht es den RezipientInnen so, ihre eigenen Schlüsse über den Charakter der rekonstruierten Ereignisse zu ziehen. Dass es die Praktiken der Redewiedergabe den RezipientInnen (vermeintlich) erlauben, zu einer eigenständigen Interpretation der rekonstruierten Äußerungen zu gelangen, vermag die wiederkehrende Verbindung dieser Praktiken mit spezifischen Aktivitätstypen erklären (Holt & Clift, 2007, S. 12–13): Zum einen wurde das Phänomen wiederholt im Kontext von Beschwerdeerzählungen über abwesende Dritte untersucht (bspw. Drew, 1998; Günthner, 1999a; Haakana, 2007; Holt, 2000; Selting, 2012). Beschwerden stellen selbst sensible und potenziell rechtfertigungsbedürftige Handlungen dar (vgl. Maynard, 2013, S. 219 ff.; Schegloff, 2005, S. 466): Durch den Einsatz von Praktiken der Redewiedergabe erhalten die RezipientInnen Zugang zu dem beschwerdewürdigen Verhalten und ihnen wird – unter Zuhilfenahme impliziter prosodischer Evaluationen – auf diese Weise die Produktion eigener (affiliativer) Bewertungen ermöglicht. Zum anderen finden sich Formen der IRS und DRS auch oftmals in humoristischen Erzählungen (bspw. Holt, 2000, 2007; Selting, 2017) oder bei der Erzählung von Witzen in Geschichtenform (bspw. Sacks, 1974, 1978;

---

von „Redewiedergabe“ sprechen. Diese Entscheidung begründet sich darin, dass aus konversationsanalytischer Perspektive bei rekonstruktiven Konservierungen von flüchtigen Ereignissen (wie Geschichtenerzählungen) grundsätzlich nicht von einem exakten Entsprechungsverhältnis ausgegangen werden kann. Für jedes vergangene Ereignis existiert eine theoretisch unendliche Anzahl rekonstruierender Konservierungen, deren konkrete Realisierungsformen unauflöslich mit den Gestaltungskonventionen der jeweiligen Gattung sowie dem lokalen Kontext der Interaktion und dem spezifischen Rezipienten-Design verwoben sind (vgl. dazu Bergmann, 1985, S. 306 f.).

<sup>59</sup> Siehe jedoch bspw. Clift (2007), Couper-Kuhlen (2007) oder Günthner (1999b, S. 701–704) für Analysen der DRS in nicht-narrativen Aktivitäten.

Norrick, 2006). Im Kontext dieser Aktivitäten werden tatsächliche oder fiktive Dialoge durch den Einsatz von Formen der Redewiedergabe rekonstruiert. Die Verwendung von DRS kann dazu beitragen, die Involvierung der RezipientInnen zu steigern (Tannen, 1989), aber sie dient – vergleichbar zu ihrer Funktion in Beschwerdeerzählungen – vor allem dazu, implizit die Evaluation des humoristischen Charakters der wiedergegebenen Äußerung zu transportieren (Holt, 2000, S. 451): Auf diese Weise wird es den RezipientInnen ermöglicht, die humoristische Bedeutung bspw. des Höhepunktes eines Witzes selbst zu interpretieren und eine angemessene, amüsierte Erwiderung zu produzieren (Sacks, 1978, S. 258–259).

Diese Übersicht über die konversationsanalytische Untersuchung verschiedener Formen der Redewiedergabe zeigt zunächst, dass es sich um einen gut erforschten Phänomenbereich handelt. Praktiken der IRS und DRS stellen verbreitete Verfahrensweisen der Hervorbringung von alltäglichen Geschichtenerzählungen dar und tragen dazu bei, unterschiedliche Figuren der Erzählung zu animieren, den sequenziellen und zeitlichen Verlauf der rekonstruierten Ereignisse für die RezipientInnen *accountable* zu machen und diesen durch implizite Evaluationen angemessene Erwiderungen zu ermöglichen. Wörtliche Zitate sind dabei durch einen Doppelcharakter zwischen Authentisierung und Fiktionalisierung gekennzeichnet: Sie erheben durch ihre Form den Anspruch einer exakten Wiedergabe der Realisierungsform der vergangenen Äußerung(en), gleichzeitig bewertet und kommentiert die SprecherIn durch prosodische Markierungen, Stilisierungen und Übertreibungen die Rede des Anderen, um die eigenen kommunikativen Ziele zu erreichen.

Im Gegensatz zu Praktiken der rekonstruierenden Wiedergabe vergangener sprachlicher Äußerungen existieren nur wenige konversationsanalytische Studien, in denen die Wiedergabe von schriftlicher Mobilkommunikation in F2F-Interaktionen untersucht wird, und keine Publikation, die diese Praktiken explizit als Elemente alltäglicher Geschichtenerzählungen fokussiert. Im Folgenden werden die Ergebnisse bisheriger Untersuchungen dieses Phänomens vorgestellt und das Desiderat identifiziert, das mit der Untersuchung der *Wiedergabe von Textnachrichten in Erzählungen* in diesem Kapitel bearbeitet wird.

Wie bereits im allgemeinen Forschungsstand diskutiert, untersuchen Raclaw et al. (2016), wie GesprächsteilnehmerInnen *mobile-supported sharing activities* einsetzen, um ihren InteraktionspartnerInnen – durch das Vorlesen von Textnachrichten oder das Zeigen von Fotografien – die Kopartizipation in der Bewertung eines Referenten zu ermöglichen. Im analytischen Fokus der AutorInnen steht also die Frage danach, wie der durch die *Affordanzen* des Smartphones ermöglichte, epistemische Zugang zu einem Referenten als Ressource eingesetzt wird, andere Anwesende zur Alignierung und Affiliation zu leiten (Raclaw et al., 2016, S. 370). Dieses

funktionale Potenzial von *mobile-supported sharing activities* wird anhand der unterschiedlichen Organisationsprinzipien von präferierten und dispräferierten Aktivitäten herausgearbeitet. Bei zwei der für die Publikation verwendeten Einzelfälle handelt es sich dabei um *smartphone-assistierte Geschichtenerzählungen* (eine humoristische und eine Beschwerdeerzählung, vgl. Raclaw et al., 2016, S. 368–370; 371–373), in denen im Hauptteil der Erzählung durch die wörtliche Zitation archivierter Textnachrichten der sequenzielle Verlauf vergangener Ereignisse rekonstruiert wird (Raclaw et al., 2016, S. 367). In Abhängigkeit davon, ob mit der Erzählung eine präferierte (Witz) oder dispräferierte (Beschwerde) Handlung realisiert werden soll, wird von den SmartphoneinhaberInnen methodisch die Prä-Sequenz der Erzählung unterschiedlich ausgestaltet (entweder mit oder ohne explizite erste Bewertung). Insbesondere im Fall der Beschwerdeerzählung zeigt sich, dass die wörtliche Zitation der Textnachrichten erst in Reaktion auf eine disaffilierende Reaktion der RezipientIn initiiert wird: Das Vorlesen wird als „interactional evidential“ (vgl. Clift, 2006, S. 585) für die Beschwerdewürdigkeit des Verhaltens der Abwesenden eingesetzt, um die RezipientInnen zur Affiliation zu leiten.

Zu ähnlichen Ergebnissen kommen Brown et al. (2018), die „text messaging in the context of co-present face-to-face-interaction“ (ebd. 2018, S. 2) untersuchen. Die AutorInnen zeigen u. a. auf, wie Textnachrichten als thematische Ressourcen für Alltagsgespräche dienen (ebd., 2018, S. 7–9), durch deren kommunikative Aneignung Sozialbeziehungen zwischen den Anwesenden lokal hervorgebracht und verhandelt werden (ebd., 2018, S. 9–13; s. a. Keppler, 2019, S. 186–187). Im Ergebnis – und in Analogie zu den Untersuchungen der DRS – stellen sie heraus, dass die SmartphoneinhaberInnen die Wiedergabe der Textnachrichten durch den Einsatz spezifischer Prä-Sequenzen, sprachlicher und prosodischer Ressourcen sowie den Wechsel zwischen IRS und DRS auf eine Weise gestalten „as to convey, both explicitly and implicitly, the reporter’s stance towards the person and content of the text-based encounter“ (Brown et al., 2018, S. 20; s. a. Raclaw et al., 2016, S. 369, 373).

Im Vergleich zu Raclaw et al. (2016) zeichnen Brown et al. (2018) ein detaillierteres Bild der sprachlichen Praktiken der Wiedergabe von Textnachrichten, die konstitutiv zur Hervorbringung von *smartphone-assistierten Erzählungen* beitragen (können). Das spezifische Forschungsdesign – die synchronisierte Aufzeichnung der sprachlichen Interaktionen und der Bildschirmaktivitäten (Brown et al., 2018, S. 4–5) – erlaubt dabei eine genaue Analyse des Verhältnisses zwischen archivierten Textnachrichten und den Formen der sprachlichen Bezugnahmen auf diese (Zusammenfassungen, wörtliche Zitation oder IRS). Auf dieser Datengrundlage lassen sich jedoch keine Aussagen über die verkörperten Handlungen treffen, die für die Integration des Smartphones zum Zweck der Wiedergabe von Textnachrichten relevant sind – Fragen

zur Blickführung, zum Timing des visuellen Zugriffes auf das Smartphonedisplay oder evtl. Zeigegesten müssen unbeantwortet bleiben. Beide Studien produzieren erste Erkenntnisse über die Praktiken der Wiedergabe von Textnachrichten in F2F-Interaktionen und liefern schlaglichtartige Betrachtungen exemplarischer Fälle dieser Praktiken im Kontext alltäglicher Geschichtenerzählungen. Eine systematische Analyse der Hervorbringung dieser multimodalen Aktivität steht jedoch noch aus. Die folgenden Analysen bauen auf den vorhandenen Erkenntnissen auf und unternehmen eine multimodale Analyse der Varianz und Systematik des Einsatzes von Praktiken des Vorlesens von Textnachrichten für die Hervorbringung von *smartphone-assistierten Geschichtenerzählungen* in alltäglichen Gesprächen.

## **5.1 Das Vorlesen von Textnachrichten in methodisch vorbereiteten Geschichtenerzählungen**

Das Ziel des gesamten Kap. 5 ist es, die Systematik und Varianz von *smartphone-assistierten Geschichtenerzählungen* zu rekonstruieren, die durch die *Wiedergabe von Textnachrichten* gestützt werden. Dies soll Schritt für Schritt anhand der Analyse auf unterschiedliche Weisen veranlasster Erzählungen erfolgen. Beginnen möchte ich den analytischen Teil damit, *smartphone-assistierte Erzählungen* zu diskutieren, die von der aktuellen SprecherIn methodisch auf eine Weise in die sprachliche Interaktion eingeführt werden, dass das Verhältnis der projizierten Erzählung zum vorangehenden Gespräch angezeigt und die Angemessenheit der Produktion der Erzählung *accountable* gemacht wird (vgl. Jefferson, 1978, S. 220). In der Analyse der sequenziellen Organisation dieser *smartphone-assistierten Geschichtenerzählungen* wird systematisch berücksichtigt, dass neben „verbalen Praktiken auch kinesische, räumliche oder materielle Ressourcen von den Erzählenden und Zuhörenden genutzt werden“ (König & Oloff, 2018b, S. 213–214), um diese als sinnhaft verständliche, multimodale Aktivität hervorzubringen. Entsprechend werde ich – insbesondere in diesen ersten analytischen Kapiteln – folgende forschungsleitende Fragestellungen mitführen, um die Spezifika des Gebrauchs von Smartphones im Kontext alltäglicher Geschichtenerzählungen rekonstruieren zu können:

- Auf welche Weise und an welchen sequenziellen Positionen wird die Integration des Smartphonegebrauchs in die F2F-Interaktion vorbereitet und realisiert?
- Welche Rolle spielen verkörperte Handlungen und die Handhabung des mobilen Endgeräts für die beobachtbare und berichtbare Hervorbringung und sequenzielle Organisation der Geschichtenerzählungen?

- Welche Teilnehmerahmen („participation framework“, C. Goodwin & M. H. Goodwin, 1992, C. Goodwin, 2007a) sind konstitutiv für die Produktion *smartphone-assistierter Geschichtenerzählungen*?
- Mittels welcher sprachlichen und verkörperten Praktiken wird das Vorlesen von Textnachrichten so gestaltet, dass sie erkennbar als konstitutive Bestandteile einer Geschichtenerzählung fungieren?

### 5.1.1 Das Vorlesen von Textnachrichten als konstitutives Element *smartphone-assistierter Geschichtenerzählungen*

Ausgehend von diesen Fragestellungen möchte ich meine Analyse der Praktiken des Vorlesens von Textnachrichten in methodisch vorbereiteten, *smartphone-assistierten Geschichtenerzählungen* anhand einer Interaktionsepisode beginnen, in der eine unterhaltsamen Geschichte erzählt wird. Der Ausschnitt stammt aus dem Beginn eines Gesprächs zwischen Julia (J) und Bianca (B), die sich gemeinsam mit zwei gerade abwesenden Freundinnen in einem Café getroffen haben, und in einer „vis-a-vis-Formation“ (Kendon, 1990, S. 213) zueinander sitzen. Unmittelbar vor Beginn der Videoaufzeichnung berichtet J amüsiert über die eigenartigen Gewohnheiten ihrer Schwester, über WhatsApp mit ihr zu kommunizieren. Das folgende Transkript setzt ein, kurz bevor J eine humoristische Geschichtenerzählung initiiert, mit der sie dieses typische Kommunikationsverhalten durch die narrative Rekonstruktion eines einzelnen Ereignisses illustriert (vgl. Schwitalla, 1991, S. 189). Die Blicke von J werden mit „&“ und ihre Gesten durch „\$“ annotiert. Gesten und Blicke von B werden durch „\*“ im Transkript kenntlich gemacht.

#### Transkript 3: 181004 „Eine Tüte Popcorn“\_Teil 1



```

j: >>--Blick nach vorne, in Raum-->>
001 J: (ich) weiß AU_nich;
002 da #kommt auch &scht (.)da $kommt #STÄndig
irgendwie sol[che sAchen;]
----->&.....----Blick SP-->020

```

SP-->013

#1

#2



003 B: [ ehehe ]he\*he

\*beugt sich nach

vorne, Blick zu J's SP-->005

004 (2.0)#(1.5)

#3



005 B: ich KOMM\* net kla;

--,,-->\*setzt sich gerade hin, Blick in Raum-->>012

[006-011: Interaktion mit Aufnahmeleitung entfernt]

Das Transkript beginnt in Z. 001 mit „Formulierungsarbeit“ (Gülich, 1994, S. 79) von J, die durch „ich weiß AU\_nich“ anzeigt, Schwierigkeiten mit der Charakterisierung des Kommunikationsverhaltens der abwesenden Schwester zu haben. In Z. 002 setzt J zu einer längeren indexikalischen Beschreibung an, die durch einen recycelten Turn-Beginn (vgl. Schegloff, 1987b) gekennzeichnet ist: Sie bricht die Produktion ihrer Äußerung („da kommt auch scht“) zunächst ab, schaut zu ihrem auf dem Tisch liegenden Smartphone, um dann nach einer Mikropause die Äußerungsproduktion („da kommt STÄNdig irgendwie solche sAchen“) fortzusetzen. Währenddessen beginnt sie, ihren Blick auf das Smartphone fixiert, dieses mit ihrer rechten Hand zu bedienen (#2). Durch die Verwendung deiktischer Formulierungen („da“, „solche Sachen“) wird zum einen erkennbar, dass sich die sprachliche Äußerung auf ein bestehendes Thema des vorangehenden Gesprächs bezieht und die vorangegangenen Äußerungen den Bezugsrahmen für ihre Interpretation bilden (Gülich & Mondada, 2008, S. 49). Zum anderen

fungiert die Äußerung in Z. 002 als „prospective indexical“ (C. Goodwin, 1996, S. 384)<sup>60</sup>: Die indexikale Charakterisierung der Ereignisse schafft für die Rezipientin einen interpretativen Rahmen, in dem es zu ihrer Aufgabe wird, im weiteren Interaktionsverlauf herauszufinden, welche Ereignisse von der Sprecherin mit dieser Formulierung bezeichnet werden und eine angemessene Erwiderung darauf zu produzieren (vgl. Sacks, 1974, S. 341). Indem die Smartphoneinhaberin den Abbruch und Neustart ihrer Äußerungsproduktion in Z. 002 mit der visuellen Fokussierung des Smartphones (und dessen anschließender Bedienung) synchronisiert, projiziert sie ihren selbst-initiierten Smartphonegebrauch (Keppler, 2019, S. 182) als etwas, das ebenfalls mit dem aktuellen Gesprächsthema verbunden ist – allerdings ohne die Hinwendung zum mobilen Endgerät sprachlich anzukündigen oder den Zweck des Gebrauchs zu explizieren. Die Gesprächspartnerin J reagiert auf die Beschreibung des Kommunikationsverhaltens in Z. 003 zunächst mit in Überlappung produziertem Lachen und markiert diese so als Lachanlass. Während des Lachens beugt sie sich nach vorne über den Tisch und schaut auf das Smartphone. Auch in der anschließend entstehenden Stille von ca. dreieinhalb Sekunden (Z. 004), in der B das Smartphone durch wiederholte Bediengesten („Scrollen“) physisch manipuliert, fokussiert J das Smartphone und scheint so die Geschehnisse auf dem Display aufmerksam zu beobachten (#3). Während der Realisierung ihrer Äußerung in Z. 005 („ich komm net kla;“), mit der sie das Kommunikationsverhalten der Schwester amüsiert bewertet, kehrt sie wieder in ihre ursprüngliche Sitzposition zurück und schaut in den Raum. Auf diese Weise zeigt sie erkennbar die Beendigung der Beobachtung der Smartphoneaktivitäten an.

Dass die Initiierung des Smartphonegebrauchs thematisch mit dem andauernden *turn-by-turn-talk* der TeilnehmerInnen verbunden ist, wird hier von der Smartphoneinhaberin durch die Produktion einer *multimodalen Gestalt* (Mondada, 2014b, S. 140) projiziert, in der sich sprachliche Äußerung, Veränderung der Blickrichtung und physische Handhabung des mobilen Endgeräts wechselseitig ihren Sinn verleihen. Dabei kann insbesondere der Neustart sprachlicher Äußerungen als methodisches Verfahren identifiziert werden, mit dem die unterschiedlichen Zeitlichkeiten von sprachlicher Äußerung und physischer Manipulation des Smartphones auf eine Weise miteinander synchronisiert werden, die diese Verbundenheit *accountable* machen kann. Mit dem Gebrauch des Smartphones im Kontext von F2F-Interaktionen geht jedoch oftmals das potenzielle Problem der *bystander inaccessibility* einher: Aufgrund der Größe des Smartphonedisplays sowie der Steuerung des Geräts über (größtenteils) stumme Bediengesten ist für

---

<sup>60</sup> Durch „prospective indexicals“ werden indexikalisch bezeichnete Referenten oder ‚Probleme‘ von einer SprecherIn als etwas konstituiert, „that has to be discovered subsequently as the interaction proceeds. Recipients are set the task of attending to subsequent events in order to find what particulars constitute ‚the problem‘ on this specific occasion“ (C. Goodwin, 2015a, S. 384).

andere Anwesende oftmals nicht ohne Weiteres erkennbar, welchem Zweck die Bedienung des mobilen Endgeräts dient (Mantere et al., 2018, S. 106). In konversationsanalytischen Studien wurden bereits einige sprachliche Methoden identifiziert, mit denen der Smartphonegebrauch in F2F-Interaktionen *accountable* gemacht wird (bspw. Brown et al., 2015, S. 512–513; Suderland, 2019). In vorliegendem Fall wird dieses Problem jedoch über verkörperte Praktiken der Teilnehmerinnen bearbeitet: Die Rezipientin verschafft sich selbst-initiiert visuellen Zugriff zum Smartphone und damit Informationen über die konkrete Nutzung des Geräts. Dieser Zugriff wird von der Smartphoneinhaberin gewährt, ohne dass eine sprachliche Aushandlung stattfindet. Die stumme, verkörperte Hervorbringung dieses neuen Teilnehmerrahmens ermöglicht der Rezipientin so das visuelle *Monitoring* des Displays, das als Methode fungiert, sich den Zweck des Smartphonegebrauchs beobachtend zu erschließen<sup>61</sup>.

Während J mit der Bedienung des mobilen Endgeräts (vermutlich zum Zweck der Suche nach einem spezifischen Medieninhalt) beschäftigt ist, kommt es zu einer kurzen Interaktion zwischen B und der Aufnahmeleitung, die aus dem Transkript entfernt wurde. Anschließend führen die zwei Freundinnen die Interaktion fort und es folgt der Hauptteil einer *smartphone-assistierten Geschichtenerzählung*, der durch das Vorlesen von Textnachrichten gekennzeichnet ist:

#### Transkript 4: 181004 „Eine Tüte Popcorn“\_ Teil 2

- 012 B: \*<<lachend> ehh h°>  
\*Blick zu J--->>
- 013 J: vor \$allen=dingen (-) SONNtag; (-)  
--->\$
- 014 <<leiernd> ich habe mir eine kleine tüte  
POP[corn geholt; >]
- 015 B: [a ha ha ha ]
- 016 J: [<<leiernd> und den rest habe ich mit einem  
Kino] [gutschein bezahlt. >]
- 017 B: [((lacht nicht hörbar))  
] [<<p> ha ha ha >] <<lachend> °hhh >
- 018 J: [<<leiernd># aber diese ] [kleine tüte hat  
&schon VIER> <<:-)> euro] [siebzig> # ]  
<<lachend> ge[kostet; > he he ]  
&...Blick B-->021
- 019 B: [ha <<lachend> (was,) ] [((nicht hörbares  
lachen) ) ] [<<p> ha ha >]  
[ ha ha ha > ]

<sup>61</sup> Wir werden dieser Praktik im weiteren Verlauf der Analysen immer wieder begegnen. Bereits anhand dieses Einzelfalls kann festgehalten werden, dass das selbst-initiierte *Monitoring* des Smartphonegebrauchs von den RezipientInnen im Unterschied zu Keppler (vgl. 2019, S. 180) und in Ergänzung zu Oloff (vgl. 2019a, S. 215) nicht unbedingt sprachlich angekündigt oder durch die SmartphoneinhaberInnen sprachlich ratifiziert werden muss.



020 J: <<lachend>und ich hab nur so DIE:\$sen;> #(-)\$  
\$Fragegeste\$

#3



021 B: hi hi TAgebuch &der:-  
j: -----,,,,,,,,-->&Blick SP-->023  
022 wie heißt deine SCHWESter,  
023 J: <<:->an&NE frank.>  
--,,,-->&Blick B-->024  
024 ((4.2 beide lachen))&((1.8 beide lachen))  
j: -----&Blick SP-->  
[response sequence]

Nach Beendigung der Interaktion zwischen B und der Aufnahmeleitung, während der J weiter mit der Bedienung des Smartphones beschäftigt war, schaut B in Z. 012 zur Smartphoneinhaberin und produziert einige Lachpartikel. Anschließend übernimmt J das Rederecht und projiziert in Z. 013 mit „vor allen=dingen (-) SONNtag (-)“ sprachlich die Verfügbarkeit einer Erzählung über ein kürzlich vergangenes Ereignis. Die Äußerung stellt retrospektiv Bezug zu ihrer einleitenden Formulierung „da kommt STÄNDIG irgendwie solche sAchen;“ aus Z. 002 her und macht für die Rezipientin erwartbar, dass im Folgenden eine Exemplifizierung stattfindet. Mit Beginn der Äußerung beendet J die physikalische Manipulation des Smartphones. Sowohl vor als auch nach der konkreten zeitlichen Verortung des Ereignisses entstehen jeweils kurze Pausen, an denen eine Übernahme des Rederechts für die Gesprächspartnerin möglich wäre. Der Verzicht der Rezipientin darauf wird von der Smartphoneinhaberin als *ticket* behandelt, mit dem sie nun die Möglichkeit zur Produktion des *big package* der Erzählung erhält und

die für den *turn-by-turn-talk* kennzeichnenden Organisationsprinzipien des Sprecherwechsels vorübergehend außer Kraft gesetzt werden (vgl. Sacks, 1974, 1978).

In Z. 014–020 produziert J nun – ihren Blick zunächst durchgehend auf das Display gerichtet – den Hauptteil der Geschichtenerzählung: Dieser besteht aus mehreren Äußerungen im Modus der DRS, mit denen die Nachrichten der abwesenden Schwester wiedergegeben werden: Diese berichtet von einem Kinobesuch und dem Kauf einer „kleinen Tüte Popcorn“. Auch wenn von der Sprecherin keine expliziten Zitationsmarker, bestehend aus Personalpronomen und Kommunikationsverb, verwendet werden (Holt, 2017, S. 173), wird durch den kombinierten Einsatz sprachlicher und prosodischer Ressourcen *accountable* gemacht, dass es sich um die Wiedergabe der Worte einer Anderen handelt: Zum einen werden in der Wiedergabe u. a. die Pronomen und Verbformen der originalen Nachrichten erhalten (vgl. Holt, 1996, S. 222); zum anderen bedient sich die Smartphoneinhaberin einer prosodischen Stilisierung – einer ‚leiernden‘ Betonung der einzelnen Silben, die sich deutlich von anderen Äußerungen der Sprecherin unterscheidet (Günthner, 1999b, S. 687). Diese wird von der Erzählerin darüber hinaus eingesetzt, um implizit ihre Haltung zu den wiedergegebenen Nachrichten anzuzeigen<sup>62</sup> und die Rezipientin so zu einer angemessenen Erwiderung zu leiten (vgl. Günthner, 1997, S. 257). Bereits gegen Ende der direkten Wiedergabe der ersten Nachricht (Z. 014: „popcorn geho:lt;“) beginnt die Rezipientin in Überlappung zu lachen (Z. 015) und orientiert sich auch im weiteren Verlauf beobachtbar durch ihr anhaltendes, teilweise lautloses Lachen (Z. 017) am Unterhaltungswert der humoristischen Erzählung.

Der Höhepunkt der Erzählung wird durch die Wiedergabe einer dritten Nachricht in Z. 018 konstituiert, in der die Schwester über den hohen Preis einer „kleinen Tüte“ Popcorn klagt. Die Wiedergabe dieser Nachricht wird erkennbar als Klimax produziert, indem die Erzählerin im Verlauf der Äußerungsproduktion die ‚leiernde‘ Intonation zunächst zu Gunsten einer Realisierung in „smile voice“ (Selting et al., 2009, S. 367) und anschließenden lachenden Aussprache aufgibt. Darüber hinaus löst sie vor der Nennung des Preises („schon VIER euro“) ihren Blick zum ersten Mal vom Display des Smartphones und schaut direkt zu B (#5). Die Wiedergabe dieser Nachricht wird von einem in Überlappung produzierten, lachend realisierten „was,“ (Z. 019) sowie durchgehendem Lachen der Rezipientin begleitet. Dass B die Äußerung als Höhepunkt der Erzählung erkennt, zeigt sich insbesondere daran, dass ihr Lachen zunehmend an Lautstärke zunimmt, bis sie in Überlappung zum projizierbaren Ende der Äußerung in lautes Gelächter ausbricht.

---

<sup>62</sup> Die ‚leiernde‘ Intonation scheint den Sachverhalt, dass die Erzählerin „ständig irgendwie solche sAchen;“ (Z. 02) von ihrer Schwester empfängt, performativ aufzugreifen.

Die Ausleitung der narrativen Rekonstruktion erfolgt durch die Erzählerin anschließend in Z. 020: In dieser lachend realisierten Äußerung rekonstruiert sie ihre Erwiderung auf die Nachrichten der Schwester. Dass es sich um die Animierung einer sinnhaft verbundenen Erwiderung aufgrund der erhaltenen Nachrichten handelt, wird durch das turn-initiale „und“ erkennbar (Deppermann, 2008, S. 65); die weitere Einleitung durch die Kombination aus Pronomen („ich“) und alternativem Zitationsmarker („hab nur so“) (vgl. Golato, 2000) macht einen Sprecherwechsel in der rekonstruierten Interaktion kenntlich. Ihre Antwort wird jedoch nicht sprachlich wiedergegeben, sondern durch eine multimodale Enaktierung (vgl. Sidnell, 2006) demonstriert: Simultan zur Realisierung des deiktischen „DIE:sen“ öffnet sie beide Arme mit den Handfläche nach oben in einer fragenden Geste (#5)<sup>63</sup>. Im Kontrast zu den ausführlichen Nachrichten der abwesenden Schwester und durch die Verwendung von „nur“ in der Einleitung der eigenen Erwiderung macht die Erzählerin abschließend noch einmal ihre eigene Haltung zur rekonstruierten Interaktionsepisode deutlich: Diese wird als Illustration einer typischen Kommunikationspraxis der Schwester präsentiert (vgl. Schwitalla, 1991, S. 189) und durch die Art der narrativen Rekonstruktion als humoristisches und befremdliches Ereignis gerahmt.

Im Anschluss an die beobachtbare und berichtbare Schließung der Erzählung findet ab Z. 021 eine Rückkehr zum *turn-by-turn-talk* statt. In der damit eröffneten *response sequence* formuliert die Rezipientin zunächst einen Witz, der den Namen der Schwester als Aufhänger nutzt, um einen Vergleich zwischen dem „Tagebuch der Anne Frank“ (Z. 021–23) und dem Kommunikationsstil der Schwester herzustellen. Auf diese Weise zeigt B zum einen ihr Verständnis der narrativen Rekonstruktion an, zum anderen setzt sie dadurch den humoristischen Charakter der Interaktion fort. Die Post-Sequenz wird durch etliche weitere Äußerungen (nicht im Transkript enthalten) fortgesetzt, in denen sich die TeilnehmerInnen über die Erzählung amüsieren und den Kommunikationsstil der abwesenden Schwester als abweichend interpretieren. Während J zu Beginn der *response sequence* noch wiederholt auf das Gerätedisplay schaut (Z. 021, 024), spielt das mobile Endgerät im weiteren Verlauf keine Rolle mehr.

Anhand der Analyse der methodischen Hervorbringung dieser Elaboration einer durch Praktiken des Vorlesens von Textnachrichten gestützten, *smartphone-assistierten Geschichtenerzählung* lassen sich bereits einige typische Merkmale dieser multimodalen Praktik identifizieren. Die Wiedergabe von Textnachrichten erfolgt hier durch die sprachliche Verfahren der DRS: Pronomen, Deixis und Verbformen der originalen Nachrichten werden erhalten, um die Form

---

<sup>63</sup> Wie sich an einem späteren Punkt des Gesprächs zeigt (nicht im Transkript enthalten), handelt es sich um die verkörperte Enaktierung eines GIFs, das die Erzählerin als Antwort auf die Nachrichten ihrer Schwester verschickt hat.

und den Inhalt der Kommunikationsbeiträge einer Anderen wiederzugeben (vgl. Holt & Clift, 2007, S. 5). Die DRS ist auch in dieser *smartphone-assistierten* Erzählung von einem Doppelcharakter zwischen Authentisierung und Fiktionalisierung gekennzeichnet: Einerseits wird durch die wörtliche Zitation ein Anspruch auf Authentizität der reproduzierten Nachrichten erhoben (vgl. Bergmann, 1998, S. 118), andererseits findet simultan eine prosodische Stilisierung dieser lediglich schriftsprachlich vorliegenden Nachrichten statt, mit denen die Erzählerin ihre Haltung anzeigt (vgl. Günthner, 1997, S. 257).

Bei der hier analysierten wörtlichen Wiedergabe von Textnachrichten handelt es sich um eine multimodale Praktik, die durch die Einnahme einer für diese Aktivität spezifischen Körperhaltung (vgl. C. Goodwin, 1984, S. 228) realisiert wird: Indem die Erzählerin während der Wiedergabe der Textnachrichten durchgehend auf das Display des Smartphones schaut, wird von ihr angezeigt, dass es sich nicht nur um eine wortgetreue Rekonstruktion, sondern um das Vorlesen des Wortlauts der Nachrichten handelt. Im Zusammenspiel mit der Handhabung des Gerätes und der visuellen Fokussierung des Smartphonedisplays wird durch die wörtliche Zitation nicht nur eine Existenz der Nachrichten unabhängig vom lokalen Kontext behauptet. Vielmehr wird die Authentizität des Wortlauts in der Materialität der auf dem Bildschirm angezeigten Textnachrichten verankert, wobei die Präsenz des Smartphones in der Interaktion diese – zumindest theoretisch – überprüfbar macht.

Auf diese Weise wird auch der für diese Art von Geschichtenerzählungen spezifische Teilnehmerahmen (C. Goodwin, 2007a) hervorgebracht: Die ZuhörerIn zeigt ihre Rezipientenschaft dadurch an, dass sie ihren Blick auf die SprecherIn gerichtet hält (vgl. C. Goodwin, 1984, S. 230) und demonstriert ihr Verständnis der im Vollzug befindlichen Erzählungen durch (para)sprachliche Erwidern, die die ErzählerIn jedoch nicht an deren Fortsetzung hindern. Im Unterschied zu nicht-mediatisierten Geschichtenerzählungen adressiert die ErzählerIn die ZuhörerIn jedoch nicht selbst mit ihrem Blick – stattdessen nimmt sie die oben beschriebene, spezifische Körperhaltung ein und schaut (in diesem Fall schon seit Beginn der Prä-Sequenz in Z. 002) auf das Display des Smartphones. Diese Konstellation ist konstitutiv für die Hervorbringung durch Vorlesen von Textnachrichten gestützter, *smartphone-assistierter Erzählungen*: Sie erlaubt der ErzählerIn das (Vor-)Lesen der Textnachrichten und das Sprechen zur RezipientIn, während es der RezipientIn das Zuhören und das visuelle *Monitoring* der SprecherIn erlaubt, um das projizierte Ende der Erzählung zu erkennen. Aufgrund der für das Vorlesen der Textnachrichten erforderlichen visuellen Fokussierung des Smartphonedisplays kommt einem Wechsel der Blickrichtung der ErzählerIn deshalb auch eine zentrale Bedeutung für die Sequenzierung und die erkennbare Beendigung der Erzählung zu: Wie noch zu zeigen ist, wird

die Beendigung methodisch durch das Zusammenspiel sprachlicher und prosodischer Ressourcen (hier eine lachend realisierte Klimax der Erzählung) und einen Blickwechsel hin zur RezipientIn *accountable* gemacht.

### **5.1.2 Varianz der Teilnehmerahmen: Das simultane (Mit-)Lesen von Textnachrichten**

Anhand der Analyse eines weiteren Falles werde ich aufzeigen, dass es sich bei den hier konstatierten Annahmen über die Organisationsprinzipien und methodischen Verfahrensweisen der Hervorbringung dieser Art *smartphone-assistierter Geschichtenerzählungen* nicht lediglich um Beobachtungen eines Einzelfalles handelt. Der folgende Fall dient sowohl dazu, die wiederkehrenden Aufgaben der Interaktionsorganisation dieser multimodalen Praktik zu identifizieren, als auch die Varianz in der methodischen Bearbeitung dieser Aufgaben herauszuarbeiten. Auch im folgenden Fall dient das Vorlesen von Textnachrichten der Hervorbringung einer *smartphone-assistierten Geschichtenerzählung*, mit der die Verhaltensweisen einer abwesenden Person illustriert wird. Die Interaktionsepisode ist einem Gespräch zwischen den zwei Studienfreunden Fritz (F) und Moritz (M) entnommen. Thema der Unterhaltung ist das Verhältnis von M zu einer abwesenden Flirtbekanntschaft, mit der er bereits einige Textnachrichten ausgetauscht hat. Im vorangegangenen Teil des Gesprächs hatte F bereits einige dieser Nachrichten auf M's Smartphone betrachtet und die Einschätzung geäußert, dass die Abwesende ebenfalls Interesse an M hat. Das folgende Transkript enthält nun eine von M initiierte, kontrastierende Beschwerdeerzählung, in der er das kommunikative Fehlverhalten der abwesenden Flirtbekanntschaft rekonstruiert (vgl. Günthner, 1999a, S. 174). Die *smartphone-assistierte Erzählung* dient in diesem Fall somit der Illustration eines singulären Ereignisses, dass eine Relevanz für die Beurteilung des Verhältnisses zwischen dem Erzähler und seiner abwesenden Flirtbekanntschaft hat.

Bevor ich die Analyse beginne, möchte ich zunächst auf zwei weitere Differenzen zum vorangegangenen Fall hinweisen: Zum einen sitzen die zwei Gesprächsteilnehmer nebeneinander auf einer Couch, d. h. sie sind in einem „side-by-side-arrangement“ (Kendon, 1990, S. 213) zueinander positioniert (s. #1). Zum anderen – und für die Frage nach der Integration des Smartphones besonders relevant – schließt die Beschwerdeerzählung unmittelbar an eine vorangegangene *mobile-supported sharing activity* an, in der das Smartphone bei der gemeinsamen Evaluation des romantischen Interesses der Abwesenden bereits im Zentrum der gemeinsamen Aufmerksamkeit der beiden Gesprächspartner stand. Wie sich das körperliche Arrangement und die Positionierung des mobilen Endgeräts im gemeinsamen Sichtfeld der Teilnehmer auf die Initiierung und Produktion der *smartphone-assistierten Beschwerdeerzählung* auswirken,

wird im Folgenden ebenfalls reflektiert. Zu Beginn des Ausschnitts hält M das Smartphone bereits in seiner rechten Hand<sup>64</sup> und beide Teilnehmer schauen auf das Display des mobilen Endgeräts. Zur besseren Lesbarkeit wird die visuelle Fokussierung des Displays deshalb im Transkript als Grundkonstellation behandelt und lediglich Blickveränderungen annotiert. Die Blicke und Gesten von M werden durch „•“ gekennzeichnet, Blicke und Gesten von F mit „+“.

### Transkript 5: 180503 „Was mich abgefickt hat“\_Teil 1



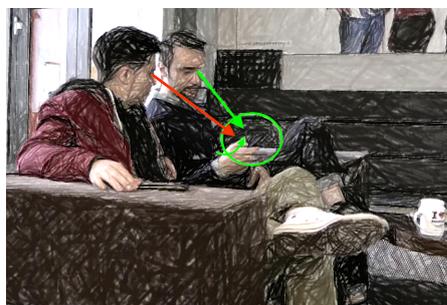
#1

```

m: >>-- hält SP in RH-->>
024 M: <<:-)würdst du Sagen; [ja?>]
025 F: [ja; ]
026 M: ah: ich WEIß •nicht; •
      •tippt Display•scrollt-->027
027 (0.3)•(0.4)
m: ---->•
028 M: aber weißt du des •was mich halt ABgefickt hat=
      =ist das die halt (mal) letzte woche gefragt hatte;
      •scrollt-->030
029 (1.0)#(0.4)

```

#2



Das Transkript setzt damit ein, dass M sich zunächst von F dessen positive Beurteilung (dass die abwesende Chatpartnerin Interesse an ihm hat) bestätigen lässt (Z. 024). Nachdem dieser

<sup>64</sup> Im Screenshot ist zudem erkennbar, dass auch F mit seiner rechten Hand, die auf der Armlehne der Couch liegt, sein Smartphone festhält. Dies wird nicht im Transkript annotiert, da sich nicht beobachten lässt, dass die Teilnehmer sich daran als relevant für die Interaktionsorganisation orientieren (vgl. Mondada, 2016b, S. 341).

dies in Überlappung zu M's Frageanhängsel („tag question“, Sacks et al., 1974, S. 718) bestätigt (Z. 025: „ja“), äußert M seine Zweifel daran, dass diese Einschätzung zutrifft (Z. 026). Während der Realisierung dieser Äußerung führt er zunächst eine einzelne, tippende Bediengeste mit seinem rechten Daumen durch und scrollt in der anschließende Gesprächspause (Z. 027) auf dem Display. Der Blick beider Teilnehmer ist dabei durchgehend auf das mobile Endgerät gerichtet.

Anschließend produziert M eine längere Äußerung (Z. 028), mit der er die Verfügbarkeit einer Erzählung projiziert: Er erwähnt zunächst, dass ihn etwas „abgefuckt“, also verärgert hat, und führt im schnellen Anschluss („rush through“, Schegloff, 1982, S. 76) eine unbestimmte Verhaltensweise der Abwesenden („das die halt (mal) letzte woche gefrAgt hatte“) als Grund für diese Verärgerung an. Auch in diesem Fall erfolgt die Projektion einer verfügbaren Erzählung durch den Einsatz spezifischer Verfahren, die die thematische Verbundenheit zum vorangegangenen *turn-by-turn-talk accountable* machen: Die Verwendung von „aber“ in der turn-initialen Position positioniert das Folgende im Kontrast zum vorangegangenen Gespräch (und damit zu F's positiver Evaluation). Die anschließende Formulierung „weist du“ stellt eine Orientierung am angenommenen Wissensstand des Gesprächspartners dar (vgl. C. Goodwin, 1979) und impliziert die Einführung eines neuen Sachverhalts. Anschließend formuliert er sein (negatives) Erleben („was mich halt abgefuckt hat“) bezüglich einer Verhaltensweise der Abwesenden, das in der jüngeren Vergangenheit („letzte woche“) stattfand. Die thematische Kohärenz zum vorangegangenen Gespräch lässt sich darüber hinaus anhand der deiktischen Personenreferenz („die“, die sich auf die abwesende Flirtbekanntschaft bezieht) erkennen. Als Begründung für M's Verärgerung wird ein Verhalten der Abwesenden benannt („das die halt (mal) letzte woche gefrAgt hatte“), das auch in diesem Fall als *prospective indexical* fungiert: Was an der Frage der Abwesenden den Sprecher verärgert hat, ist für den Rezipienten noch nicht verfügbar und muss im weiteren Vollzug der Interaktion erschlossen werden.

Darüber hinaus wird angezeigt, dass der Gebrauch des Smartphones ebenfalls thematisch im Bezug zur andauernden sprachlichen Interaktion steht, im Unterschied zum vorangegangenen Fall erscheint dies jedoch weniger augenscheinlich: Da der Sprecher das Smartphone bereits in seiner rechten Hand hält und beide Gesprächsteilnehmer auf das Display schauen, ist die erneute Initiierung des Smartphonegebrauchs lediglich daran zu erkennen, dass M während der Formulierung des ‚beschwerdewürdigen Verhaltens‘ in Z. 028 beginnt, auf dem Bildschirm zu scrollen. Auch hier findet die (Wieder-)Aufnahme der Gerätebedienung simultan zur sprachlichen Projektion einer Erzählung statt, sodass durch die Synchronisierung dieser Aktivitäten eine *multimodale Gestalt* hervorgebracht wird, die die Relevanz des Smartphonegebrauchs für

die projizierte Erzählung erwartbar macht<sup>65</sup>. Aufgrund der Kameraperspektive lässt sich jedoch nicht beurteilen, welche Funktion das Scrollen übernimmt – es lässt sich lediglich vermuten, dass dieses der Suche nach einem für die projizierte Erzählung relevanten Medieninhalt gilt. Erkennbar ist darüber hinaus, dass sich F beobachtbar an der projizierten Relevanz des Smartphonegebrauchs orientiert: Er fokussiert das Display und scheint M's Bedienung des Geräts aufmerksam zu beobachten. Das Scrollen hält auch während der entstehenden Gesprächspause in Z. 029 an und wird durchgehend von F beobachtet. Auch hier finden sich also die verkörperten Praktiken des *Monitorings*, mit der sich der Rezipient durch visuellen Zugriff zum Display Informationen über den Zweck des Smartphonegebrauchs verschafft und das von dem Smartphoneinhaber durchgehend durch die Positionierung des Geräts im gemeinsamen Sichtfeld der Teilnehmer ermöglicht wird.

Die Interaktion in Z. 028-029 konstituiert die Prä-Sequenz dieser smartphone-gestützten Erzählung, in der sprachlich die Verfügbarkeit einer Beschwerdeerzählung und die Relevanz des Smartphonegebrauchs für diese Aktivität durch eine *multimodale Gestalt* projiziert wird. Die Alignierung des Rezipienten mit der projizierten Aktivität wird auch hier durch dessen verkörpertes Verhalten angezeigt: Indem M in Z. 029 auf die Übernahme des Rederechts verzichtet und durch seine körperliche Ausrichtung die Fokussierung des Smartphonedisplays anzeigt, nimmt er eine verkörperte Haltung (C. Goodwin, 2007a, S. 61f.) ein, die für den Sprecher seine Alignierung mit der projizierten Erzählung *accountable* macht. Auch hier wird diese verkörperte Ausrichtung und der Verzicht auf die Übernahme des Rederechts vom Smartphoneinhaber als *ticket* für die anschließende Produktion der Erzählung behandelt.

Durch die Formulierung einer explizit negativen Evaluation des Verhaltens der Abwesenden zeigt der Smartphoneinhaber seine Erwartung an den Rezipienten, sich mit dieser Haltung zu affiliieren und die Beschwerdewürdigkeit des zu rekonstruierenden Verhaltens anzuerkennen (vgl. Traverso, 2009, S. 2386). Im folgenden Hauptteil der Geschichtenerzählung steht der Erzähler also vor der interaktiven Aufgabe, eine Sequenz vergangener Ereignisse auf eine Weise zu rekonstruieren, dass das Verhalten der Abwesenden für den Rezipienten als beschwerdewürdiges ‚Fehlverhalten‘ interpretierbar wird (vgl. Drew, 1998).

---

<sup>65</sup> Im Sinne der multimodalen KA lässt sich argumentieren, dass die sprachliche Bezeichnung der ‚beschwerdewürdigen‘ Aktivität des Fragens hier vom Rezipienten aufgrund des Zusammenspiels der unterschiedlichen semiotischen Ressourcen als ‚über das Smartphone vermittelt gefragt hatte‘ bzw. konkreter ‚sie hat mich per WhatsApp gefragt‘ gehört werden kann. Da der Rezipient visuellen Zugriff zum Smartphone hat, stellt das Display ein semiotisches Feld dar (C. Goodwin, 2007a), das dahingehend interpretiert werden kann, wie es den Sinn der sprachlichen Äußerung ergänzt. Aus dieser Kameraperspektive lässt sich jedoch nicht rekonstruieren, was der Rezipient sehen kann.

Transkript 6: 180503 „Was mich abgefickt hat“\_Teil2

030 M: und •dann hab ich halt geschrieben magst heute  
abend vorBEikommen,

---->•

031 (0.6)#

032 M: •und <<lachend> #deine referate vorste+llen,>•

•.....-----blickt zu F-----,//////////•

f: +:-)--->034

#3

#4



033 und dann •meinte [sie einfach ] NUR so-•

m: •-----scrollt-----•

034 F: +[<<:-)hjaja;>]

f: ----->+

035 (0.4)

036 M: ä::h ah doch ich glaub da SCHO:N dran das  
könnten die schon noch schaffen;

037 ä:h •das war noch• (spiel) für die BAYern-

• scrollt •

038 °h ich MUSS dir wohl ABSagen ich schaffe es  
heute wirklich nicht;•

•scrollt-->039

039 muss:• (.) noch so viel erledigen dass ich nich  
raffe;

-->•

040 •da +denk# ich mir AUCH so-

•.....schaut nach vorne-->041

f: +langsames kopfnicken-->041

#5



041 oh ke isses jetzt so das sie einfach  
all•geMEIN keine oder wenig zeit hat (.) äh  
innerhalb der nächsten wochen;•+ oder • is es halt  
nur,  
-->•...Blick zu F-----  
-----,,,,•scrollt•Blick SP->042  
f: ----->+  
042 F: ja vielleicht (.)+ hat sie die •anspielung soFORT  
schon verstanden und wollte dringend was für\_s  
referAT machen;•  
+....Blick zu M-->044  
m: ----->•.....Blick zu F  
-----,,•Blick SP-->044  
043 M: JA; •ha [ha ha +ha• hahe]  
044 F: [ha ha +ha• hahe]  
-->+  
m: --->•Blick zu F •  
045 M: [<<lachend> vielleicht hast du RECHT> •ja; •]

Nachdem in der Prä-Sequenz ein interpretativer Rahmen geschaffen wurde, in dem die negative Haltung des Erzählers explizit formuliert wurde, erfolgt nun – typisch für Beschwerdeerzählungen (vgl. Mandelbaum, 1991, S. 135) – im Hauptteil eine im Vergleich ‚neutralere‘ Rekonstruktion der relevanten Ereignisse. Die Erzählung beginnt damit, dass M in Z. 030 seine Antwort auf die eingangs erwähnte Frage der Abwesenden wiedergibt, die sequenziell als nächster Schritt („und dann“) in einer Abfolge von Ereignissen eingeführt wird (vgl. Deppermann, 2008, S. 65). In der bereits bekannten Form wird diese Erwiderung durch den simultanen Einsatz sprachlicher Praktiken – den Einsatz eines Zitationsmarkers („hab ich halt geschrieben“) und den Modus der DRS („magst heute abend vorBEIkommen“) – und verkörperter Ressourcen – die visuelle Fokussierung des Displays und die Handhabung des mobilen Endgeräts – multi-modal als Vorlesen der Textnachricht hervorgebracht. In der anschließenden Gesprächspause (Z. 031) schaut M zunächst weiter auf das Display seines Smartphones (#3), bevor er in seiner nächsten Äußerung (Z. 032) den Zweck seiner Einladung formuliert: Er bot seiner

Chatpartnerin an, ihm ihre „Referate vorstellen“ zu können. Hierbei handelt es sich nun um eine direkte Redewiedergabe, denn während dieser lachend produzierten Äußerung schaut M zum ersten Mal direkt zu F (#4). Mit seinem Lachen orientiert sich M am impliziten Sinngehalt seiner Einladung<sup>66</sup> und schaut zu F, um dessen Verständnis der wiedergegebenen Nachricht zu überprüfen. Dieser beginnt zunächst in Z. 032 vor Beendigung der Äußerung zu lächeln, sein Blick bleibt jedoch auf das Display des Smartphones gerichtet.

M's nächste Äußerung in Z. 033 bereitet die Wiedergabe der Antwort der Abwesenden vor: Erneut werden die Ereignisse in ihrer sequenziellen Abfolge verknüpft („und dann“) und der Sprecherwechsel zwischen den animierten Figuren der Erzählwelt durch einen Zitationsmarker („meinte sie“) *accountable* gemacht. Durch die Formulierung „einfach nur so“ projiziert der Erzähler sprachlich eine Diskrepanz zwischen der nun folgenden Antwort der Abwesenden und seinen Erwartungen – und rahmt die Erwiderung dadurch als zu ‚beiläufig‘ oder ‚lapidar‘. In Überlappung zu dieser Äußerung (simultan zum Beginn von „meinte“ in Z. 033) produziert F in Z. 034 mit *smile voice* einen „continuer“<sup>67</sup> (vgl. C. Goodwin, 1986, S. 210), mit dem er nun auch sprachlich sein Verständnis des Subtexts der Einladung von M anzeigt<sup>68</sup>. Nach einer kurzen Stille (Z. 035) animiert M in Z. 036 die nächste Erwiderung seiner Chatpartnerin: Die Äußerung beginnt zunächst mit einer Stockung („äh ah“), bevor der Sprecher eine weitere Nachricht vorliest („doch ich glaub da SCHO:N dran das könnten die schon noch schaffen;“), die zwar an der Textoberfläche syntaktische Markierungen enthält (bspw. Pronomina, Deixis), die kohäsive Beziehungen zu vorangegangenen Äußerungen anzeigen (vgl. Deppermann, 2008, S. 64), jedoch nicht unmittelbar als sinnhafter nächster Zug des archivierten Chats verständlich scheint. Anschließend produziert M eine kontextualisierende Äußerung, die mit Hilfe der Deplatzierungsmarkierung (vgl. Schegloff & Sacks, 1973, S. 320) „das war noch“ kenntlich

---

<sup>66</sup> Mit seinem Lachen scheint sich der Erzähler daran zu orientieren, dass dieser Zweck des Treffens von ihm als unverfänglicher Vorwand für eine eigentlich romantisch oder sexuell motivierte Verabredung verwendet wurde. Es scheint sich also um einen Fall impliziter und abstreitbarer Flirtkommunikation zu handeln, in dem dieser Vorwand als „cautious strategy for handling the high risk of potential personal rejection, which is an inherent part of initial romantic encounters“ (Mortensen, 2017, S. 583) fungiert. Dass sich beide TeilnehmerInnen an eben diesem impliziten Charakter der (Online-)Flirtkommunikation orientieren, wird im weiteren Verlauf der Interaktion eindeutig erkennbar.

<sup>67</sup> Dem Vorschlag von Egbert (2009, S. 184) folgend wird der Begriff des „continuers“ nicht übersetzt. Als *continuer* werden Handlungen verstanden, die das Verständnis des Rezipienten anzeigen, „that an extended turn at talk is in progress but not yet complete, while simultaneously collaborating in the achievement of that multi-unit utterance by passing the opportunity to either (a) produce a more extended turn of their own or (b) initiate repair on the talk just heard.“ (C. Goodwin, 1986, S. 207). *Continuer* signalisieren Verstehen seitens der RezipientIn, ohne dieses Verständnis zu explizieren (vgl. Schegloff, 1982, S. 81)

<sup>68</sup> Die Interpretation, dass F mit dem *continuer* in Z. 034 auf eine frühere und nicht die aktuell im Vollzug befindliche Äußerung reagiert, wird damit begründet, dass eine überlappende Realisierung der Äußerung erfolgt, bevor M's Turn (Z. 033) einen erkennbaren „transition-relevance place“ (Sacks et al., 1974, S. 703) erreicht. Weder lässt sich „und dann meinte“ (Z. 033) bereits als eigenständige Turnkonstruktionseinheit begreifen, noch ist davon auszugehen, dass die weitere Entwicklung für den Rezipienten hier bereits projizierbar ist (vgl. ebd., 1974, S. 720).

macht, dass diese Nachricht sich auf frühere und nicht die unmittelbar vorher wiedergegebenen Chatbeiträge bezieht. Durch die Verwendung der Vergangenheitsform und die Formulierungsarbeit (Gülich, 1994, S. 79), mit der der Erzähler die deiktischen Verweise dieser Nachricht durch explizite Referenzen („ä:h das war noch (spiel) für die BAYern“) ersetzt, macht er erkenntlich, dass diese nicht selbst für die Rekonstruktion des beschwerdewürdigen Verhaltens relevant ist. Während der Produktion dieser Äußerung führt der Erzähler erneut scrollende Bediengesten durch und liest anschließend in Z. 038-039 nach einem hörbaren Einatmen weitere Nachrichten der Abwesenden vor: Lautsprachlich realisiert er zunächst die als dispräferierte Handlung gestaltete Absage („ich MUSS dir wohl ABSagen ich schaffe es heute wIrklich nicht;) und eine anschließende Erklärung als Rechtfertigung (vgl. Heritage, 1988) („muss: (.) noch so viel erledigen dass ich nich raffe;“).

Ab Z. 040 wird auch in diesem Fall das Ende der narrativen Rekonstruktion durch das Zusammenspiel verkörperter Handlungen und sprachlicher Äußerungen des Erzählers *accountable* gemacht. Mit Beginn seiner Äußerung in Z. 040 blickt er vom Smartphone auf und schaut in den Raum. Seine linke Hand leicht fragend geöffnet (#5), kehrt der Erzähler sprachlich zum ‚Hier und Jetzt‘ zurück, indem er – nun im Präsens – seine Gedanken zu den Erwidern der Abwesenden formuliert („da denk ich mir AUCH so“). Wie Haakana (2007) zeigt, wird die Wiedergabe von Gedanken in Beschwerdeerzählungen oftmals sequenziell nach der Komplikation platziert und methodisch eingesetzt, um die Evaluation dieses Verhaltens durch die ErzählerIn kenntlich zu machen und die RezipientIn zu einer affiliativen Erwidern zu führen (ebd. 2007, S. 176). Während dieser Äußerung beginnt F wiederholt langsam und bedächtig zu nicken (Z. 040).

Während der Elaboration dieser Erzählung hat der Rezipient – abgesehen von dem verständnisanzeigenden *continuer* in Z. 034 – bis jetzt keine weiteren sprachlichen Äußerungen produziert, insbesondere keine Anzeigen von Affiliation mit dem Erzähler. Seine Alignierung mit der Aktivität der interaktiven Hervorbringung der Geschichtenerzählung wird überwiegend durch die Aufrechterhaltung einer verkörperten instrumentellen und epistemischen Haltung (C. Goodwin, 2007a, S. 61–62) angezeigt: Er hört der narrativen Rekonstruktion des Sprechers zu, während er gleichzeitig seine visuelle Aufmerksamkeit auf das Display des Smartphones fokussiert und so in der Lage ist, den Chat mitzulesen. Dieses Ausbleiben affiliativer Erwidern, mit der der Rezipient anerkennt, dass es sich um eine beschwerdewürdige Situation handelt und die (negativen) Gefühle des Beschwerdeproduzenten gerechtfertigt sind (vgl. Traverso, 2009, S. 2386), scheint dazu zu führen, dass M in der Ausleitung der Erzählung vor allem seine Schwierigkeiten, die Bedeutung der Absage zu interpretieren, betont: Er fragt (sich) – den Blick

währenddessen auf F gerichtet –, ob die Abwesende „allgeMEIN keine oder wenig zeit hat“ (Z. 041) oder – den Blick nun wieder auf das Smartphone gerichtet – ob es einen anderen, nicht explizit benannten Grund für die Absage geben könnte (Z. 042). Indem M eine der möglichen Interpretationen expliziert, die alternative Erklärung jedoch unausgesprochen lässt, werden Folgerwartungen hinsichtlich des präferierten Verständnisses der Erzählung geschaffen (vgl. Jefferson, 1972, S. 315): Es legt dem Rezipienten nahe, die ausformulierte Begründung als Interpretation der Geschichte zu bestätigen, oder stattdessen die Arbeit zu leisten, selbst eine – bis jetzt implizit gebliebene – Erklärung zu finden<sup>69</sup>. Darüber hinaus macht es dieses offene Ende der Erzählung nun zur Aufgabe des Rezipienten, die Bedeutung der narrativen rekonstruierten Ereignisse zu formulieren (vgl. Drew, 1984, S. 133).

In der in Z. 042 beginnenden *response sequence* produziert F nun seine Interpretation der Gründe für die Absage. Diese wird methodisch auf eine Weise hervorgebracht, dass sie als subtile Nicht-Beachtung der Beschwerde („subtle disattending of complaints“, Mandelbaum, 1991, S. 105) erkennbar wird: Er vermutet, dass die Abwesende die in der Einladung enthaltene Anspielung direkt verstanden hat und tatsächlich an ihrem Referat arbeiten wollte. Der Rezipient vermeidet somit eine Affiliation mit der Beschwerde „by picking up on other available elements [of the story, Anm. DS] and developing them“ (Mandelbaum, 1991, S. 116): In seiner Äußerung greift er zum einen die vom Erzähler präferierte Erklärung („wirklich arbeiten müssen“) für die Absage auf, darüber hinaus expliziert er nun auch sein eigenes – vorher implizit gebliebenes – Verständnis der eigentlichen Absichten der Einladung und integriert dieses in den *Account* für die Absage. Diese Deutung ist insofern disaffiliativ, als dass sie dem kommunikativen Verhalten des Erzählers ebenfalls eine Verantwortung für den Ausgang der Verabredungsanfrage zuschreibt und damit die Beschwerdewürdigkeit des Verhaltens der Abwesenden in Frage stellt. Auffällig ist, dass der Rezipient erst kurz nach Beginn der Explikation seines Verständnisses seinen Blick vom Smartphonedisplay löst und zu M schaut, der diesen Blick kurz darauf erwidert und ihn bis zum Ende der Äußerung aufrechterhält<sup>70</sup>. Auf diesen Umstand wird im Folgenden noch näher eingegangen.

---

<sup>69</sup> Im Kontext einer Beschwerdeerzählung über das Verhalten einer Flirtbekanntschaft liegt es nahe, dass bspw. eine Absage aus mangelndem Interesse aus Sicht des Erzählers eine dispräferierte Interpretation darstellt. Damit würde die Absage als ‚Absage unter Vorgabe vorgeschobener Gründe‘ und somit als beschwerdewürdiges Verhalten interpretiert werden.

<sup>70</sup> Bemerkenswert erscheint dies insofern, als dass sich SprecherInnen in (nicht-mediatisierten) F2F-Interaktionen normalerweise an der Regel „*When a speaker gazes at a recipient he should make eye contact with that recipient*“ (C. Goodwin, 1979, S. 106, Herv. i. O.) orientieren. Wenn SprecherInnen eine RezipientIn adressieren, die sie nicht anschaut, kommt es oftmals zur Produktion von Stockungen, Pausen oder Neustarts, bis die visuelle Aufmerksamkeit und damit die Zuhörerschaft der RezipientIn gesichert ist (vgl. C. Goodwin, 1981, S. 60–77). In der hier fokussierten Passage, aber auch in der Gesamtschau der Prä-Sequenzen und Hauptteile der untersuchten, smartphone-assistierten Erzählungen zeigt sich in den beobachtbaren Orientierungen, dass diese Regel während

In der anschließenden Rückkehr zum *turn-by-turn-talk* zeigt sich, dass der Erzähler den Versuch, den Rezipienten zur Affiliation mit der Beschwerdeerzählungen zu bewegen, nicht weiterverfolgt (Mandelbaum, 1991, S. 135): Er nimmt das Deutungsangebot von F in Z. 043 lachend an, woraufhin beide Gesprächspartner beginnen zu lachen (Z. 043-044). Bereits gegen Ende dieses gemeinsamen Lachens beginnt M wieder auf sein Smartphone zu schauen und dieses durch Scrollen zu bedienen. In Z. 045 produziert er anschließend eine (vorsichtige) sprachliche Bestätigung des Interpretationsangebots. Das Zusammenspiel aus gemeinsamen Lachen, dieser sprachlichen Bestätigung und dem verkörperten Verhalten projiziert eine potenzielle Beendigung der *response sequence*, die anschließend auch durch eine Schließungssequenz („closing section“, Schegloff & Sacks, 1973, S. 300) kollaborativ zu Ende gebracht wird (nicht mehr im Transkript enthalten).

In der Analyse dieses zweiten Falles lassen sich einige bereits identifizierte Praktiken und Organisationsprinzipien *smartphone-assistierter Geschichtenerzählungen*, die durch das Vorlesen von Textnachrichten gekennzeichnet sind, erneut beobachten: Insbesondere die Projektion der Smartphoneintegration durch eine *multimodale Gestalt*, das *Monitoring* des Smartphonegebrauchs des Rezipienten, das als *ticket* für die Elaboration der Erzählung behandelt wird, das Vorlesen von Textnachrichten, das durch das Zusammenspiel sprachlicher Praktiken der DRS und die Einnahme der typischen Körperhaltung hervorgebracht wird, sowie der Einsatz eines Blickwechsels, mit dem das Erreichen der Komplikation der narrativen Rekonstruktion angezeigt wird. Ich werde im Fazit des Kapitels ausführlich auf diese typischen Praktiken derartiger Geschichtenerzählungen eingehen.

Zunächst möchte ich jedoch einige fallspezifische Besonderheiten des Falles reflektieren: Aufgrund des *side-by-side-arrangements* der Teilnehmer und der Präsenz des Smartphones bereits zu Beginn der Gesprächsepisode ist aus der konversationsanalytischen Beobachterperspektive die Relevanz verkörperten Praktiken für die Hervorbringung der Geschichtenerzähler weniger augenscheinlich: Trotzdem lässt sich die Sequenzierung anhand des Zusammenspiels sprachlicher Äußerungen, stummer Bediengesten und Veränderungen in der Blickführung rekonstruieren. Dabei darf nicht außer Acht gelassen werden, dass der Rezipient während der Prä-Sequenz und dem Hauptteil der Erzählung durchgehend visuellen Zugriff zum Smartphone hat und ihm das semiotische Feld des Displays eine Ressource bietet, die Sinnhaftigkeit des Smartphonegebrauchs und der darauf bezogenen sprachlichen Äußerung zu erschließen.

---

der Phasen des simultanen Smartphonegebrauchs ausgesetzt oder zumindest erheblich gelockert wird. Ich werde auf diesen Umstand zurückkommen.

Der *Affordanz* des Smartphones, den RezipientInnen durch die interaktive Hervorbringung spezifischer Teilnahmerahmen visuellen Zugriff zur archivierten Textnachrichtenkommunikation zu ermöglichen, kommt im Kontext dieser Beschwerdeerzählungen eine spezifische Funktion zu: Wie in zahlreichen Studien gezeigt wurde, werden die Praktiken der DRS in Beschwerdeerzählungen eingesetzt, um den RezipientInnen Zugang zur vergangenen Interaktion zu bieten und ihnen so zu ermöglichen, zu einer eigenen Interpretation des Verhaltens der Abwesenden zu gelangen (bspw. Drew, 1998; Holt, 1996, S. 229, 2000; Holt & Clift, 2007, S. 12–13). Mithilfe des Smartphones wird dem Rezipienten hier ein erweiterter epistemischer Zugang zu einem bewertbaren Referenten (vgl. Raclaw et al., 2016, S. 377) gewährt, der es ihm erlaubt, die archivierten Originale der Nachrichten im Verlauf der narrativen Rekonstruktion mitzulesen. In Kombination damit, dass der Erzähler – nach der negativen Rahmung in der Prä-Sequenz – in der Elaboration eine eher neutrale Version der Ereignisse produziert (er verwendet keine prosodische Stilisierung beim Vorlesen und lediglich implizite sprachliche Markierung seiner Haltung (Z. 033, 040–41)), wird dem Rezipienten eine eigenständige Interpretation ermöglicht. Dass der Rezipient sein Deutungsangebot als unabhängig von der Haltung des Erzählers hervorbringt, zeigt sich nicht nur am disaffiliativen Charakter seiner sprachlichen Erwiderung, in der ein implizit gebliebener Aspekt der Erzählung als Grund für die Absage eingeführt wird. Es lässt sich darüber hinaus argumentieren, dass er durch die durchgehende visuelle Fokussierung des Displays, die erst nach dem Beginn der Produktion seines Deutungsangebots aufgelöst wird, seine Interpretation als in seiner lesenden Aneignung des archivierten Chats begründet darstellt. Die fallspezifische Komplexität dieser *smartphone-assistierten Erzählung* zeigt somit eindrücklich, dass „the ability to go back over archived text and images allows for different sorts of reporting on those conversations . . . each of them designed with respect to local arrangements” (Brown et al., 2018, S. 23). Im vorliegenden Fall bezieht sich diese kontextsensitive Partikularisierung nicht nur auf eine Anpassung der sprachlichen Ausgestaltung der Wiedergabe der Textnachrichten, sondern auch auf die Ermöglichung des Mitlesens für den Rezipienten, das eine spezifische Funktion im Kontext einer Beschwerdeerzählungen über abwesende Dritte übernimmt.

### **5.1.3 Fazit: Praktiken des Vorlesens von Textnachrichten in methodisch vorbereiteten Geschichtenerzählungen**

Die beiden in diesem Kapitel analysierten Fälle unterscheiden sich u. a. insofern voneinander, als dass es sich einerseits um eine humoristische und andererseits um eine Beschwerdeerzählung handelt. Aufgrund der fallspezifischen Unterschiede ließ sich zeigen, wie durch Vorlesen

von Textnachrichten gestützte Erzählungen an den spezifischen lokalen Kontext konkreter Interaktionsepisoden angepasst werden (vgl. Brown et al., 2018, S. 23). Im Folgenden sollen nun die typischen Merkmale und Methoden der Hervorbringung methodisch vorbereiteter, *smartphone-assistierten Erzählungen* dargestellt werden. Die Darstellung der zentralen Ergebnisse folgt dabei dem sequenziellen Verlauf von Geschichtenerzählungen in F2F-Interaktionen.

In der Prä-Sequenz *smartphone-assistierter Geschichtenerzählungen* der hier diskutierten Variante stehen die potenziellen ErzählerInnen nicht nur vor der Aufgabe, die Verfügbarkeit einer Erzählung zu projizieren und deren Verbindung zum vorangegangenen *turn-by-turn-talk* kenntlich zu machen. Sie müssen auch die Integration des Smartphones in die F2F-Interaktion vorbereiten und die Verbindung des Smartphonegebrauchs zu den aktuellen Gesprächsaktivitäten *accountable* machen. Die Verfügbarkeit einer *smartphone-assistierten Erzählung* wird in den untersuchten Fällen erkennbar durch die Hervorbringung *multimodaler Gestalten* projiziert: Die sprachlichen Äußerungen und die simultane – oftmals methodisch synchronisierte – visuelle Fokussierung und physische Handhabung des mobilen Endgeräts verleihen sich wechselseitig ihren Sinn und zeigen für die TeilnehmerInnen an, dass der Gebrauch des mobilen Endgeräts eine praktische Relevanz für die lokale Interaktion besitzt. Insbesondere in thematisch verbundenen Geschichtenerzählungen erfolgt die Initiierung des Smartphonegebrauchs dabei oftmals ohne explizite sprachliche Ankündigung (vgl. Keppler, 2019, S. 179), sodass seine sinnhafte Relevanz für die lokale Interaktion für andere Anwesende lediglich inferentiell (Sacks, 1995, Vol I, S. 113ff.) zu erschließen ist. Die von den (potenziellen) ErzählerInnen produzierten *prospective indexicals* machen es für die RezipientInnen nicht nur zur Aufgabe, im Verlauf der Interaktion herauszufinden, auf welche Ereignisse die SprecherIn in der Prä-Sequenz konkret Bezug nimmt, sondern auch zu ergründen, welche Relevanz die begonnene Bedienung des Smartphonegebrauchs für die projizierte Erzählung besitzt.

In den Reaktionen der RezipientInnen zeigt sich jedoch, dass eine derart inferentiell projizierte Smartphoneintegration nicht immer als beobachtbare und berichtbare Handlung erkennbar ist. Sie orientieren sie sich mit ihren verkörperten Handlungen an der potenziellen *bystander inaccessibility* des Smartphonegebrauchs: Da der praktische Zweck der Gerätebedienung aufgrund der materiell-technischen Eigenschaften der mobilen Endgeräte (Größe des Geräts, Displaypositionierung, Steuerung über stumme Bediengesten) nicht ohne Weiteres für andere Anwesende ersichtlich ist, verschaffen sich die RezipientInnen selbst-initiiert visuellen Zugriff zum Display des Smartphones und damit zu Informationen über den Zweck des Gebrauchs. Neben den bereits in der Literatur identifizierten sprachlichen Praktiken, mit denen der Smartphonegebrauch für andere Anwesende *accountable* gemacht wird (bspw. *kollaborative Suche* und

*suchbegleitende Kommentierungen*, s. Brown et al., 2015; Suderland, 2019), lassen sich hier mithilfe einer multimodale Analyse verkörperte Praktiken des *Monitorings* identifizieren, die der Bearbeitung des gleichen Problems dienen. Die Realisierung dieser Beobachtung ist dabei auf die lokale Hervorbringung eines spezifischen Teilnehmerrahmens angewiesen, in der das Smartphone im Sichtfeld aller InteraktionspartnerInnen platziert wird. Die Initiierung des visuellen Zugriffs durch die RezipientInnen wird dabei so gut wie nie sprachlich angekündigt – die Ratifikation der Angemessenheit des *Monitorings* durch die SmartphoneinhaberInnen erfolgt ebenfalls durch verkörperte Handlungen, oftmals lediglich dadurch, dass die Positionierung des Gerätes nicht verändert oder so justiert wird, dass das Display für die RezipientInnen einsehbar ist<sup>71</sup>.

Die Praktiken des *Monitorings* lassen sich zum einen als kontext-unabhängige Methoden begreifen, die zur *Accountability* des selbst-initiierten Smartphonegebrauchs in F2F-Interaktionen beitragen. Im spezifischen Kontext von Prä-Sequenzen *smartphone-assistierter Geschichten-erzählungen* übernimmt das *Monitoring* eine doppelte Funktion („double duty“, Schegloff, 2007a, S. 169): Die stumme, fokussierte Beobachtung des Displays durch die RezipientInnen wird von den SmartphoneinhaberInnen als verkörperte Alignierung mit der projizierten Erzählung und – wenn die RezipientInnen nicht das Rederecht ergreifen – als *ticket* für die Produktion des Hauptteils der Erzählung behandelt.

Der spezifische Teilnehmerrahmen des Hauptteils derartiger, *smartphone-assistierter Erzählungen* wird dadurch hervorgebracht, dass die SmartphoneinhaberIn eine spezifische Körperhaltung (vgl. C. Goodwin, 1984, S. 222) einnimmt: Während sie zu den RezipientInnen spricht, hält sie das Smartphone so, dass sie das Display während der wörtlichen Wiedergabe der Textnachrichten durchgehend fokussieren kann. Die Produktion der Erzählung ist also durch eine „intrapersonelle Multiaktivität“ (vgl. Deppermann, 2014, S. 249) gekennzeichnet, in der die

---

<sup>71</sup> In dieser Hinsicht erweitern diese Beobachtungen die Ergebnisse von Oloff (2019a), die zeigen konnte, dass „ein Teilen visueller Inhalte auch a priori ein Teilen bzw. die Übergabe des Geräts beinhaltet und die Initialisierung einer Zeigesequenz das Smartphone aus dem persönlichen Territorium seines/seiner Besitzers/Besitzerin ‚herauslöst‘“ (2019a, S. 213). Dies scheint – zumindest im vorliegenden Korpus – ebenfalls auf die Phasen der Vorbereitung des Vorlesens von Textnachrichten zuzutreffen, wenn der Gebrauch des Smartphones als mit den aktuellen Gesprächsaktivitäten verbunden projiziert wurde und sich RezipientInnen selbst-initiiert einen visuellen Zugang zum Smartphone verschaffen. Oloffs (2019a) und meine Ergebnisse zeigen somit eine entscheidende Differenz zwischen einem multimodal-konversationsanalytischen Vorgehen und der Analyse von Audioaufzeichnungen, denn Keppler (2019) kommt in ihren Analysen des Smartphonegebrauchs in Alltagsgesprächen zu dem Ergebnis, dass „der Blick auf das Smartphone einer anderen Person grundsätzlich zustimmungs- bzw. erlaubnispflichtig“ (2019, S. 180) sei. Diese widersprüchlichen Einschätzungen mögen darin begründet liegen, dass basierend auf der Analyse von Audioaufzeichnungen lediglich diejenigen Erscheinungsformen des Phänomens rekonstruiert werden können, in denen der Smartphonegebrauch und damit verbundene Aktivitäten explizit sprachlich thematisiert werden. Schon die ersten zwei, hier präsentierten Fällen zeigen jedoch, dass die Sinnhaftigkeit der projizierten Smartphoneintegration nicht zwangsläufig durch explizite lexiko-semantische Bezüge auf die Verwendung der mobilen Endgeräte, sondern durch das komplexe Zusammenspiel verschiedenster Ressourcen in *multimodalen Gestalten* hervorgebracht wird.

SmartphoneinhaberIn ihre visuellen und gestischen Ressourcen auf das Smartphone fokussiert und mit ihren sprachlichen Äußerungen die RezipientIn adressiert. Auch die RezipientInnen müssen eine instrumentelle Haltung einnehmen, die es ihnen ermöglicht, für die im Vollzug befindliche, smartphone-assistierte Erzählung relevante Handlungen produzieren zu können (vgl. C. Goodwin, 2007a, S. 61). Dies bedeutet zunächst, dass sie sich mindestens so positionieren, dass sie die Äußerungen der ErzählerIn hören können. Visuell fokussieren die RezipientInnen entweder die SprecherIn oder – um einen erweiterten epistemischen Zugang zu den wiedergegebenen Textnachrichten zu erhalten – das Smartphonedisplay. Die Betrachtung des Displays während des Hauptteils der Erzählung ist dabei keinesfalls als ‚Nebenprodukt‘ der körperlichen Arrangements der TeilnehmerInnen zu verstehen, sondern wird von ihnen methodisch eingesetzt, um spezifische Aufgaben der Interaktionsorganisation zu bearbeiten – bspw. um zu einer eigenständigen Beurteilung der rekonstruierten, ‚beschwerdewürdigen‘ Ereignisse zu gelangen, aber auch um Probleme des Verstehens zu reparieren, wie das folgende Kapitel zeigen wird.

Die Elaboration der durch das Vorlesen von Textnachrichten realisierten Erzählungen erfolgt durch die sprachlichen Praktiken der DRS. Die persönliche, räumliche und zeitliche Deixis der Nachrichten wird aus Sicht der ursprünglichen AutorIn wiedergegeben, um die Form und den Inhalt der Nachricht zu bewahren (vgl. Holt & Clift, 2007, S. 5). Der Wechsel in der Wiedergabe von Textnachrichten verschiedener TeilnehmerInnen der archivierten Interaktionen wird hauptsächlich durch die Verwendung von Zitationsmarkern (Personalpronomen und *verbum dicendi*) kenntlich gemacht. Außerdem werden Variationen in der Stimmqualität und Prosodie eingesetzt, um anzuzeigen, wessen Nachrichten wiedergegeben werden. Als multimodale Praktik wird diese Wiedergabe der Nachrichten durch die simultane, visuelle Fokussierung des Smartphonedisplays als Vorlesen von Textnachrichten hervorgebracht. Dies fügt derartigen smartphone-assistierten Erzählungen eine zusätzliche Dimension epistemischer Autorität hinzu (vgl. Raclaw et al., 2016, S. 377), denn die Verankerung des Wortlauts der Nachrichten in der Materialität der auf dem Bildschirm angezeigten Textnachrichten verbürgt für die Authentizität der Wiedergabe und macht diese potenziell überprüfbar. Diese gesteigerte Authentizität ist jedoch nicht gleichbedeutend damit, dass es sich beim Vorlesen von Textnachrichten um eine neutrale Wiedergabe handelt: Die ErzählerInnen produzieren spezifisch designte Prä-Sequenzen und setzen sprachliche (bspw. „nur“), para-sprachliche (bspw. Lachpartikeln) „stance-marker“ und prosodische (bspw. das „leiernde Vorlesen“ im ersten Fall) Ressourcen in der lautsprachlichen Realisierung ein, um implizit und explizit ihre Erzählhaltung zu den AutorInnen und Inhalten der Textnachrichten anzuzeigen (Brown et al., 2018, S. 20; s. a. Raclaw et al.,

2016, S. 369, 373). Auf diese Weise machen sie für die RezipientInnen erkennbar, wie diese Geschichtenerzählung zu verstehen ist.

Der Doppelcharakter der wörtlichen Wiedergabe zwischen Authentisierung und Fiktionalisierung wird von den SmartphoneinhaberInnen abhängig von der spezifischen narrativen Aktivität und dem Verhalten der RezipientInnen moduliert. Insbesondere im Fall von narrativen Aktivitäten (bspw. Beschwerdeerzählungen über abwesende Dritte), die selbst potenziell rechtfertigungsbedürftig sind (Maynard, 2013, S. 199), bietet die wörtliche Zitation der Textnachrichten einen Zugang zu den als beschwerdewürdig gerahmten Verhaltensweisen der Abwesenden (vgl. Drew, 1998; Holt, 2000). Die *Affordanzen* des Smartphones ermöglichen es darüber hinaus, den RezipientInnen visuellen Zugriff zum Display zu gewähren, sodass diese die sprachlich wiedergegebenen Textnachrichten simultan (mit-)lesen können. Damit erhalten die RezipientInnen einen erweiterten epistemischen Zugriff zu einem bewertbaren Referenten, der eingesetzt wird, um sie zur Alignierung und Affiliation zu leiten (Raclaw et al., 2016, S. 377), da dieser ihnen vermeintlich erlaubt, zu einer ‚eigenständigen‘ Interpretation der archivierten Interaktion zu gelangen (vgl. Holt & Clift, 2007, S. 12).

Die Aushandlung der Bedeutung der Erzählung für die lokale Interaktion erfolgt nach der beobachtbaren und berichtbaren Beendigung der narrativen Rekonstruktion. Spezifisch für smartphone-gestützte Erzählungen ist dabei, dass die ErzählerInnen die Ausleitung aus der Erzählung nicht nur sprachlich *accountable* machen: Eine zentrale Rolle spielt die Veränderung der Blickrichtung, denn indem die ErzählerInnen ihren Blick vom Smartphonedisplay lösen und zu den RezipientInnen oder in den Raum schauen, machen sie erkennbar, dass das Vorlesen beendet ist. Oftmals wird dies darüber hinaus von einer Beendigung der physikalischen Manipulation des Geräts sowie die Verwendung erzählrelevanter Gesten begleitet. Die anschließende *response sequence* ist durch eine Rückkehr zum *turn-by-turn-talk* gekennzeichnet, in der die RezipientInnen durch ihre Erwidernungen ihr Verständnis der smartphone-assistierten Erzählung anzeigen und anschließend die Bedeutung der Geschichte für die lokale Interaktion interaktiv ausgehandelt wird. In dieser Phase kommt es nun vermehrt zur wechselseitigen Blickverschränkung zwischen den TeilnehmerInnen. Die Bedienung des Smartphones – obwohl nicht mehr unmittelbar für diese Sequenz relevant – wird jedoch oftmals nicht direkt beendet, sodass der Blick der SmartphoneinhaberIn zwischen Gerätedisplay und GesprächspartnerInnen wechselt.

Betrachtet man nun die interaktive Hervorbringung derartiger *smartphone-assistierter Geschichtenerzählungen* über alle Phasen hinweg, so lässt sich bereits hier argumentieren, dass die Beteiligungsrollen der TeilnehmerInnen mit den ‚kanonischen‘ Bezeichnungen ErzählerIn

und ZuhörerIn (vgl. bspw. Quasthoff, 2001, S. 1300) nur unzureichend erfasst werden können: Aus Perspektive der multimodalen KA wird grundsätzlich betont, dass die ZuhörerInnen einer Erzählung immer „gleichzeitig auch ZuschauerInnen und Mitakteure einer audio-visuellen . . . Inszenierung“ (König & Oloff, 2018a, S. 303) der narrativen Rekonstruktion sind. Für die hier diskutierte Variante von Erzählungen lassen sich die Beteiligungsrollen dahingehend spezifizieren, dass die RezipientInnen, indem sie sich durch ihre sprachlichen Äußerungen und verkörperten Handlungen als ZuhörerInnen, BeobachterInnen des Smartphonegebrauchs und – in einigen Fällen – als LeserInnen der archivierten Textnachrichten positionieren, konstitutiv zur Hervorbringung eines intersubjektiv verständlichen, smartphone-assistierten *tellings* beitragen. Die Story-ProduzentInnen wiederum agieren als ErzählerInnen, Smartphone-BedienerInnen und – insbesondere im Hauptteil der Geschichte – als VorleserInnen der archivierten Textnachrichten.

Dies verdeutlicht, dass es sich bei *smartphone-assistierten Geschichtenerzählungen* um eine multimodale Praktik handelt, die durch das komplexe Zusammenspiel sprachlicher, verkörperter, räumlicher und materieller Ressourcen hervorgebracht wird (vgl. König & Oloff, 2018b, S. 209). Die Analyse dieser grundlegenden Multimodalität der Praktiken *smartphone-assistierter Geschichtenerzählungen* möchte ich im Folgenden anhand ‚unvorbereitet‘ eröffneter Erzählungen vertiefen. Die Untersuchung dieser Variante verdeutlicht zum einen erneut, wie sehr die Hervorbringung *smartphone-assistierter Geschichtenerzählungen* von der *Affordanz* der mobilen Endgeräte als *Archive sozialer Originale* geprägt ist. Zum anderen werde ich aufzeigen, wie die TeilnehmerInnen die Ermöglichung eines visuellen Zugriffs zum Smartphonedisplay methodisch zur Bearbeitung spezifischer Aufgaben der Interaktionsorganisation einsetzen.

## 5.2 Das Vorlesen von Textnachrichten in ‚unvorbereiteten‘ Geschichtenerzählungen und das Auftreten von interaktionalen Problemen

Im Fokus des vorangegangenen Unterkapitels standen *smartphone-assistierte Geschichtenerzählungen*, die von der aktuellen SprecherIn methodisch in das Gespräch eingeführt werden: Durch die Hervorbringung einer Prä-Sequenz werden sie nicht nur eindeutig vom vorangehenden *turn-by-turn-talk* abgegrenzt, sie dient auch dazu, die Verfügbarkeit einer Geschichtenerzählung und ihre thematische Verbundenheit zum Gespräch für andere Anwesende *accountable* zu machen. Im Folgenden werde ich eine weitere Variante von Geschichtenerzählungen analysieren, die durch das Vorlesen von Textnachrichten gestützt werden. Dabei nutzt die GeräteinhaberIn ihre diskursive Identität (Zimmerman, 1998) als ZuhörerIn, um die *Affordanzen* des Smartphones auf spezifische Weise als Ressource für die Hervorbringung einer Geschichtenerzählung zu nutzen: Sie behandelt es als *Archiv sozialer Originale* vergangener Kommunikationsverläufe, die jederzeit ‚zur Hand‘ sind, spontan aufgerufen und unvorbereitet – d.h. ohne methodische Einführung mittels einer Prä-Sequenz – zum Thema der F2F-Interaktion gemacht werden können (vgl. Kotthoff, 2017, S. 13). Im Gegensatz zu methodisch vorbereiteten Geschichtenerzählungen wird diese Variante durch einen regulären Redebeitrag initiiert und „erst in der Folge zu einer Erzählung ausgearbeitet“ (König & Oloff, 2018a, S. 303). Damit wird die Art der Initiierung der *smartphone-assistierten Geschichtenerzählung* und ihr Verhältnis zum *turn-by-turn-talk* zum Differenzkriterium der in diesem Kapitel diskutierten Variante.

Anhand der folgenden Analysen wird deutlich werden, dass die GeräteinhaberInnen ihre Smartphones als *Archive sozialer Originale* behandeln, deren Inhalte ihnen hochgradig vertraut sind. Die gespeicherten, text-basierten Interaktionen sind Teil ihrer fortlaufenden Soziabilität mit Anderen und erhalten ihre Relevanz im Zusammenhang mit anderen (medial-vermittelten und kopräsenten) Interaktionen (vgl. Brown et al., 2018, S. 2; O’Hara et al., 2014). Dies kann dazu führen, dass sie von den SmartphoneinhaberInnen im Kontext von Gesprächen als bedeutungsvoll erachtet werden und bspw. eine Äußerung einer GesprächspartnerIn oder die Erinnerung an ein früheres Thema der Unterhaltung sie dazu veranlasst, einen ‚passenden‘ Nachrichten-Thread aufzurufen und sprachlich zu thematisieren.

Wie im Verlauf der Analysen zu zeigen sein wird, kann eine unvorbereitete Integration von Textnachrichten andere Anwesenden jedoch vor einige Herausforderungen stellen, die für die GeräteinhaberInnen scheinbar so ‚offensichtliche‘ Verbindung der Medieninhalte zum vorangehenden Gespräch und ihre Relevanz für die lokale Interaktion zu erkennen. Die unvorbereitete Eröffnung *smartphone-assistierter Geschichtenerzählungen* macht diese anfällig für das Auftreten von Problemen des Verstehens und Missverständnissen hinsichtlich der Bedeutung

der sequenziellen Abfolge („sequential implicativeness“, Schegloff, 1987c, S. 206), die von den TeilnehmerInnen interaktiv bearbeitet werden müssen.

Um die sequenzielle Struktur und spezifischen Aufgaben der Interaktionsorganisation dieser Variante *smartphone-assistierter Geschichtenerzählungen* rekonstruieren zu können, werde ich auf den Ergebnissen des ersten Analysekapitels aufbauen und folgenden, forschungsleitenden Fragen nachgehen:

- Auf welche Weise und an welchen sequenziellen Positionen erfolgt die Initiierung des Smartphonegebrauchs in unvorbereiteten Geschichtenerzählungen?
- Welche spezifischen Herausforderungen und Folgeprobleme entstehen dadurch, dass die Erzählungen ohne elaborierte sprachliche Vorbereitung initiiert werden, mit der thematische Verbundenheit des *tellings* zum vorangehenden Gespräch *accountable* gemacht wird?
- Wie werden die dadurch auftretenden Verstehensprobleme und Missverständnisse in der Bedeutung der sequenziellen Abfolge interaktiv bearbeitet?
- Welche Relevanz besitzt die methodische Ermöglichung eines visuellen Zugriffs zum Smartphonedisplay als Ressource in der Bearbeitung von Verstehensproblemen und Missverständnissen?

### **5.2.1 Die ‚unvorbereite‘ Eröffnung smartphone-assistierter Geschichtenerzählungen und die Bearbeitung von Folgeproblemen**

Beginnen möchte ich mit der Analyse eines Ausschnitts aus einem Gespräch zwischen den KommilitonInnen Francesca (F) und Marc (M). Die beiden unterhalten sich zu Beginn der Aufnahme über ihr Verständnis von Freundschaft: F äußert dabei ihre Wertschätzung für einen Freundeskreis, in dem Belangloses und Lustiges besprochen wird, was ihr ein Abschalten vom stressigen Alltag ermöglicht. M betont hingegen, wie wichtig ihm tiefgründige und ernste Gespräche mit seinen engsten Freunden sind. Während der Behandlung dieses Themas beginnt F aus ihrer Rolle als ZuhörerIn heraus die Bedienung ihres Smartphones – zunächst jedoch ohne diese divergierende Nutzung sprachlich zu kommentieren (vgl. DiDomenico et al., 2018, S. 9). Erst nach einer Weile der Bedienung fordert M einen *Account* für die Beschäftigung mit dem mobilen Endgerät ein. In Erwiderung darauf kommt es zu einer expliziten Thematisierung von erhaltenen Textnachrichten in der sprachlichen Interaktion (vgl. DiDomenico et al., 2018, S. 19 ff.) und einer unmittelbar anschließenden initialen Zeigesequenz (Oloff, 2019a), mit der die SmartphoneinhaberIn dem Rezipienten visuellen Zugriff zu den Nachrichten eröffnet. Dies

veranlasst die kommunikative Aneignung (vgl. Keppler, 2019, S. 180) der Textnachrichten eines gemeinsamen abwesenden Bekannten.

Während dieser kommunikativen Aneignung ruft die SmartphoneinhaberIn F nun aus ihrer Rolle als ZuhörerIn einen weiteren archivierten Nachrichten-Thread auf, um diesen zum Thema der F2F-Interaktion zu machen. Der folgende Transkriptausschnitt setzt kurz vor der Eröffnung dieser zweiten *mobile-supported sharing activity* ein, und zeigt, wie diese unvorbereitet von der GeräteinhaberIn initiiert und im Interaktionsvollzug zu einer *smartphone-assistierten Geschichtenerzählung* entwickelt wird. Da beide TeilnehmerInnen während der Aneignung der Textnachrichten des abwesenden gemeinsamen Freundes bereits auf das Display des Smartphones von F schauen (#1), wird die visuelle Fokussierung des Displays zur besseren Lesbarkeit des Transkripts als „home position“ (vgl. Sacks & Schegloff, 2002) behandelt: D. h., dass die Blicke der TeilnehmerInnen nur dann annotiert werden, wenn sie vom Display gelöst werden. Die Blicke von F werden mit „+“ und ihre Gesten mit „\*“ annotiert; die Blickführung von M wird durch „%“, seine Gesten werden durch „&“ kenntlich gemacht.

#### Transkript 7: 181008 „Lästerei“\_Teil 1



```

#1
075  M:  #wir [ sind ALle nicht dumm]
076  F:      [ war*te (.) wir ham ]
      >>scrollt>*

077  M:      in der gru[ppe; ]
078  F:      [o:*ke] +wir
      ham auch_n bisschen gel+ä*stert=
                                     *...dreht
SP nach L-----, *...dreht SP zu M-->109
                                     +zu
M-----+
079  =+also_s_is *&# das_is +m*ein MÄ#dels *chat; *
      *LH zeigt zu SP*                *swipe R*
      +..Blick M-----,,,,+,
m:      &....lehnt sich nach vorne--->086
      #2                                #3

```



haben. Dabei dreht sie ihr Smartphone zu Beginn der Äußerung stärker zu sich und blickt zu M. Mit dem Ende von „gelästert“ dreht sie es wieder zu ihm und schaut selbst auf das Display. Im schnellen Anschluss expliziert sie in Z. 079 durch eine präzisierende Selbstreformulierung (Deppermann & Helmer, 2013, S. 8), das auf dem Display nun ihr „Mädelschat“ zu sehen ist<sup>74</sup>. Während der Äußerungsproduktion schaut F zunächst zu M, zeigt dann synchron zum Neustart des Turns („also\_s\_is das\_is“) mit ihrer rechten Hand auf das Gerätedisplay (#2) und richtet ihren Blick wieder auf den Bildschirm. Diese multimodale Aufforderung zur Displayfokussierung (Oloff, 2019a, S. 214) veranlasst M dazu, sich nach vorne zu beugen (#2) und erneut den Bildschirm des Smartphones zu betrachten. Dieses beobachtbare *Monitoring* des Displays durch den Rezipienten wird von der Smartphoneinhaberin auch in diesem Fall als verkörperte Alignierung mit der unvorbereitet eröffneten, smartphone-assistierten Aktivität behandelt und diese im Folgenden fortgesetzt.

F produziert in Z. 080 eine multimodale Demonstration, bei der der deiktische Verweis, dass es „dann so“ in ihrem „Mädelschat“ läuft, von ihrem sichtbaren Scrollen durch den archivierten Chat begleitet wird, sodass die auf dem Display angezeigten Nachrichten ‚typische Aktivitäten‘ im „Mädelschat“ exemplifizieren. Im Anschluss an die kurze Stille in Z. 081, in der F weiter das Gerät manipuliert und M das Display aufmerksam betrachtet, produziert der Rezipient in Z. 082 mit fragender Intonation eine Teilwiederholung von F’s Äußerung („mädelschat“) und zeigt so ein Verstehensproblem mit dem verwendeten Referenten an (vgl. Schegloff, 1987c, S. 203). Dies veranlasst F dazu, ihre fast simultan begonnene Äußerung in Z. 083 abzurechnen und – nach einer erneuten Gesprächspause – eine Antwort auf M’s Frage zu formulieren (Z. 085). Sie unternimmt einen Reparaturversuch in dritter Position, indem sie die Aktivitäten der Gruppe beschreibt: Es werden Bilder verschickt und über diese gelästert. In Überlappung zum Ende der Erklärung beginnt M leise zu lachen (Z. 086), worauf F mit einem lauterem Lachen reagiert. M’s Ratifikation („ja“) in Z. 088, mit der er sein Verständnis anzeigt, schließt die Reparatursequenz.

Der sequenzielle Kontext, indem die Smartphoneinhaberin hier eine zweite *mobile-supported sharing activity* initiiert, ist vergleichbar zum ersten Fall des vorangegangenen Kapitels („was mich abgefickt hat“), denn auch hier wird das Gerätedisplay bereits von beiden TeilnehmerInnen fokussiert. Die Integration neuer Textnachrichten findet hier jedoch unvorbereitet statt und

---

S. 108). Wie sich anhand des Gesprächsverlauf zeigt, ist dem Rezipient M die auf dem Foto abgebildete Person jedoch nicht bekannt, sonst wäre sie anders in das Gespräch eingeführt worden.

<sup>74</sup> Mithilfe dieser Formulierung macht die Smartphoneinhaberin sprachlich *accountable*, dass ihre Verwendung von „wir“ in den Z. 076 und 078 sich nicht auf den gleichen Referenten wie M’s „wir“ in Z. 075 bezieht – und zeigt somit die Aussetzung des Prinzips der lokalen Kohärenz an.

wird von der Smartphoneinhaberin nicht methodisch vorbereitet. Stattdessen nutzt sie die *Affordanzen* des Smartphones, um aus ihrer diskursiven Identität als ZuhörerIn mit einigen Bedingtesten einen weiteren archivierten Chat aufzurufen. Sie versucht diesen zunächst dadurch ins Gespräch einzuführen, dass sie in Überlappung eine „Warte“-Aufforderung<sup>75</sup> produziert, mit der sie den Rezipienten dazu anhält, die Progression der aktuell im Vollzug befindlichen Sequenz zu suspendieren. Als dieser seinen Turn jedoch unbeirrt fortsetzt, initiiert sie den Themenwechsel im Anschluss an ein *acknowledgment token*. Die Eröffnung einer weiteren smartphone-gestützten Zeigesequenz wird durch eine *multimodale Gestalt*, bestehend aus sprachlichen Äußerungen, Blickführung und der Manipulation des Geräts, *accountable* gemacht. Eine zentrale Rolle spielt dabei der visuelle Zugriff des Rezipienten zum Display – dieser wird in Z. 078 zunächst eingeschränkt und mit der sprachlichen Formulierung der relevanten Aktivität („gelästert“) erneut eröffnet. Auffällig ist, dass dabei keine Deplatzierungsmarkierungen (vgl. Schegloff & Sacks, 1973, S. 319–320) eingesetzt werden, die erkennbar machen, dass das Prinzip der lokalen Kohärenz außer Kraft gesetzt werden soll. Vielmehr wird sprachlich zunächst eine kohäsive Beziehung der neuen Zeigesequenz zum lokalen Kontext angezeigt („wir ham auch“). Dass die deiktische Referenz („wir“) in Z. 078 im aktuellen sequenziellen Kontext der Interaktion dabei potenziell missverständlich ist, wird durch die im unmittelbaren Anschluss produzierte, retrospektiv orientierte Selbstreformulierung und das Zeigen auf die Nachrichten bearbeitet (Z. 079).

Wie auch im vorangegangenen Kapitel wird die Einnahme einer verkörperten epistemischen und instrumentellen Haltung (C. Goodwin, 2007a, S. 61 f.) des Rezipienten, die ihm das *Monitoring* des Displays ermöglicht, als Alignierung mit der smartphone-assistierten Sequenz behandelt. Anhand des Interaktionsverlaufs zeigt sich jedoch, dass „der visuelle Zugang zum Display und die Möglichkeit, den Inhalt zu bewerten, einige Zeit in Anspruch nehmen können“ (Oloff, 2019a, S. 215): Der Rezipient macht durch die Produktion einer reparatur-initiierten Teilwiederholung (Z. 082) erst mit einiger Verzögerung *accountable*, dass das Design der *multimodalen Gestalt*, mit der das erneute Zeigen eröffnet wurde, für ihn kein hinreichendes Verständnis der integrierten Chatkommunikation ermöglicht. Dieser Verlauf illustriert zum einen, dass aufgrund der unvorbereiteten Eröffnung der smartphone-assistierten Erzählung Probleme des Verstehens dadurch entstehen können, dass die thematischen Bezüge zum vorangegangenen Gespräch und die Referenten indexikaler Bezeichnungen nicht intersubjektiv verständlich

---

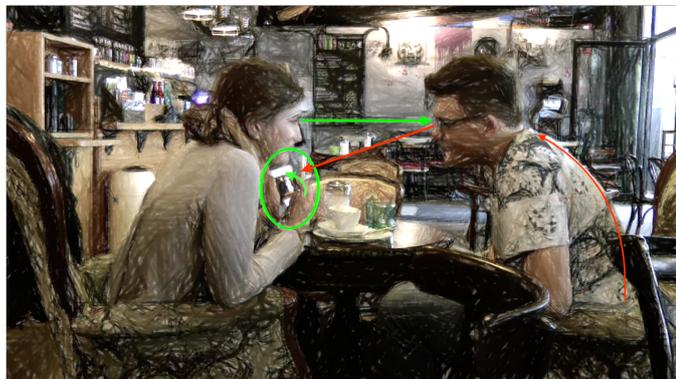
<sup>75</sup> „Warte“-Aufforderungen werden auch von Keppler (2019, S. 179) untersucht. Im vorliegenden Fall dient diese Aufforderung jedoch weniger dazu, prospektiv *accountable* zu machen, dass die SmartphoneinhaberIn sich zum Zweck der Suche nach einem Medieninhalt aus dem Gespräch zurückzieht und für den Zeitpunkt der Beendigung des Suchprozesses das Rederecht beansprucht.

sind. Zum anderen verweist er darauf, dass mit der diskursiven Identität des Betrachtenden die Verantwortung einhergeht, Probleme im Verständnis der eröffneten smartphone-gestützten Erzählung sprachlich anzuzeigen. Die Reparatur-Initiierung führt zu einer Suspension der im Vollzug befindlichen Aktivität und zur Eröffnung einer Einschubsequenz, die der Reparatur der Problemquelle dient. Dass die smartphone-gestützte Sequenz erst fortgesetzt wird, als der Rezipient die Erklärung der Erzählerin ratifiziert, zeigt, dass die TeilnehmerInnen die Wiederherstellung eines intersubjektiven Verständnisses gegenüber der Fortsetzung der smartphone-gestützten Erzählung priorisieren (vgl. Kamunen, 2019, S. 423).

Nach der erfolgreichen Reparatur der Problemquelle beginnt die Smartphoneinhaberin im folgenden Ausschnitt nun, die narrative Rekonstruktion der „Lästerei“ durch eine Elaboration kontextualisierender Hintergrundinformationen zu entwickeln:

### Transkript 8: 181008 „Lästerei“\_Teil 2

089 F: [und] dann ähm (.) \*&das \*wa? \*+(.) #DAS is zum  
 beispiel die die hat\_n abi von null komma acht&  
 bei uns gem[acht-]  
 ----->\* \*tippt\*  
 +....zu M-->093  
 m: &...beugt sich etwas vor-->102  
 #4



090 M: [mh\_hm] oh [ja;]  
 091 F: [die] ist (.) ME:ga (.)  
 intelligent;  
 092 M: O[kay;]  
 093 F: [%die] ist] (.)% VOLL der (.) intelligenz+  
 +[bol ]zen,  
 ----->+  
 +zu Tür-->096  
 m %zu F %  
 094 M [mhm;]  
 095 (0.24)  
 096 F: %aber die s%ieht halt aus+ wie so ne bu?  
 bulgarische PROSti[tuierte (a che)+ he]°h



die Smartphoneinhaberin mehrere explizit positive Beurteilungen der Intelligenz der abgebildeten Person (Z. 091, 093). Diese werden – nach einer kurzen Stille (Z. 095) – mithilfe einer Kontraststruktur (Smith, 1978, S. 38–47) einer Beschreibung des Aussehens der Schulkameradin gegenüberstellt: Diese sähe aus wie eine „bulgarische Prostituierte“ (Z. 096). Der zu M gewandte Blick, der initiale Abbruch dieser Kategorisierung sowie das gemeinsame Lachen beider GesprächsteilnehmerInnen zum Ende der Äußerung (Z. 097–098) zeigen dabei eine implizite Orientierung an der ‚Ungehörigkeit‘ der verwendeten Kategorisierung sowie ein gewisses Vergnügen am „Spielen mit verpönten Ausdrucksweisen . . ., die den guten Ton verletzen“ (vgl. Bergmann, 1998, S. 120) an. Dass M sich auf die Lästerei einlässt, zeigt sich zunächst an seiner sprachlichen Zustimmung zu der potenziell ‚beschwerdewürdigen‘ Kategorisierung (Z. 098) und seiner anschließenden Einschätzung (Z. 102), dass nur eine „Plüschjacke mit Fell“ fehle, um den Eindruck einer „Prostituierten“ zu vervollständigen.

Während M die Interpretation der Fotografie ausbaut und sich dabei wieder aufrechter hinsetzt (#5), versucht die Smartphoneinhaberin wiederholt, das Rederecht wiederzuerlangen („und“, Z. 099, 101). Dies gelingt ihr jedoch erst, als sie in Überlappung zu M’s Äußerung (Z. 102) mit „hier so“ in turn-initialer Position eine erneute Aufforderung zur Aufmerksamkeitsfokussierung (Oloff, 2019a, S. 214) und nach einem „Herausplatzer“ (vgl. Goffman, 1978, S. 813) („alter“) zwei als W-Fragen formulierte Äußerungen realisiert (Z. 103, 104). Diese lenken die Aufmerksamkeit des Rezipienten auf die „Oberweite“ der abgebildeten Person. Mit kleiner Verzögerung zu dieser Aufforderung lehnt sich M erneut Richtung Smartphone und fokussiert das Display (#6). Dass die syntaktisch und prosodisch als Fragen produzierten Äußerungen nicht als Handlungen behandelt werden, die eine (ausgearbeitete) Antwort verlangen, sondern als Methoden, die Aufmerksamkeit des Rezipienten zu rekrutieren und eine Verifikation zu veranlassen (vgl. C. Goodwin, 1987, S. 124), zeigt sich an der Produktion von M’s *continuer* (Z. 105): Mit diesem ratifiziert er die ‚Wahrnehmungswürdigkeit‘ der „Oberweite“ der abgebildeten Person, überlässt der Smartphoneinhaberin aber nun wieder das Rederecht (ebd., 1987, S. 124) zur Fortsetzung der smartphone-assistierten Erzählung.

Der erste Ausschnitt zeigte, wie es im Anschluss an die unvorbereitete Eröffnung einer smartphone-assistierten Zeigesequenz zunächst zu einem Problem des Verstehens kommt, das durch die Hervorbringung einer Einschubsequenz repariert wird. Im zweiten Ausschnitt findet nun eine beobachtbare Orientierung der TeilnehmerInnen an der Multimedialität der archivierten IM-Kommunikation statt, die durch das Zusammenspiel schriftsprachlicher und visueller Elemente (hier einer Fotografie) konstituiert wird: Die Smartphoneinhaberin integriert die Fotografie einer abwesenden Klassenkameradin als visuellen Medieninhalt, der eine prospektive

Relevanz für die initiierte Geschichtenerzählung besitzt. Im Verlauf der gemeinsamen kommunikativen Aneignung<sup>76</sup> kommt es jedoch zu einem Missverständnis hinsichtlich der Bedeutung der sequenziellen Abfolge (Schegloff, 1987c, S. 206)<sup>77</sup> der begonnenen *mobile-supported sharing activity*. Es zeigt sich, dass aufgrund der unvermittelten Eröffnung der smartphone-assistierte Sequenz zwischen den TeilnehmerInnen noch kein intersubjektiv geteiltes Verständnis hinsichtlich der erwarteten Erwidern des Rezipienten etabliert wurde.

M behandelt das Zeigen der digitalen Fotografie als Aktivität ‚in eigenem Recht‘, die seine Einschätzung des Bildes konditionell relevant macht: Er produziert zunächst affiliative Lachpartikel (Z. 097) sowie eine explizit zustimmende Äußerung (Z. 098) und anschließend seine Interpretation des Aussehens der Schulkameradin (Z. 102). Anhand des Umgangs der Smartphoneinhaberin mit den Erwidern des Rezipienten wird jedoch erkennbar, dass F das Zeigen und Kontextualisieren der Fotografie als Vorarbeit (vgl. Schegloff, 1980, S. 113) für eine *smartphone-assistierte Geschichtenerzählung* behandelt: Wiederholt versucht sie, das Rederecht zu übernehmen und produziert dann in Überlappung eine Aufforderung zur Aufmerksamkeitsfokussierung. Sie nutzt ihre exklusiven kommunikativen Rechte als Smartphoneinhaberin (Keppler, 2013, S. 95), um zu verhindern, dass das Interpretationsangebot des Rezipienten zum Fokus der sprachlichen Interaktion werden kann (M. H. Goodwin, 1997, S. 98–99), und seine Aufmerksamkeit auf einen spezifischen Aspekt der Fotografie (die „Oberweite“ der abgebildeten Person) zu lenken. Nach dieser kompetitiven Übernahme des Rederechts zeigt der Rezipient durch die Produktion eines *continuers* sein Verständnis davon an, dass sich die Produktion eines *multi-unit-turns* im Vollzug befindet, und trägt zu dessen Hervorbringung kollaborativ bei, indem er auf die Übernahme des Rederechts verzichtet (vgl. C. Goodwin, 1986, S. 207).

Dies ermöglicht der Smartphoneinhaberin im nun folgenden Abschnitt die Ausgestaltung des Hauptteils einer narrativen Rekonstruktion, mit der sie die typische Kommunikation in ihrem „Mädelschat“ illustriert (vgl. Schwitalla, 1991, S. 189) und die erneut konstitutiv durch die Praktiken des Vorlesens von Textnachrichten verfertigt wird:

---

<sup>76</sup> Die Integration eines visuellen Medieninhalts, der in den archiviertem Thread eingebettet ist (Brown et al., 2018, S. 13), unterscheidet diesen Fall von den bis hierhin diskutierten Exemplaren *smartphone-assistierter Geschichtenerzählungen*. Hier zeigt sich bereits deutlich, dass „offering visual access to a shared image entails a layer of interpretation [...] is of clear interactional consequence to participants“ (Raclaw et al., 2016, S. 377). Ich werde die Prozesse der kommunikativen Aneignung visueller Medieninhalte an dieser Stelle jedoch (noch) nicht als ‚Phänomen in eigenem Recht‘ behandeln, da hier der Zusammenhang zwischen der unvorbereiteten Smartphoneintegration und Problemen in der Bedeutung der sequenziellen Abfolge im Zentrum der Analyse steht. Ich werde mich der Integration digitaler Fotografien in *smartphone-assistierte Geschichtenerzählungen* ausführlich in Kap. 6 widmen.

<sup>77</sup> Die Problematik bei Missverständnissen in der sequenziellen Abfolge besteht in den Implikationen eines Turns bzw. einer ganzen Äußerung: „the issue is what action is being done by/through some turn or turn component, and accordingly what type of talk/action is relevant or appropriate next“ (Schegloff, 1987c, S. 206).

Transkript 9: 181008 „Lästerei“\_Teil 3

106 F: ähm (-)\* und (.) \* un\_so (.) WOW sieht aus als wär  
 <<:-)> sie so ne ukrainerin die mit ihren freunden  
 in\* so russen dis[kos \*geht >,  
 ---->\*scrollt \*-----fährt mit R Zeigefinger von  
 links nach rechts über Textzeilen-----  
 --\*scrollt--->110

107 M: [he ch ] ch ch;  
 108 F: <<:-)> STIMMT\_s?>  
 109 he he he °h=  
 110 =<<:-)> dE dE >\* OHne witz;  
 ---->\*

111 <<:-)> \*#eh h h a\*+lso #SO->&  
 \*wischt nach R\*  
 +...zu M--->112

m: ////////////////--->&  
 #7 #8



112 M: <<p>mh h hi> hi+ hi  
 f: ----->+

113 F: so richtig (.) \*mä[dels ]  
 \*scrollt-->116

114 M: [mädels]GRUPpen;  
 115 F: so RIChtig  
 mädels talk halt;  
 116 (0.5)\*(2.5)\*(0.38)  
 f: \*dreht SP zu M-->>

117 F: oder das (.) \*DAS ist (.) das is cleo das\* ne  
 <<:-) \*freundin von \* \*(mir)-  
 \*..---zeigt auf Foto----, , , , \*  
 \*wischt nach R\* \*zeigt auf Foto --->>

F's Äußerung in Z. 106 beginnt erneut mit Formulierungsarbeit, während der sie die sprachliche Äußerung mit der durch scrollende Bediengesten durchgeführten Suche nach den relevanten Textnachrichten synchronisiert. Mit der sichtbaren Beendigung der Gerätemanipulation beginnt sie mit „un\_so“ die wörtliche Wiedergabe der Textnachrichten aus ihrem „Mädelschat“,

die als Erwidernngen auf die digitale Fotografie versendet wurden. Der verwendete Zitationsmarker („un\_so“) erscheint dabei als fragmentierte Version des „und ich so/er so“-Formats (Golato, 2000), identifiziert jedoch nicht namentlich die Autorin der originalen Textnachricht. Das Vorlesen der Textnachrichten wird auch in diesem Fall durch das Zusammenspiel sprachlicher Äußerungen und die Einnahme einer spezifischen Körperhaltung als solches *accountable* gemacht. Darüber hinaus erscheint für die Erkennbarkeit des Vorlesens konstituierend, dass F während der Äußerungsproduktion mit ihrem rechten Zeigefinger von links nach rechts und oben nach unten über den Bildschirm des mobilen Geräts fährt (ohne diesen zu berühren). Auf diese Weise scheint sie dem Text der Nachrichten zu folgen, während sie den Wortlaut lautsprachlich realisiert. Die Fingerbewegungen fungieren somit als „environmentally coupled gestures“ (C. Goodwin, 2007b), die auf dem Bildschirm diejenigen Elemente relevant und für den Rezipienten nachvollziehbar machen, die sprachlich von der Smartphoneinhaberin wiedergegeben werden.

Die erste wiedergegebene Nachricht enthält eine elaborierte Interpretation des Äußeren der abgebildeten Person, die aussähe wie eine „ukrainerin die mit ihren freunden in so rUssendiskos geht“ (Z. 106). Diese zugeschriebene Zugehörigkeit zu einer ethnischen Gruppe und Unterstellung der Teilnahme an einer stereotypisierenden Aktivität<sup>78</sup> wird von der Erzählerin mit „smile-voice“ (Selting et al., 2009, S. 367) vorgetragen und die „laughability“<sup>79</sup> dieser Kategorisierung von M durch sein Lachen (Z. 107) ratifiziert. Die nächste, ebenfalls in *smile-voice* wiedergegebene Nachricht (Z. 108) scheint von derselben Autorin verfasst und fordert die Chatteilnehmerinnen durch „STIMMT\_s?“ zur Verifikation dieser Interpretation auf. Nach ihrem Lachen in Z. 109 gibt die Erzählerin eine Nachricht wieder, die durch „de de“ (Z. 110) auf die ‚Körbchengröße‘<sup>80</sup> der abgebildeten Person Bezug nimmt, und der mit einem nachdrücklichen „ohne Witz“ eine äußerungsmodalisierende Kommentarphrase (Torres Cajo, 2017) nachgestellt ist. Auch in diesem Fall findet also keine neutrale Wiedergabe der Textnachrichten statt – viel mehr zeigt die Smartphoneinhaberin durch prosodische Modulation (*smile-voice*) und Lachpartikeln

---

<sup>78</sup> Diese Geschichtenerzählung stellt einen spannenden Fall für eine „Membership Categorization Analysis“ (MCA) dar. Innerhalb dieses von Sacks (1972) entwickelten Programms wird die „in situ work displayed and deployed in members’ organization and animation of taken-for-granted social categories and their related actions“ (Fitzgerald, 2012, S. 306) untersucht. Die situative Identifikation der fotografisch repräsentierten Person als Angehörige einer ethnischen Kategorie und die Verknüpfung dieser mit einer „category-bound activity“ (Sacks, 1972; Schegloff, 2007b, S. 470–471) scheint hier den Anlass für das Amusement der TeilnehmerInnen darzustellen. Die interaktive Organisation dieses als selbstverständlich angenommenen Wissens über soziale Kategorien und ihrer Eigenschaften und Tätigkeiten kann und soll hier jedoch nicht selbst zum Gegenstand der Analysen werden.

<sup>79</sup> Auch für diesen konversationsanalytischen Begriff existiert keine ‚zufriedenstellende‘ deutsche Übersetzung. König und Oloff (2018b, S. 288) folgend verwende ich deshalb den englischen Begriff, um die Eigenschaft zu bezeichnen, dass etwas zum ‚Lachanlass‘ werden kann.

<sup>80</sup> „Doppel D“ – eine große Körbchengröße.

ihre Haltung an und legt dem Rezipienten so ein humoristisches Verständnis des rekonstruierten Chatverlaufs nahe.

Die Ausleitung der *smartphone-assistierten Geschichtenerzählung* erfolgt auch hier durch das Zusammenspiel sprachlicher Äußerungen und verkörperter Handlungen der Smartphoneinhaberin: In Z. 111 lacht F zunächst verhalten und produziert anschließend mit „also SO“ einen „return home“ (Jefferson, 1978, S. 232), der eine verkürzte Reformulierung ihrer Ankündigung aus Z. 080 („so läuft das bei uns ab im Mädelschat“) darstellt. Während der Äußerungsproduktion schaut sie zunächst noch auf ihr Smartphone (#7), verändert die Displayanzeige durch eine wischende Bediengeste und blickt dann vom Display zum Rezipienten (#8). In der anschließenden *reponse sequence* erfolgt die Würdigung der Erzählung durch ein zunehmend lauter werdendes Lachen von M (Z. 112), wobei F schon währenddessen wieder beginnt, auf ihr Smartphone zu schauen. Die Beendigung der Sequenz erfolgt in den Z. 113–115, in denen F eine Äußerung beginnt („so richtig (.) mädels“), die M in Z. 114 in Überlappung versucht zu komplettieren („mädelsGRUPpen“). Die vorgeschlagene Kategorisierung wird jedoch von der Erzählerin nicht angenommen, stattdessen produziert sie einen recycelten Turn-Anfang und formuliert ihr Fazit der Geschichte („mädelsstalk“), die als alternative Bezeichnung des anfangs verwendeten „Mädelschats“ verstanden werden kann. Anschließend entsteht eine mehrsekündige Pause (Z. 116), in der F das Display des Smartphones erneut stärker zu M dreht. Ihre nächste Äußerung in Z. 117 macht durch ein kontrastierendes „oder“ in der turn-initialen Position sowie die anschließende Erklärung, wer auf dem Display zu sehen ist, ersichtlich, dass sie eine weitere *smartphone-assistierte Zeigesequenz* eröffnet, die die Chatkommunikation mit ihren Freundinnen weiter exemplifiziert.

Dieser erste Fall einer unvorbereitet eröffneten, *smartphone-assistierten Geschichtenerzählung* ist in eine Serie von *mobile-supported sharing activities* eingebettet. Die Smartphoneinhaberin nutzt die *Affordanzen* des Smartphones als *Archiv sozialer Originale*, um mittels einiger Bediengesten einen gespeicherten Chatverlauf aufzurufen, und ihre exklusiven kommunikativen Rechte, um einen Themenwechsel herbeizuführen. Aufgrund der – für den Rezipienten – unzureichenden thematischen Vorbereitung der Integration der Textnachrichten müssen jedoch im Vollzug der Interaktion Verstehensprobleme und Missverständnisse in der Bedeutung der sequenziellen Abfolge bearbeitet werden, bevor die Geschichte auf bekannte Weise ausgestaltet werden kann:

Zunächst werden in einer Reparatursequenz Probleme im Verständnis eines Referenten bearbeitet, anschließend müssen Missverständnisse hinsichtlich der sequenziellen Abfolge ausgeräumt werden. Diese entstehen dadurch, dass die Smartphoneinhaberin eine Fotografie in die

Interaktion einführt, deren Kontextualisierung als Hintergrunddetails für die Rekonstruktion der in Erwiderung darauf versendeten Textnachrichten dienen soll. Die Integration der Fotografie dient vor allem dazu, die lokale Kohärenz der originalen Chatnachrichten *accountable* zu machen und die Aufmerksamkeit des Rezipienten auf diejenigen Aspekte der Fotografie zu lenken (‚Kleidungsstil‘ und ‚Oberweite‘), an denen sich auch die TeilnehmerInnen der Chatkommunikation in den archivierten Textnachrichten orientieren. Der Rezipient behandelt das Zeigen des Bildes jedoch zunächst als eigenständige Aktivität, die seine Interpretationen konditionell relevant macht, sodass die Smartphoneinhaberin das Rederecht wieder an sich nehmen muss, um die Geschichte entwickeln zu können.

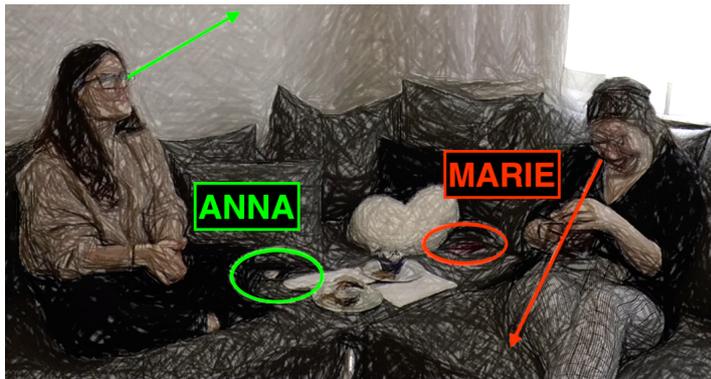
Die Elaboration dieser *smartphone-assistierten Geschichtenerzählung* erfolgt auch in diesem Fall wieder durch Praktiken des Vorlesens von Textnachrichten, die einen exemplarischen Fall der Chatkommunikation in der „Mädelsgruppe“ der Smartphoneinhaberin illustrieren. Im Vergleich zu den Fällen des vorangegangenen Kapitels wird das Vorlesen der Nachrichten hier jedoch nicht in gleichem Maße bereits an der linguistischen Gesprächsoberfläche als DRS erkennbar, denn es kommen kaum Zitationsmarker zum Einsatz. Die sprachlichen Äußerungen werden viel mehr im Zusammenspiel mit den *environmentally coupled gestures* der Smartphoneinhaberin als wörtliche Zitate *accountable* gemacht: Indem sie den Nachrichten während ihrer lautsprachlichen Realisierung Wort-für-Wort mit ihrem Zeigefinger folgt, macht sie beobachtbar und berichtbar, wessen Nachrichten sie gerade wiedergibt. Konstitutiv für die Erkennbarkeit dieser *multimodalen Gestalt* ist dabei das *Monitoring* des Displays durch den Rezipienten, der sich den Sinn der Zeigegesten nur durch das reflexive Zusammenspiel der sprachlichen Äußerungen mit dem semiotischen Feld des Displays erschließen kann (vgl. C. Goodwin, 2015b, S. 418–419). Die Ausleitung aus der Erzählung erfolgt auch hier über das Zusammenspiel eines sprachlichen *return home*, den Blickwechsel vom Display des Smartphones zum Rezipienten und Bediengesten, die die Bildschirmanzeige verändern. Diese Manipulation der Displayanzeige erfüllt dabei eine Doppelfunktion, denn sie trägt nicht nur dazu bei, die Beendigung der *smartphone-assistierten Erzählung accountable* zu machen, sondern dient gleichzeitig als Vorbereitung der nächsten *smartphone-gestützten* Sequenz, die im unmittelbaren Anschluss eröffnet wird.

### 5.2.2 Der visuelle Zugriff zum Smartphonedisplay als Ressource für die Bearbeitung von interaktionalen Problemen

Anhand der Analyse eines weiteren Falles möchte ich im Folgenden zum einen aufzeigen, dass Verstehensprobleme und Missverständnisse hinsichtlich der sequenziellen Abfolge wiederkehrende Phänomene in unvorbereitet eröffneten, *smartphone-assistierten Geschichtenerzählungen* sind. Diese entstehen dadurch, dass für die RezipientInnen die Inhaltsrelevanz und Verbindung der Nachrichten zum Gesprächskontext nicht hinreichend *accountable* werden. Zum anderen wird weiter untersucht, wie die TeilnehmerInnen diese Probleme des intersubjektiven Verständnisses mithilfe der durch das Smartphone verfügbaren Ressourcen interaktiv bearbeiten. Ziel ist es dabei, die spezifische Funktion des visuellen Zugriffs zum Smartphonedisplay als Ressource für die interaktive Hervorbringung *smartphone-assistierter Geschichtenerzählungen* genauer zu definieren und neben den exklusiven Rechten der SmartphoneinhaberInnen auch ihre Verpflichtungen (vgl. Oloff, 2019a, S. 215) näher zu beleuchten.

Auch in folgendem Fall integriert die Smartphoneinhaberin aus ihrer Rolle als ZuhörerIn heraus das mobile Endgerät ohne elaborierte Vorbereitung in das Gespräch. Im Unterschied zum vorangegangenen Fall wird die Smartphoneintegration hier jedoch durch eine vorangegangene Äußerung den GesprächspartnerIn ausgelöst („triggered“, Jefferson, 1978, S. 220) und ist thematisch zum aktuellen Thema der Unterhaltung verbunden. Aufgrund der unvorbereiteten Eröffnung ist diese Verbindung jedoch für die RezipientIn nicht *accountable*, sodass erneut Verstehensprobleme und Missverständnisse bearbeitet werden müssen. Der Ausschnitt stammt aus einem Gespräch zwischen den zwei Freundinnen Anna (A) und Marie (M), die in einem L-Arrangement (Kendon, 1990, S. 213) auf einer großen Couch sitzen (#1) und sich über die Möglichkeit eines gemeinsamen Besuchs bei der Familie von M in Südtirol unterhalten. Scherzhaft fordert A, dass M den Dialekt der Familie im Falle eines Besuchs für sie ‚übersetzen‘ müsse. Das Transkript setzt ein, als beide GesprächsteilnehmerInnen ihre Einschätzung dazu äußern, wie gut bzw. schlecht A den Dialekt der Familie verstehe. Der Blick von A wird im Transkript mit „\*“ annotiert, die Gestik ihrer linken Hand (LH) mit „+“, die ihrer rechten Hand (RH) mit „†“. Die Blicke von M werden mit „&“ und ihre Gestik mit „%“ kenntlich gemacht.

Transkript 10: 191024 „Gnumma?“\_Teil 1



#1

m: -->>Blick vor sich -->010  
 022 M: °hh n[ee: das ]  
 023 A: [aber wenn ihr ] SCHNELL redet vesteh ich  
 nichts.  
 024 M: nja &aber NEE (.) du verstehst uns& schon;+\*#  
 -->& ...----Blick vorne oben--,,, & A-->25  
 a: \*  
 ..nach links zu SP-->026  
 a: +..  
 .LH Griff SP-->026

#2



025 ju\*LIA is † sowieso#†† german\*istin# °h  
 germanistik st[uden]tin ne †&die scht²† &redet  
 sowieso schön;

026 A: [mmm ]  
 a: -->& SP  
 vor sich-->030  
 a: \* zu M \*zu SP-->>035  
 a: †zeigt auf†+SP vor Körper-->>038  
 †scrollt †

#3

#4



027 (.)  
 028 M: lukas †AU:CH;  
 a: †scrollt-->034  
 029 (1.0)  
 030 M: PAT&rick ist dann nochmal was anderes als  
 südtirolerisch weil er ist ja vo\_INSchbruck;  
 -->& zu A-->040  
 031 A: mmm (-) warTE;  
 032 (2.0)  
 033 M: mh[m ]

In Z. 023 äußert A ihre Einschätzung, dass sie bei hoher Sprechgeschwindigkeit „nichts“ des Dialekts verstehe, woraufhin ihr M in Z. 024 widerspricht. Am Ende von M's Äußerung schaut A nach links auf die Couch und greift mit ihrer linken Hand nach ihrem dort liegenden Smartphone (#2). Dass M sich an ihrem Widerspruch als dispräferierte Erwiderung orientiert, zeigt sich in Z. 024 bereits an der Einleitung mit teilweiser Zustimmung („nja“) sowie an der Abwendung ihres Blicks während der Äußerungsproduktion (vgl. Kendrick & Holler, 2017). Noch ausdrücklicher wird dies, als sie in Z. 025 beginnt, Erklärungen als *Account* für ihre nicht übereinstimmende Einschätzung zu produzieren: In mehreren Äußerungen beschreibt und bewertet M den Dialekt einiger Personen – vermutlich ihrer Familienmitglieder<sup>81</sup>. Während M's erster Einschätzung in Z. 025, in der sie betont, wie „schön“ Julia aufgrund ihres Germanistikstudiums „redet“, hebt A ihre rechte Hand mit ausgestrecktem Zeigefinger in einer ‚Meldegeste‘ (#3) und blickt zunächst zu M, wendet sich dann jedoch der Bedienung ihres Smartphones zu, das sie nun mit beiden Händen vor sich hält (#4). In Z. 026 produziert sie in Überlappung zu „studentin“ zunächst einen *continuer*, mit dem sie – trotz der Initiierung des divergenten Smartphonegebrauchs – ihre Alignierung mit der Unterhaltung anzeigt und die aktuelle Sprecherin animiert, ihre Erklärungen fortzusetzen. Im Anschluss folgen zwei weitere Einschätzungen (Z. 028, 030) von M, die von einer einsekündigen Stille (Z. 029) unterbrochen werden, in der

<sup>81</sup> Auch wenn sich nicht aufklären lässt, um wen es sich handelt, müssen diese Personen beiden Gesprächspartnerinnen bekannt sein, da sie von M lediglich mit ihren Vornamen identifiziert werden und die Verwendung dieser „recognitionals“ als Minimalform der Referenz auf Personen (vgl. Schegloff & Sacks, 1979, S. 17) von der Rezipientin A nicht als problematisch behandelt wird.

sie abzuwarten scheint, ob A das Wort ergreift. In Erwiderung auf M's Äußerung produziert A in Z. 031 zunächst einen weiteren *continuer*, dem sie jedoch nach einer kurzen Pause eine *Warte-Aufforderung* folgen lässt. Diese wird von der Rezipientin als Aufforderung zur Suspension der gerade im Vollzug befindlichen Sequenz behandelt, da es im Anschluss zur Hervorbringung einer Stille von zwei Sekunden (Z. 032) kommt, in der A weiterhin sichtbar mit ihrer rechten Hand scrollende Bediengesten ausführt. In Z. 033 produziert M einen *continuer*, mit dem sie ihre Bereitschaft anzeigt, auf die Übernahme des Rederechts zu verzichten und A die Entwicklung einer smartphone-assistierten Aktivität zu ermöglichen.

Auch in diesem Fall wird der Smartphonegebrauch von der Geräteinhaberin aus ihrer diskursiven Identität als ZuhörerIn initiiert. Die Initiierung erfolgt zunächst unkommentiert (Z. 025) und wird kurz darauf von einer verkörperten Ankündigung (durch eine ‚Meldegeste‘) gefolgt (Z. 026). Eine eventuelle thematische Verbundenheit des Smartphonegebrauchs mit dem Gespräch lässt sich höchstens inferentiell anhand des Timings der Initiierung des Gerätegebrauchs im Verhältnis zum *turn-by-turn-talk* erschließen<sup>82</sup>: Es findet jedoch zunächst weder eine sprachliche Darstellung der Verbindung zum Gesprächskontext durch die Smartphoneinhaberin noch eine beobachtbare Reaktion der Gesprächspartnerin statt. Der sequenzielle Verlauf des Gesprächs zeigt, dass der divergente Gebrauch mobiler Endgeräte in F2F-Interaktionen – selbst, wenn er nicht sprachlich angekündigt wird – nicht per se als rechtfertigungsbedürftig behandelt wird (Oloff, 2019b, Z. 785–786): Vielmehr entsteht vorübergehend ein Setting *intrapersonaler Multiaktivität*, in dem die Smartphoneinhaberin das Gerät bedient, während sie durch *continuer* zunächst ihre fortgesetzte Alignierung mit dem Gespräch anzeigt (Z. 026). Erst nach einer Weile fordert die Smartphoneinhaberin die Sprecherin durch „Warte“ zur Suspension der Progression der aktuellen Sequenz auf und kündigt gleichzeitig an, sich als Gesprächsteilnehmerin zurückzunehmen (Keppler, 2019, S. 180). Die *Warte-Aufforderung* veranlasst die interaktive Hervorbringung von Stille (vgl. Porcheron et al., 2016b, S. 212), die es der Geräteinhaberin ermöglicht, die Manipulation des Smartphones – vermutlich ein Suchprozess – als notwendige Vorarbeit der Integration zu beenden (Suderland, 2019, S. 101). Auf diese Weise wird das Multiaktivitätssetting nun in einer „exclusive order“ (Mondada, 2014a, S. 70)

---

<sup>82</sup> Durch die wiederholte Betrachtung des Videos lässt sich das exakte Timing der Initiierung des Smartphonegebrauchs identifizieren: Der Griff zum Smartphone scheint durch M's Einschätzung, dass A den Dialekt „schon versteht“ (Z. 024) ausgelöst zu werden und anzuzeigen, dass A ‚just dazu‘ etwas einfällt (vgl. Sacks, 1995, Vol II, S. 442). Es ließe sich also aus BeobachterInnenperspektive argumentieren, dass bereits durch das verkörperte Verhalten eine thematische Kohärenz ihrer anstehenden Äußerungen zu M's Einschätzung angezeigt wird (vgl. Morrek, 2018, S. 316). Diese Argumentation erscheint jedoch insofern problematisch, dass sich aufgrund einer fehlenden beobachtbaren Orientierung – bspw. durch sprachliche Bezugnahme oder sichtbares *Monitoring* – der Rezipientin nicht eindeutig feststellen lässt, ob für diese die Initiierung des Smartphonegebrauchs als mit dem Gesprächsverlauf verbunden *accountable* ist.

organisiert, in der das Gespräch zunächst zugunsten der Smartphonebedienung unterbrochen wird. Nach Beendigung des Suchprozesses eröffnet A im folgenden Ausschnitt eine narrative Rekonstruktion, die durch Praktiken des Vorlesens von Textnachrichten gestützt wird:

**Transkript 11: 191024 „Gnumma?“\_Teil 2**

033 M: mh[m ]  
 034 A: [au]f †EINmal kommt- (-)  
 ----, , , , †  
 035 °h hoffentlich (---) n\_bissel: GNUM\*ma; h°  
 -----, , , , ---->\*zu M-->037  
 036 (0.5)  
 037 A: GNUM\*ma?  
 ,-->\* zu SP-->038  
 038 A: und [ich so +#ge\*]  
 039 M: [w\_%was? ]  
 a: ----->\*zu M -->056  
 a: +...streckt LA mit SP zu M-->042  
 m: %.....streckt RA zu A aus-->041  
 040 A: **NOMmen?**  
 #5



041 A: hhe TI&%#na so;  
 m: & SP-Display -->054  
 m: %....greift nach SP-->043  
 #6



042 °h und ich [so +was geht ]  
 ...-----+, , , ,  
 043 M: [du m%usst mir\_s-]  
 %....nimmt SP an-->058  
 044 A: mit EUCH,  
 045 M: du muscht mir des aus\_m KONText;

In Überlappung zu M's *continuer* in Z. 033 eröffnet A die narrative Rekonstruktion eines vergangenen Ereignisses. Eingeleitet wird diese mit einem *prospective indexical* („auf einmal kommt“), mit dem die Erlebnisqualität und damit die Erzählwürdigkeit des Folgenden etabliert wird: Für die Rezipientin wird so zum einen erwartbar, dass die folgenden Ereignisse nicht den ‚Normalitätserwartungen‘ der Sprecherin entsprechen<sup>83</sup>. Zum anderen wird es ihr zur Aufgabe gemacht, herauszufinden, welche Ereignisse eingetreten sind. Während der Äußerungsproduktion beendet A die physische Manipulation des Geräts. Nach einem hörbaren Einatmen folgt in Z. 035 das Vorlesen (von Teilen) einer einzelnen Textnachricht: „hoffentlich (---) n\_bIssel: GNUmma“; gefolgt von einem hörbaren Ausatmen. Die längere Pause nach „hoffentlich“ macht dabei für die Rezipientin erkennbar, welcher Teil dieser Nachricht aus Perspektive der Erzählerin unerwartet ist. Gegen Ende der Äußerung schaut A vom Display zu ihrer Gesprächspartnerin hoch und beobachtet ihre Reaktion. Als diese nicht das Rederecht übernimmt (Z. 036), wiederholt die Smartphoneinhaberin in Z. 037 mit Frageintonation das problematische Element der Nachricht („GNUmma?“) und richtet ihren Blick erneut auf das Smartphone. Anschließend fährt sie mit der narrativen Rekonstruktion der Chatkommunikation fort: Diese wird durch den Zitationsmarker „und ich so“ (vgl. Golato, 2000) als DRS ihrer anschließenden Erwiderung eingeführt, die aus ihrem Verstehensversuch („try-marked“, Schegloff & Sacks, 1979, S. 20) der erhaltenen Nachricht – einem mit Frageintonation realisierten „geNOMMen“ (Z. 037, 039) – besteht. In Überlappung zur DRS dieser Nachricht produziert M in Z. 039 mit „w\_was?“ ein Fragewort, mit der sie anzeigt, Probleme mit dem Verständnis der vorangegangenen Äußerung zu haben<sup>84</sup>. Währenddessen streckt sie ihren rechten Arm zur Smartphoneinhaberin aus (#5). Das Zusammenspiel aus sprachlicher Äußerung und verkörperter Handlung veranlasst A dazu, zu M zu schauen und ihrer Gesprächspartnerin ihr Smartphone zu reichen (#6).

Während die Teilnehmerinnen die verkörperte Übergabe des Smartphones realisieren, setzt A die narrative Rekonstruktion der Chatkommunikation im Modus der DRS fort: In Z. 042 führt sie lachend mit dem Zitationsmarker „Tina so“ eine weitere ‚Autorin‘ in die Erzählung ein,

---

<sup>83</sup> Diese Weise, die dialektal gefärbte Nachricht in die narrative Rekonstruktion einzuführen, scheint eine fragmentierte Version des „I was just doing X when Y“-Formats (vgl. Wooffitt, 1992) darzustellen. Auf diese Weise wird die eigene Involvierung in eine alltägliche Beschäftigung mit dem berichteten Ereignis kontrastiert, das so als ungewöhnlich und damit als erzählwürdig dargestellt wird (vgl. a. Bergmann, 1993, S. 311 f.). In diesem Fall erfolgt jedoch keine Benennung der Aktivität, in die die Erzählerin beim Erhalt der zu rekonstruierenden Nachrichten involviert war.

<sup>84</sup> Aufgrund des Timings der Überlappung der beiden Äußerungen erscheint es nicht plausibel, dass sich M's Äußerung auf A's im Vollzug befindlichen Turn in Z. 038 bezieht: Die Überlappung setzt bereits nach „und“ (Z. 038) ein – einer Stelle, an der weder das Ende der Äußerung noch eine handlungskonstituierende Turnkonstruktionseinheit projizierbar ist.

ohne jedoch deren Nachricht wiederzugeben. Währenddessen nimmt M das Handy entgegen und beginnt, auf das Display zu schauen. Als A durch einen weiteren Zitationsmarker („und ich so“) die Wiedergabe einer weiteren Nachricht des Chats (Z. 042) beginnt („was geht“), produziert M in Z. 043 in Überlappung eine deontisch formulierte Aufforderung („du musst mir\_s). Sie bricht die Turnkonstruktion jedoch zunächst ab und wartet, bis die Erzählerin die Wiedergabe ihrer Antwort („mit EUCH“) beendet hat. Anschließend formuliert M einen dialektal gefärbten, recycelten Turnanfang („muscht“) und macht *accountable*, dass für ihr Verständnis der wiedergegebenen Nachrichten eine weitere Kontextualisierung erforderlich ist (vgl. Keppler, 2019, S. 180).

Dieser Ausschnitt zeigt, wie die Smartphoneinhaberin nach Beendigung des Suchprozesses mit einer einzelnen Äußerung („auf einmal kommt“) eine narrative Rekonstruktion initiiert: Aufgrund der simultanen Handhabung des Geräts lässt sich darauf schließen, dass sich „kommt“ auf das Empfangen einer Nachricht bezieht. Abgesehen von diesem *prospective indexical*, das es zur Aufgabe der Rezipientin macht, im weiteren Verlauf der Interaktion herauszufinden, welches Ereignis nicht den Normalitätserwartungen der Erzählerin entspricht, findet keine Bearbeitung der für die Prä-Sequenz einer Geschichtenerzählung typischen Aufgaben statt: Weder wird das Verhältnis des Folgenden zum vorangegangenen Gespräch methodisch *accountable* gemacht, noch werden Hintergrunddetails (C. Goodwin, 2015a, S. 199), wie der Zeitpunkt des Ereignisses oder beteiligte Personen, benannt. Die Smartphoneinhaberin holt sich auch kein *ticket* für die Produktion eines *multi-unit-turns* ein – viel mehr behandelt sie die vorangegangene Suspension des Gesprächs durch die Rezipientin und ihr ‚Warten‘ als Alignment mit einer bis zu diesem Punkt noch nicht spezifizierten smartphone-assistierten Aktivität.

Die Elaboration der Geschichte erfolgt zunächst in der bereits bekannten Form durch das Vorlesen von Textnachrichten, wobei die Smartphoneinhaberin mit der Realisierung des ungewöhnlichen Ausdrucks ihren Blick auf M richtet, den Begriff wiederholt und die Reaktion der Rezipientin beobachtet (Z. 035–037). Als sie gerade beginnt, die Erzählung fortzusetzen, produziert die Rezipientin eine Frage (Z. 039), mit der sie Verstehensprobleme mit diesem fokalen Element anzeigt. Erneut ist es die Rezipientin, die Probleme im Verständnis einer vorangehenden Äußerung der Erzählerin sprachlich relevant macht. Im Unterschied zu allen vorher diskutierten Fällen wird der visuelle Zugriff hier durch die Rezipientin durch das Zusammenspiel aus Frage und dem Ausstrecken ihres Arms ‚eingefordert‘: Die unmittelbare – und unkommentierte – Übergabe des mobilen Endgeräts durch die Smartphoneinhaberin zeigt, dass diese *multimodale Gestalt* als ‚Aufforderung zur Geräteübergabe‘ *accountable* ist und als angemessen behandelt wird.

Dieser Interaktionsverlauf lässt sich in Bezug zu Oloffs (2019a) Analysen initialer Zeigesequenzen setzen: Zum einen scheint es – zumindest unter FreundInnen<sup>85</sup> – Situationen zu geben, in denen die Übergabe des Smartphones nicht rechtfertigungsbedürftig ist (vgl. Oloff, 2019a, S. 215). Zum anderen zeigt sich auch hier, dass RezipientInnen sprachliche Aufforderungen (bspw. „zeig ma“) nutzen, „wenn der/die Gerätehalter/in die Bereitstellung des Displays verzögert“ und diese so gleichzeitig „an dessen/deren Verantwortung erinnern, ihnen das Angekündigte zur Betrachtung schnell bereitzustellen“ (Oloff, 2019a, S. 215). Im Kontext narrativer Rekonstruktionen von Textnachrichten können derartige Aufforderungen der Story-RezipientIn – sowie die anschließende Übergabe des Geräts durch die ErzählerIn – eine Methode darstellen, Probleme im Verständnis zu bearbeiten, in der die Möglichkeit des Lesens der Textnachrichten als Ressource für die Reparatur fungiert. Vergleichbar zum ersten Fall dieses Kapitels führt die Problemmarkierung durch die Rezipientin auch hier zur Eröffnung einer Einschubsequenz, in der die Problemquelle interaktiv repariert werden soll. Im Unterschied findet das Lesen der Textnachrichten durch die Story-Rezipientin jedoch zunächst simultan zur Fortsetzung der Geschichtenerzählung statt, sodass ein Multiaktivitätssetting entsteht, in dem die Rezipientin Textnachrichten liest und die Erzählerin gleichzeitig weitere Textnachrichten im Modus der DRS rekonstruiert<sup>86</sup>.

Die deontische Äußerung der Rezipientin macht jedoch am Ende des Ausschnitts (Z. 043, 045) sprachlich explizit, dass eine erfolgreiche Reparatur der Problemquelle nicht ausschließlich durch das Lesen des Threads erfolgen kann, sondern weitere Kontextualisierungen durch die Erzählerin erfolgen müssen. Mit dieser Äußerung wird die *intrapersonale Multiaktivität* – zusammengesetzt aus simultanem Lesen des Chats und dem Zuhören der Erzählung – abgebrochen und das Multiaktivitätssetting in einer *exclusive order* organisiert, in der die Progression der Geschichtenerzählung zunächst suspendiert wird, um die Reparatur und Wiederherstellung eines intersubjektiven Verständnisses als zentrale Aufgabe der Interaktionsorganisation bearbeiten zu können. Der Sequenzverlauf ist insgesamt durch eine zunehmend explizite Orientierung an der normativen Verantwortung der Smartphoneinhaberin geprägt, die Integration von

---

<sup>85</sup> Konversationsanalytisch kann hier nicht von einem determinierenden Verhältnis zwischen der Sozialbeziehung der InteraktionsteilnehmerInnen und der lokalen Handhabung des Smartphonegebrauchs ausgegangen werden. Zum einen muss in jeder einzelnen Instanz die *Accountability* und Angemessenheit der Übernahme bzw. -gabe des mobilen Endgeräts „each next first time“ (vgl. Rawls, 2002, S. 30) interaktiv hervorgebracht werden. Zum anderen kann die Sozialbeziehung der GesprächspartnerInnen (bspw. als Freunde) nicht als ‚gegeben‘, sondern muss selbst als interaktive Hervorbringungsleistung der TeilnehmerInnen verstanden werden – sodass bspw. ‚der GesprächspartnerIn das Smartphone unkommentiert übergeben‘ als Praxis erscheint, mit der eine freundschaftliche oder vertrauensvolle Beziehung überhaupt erst (re-)produziert wird.

<sup>86</sup> Auffällig ist der stark fragmentierte Charakter der Erzählung: Nur teilweise werden die AutorInnen der wiedergegebenen Textnachrichten durch Zitationsmarker explizit kenntlich gemacht. An einer Stelle wird eine Chatteilnehmerin sogar namentlich in die Erzählung eingeführt, ohne dass ihre Nachricht rekonstruiert wird.

Medieninhalten so zu kontextualisieren (vgl. Keppler, 2019, S. 180), dass ein intersubjektives Verständnis ermöglicht wird. Die Bearbeitung der Probleme mit der Verständlichkeit des rekonstruierten Chatverlaufs erfolgt im folgenden Ausschnitt, in dem M – ihren Blick auf das Smartphone gerichtet – die Aneignung der archivierten IM-Nachrichten beginnt:

**Transkript 12: 191024\_„Gnumma?“\_Teil 3**

042 °h und ich [so +was geht ]  
 ...-----+,,,,,  
 043 M: [du m%usst mir\_s-]  
 %....nimmt SP an, hält es vor  
 Gesicht-->058  
 044 A: mit EUCH,  
 045 M: du muscht mir des aus\_m KONTtext;  
 046 WO denn?  
 047 (-)  
 048 A: na mittendr+in: (.) +da weiter unten hinter  
 dem langen text °hh <<leicht lachend> GNUMm%a;>  
 +LH zeigt M+  
 m: %be-  
 wegt lautlos ihre Lippen-->051  
 049 (--)  
 050 A: genommen mmhihihi-  
 051 (1.2)% (0.3)  
 m: --->%  
 052 M: <<f>ACH tse-  
 053 A: +<<schrill>das &war [so WITzig#?>]  
 +.....streckt LA zu M aus-->054  
 m: -----,,-->& zu A-->054

#7



054 M: [ich GLAU:: ]B& +ich ha²  
 sie #hat\_s n bisschen persönlich gE&nommen;  
 -----,,,,,-->& SP -  
 Display-----& zu M -->055  
 a: -----,,,,,,+senkt LA-->056  
 #8



zu lachen. Auch diese Äußerung wird von M, die weiter mit dem Lesen des Chats beschäftigt ist, nicht erwidert: Sie beendet in der mehrsekündigen Stille in Z. 052 die lautlosen Lippenbewegungen und produziert kurz darauf in erhöhter Lautstärke mit „ACH tse“ einen „change-of-state token“ (Heritage, 1984b). Dieser wird von A als Display von Verstehen – und damit als Anzeige einer erfolgreichen Reparatur der Problemquelle – behandelt, denn sie beginnt in Z. 053 unmittelbar ihren linken Arm auszustrecken (#7), während sie mit „das war so witzig“ eine Bewertung der archivierten Chatkommunikation vornimmt. Mit dieser Bewertung, die anschließend die Produktion einer zweiten Bewertung konditionell relevant macht (vgl. Pomerantz, 1984, S. 61), orientiert sich A daran, dass ihre Interaktionspartnerin nun durch das Lesen des Chats über das notwendige Wissen verfügt, eine eigene Einschätzung des Ereignisses vorzunehmen (vgl. Raclaw et al., 2016, S. 377).

In Überlappung zu „witzig“ produziert M – ihren Blick auf A gerichtet – nun ihre eigene, durch „ich GLAU::B“ epistemisch herabgestufte (vgl. Heritage & Raymond, 2005, S. 19) Übersetzung der dialektal gefärbten Nachricht ins Hochdeutsche. Sie zeigt dabei keinerlei beobachtbare Orientierung an A's verkörperter Übergabeaufforderung, sodass diese ihren Arm wieder senkt (#8). M produziert ihr Interpretationsangebot – ihren Blick wieder auf das Display gerichtet – zunächst im Modus der DRS (das Pronomen „ich“ der Originalnachricht bleibt erhalten), bricht die Äußerung jedoch mit Glottalverschluss ab, um ihre Übersetzung zu formulieren („sie hat\_s n bisschen persönlich genommen“).

An A's Erwiderung in Z. 055 („ja ich WEIß ich hab\_s verstanden; hh“) zeigt sich anschließend jedoch, dass sich auch in dieser Interaktionsepisode ein Missverständnis hinsichtlich der angemessenen Erwiderungen durch die Rezipientin entwickelt hat: Die Rezipientin behandelt die Geschichte als „troubles telling“ (Jefferson, 1988), in der die Erzählerin ihre Schwierigkeiten mit dem Verständnis des Dialekts rekonstruiert. Entsprechend orientiert sich M in ihrer Aneignung des archivierten Chats daran, dass derartige Problemerkzählungen normalerweise einen Ratschlag („advice“, Jefferson & Lee, 1981) durch die Rezipientin konditionell relevant machen, und versucht durch ihre Übersetzung zur Aufklärung der Verstehensprobleme der Erzählerin beizutragen<sup>87</sup>. Die Erzählerin rahmt durch ihr Lachen (Z. 050) und ihre Bewertung („witzig“, Z. 053) das *telling* jedoch als humoristische Erzählung über ein unerwartetes Ereignis, mit der sie illustriert, den Heimatdialekt von M zu verstehen. Eine derartige Rahmung impliziert sequenziell wiederum die Anerkennung des amüsanten Charakters der Geschichtenerzählung

---

<sup>87</sup> Retrospektiv wird somit erkennbar, dass die Einschubsequenz, in der die Rezipientin visuellen Zugriff auf das Smartphonedisplay einfordert, auf der Grundlage eröffnet wird, dass sie nicht über hinreichende Informationen zum Kontext verfügte, um der Erzählerin einen Lösungsvorschlag für die Bearbeitung ihrer Schwierigkeiten im Verständnis der fokalen Nachricht des archivierten Chats zu unterbreiten.

durch die Rezipientin (vgl. Schegloff, 1987c, S. 206). Indem die Erzählerin durch ihre Reparatur in dritter Position (Z. 055) betont, dass sie die erhaltene Nachricht verstanden hat, macht sie nun explizit, dass die RezipientIn die sequenziellen Implikationen ihrer Erzählung missverstanden hat und ihr ‚Hilfe‘ bei der Übersetzung der fokalen Nachricht keine angemessene Erwiderung darstellt.

In Überlappung zu dieser Äußerung produziert M in Z. 056 zunächst ein gedehnt realisiertes „ja:::“, das sich als „post-completion stance marker“ (Schegloff, 1996b, S. 92) noch auf ihre eigene vorangegangene Äußerung zu beziehen scheint und retrospektiv als eine Art der Selbstbestätigung ihres Interpretationsangebots dient. Während dieser Überlappung findet die – erneut rein durch verkörperte Handlungen koordinierte – Rückgabe des Smartphones statt (#9). Sobald die Erzählerin das mobile Endgerät übernommen hat<sup>88</sup>, fokussiert sie wieder das Display. Anschließend wiederholt M in Z. 057 erneut das fokale Element der dialektalen Nachricht, dem sie ein ‚verächtliches‘ „tse“ anfügt, und auf diese Weise eine sequenzschließende Äußerung produziert. Die SmartphoneinhaberIn beginnt nun – ihren Blick auf ihr Smartphone fixiert – in Überlappung zu dieser Äußerung die Geschichtenerzählung fortzusetzen: Mit „aber dann“ projiziert sie in Z. 058 zunächst die Rekonstruktion weiterer sequenziell verknüpfter Ereignisse und rahmt diese durch „das war so WITzig“ erneut als humoristisch. Auf diese Weise wirkt A prospektiv dem wiederholten Auftreten eines Missverständnisses entgegen und ermöglicht der Rezipientin die Produktion angemessener Erwiderungen. Während des hörbaren Einatmens der Erzählerin in Z. 058 blickt die Rezipientin direkt zur Sprecherin und aligniert sich so wieder mit der projizierten Elaboration der Ereignisse. Die *smartphone-assistierte Geschichtenerzählung* wird im Folgenden von der GeräteinhaberIn auf bekannte Weise durch das Vorlesen von Textnachrichten fortgesetzt (nicht im Transkript enthalten).

Die Analyse dieses zweiten Falles zeigt erneut, wie es aufgrund der unvorbereiteten Initiierung einer *smartphone-assistierten Erzählung* zu Verstehensproblemen und Missverständnissen in der Bedeutung der sequenziellen Abfolge kommt, da die Verbindung der Erzählung zum Gesprächskontext für die Rezipientin nicht hinreichend *accountable* ist. Erneut wird die Progression der Geschichtenerzählung unterbrochen, um in Einschubsequenzen die aufgetretenen Verstehensprobleme zu reparieren und die Wiederherstellung eines intersubjektiven Verständnisses als zentrale Aufgabe der Interaktionsorganisation bearbeiten zu können (vgl. Kamunen,

---

<sup>88</sup> Die Art der Reparatur der Problemquelle zeigt eindeutig, dass mobile Endgeräte in Interaktionen zwischen FreundInnen keinesfalls ausschließlich von einer Person genutzt und eingesehen werden können (vgl. a. Weilenmann & Larsson, 2002, S. 97). Indem die SmartphoneinhaberIn unmittelbar nach Beendigung der lesenden Aneignung durch das Ausstrecken ihres Arms die Rückgabe des Smartphones einfordert, behandelt sie das mobile Endgerät als konstitutive Ressource in der Hervorbringung ihres *tellings*.

2019, S. 423). Missverständnisse in der Bedeutung der sequenziellen Abfolge werden dagegen in dritter Position von der ErzählerIn repariert, wodurch die Erwidern der RezipientIn als ‚unangemessen‘ behandelt werden.

Dieser Fall stellt insofern einen *abweichenden Fall* in der Kollektion dar, als dass er der einzige ist, indem die Smartphoneinhaberin von der RezipientIn aufgefordert wird, den visuellen Zugriff zum Display zu ermöglichen, und in der eine RezipientIn sprachlich explizit ihre Erwartungen an Produzentin der *smartphone-assistierten Geschichtenerzählung* formuliert. Aufgrund dieses Verlaufs ist der Fall besonders geeignet, Aufschluss über die Verantwortung der GeräteinhaberInnen sowie das Funktionspotenzial des visuellen Zugriffs zum Smartphonedisplay zu liefern: Die TeilnehmerInnen *smartphone-assistierter Geschichtenerzählungen* orientieren sich beobachtbar an einer normativen Verantwortung der Geräteinhaberin, integrierte Medieninhalte für alle praktischen Zwecke und rezipientenspezifisch hinreichend zu kontextualisieren, sodass ihre sinnhafte Einbettung in die Erzählung *accountable* wird. Dass die Rezipientin in diesem Fall die Story-Produzentin zunächst ohne sprachliche Erklärungen zur Übergabe des Smartphones auffordert und diese ihr das mobile Endgerät – ohne sprachliche Kommentierung – übergibt, zeigt eindeutig, dass sich die TeilnehmerInnen an dieser normativen Verantwortung der Smartphoneinhaberin als handlungsleitend orientieren<sup>89</sup>.

Die Ermöglichung des visuellen Zugriffs zum Smartphonedisplay, der den RezipientInnen eine lesende Aneignung der Textnachrichten erlaubt, stellt eine Methode dar, auftretende Verstehensprobleme zu bearbeiten. Das Zeigen der vorgelesenen Textnachrichten kann also gezielt als Ressource eingesetzt werden, um in Einschubsequenzen ein problematisch gewordenes, intersubjektives Verständnis der integrierten Medieninhalte zu reparieren. Vor dem Hintergrund der Analysen dieses Kapitels wird nun auch deutlicher, welche Relevanz das simultane Mitlese-Lassen während des Vorlesens der Textnachrichten besitzt: In vielen *smartphone-assistierten Geschichtenerzählen* findet sich eine prospektive Orientierung an diesem Funktionspotenzial des visuellen Zugriffs zum Display. Die kollaborative Hervorbringung eines

---

<sup>89</sup> Durch die Übergabe des Smartphones entsteht ein Teilnehmerrahmen, indem vorübergehend die Rezipientin das mobile Endgerät der Erzählerin handhabt. Obwohl die Smartphoneinhaberin nicht länger visuellen Zugriff zum Display des Geräts hat, wird sie weiterhin als wissende Teilnehmerin adressiert, die durch ihre sprachlichen Äußerungen die Lokalisierung des gesuchten Objekts anleitet. Interaktiv wird sie so weiterhin als Inhaberin des mobilen Endgeräts hervorgebracht, der eine epistemische Autorität (Heritage, 2012) hinsichtlich der Lokalisierung des gesuchten Objekts im Archiv ihres Smartphones zugeschrieben wird. Dies verweist darauf, dass die SmartphonebesitzerIn, wenn es sich bei einem „searchable object“ (Brown et al., 2015) um einen lokal gespeicherten Medieninhalt handelt, über exklusives Wissen, epistemische Autorität und die praktischen Kompetenzen verfügen muss, diese zu suchen-und-zu-finden. Anhand des Gesprächsverlaufs zeigt sich hier die „procedural consequentiality“ (Schegloff, 1991, S. 52–57) der diskursiven Identität der Smartphonebesitzerin, die als Interaktionsteilnehmerin mit exklusiven Rechten und Verantwortungen hinsichtlich dieses Wissens (Heritage & Raymond, 2005, S. 16; vgl. auch Sacks, 1985) behandelt wird und agiert.

Teilnahmerahmens, in dem das Display für Story-ProduzentIn und -RezipientIn einsehbar ist, bevor es zum Auftreten von Verstehensproblemen kommt, dient als intersubjektivitätssicherndes Verfahren, denn es erlaubt den RezipientInnen, sich den Wortlaut und Kontext des archivierten Chats simultan zum Zuhören auch lesend anzueignen. In methodisch vorbereiteten *smartphone-assistierten Geschichtenerzählungen* manifestiert sich diese prospektive Orientierung darüber hinaus durch die Produktion einer sprachlichen Prä-Sequenz, in denen die Verfügbarkeit einer Erzählung projiziert, deren Ereignis- und Erlebnisqualität etabliert, sowie die Verbindung des integrierten Nachrichten-Threads zum Gesprächskontext *accountable* gemacht wird.

### **5.2.3 Fazit: ‚Unvorbereitete‘ Geschichtenerzählungen und die Bearbeitung interaktionaler Probleme**

Mit unvorbereitet eröffneten, *smartphone-assistierten Geschichtenerzählungen* wurde eine Variante *smartphone-assistierter Geschichtenerzählungen* fokussiert, die sich hinsichtlich der Initiierung des Smartphonegebrauchs und ihrer Eröffnung maßgeblich von den methodisch vorbereiteten Erzählungen des vorangegangenen Kapitels unterscheidet. In der unvorbereiteten Variante werden archivierte Threads von den GeräteinhaberInnen aus ihrer diskursiven Identität als ZuhörerInnen während der Äußerungen ihrer InteraktionspartnerInnen gesucht und aufgerufen. Die SmartphoneinhaberInnen behandeln ihre Smartphones dabei als *Archive sozialer Originale*, aus denen sich spontan und mittels weniger Bediengesten Nachrichten-Threads aufrufen lassen, die mit dem aktuellen oder einem vergangenen Thema der sprachlichen Interaktion sinnhaft verbunden sind. Im Unterschied zu methodisch vorbereiteten Erzählungen wird die Verbindung zum *turn-by-turn-talk* jedoch nicht durch die Hervorbringung einer Prä-Sequenz *accountable* gemacht, sondern lediglich durch einzelne Redebeiträge inferiert und diese im Interaktionsverlauf zu einer Erzählung ausgebaut (vgl. König & Oloff, 2018a, S. 303).

Die ErzählerInnen bedienen sich dabei unterschiedlicher sprachlicher (bspw. „acknowledgement tokens“ und anschließender Themenwechsel) und verkörperter Verfahren (bspw. Meldegesten), um einen sequenziellen Slot im Gespräch zu schaffen, den selbst-initiierten Smartphonegebrauch zum Thema der gemeinsamen Interaktion zu machen. Unabhängig davon, ob mit der *smartphone-assistierten Erzählung* ein Themenwechsel herbeigeführt oder eine thematisch verbundene Episode entwickelt wird, setzen die GeräteinhaberInnen wiederholt *Warte-Aufforderungen* ein, mit denen sie andere Anwesende zur Suspension der gerade im Vollzug befindlichen Sequenz anhalten. Wird diese Aufforderung von den InteraktionspartnerInnen anerkannt und die aktuelle Aktivität unterbrochen, ergibt sich anschließend die Möglichkeit, – falls erforderlich – einen *smartphone-gestützten Suchprozess* als notwendige Vorarbeit für die Integration

abzuschließen (Suderland, 2019, S. 101) und den aufgerufenen Medieninhalt in der Interaktion zu thematisieren.

Anhand der in diesem Kapitel analysierten Fälle wurde gezeigt, dass *smartphone-assistierte Geschichtenerzählungen*, die durch einzelne Äußerungen eröffnet werden und deren Verbindung zum andauernden Gespräch lediglich inferentiell erschließbar ist, anfällig für das Auftreten von Folgeproblemen sind. Aufgrund der fehlenden methodischen Vorbereitung, durch die die Verfügbarkeit eines narrativen *big packages* und dessen Verbindung zum Gesprächskontext interaktiv etabliert wird, können sowohl Probleme des Verstehens als auch Missverständnisse in der Bedeutung der sequenziellen Abfolge entstehen<sup>90</sup>. Die Formen der interaktiven Bearbeitung dieser Probleme und ihr Einfluss auf den sequenziellen Verlauf der Geschichtenerzählungen sollen hier noch einmal zusammengefasst werden: Probleme des Verstehens einzelner Ausdrücke oder Äußerungen werden von den RezipientInnen durch die Produktion von Fragen sprachlich angezeigt (bspw. durch Teilwiederholungen oder Frageworte). Als erste Paarteile von Adjazenzpaaren veranlassen diese anschließend die Produktion fremd-initiiertes Reparatursequenzen (Schegloff et al., 1977, S. 367), in denen eben diese Probleme interaktiv bearbeitet werden, bevor die narrativen Rekonstruktionen fortgesetzt werden können. Missverständnisse hinsichtlich der Bedeutung der sequenziellen Abfolge werden wiederum durch die ErzählerInnen selbst in dritter Position angezeigt. Sie orientieren sich an den Äußerungen der RezipientInnen als unpassenden Erwidern auf die *smartphone-assistierte* Erzählung und zeigen so an, dass ihre Handlungsabsicht von den RezipientInnen nicht erkannt wurde (vgl. Egbert, 2009, S. 130): Die ErzählerInnen reagieren auf das Auftreten derartiger Missverständnisse in der Bedeutung der sequenziellen Abfolge durch die Initiierung und Durchführung einer Reparatur, bspw. indem sie das Rederecht kompetitiv übernehmen, um zu verhindern, dass die Äußerungen der RezipientInnen zum Fokus der Interaktion werden, oder indem sie explizieren, wie die Geschichte von ihnen gemeint war. Auch in diesen Fällen wird also die Progression der Geschichtenerzählung unterbrochen, um Probleme im intersubjektiven Verständnis bearbeiten und reparieren zu können.

---

<sup>90</sup> Diese Argumentation soll weder implizieren, dass es immer zu derartigen Folgeproblemen kommt, noch dass in methodisch vorbereiteten *smartphone-assistierten Geschichtenerzählungen* keine Verstehensprobleme oder Missverständnisse auftreten. Die Anfälligkeit unvorbereiteter Erzählungen für derartigen Probleme verweist lediglich darauf, welche Funktionen durch die methodische Hervorbringung einer Prä-Sequenz erfüllt werden und welche Voraussetzungen für ein intersubjektives Verständnis der im Vollzug befindlichen Aktivität und der Geschichtenerzählungen geschaffen werden. Wie bereits in Kap. 3.1. dargestellt, agieren die TeilnehmerInnen dabei grundsätzlich unter der Annahme einer Konsensunterstellung (Garfinkel, 1959, S. 62), wobei sich im Interaktionsvollzug jederzeit anhand einer nächsten Äußerung zeigen kann, dass die TeilnehmerInnen sich gegenseitig nicht ‚richtig‘ verstehen.

Die durch diese Reparaturen entstehende Varianz im sequenziellen Verlauf verdeutlicht zum einen, dass narrative Rekonstruktionen kollaborative Hervorbringungsleistungen darstellen, die maßgeblich durch die Beiträge der RezipientInnen geprägt sind (vgl. Mandelbaum, 2013, S. 501). Zum anderen verweist sie auf die spezifischen Verantwortlichkeiten, die mit den diskursiven Identitäten der TeilnehmerInnen an *smartphone-assistierte Geschichtenerzählungen* verbunden sind: Die Story-ProduzentInnen behandeln die Einnahme einer instrumentellen Haltung der RezipientInnen (durch Zuhören und *Monitoring* des Displays) als Alignierung mit der im Vollzug befindlichen smartphone-gestützten Aktivität. Es liegt in der Verantwortung der RezipientInnen, an passenden Stellen der Interaktion angemessene Erwidern zu produzieren oder Verstehensprobleme sprachlich relevant zu machen. Werden Verstehensprobleme angezeigt, orientieren sich die TeilnehmerInnen anschließend beobachtbar an der Verantwortung der SmartphoneinhaberInnen, die Medieninhalte rezipientenspezifisch zu kontextualisieren (Keppler, 2019, S. 180), sodass diese hinsichtlich ihrer Relevanz im Kontext einer Geschichtenerzählung und ihrer sequenziellen Implikationen *accountable* werden.

Im Kontext von Reparatursequenzen zeigt sich besonders deutlich, welche Relevanz der visuelle Zugriff zum Smartphonedisplay für die (Wieder-)Herstellung eines intersubjektiven Verständnisses der integrierten Medieninhalte besitzt: Zeigen sich Probleme in der Verständlichkeit eines Ausdrucks oder einer Äußerung, wird die Ermöglichung (bzw. Aufrechterhaltung) des visuellen Zugriffs zum Display von den TeilnehmerInnen methodisch als Ressource eingesetzt, um die Problemquelle zu bearbeiten: Sei es, um kommunikative Aktivitäten beispielhaft zu demonstrieren oder die Schreibweise eines Wortes zu zeigen. Im Zuge der lesenden Aneignung erhalten die RezipientInnen einen epistemischen Zugriff zu den problematisch gewordenen Referenten (vgl. Raclaw et al., 2016, S. 370) und die Möglichkeit, sich deren Sinn zu erschließen. Die Analysen der smartphone-assistierte Bearbeitung von interaktionalen Problemen zeigen somit in Erweiterung zu McGregor (2020, S. 149), dass Sprache als eine zentrale, jedoch nicht immer als primäre Ressource für die Reparaturen fungiert. Die materiell verfestigten Handlungsmöglichkeiten der mobilen Endgeräte und integrierten Medieninhalte werden von den TeilnehmerInnen ebenfalls als Ressourcen für Reparaturen eingesetzt: Die auf dem Smartphonedisplay angezeigten Medieninhalte können – im Zusammenspiel mit sprachlichen Äußerungen und verkörperten Handlungen – als Ressourcen für die Bearbeitung der Problemquellen genutzt werden, indem die SmartphoneinhaberInnen visuellen Zugriff zum Display und damit eine lesende oder betrachtende Aneignung der Medieninhalte ermöglichen. Die (erfolgreiche) Bearbeitung der Problemquellen erfolgt jedoch nicht ausschließlich lesend, sondern immer in wechselseitiger Bezugnahme aufeinander und durch das Zusammenspiel sprachlicher

Äußerungen, verkörperter Handlungen und dem semiotischen Feld des Gerätedisplays. Auf diese Weise werden komplexe *multimodale Gestalten* hervorgebracht, mithilfe derer der spezifische soziale Sinn der Nachricht(en) für die im Vollzug befindliche Aktivität etabliert wird.

Die durch die Materialität der archivierten Threads ermöglichte lesende Aneignung verweist auf zwei zentrale Differenzen zu narrativen Rekonstruktionen flüchtiger sozialer Situationen: Diese können weder wiederholt betrachtet werden, noch können sie anders als durch den Einsatz sprachlicher und verkörperter Ressourcen rekonstruiert werden. Die Analyse der Bearbeitung von Verstehensschwierigkeiten und Missverständnissen in *smartphone-assistierten Geschichtenerzählungen* verdeutlicht jedoch auch, dass obwohl die technischen Möglichkeiten der mobilen Endgeräte es zuließen, dass sich die GesprächsteilnehmerInnen archivierte Textnachrichten (ausschließlich) lesend aneigneten, die medial-vermittelte Kommunikation immer wieder durch Praktiken des Vorlesens in narrative Rekonstruktionen integriert wird: „Talking about text is therefore not just an activity in its own right, it is about establishing what texting is and the meaning texts have“ (Brown et al., 2018, S. 21). Eines der Vehikel, mit dem die archivierten Threads als bedeutungsvolle, vergangene Ereignisse im Alltag der NutzerInnen etabliert werden, ist das Erzählen smartphone-assistierter Geschichten mit ihren für diese Aushandlung erforderlichen Reparatursequenzen, in denen ihr intersubjektiver Sinn explizit ausgehandelt wird.

Durch ihre Einbettung in narrative Rekonstruktionen werden die gespeicherten Threads als erzählwürdige Ereignisse hervorgebracht, durch die unterschiedlichste soziale Handlungen in F2F-Interaktionen realisiert werden können. Gleiches gilt in veränderter Form auch für *smartphone-assistierte Geschichtenerzählungen*, die durch den Erhalt einer Textnachricht veranlasst und in denen noch nicht abgeschlossene Ereignisse rekonstruiert werden. Mit der Fokussierung dieser Variante im folgenden Kapitel rückt nun die *Affordanz* der mobilen Endgeräte, orts- und zeitunabhängig Nachrichten Abwesender zu empfangen und an sie zu versenden, in den Fokus der Analysen.

### 5.3 „Stories in the making“: durch eingehende (Text-)Nachrichten veranlasste Geschichtenerzählungen

In den vorangegangenen Kapiteln wurden *smartphone-assistierte Geschichtenerzählungen* analysiert, in denen archivierte Episoden medial-vermittelter Kommunikation (bspw. WhatsApp-Chats) durch das Vorlesen von Textnachrichten narrativ rekonstruiert wurden. Für die Hervorbringung derartiger Erzählungen spielt insbesondere die *Affordanz* des Smartphones als *Archiv sozialer Originale* eine konstitutive Rolle: Die TeilnehmerInnen nutzen Nachrichtenthreads als Ressource in der Produktion der Geschichtenerzählungen. In allen Erzählungen werden die archivierten Threads dabei als für alle praktischen Zwecke abgeschlossene Interaktionsepisoden behandelt. In diesem Kapitel möchte ich nun eine Variante *smartphone-assistierter Geschichtenerzählungen* fokussieren, die davon ausgelöst wird, dass die GeräteinhaberIn den Erhalt einer neuen (Text-)Nachricht bemerkt und auf eine Weise sprachlich relevant macht, dass es zur interaktiven Hervorbringung einer Geschichtenerzählung kommt. In dieser Variante spielt die *Affordanz* des Smartphones zur permanenten<sup>91</sup> Erreichbarkeit für Telekommunikation eine konstitutive Rolle für die Veranlassung und Elaboration der Geschichtenerzählung.

Aufgrund dieser grundsätzlichen Differenz zu den vorangegangenen Varianten smartphone-gestützter Erzählungen möchte ich den Analysen einige Überlegungen voranstellen: Dass das Empfangen von (Text-)Nachrichten als Anlass, Thema *und* Ressource für die interaktive Hervorbringung von Geschichtenerzählungen dienen kann, verweist darauf, dass die durch den Gebrauch mobiler Endgeräte mediatisierten F2F-Interaktionen mit Knorr Cetina (2009) als „synthetische Situationen“ verstanden werden können: Als synthetisch werden Situationen verstanden, „that include electronically transmitted on-screen projections that add informational depth and new response requirements to the ‚ecological huddle‘ (Goffman, 1964, S. 134) of the natural situation“ (Knorr Cetina, 2009, S. 61). Im vorliegenden Fall werden die alltäglichen F2F-Interaktionen durch die Präsenz des Smartphones, der über sie empfangbaren und auf

---

<sup>91</sup> Diese Formulierung soll nicht implizieren, dass NutzerInnen lediglich aufgrund des Mitführens eines mobilen Endgeräts permanent erreichbar wären. Erreichbarkeit über das Smartphone ist zum einen von der technologischen Infrastruktur – bspw. der Netzabdeckung oder des Internetempfangs – abhängig. Zum anderen wird Erreichbarkeit aus Perspektive der KA als praktische Hervorbringungsleistung verstanden: Es muss empirisch rekonstruiert werden, ob und wie die NutzerInnen mobiler Endgeräte für kopräesente Andere *accountable* anzeigen, dass die technische Möglichkeit zur Erreichbarkeit eine praktische Relevanz für ihre lokalen Handlungen besitzt. Darüber hinaus verfügen die SmartphonennutzerInnen über Methoden, Erreichbarkeit zu vermeiden, bspw. indem die Geräte auf „Airplane-Mode“ oder „stumm“ geschaltet werden und außerhalb des Sichtbereichs der NutzerInnen aufbewahrt werden, sodass eingehende Anrufe oder Textnachrichten weder visuell noch akustisch für Ihre NutzerInnen wahrnehmbar sind. Es kann also keinesfalls a priori von einem Zustand der Smartphone-NutzerInnen als „permanently online, permanently connected“ (bspw. Vorderer et al., 2017) ausgegangen werden.

dem Gerätedisplay angezeigten Informationen erweitert. Indem die TeilnehmerInnen ihre Smartphones so handhaben, dass sie auf akustische oder visuelle Signale eingehender Benachrichtigungen (oder Telefonate) reagieren können, orientieren sie sich beobachtbar daran, dass die telekommunikativen *Affordanzen* der mobilen Endgeräte für sie eine gewisse Erwiderverantwortung („response presence“ Knorr Cetina, 2009, S. 69) erzeugen. Als beobachtbare Orientierung an einer Erwiderverantwortung werden hierbei nicht nur im ‚engen Sinne‘ antwortende Handlungen (wie das Annehmen eines Telefonats oder das Verfassen einer Nachricht), sondern all diejenigen Aktivitäten der TeilnehmerInnen verstanden, mit denen diese das Display ihres Handys ‚checken‘ (Oulasvirta et al., 2012) und es auf neue Informationen überprüfen<sup>92</sup>. In und durch derartige Aktivitäten wird das Smartphone von den TeilnehmerInnen als Technologie hervorgebracht, über deren Display neue, potenziell (interaktions-)relevante Informationen empfangen werden können (vgl. Knorr Cetina, 2009, S. 69–70) und deren Beobachtung zu einer (mehr oder minder) kontinuierlichen Aufgabe werden kann (vgl. ebd., S. 70–72).

Eingehende (Text-)Nachrichten können von den TeilnehmerInnen also als Ereignis behandelt werden, das eine lesende Aneignung oder medial-vermittelte Erwidern erforderlich macht. Sie können darüber hinaus auch sprachlich zum Thema der F2F-Interaktion gemacht werden (vgl. Brown et al., 2018, S. 7–9, 15–19; DiDomenico et al., 2018, S. 19–22)<sup>93</sup> und so als Anlass für die Hervorbringung einer *smartphone-assistierten Geschichtenerzählung* dienen. Wie zu zeigen sein wird, handelt es sich dabei – im Unterschied zu den Varianten der vorangegangenen Kapitel – um narrative Rekonstruktionen von noch nicht abgeschlossenen Ereignissen. Derartige Geschichtenerzählungen werden von Georgakopolou (2008) als „stories-in-the-making“ – *im Entstehen befindliche* Geschichten – bezeichnet, in denen die in situ empfangenen Textnachrichten vorgelesen werden und dargestellt wird, welche Konsequenzen diese für zukünftige Handlungsverläufe oder projizierte mögliche Ereignisse haben (vgl. Georgakopoulou, 2008, S. 601)<sup>94</sup>.

---

<sup>92</sup> Siehe dazu auch die überblicksartige Darstellung des Smartphonegebrauchs über den gesamten Korpus in Kap. 4.3.

<sup>93</sup> In beiden Studien wird gezeigt, dass gerade erhaltene Textnachrichten von den SmartphoneinhaberInnen sprachlich thematisiert und so zum Fokus der F2F-Interaktionen werden können. Es werden jedoch keine Fälle untersucht, in denen durch das Vorlesen gerade erhaltener Textnachrichten Geschichtenerzählungen hervorgebracht werden.

<sup>94</sup> Derartige Formen narrativer Rekonstruktionen wurden von Georgakopolou (2008) im Rahmen ihres „small-stories“-Forschungsansatzes (bspw. Georgakopoulou, 2007; Bamberg & Georgakopoulou, 2008) in konversationsanalytisch inspirierten, ethnographischen Studien Anfang der 2000er-Jahre identifiziert – also zu einem Zeitpunkt, zu dem Smartphones noch nicht entwickelt waren. Das Konzept von im *Entstehen befindlichen* Geschichtenerzählungen erscheint für das hier untersuchte Phänomen unmittelbar anschlussfähig; Georgakopoulou untersucht in ihren eigenen Analysen jedoch die Zusammenhänge zwischen alltäglichen Erzählungen und den Identitäten der Beteiligten aus narrationsanalytischer Perspektive (Georgakopoulou, 2007, S. 89–118).

Die Analyse von *smartphone-assistierten Geschichtenerzählungen*, die durch eingehende Textnachrichten veranlasst werden, wird im Zentrum dieses Kapitels stehen. Gemeinsam ist diesen Erzählungen, dass sie von der SmartphoneinhaberIn durch ein „noticing“ (Schegloff, 2007a, S. 219) der eingegangenen Textnachricht initiiert werden: Als *noticing* werden sprachliche oder verkörperte Handlungen bezeichnet, durch die ein Aspekt des wahrnehmbaren Settings – einschließlich vorangegangener Äußerungen – relevant gemacht wird, der vorher von den InteraktionsteilnehmerInnen nicht als relevant behandelt wurde (vgl. Schegloff, 2007a, S. 219)<sup>95</sup>. Bei *noticings* handelt es sich um multimodale, emergente Projekte, die mit einer beobachtbaren Veränderung im verkörperten Verhalten einer TeilnehmerIn beginnen und so interaktionsöffentlich *accountable* machen, dass etwas wahrgenommen wurde (vgl. Kääntä, 2014, S. 89). Diese verkörperte Orientierung deutet bereits eine bevorstehende Handlung der Bemerkenden an. Interaktionsöffentlich relevant wird das *noticing* jedoch erst, wenn es ihr gelingt, die Aufmerksamkeit anderer Anwesender zu rekrutieren, auf das ‚noticeable‘ zu lenken und eine anschließende Erwiderung auf das *noticing* zu erhalten (Helisten, 2019, S. 339). Die Aufmerksamkeitsrekrutierung erfolgt dabei oftmals durch sprachliche Äußerungen, mit denen das Bemerkte für andere expliziert wird. Diese werden häufig als unmittelbar konstruiert und zeigen so an, dass die SprecherIn es gerade ‚im Moment‘ (vgl. Sacks, 1995, Vol. II, S. 90–92) in ihrer wahrnehmbaren Umgebung bemerkt hat (vgl. a. Drew & Chilton, 2000).

In den folgenden Analysen durch eingehende Textnachrichten ausgelöster Geschichtenerzählungen werde ich folgenden Forschungsfragen nachgehen:

- An welchen sequenziellen Positionen werden die *noticings* platziert, um die Aufmerksamkeit der Anwesenden auf den Erhalt einer (Text-)Nachricht zu lenken?
- Wie werden in situ erhaltene (Text-)Nachrichten methodisch als Anlass und Ressource für die Entwicklung *smartphone-assistierter Geschichtenerzählungen* hervorgebracht?
- Wie wird der Erhalt von (Text-)Nachrichten als erzählwürdiges Ereignis dargestellt und wie unterscheiden sich die entwickelten Geschichten von narrativen Rekonstruktionen abgeschlossener Ereignisse?

---

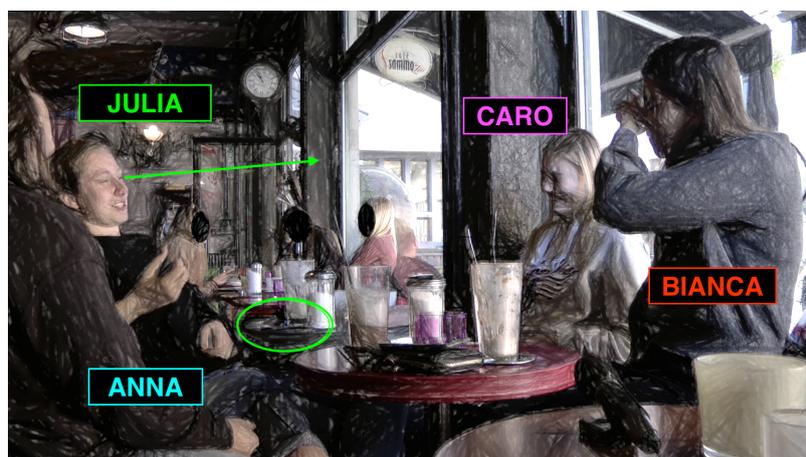
<sup>95</sup> Aus Perspektive der KA bezieht sich der Begriff des *noticings* dabei nicht auf die kognitive Wahrnehmung eines Ereignisses, sondern auf das beobachtbare Verhalten, das für andere Anwesende das Bemerkende eines Ereignisses anzeigt (Schegloff, 2007a, S. 87).

### 5.3.1 Die Thematisierung von Textnachrichten an sequenziell passenden Positionen

Anhand der Analyse des ersten Falles möchte ich exemplarisch aufzeigen, wie die SmartphoneinhaberIn aus ihrer diskursiven Identität als RezipientIn auf den visuell wahrnehmbaren Empfang einer (Text-)Nachricht reagiert und simultan zur sprachlichen Interaktion der Anwesenden ihr Smartphone ‚checkt‘. Die SmartphoneinhaberIn nutzt dabei unterschiedliche multimodale Ressourcen und die *Affordanzen* des mobilen Endgeräts, um den Smartphonegebrauch mit den Anforderungen der F2F-Interaktion zu koordinieren: Sie vermeidet eine potenzielle Unterbrechung des Gesprächs und wartet die beobachtbare Beendigung des aktuellen Themas ab, um an einer sequenziellen ‚Übergangsposition‘ (vgl. Drew & Chilton, 2000) die erhaltene Nachricht sprachlich relevant zu machen. Der folgende Ausschnitt ist Teil einer längeren Gesprächsepisode, in denen sich die Freundinnen Anna (A), Julia (J), Caro (C) und Bianca (B) über die Eigenheiten einer Lehrerin aus der gemeinsamen Schulzeit amüsieren. Im Fokus steht hier jedoch nicht das Thema des Gesprächs, sondern das Verhalten von B (hinten links) die während der Unterhaltung beginnt, sich ihrem Smartphone zuzuwenden.

Zur besseren Lesbarkeit des Transkripts werden zunächst ausschließlich Julias Blicke („+“) und verkörperte Handlungen („†“) festgehalten, die im Bezug zur Handhabung des mobilen Endgeräts stehen. Wird ihr Blick nicht annotiert, so ist er auf die Anwesenden, insbesondere die jeweils primäre Sprecherin gerichtet.

#### Transkript 13: 181004 „Saugroboter Worschter“\_Teil 1



```
#1
j: >>+Blick schräg vorne--->Z.004
001 J: #<<:-> [oder so] ne vierte [HAND] dann die ich auch
so (in [die da hoch ausfahrt) >, + ]
002 C: [ha ha ]
003 B: [ha ]
004 A: [<<:-> (und dann) (.) und +die KR]Eide hoch
wirft >;
```

j: -----+Blick A--->006  
 005 B: ha [ha ha] ((klatscht sich auf Bein))  
 006 J: [ha #ha] ha ha [+ #ha ha ]  
 +Blick SP--->012

#2

#3



007 A: [für NIX ] die schreIbt nicht mal  
 #+was; [he]

j: #+tippt auf SP-Display--->013  
 #4



008 C: [ha] (.) °hh

009 (.)

010 A: sogar im BAde[anzug (xxx xxx) [mehr ];

011 C: [ja: ] [(xxx)]

012 C: <<stöhnend> ha: ja (.) ich + #find\_s EINFach; (.)

j: -----,,,--->+Blick C--->>

#5



013 vor talem ich dacht ich wirklich die ganze zeit  
 dass die irgendwie Anfangen wird zu SCHREIben;

j: --->†

014 und (dann) [auf einmal läuft die so] [(hüpfend) (.)  
 und dann wieder läuft die dann so LOS.]

Anhand dieses Ausschnitts zeigt sich, dass Mehrpersonengespräche den TeilnehmerInnen die Möglichkeit bieten, ihre Teilnahme an der F2F-Interaktion vorübergehend zu reduzieren, um sich der Bedienung ihres mobilen Endgeräts zuzuwenden (vgl. DiDomenico et al., 2018, S. 679–681): Die Smartphoneinhaberin J setzt in Z. 001 durch ihre sprachliche Äußerung – begleitet durch das Lachen der Anwesenden (Z. 002-003) – zunächst die gemeinsame Belustigung über die Eigenheit der Lehrerin fort und blickt anschließend zu A, die diese in Überlappung und im Modus der Fiktionalisierung fortsetzt (Z. 004). Durch die visuelle Fokussierung der nächsten Sprecherin zeigt B ihre Rezipientenschaft (C. Goodwin, 1981, S. 58) und ihr exklusives Engagement (C. Goodwin, 1984, S. 228) an der sprachlichen Interaktion an. Während des durch A's Äußerung veranlassten Lachens (Z. 005-006) blickt J zunächst noch zu A (Z. 006, #2) und beginnt dann auf das Display ihres auf dem Tisch liegenden Smartphones zu schauen (#3). Simultan zu A's nächster Äußerung initiiert sie anschließend die Bedienung des Geräts mit ihrer rechten Hand (Z. 007, #4). Während die anderen Teilnehmerinnen die humoristische Kommemoration der gemeinsamen Schulzeit erneut aufnehmen (Z. 010 ff., nicht im Transkript enthalten), setzt J die – ohne sprachliche Kommentierung initiierte – divergente Bedienung des mobilen Endgeräts fort.

Die Veränderungen im verkörperten Verhalten der Smartphoneinhaberin, die aus ihrer Rolle als ZuhörerIn agiert, machen interaktionsöffentlich beobachtbar, dass sie ein Ereignis auf dem Display ihres Smartphones bemerkt hat (vgl. Helisten, 2019, S. 339). Indem sie visuell das Display des Geräts fokussiert und dieses physisch manipuliert, suspendiert sie vorübergehend ihre Rezipientenschaft und zeigt ihre primäre Involvierung in den Smartphonegebrauch an (DiDomenico et al., 2018, S. 679). Das dadurch entstehende Multiaktivitätssetting wird von den TeilnehmerInnen auf divergierende Weise gemanagt: Während B durch ihr verkörpertes Verhalten den Smartphonegebrauch vorübergehend als ihre Hauptaktivität hervorbringt<sup>96</sup>, sind die anderen Anwesenden auf das gemeinsame Gespräch fokussiert und führen dieses fort – ohne den selbst-initiierten Smartphonegebrauch von B durch Kommentare oder Beobachtung als *accountable* zu behandeln (vgl. Oloff, 2019b, Z. 452–454). Aufgrund dieses Managements des Multiaktivitätssettings entfalten die beobachtbaren Veränderungen im verkörperten Verhalten der Bemerkenden zunächst keine weiteren Implikationen für die F2F-Interaktionen (vgl. Helisten, 2019, S. 339): Die Bedienung des mobilen Endgeräts durch J und die sprachliche Interaktion der anderen Anwesenden laufen zunächst parallel zueinander ab: Als C in Z. 012

---

<sup>96</sup> Zur Erinnerung: Im Unterschied zu Goffman (1963, S. 43) wird aus Perspektive der KA nicht davon ausgegangen, dass sich a priori unterscheiden lässt, welche Aktivität eine Haupt- und welche eine Nebentätigkeit konstituiert, sondern dass dies dynamisch und lokal im Vollzug der Interaktion von den TeilnehmerInnen organisiert und ausgehandelt wird (vgl. Mondada, 2014a, S. 46).

das Rederecht ergreift und eine thematisch verbundene narrative Rekonstruktion initiiert, blickt J wieder vom Gerätedisplay auf und zur neuen primären Sprecherin (Z. 012, #5); kurz danach löst sie auch die Hände vom Gerät, um ihre Arme auf der Rückenlehne abzulegen (Z. 013). Auf diese Weise macht sie die Beendigung des Smartphonegebrauchs und die Wiederaufnahme ihrer Rezipientenschaft *accountable*. Nachdem J die Geschehnisse auf dem Display ihres Smartphones ‚gecheckt‘ und eine empfangene Textnachricht gelesen hat, gilt ihr exklusives Engagement nun also wieder der Teilnahme an der sprachlichen Interaktion.

Dass die Smartphoneinhaberin sich in diesem Ausschnitt vorübergehend aus dem Gespräch zurücknimmt, den Gebrauch des mobilen Endgeräts als Hauptaktivität behandelt und anschließend ihre Rezipientenschaft in der F2F-Interaktion wieder aufnimmt, wird durch die *Affordanzen* des Smartphones sowohl veranlasst als auch ermöglicht: Die technischen Möglichkeiten zur Telekommunikation – hier des orts- und zeitunabhängigen Empfangs von Kurznachrichten – führen zunächst dazu, dass auf dem Display aktualisierte Informationen angezeigt werden. Aufgrund der Platzierung des Geräts im Sichtfeld der Smartphoneinhaberin wird diese visuell auf neu eingegangene Nachrichten ‚hingewiesen‘, was sie dazu veranlasst, sich ihrem Smartphone zuzuwenden und die neuen Informationen zu ‚checken‘. Anschließend sind es die spezifischen *Affordanzen* der handyvermittelten Textnachrichtenkommunikation, die der Geräteinhaberin das Management paralleler Aktivitäten und die anschließende ‚Rückkehr‘ zur lokalen F2F-Interaktion ermöglichen: Asynchrone<sup>97</sup> Kommunikationsformen wie SMS, WhatsApp oder E-Mail erlauben es den NutzerInnen, eingegangene Nachrichten zu lesen und zu einem (grundsätzlich) beliebigen Zeitpunkt zu beantworten<sup>98</sup>. Die Smartphoneinhaberin orientiert sich an eben dieser *Affordanz*, wenn sie sich nach dem Bemerkten und Lesen der erhaltenen Nachricht wieder der F2F-Interaktion zuwendet, in der weiter Erinnerungen an die Lehrerin ausgetauscht werden (Z. 015–047, nicht im Transkript enthalten).

---

<sup>97</sup> Es wird immer wieder darauf hingewiesen, dass moderne IM-Applikationen eine ‚quasi-synchrone‘ Kommunikation erlauben, da sie den Austausch unmittelbar aufeinander folgender Nachrichten ermöglichen und den NutzerInnen durch ‚Online-‘ und ‚Aktivitäts‘-Anzeigen „zumindest Kenntnis über die generelle Schreibaktivität des Dialogpartners“ (Graffe, 2014, S. 45) vermitteln. Der entscheidende Punkt ist jedoch, dass IM-Applikationen keine Möglichkeit einer ‚echten‘ wechselseitigen Beobachtung bieten: Das Turn-Taking unterliegt gänzlich anderen Bedingungen als in mündlichen Gesprächen, denn die Produktion von Turnkonstruktionseinheiten ist nicht beobachtbar und es gibt keine Möglichkeiten der Rederechtsübernahme durch Unterbrechung. Aus diesem Grund erzeugen ‚quasi-synchrone‘ Kommunikationsformen andere Zugzwänge als mündliche Gespräche und stellen verringerte Anforderungen an die wechselseitigen Reaktionen der TeilnehmerInnen aufeinander (Dürscheid, 2016, S. 439; s. a. Wyss & Hug, 2016).

<sup>98</sup> Wie im Folgenden noch zu zeigen sein wird, ist die technische Möglichkeit, zu einem beliebigen Zeitpunkt zu antworten, keinesfalls ein determinierender Faktor im Umgang mit empfangenen Kurznachrichten: Die Teilnehmer orientieren sich beobachtbar an der sozialen Relevanz der empfangenen Botschaften und behandeln diese als dringliche Angelegenheit, die einer unmittelbaren Antwort bedarf, oder als etwas, das momentan keine praktische Handlungskonsequenzen für sie hat.

Im folgenden Ausschnitt zeigt sich nun, dass J zunächst die beobachtbare Beendigung des aktuellen Themas abwartet, bevor sie die erhaltene Nachricht als Anlass zur Initiierung einer *smartphone-assistierten Geschichtenerzählung* nutzt. Erneut werden die Blicke der Smartphoneinhaberin (J) mit „+“ und ihre verkörperten Handlungen mit „†“ annotiert. Blicke und verkörperte Handlungen von A werden mit „\*“, von C mit „%“ und von B mit „\$“ transkribiert. Sobald J in Z. 052 das Wort ergreift, richten die anderen Teilnehmerinnen ihren Blick auf sie und zeigen so ihre Zuhörerschaft an. Zur besseren Lesbarkeit werden anschließend lediglich Änderungen im Blickverhalten der Zuhörerinnen annotiert.

#### Transkript 14: 181004\_„Saugroboter Worschter“\_Teil 2

048 ((2.0 A, B, C, und D lachen still))

049 J: °h °hh (.) [°hh ]

050 C: [ha ha]

051 (0.6)#+(0.8)

j: +Blick SP--->054  
#6



052 J: °h ich <<:-)> fin[d#† das so SÜ:SS] >=

+tippt auf SP-Display--->054  
#7



053 C: [ha ha ha ]

054 J: =wir +†ham doch so nen stau? (.) äh so nen SAUGrobot†ter;

,,--->+Blick B--->057

----->†...kreisförmige Gesten rechte Hand---

,,,--->†

055 (0.69)

056 C: ha ha w[as, (.) ha ha ]

057 B: [ja: +hat †mein] papa %AUCH (.) ich find

den voll [putzig- ]  
 j: -----,,,-->+Blick SP--->058  
 j: †...RH hebt SP vor Gesicht-->064  
 c: %..beugt sich vor->077

#8



058 J: [ <<lachend> und dann ] (.) dann steht > da-

Der Beginn dieses Ausschnitts (Z. 048) zeigt zunächst noch das gemeinsame Lachen der Teilnehmerinnen als Anerkennung der humoristischen Kommemoration der ehemaligen Lehrerin. Dieses kommt nach einigen hörbaren Lachpartikeln von J und C (Z. 049–050) zum Erliegen und beendet so auf beobachtbare Weise das aktuelle Thema (vgl. Holt, 2010, S. 1524). In der anschließenden Stille beginnt J zunächst auf das Display ihres Smartphones zu schauen (#6) und ergreift in Z. 052 mit einem hörbaren Einatmen das Wort. Sie produziert mit *smile-voice* ein *prospective indexical*: sie findet etwas – das in dieser Äußerung nicht weiter spezifiziert wird – „so SÜ:SS“. Ihren Blick weiter auf das Display gerichtet, beginnt sie mit ihrer rechten Hand die Bedienung des Smartphones (#7). In Überlappung produziert C in Z. 053 ein hörbares Lachen, das jedoch von niemandem aufgenommen wird<sup>99</sup>. In ihrem nächsten Turn (Z. 054) produziert J einige Hintergrunddetails (vgl. C. Goodwin, 2015a, S. 199), die das neue Thema explizieren: „wir ham doch so nen stau? (.) äh so nen saugroboter“. Mit Beginn dieser Äußerung suspendiert J die Bedienung des Smartphones wieder: Sie blickt vom Display zu B und selektiert diese so als primäre Adressatin (vgl. Kidwell, 2013, S. 109), außerdem produziert sie mit der rechten Hand einige kreisförmige Gesten, die die Form eines handelsüblichen Saugroboter demonstrieren<sup>100</sup>. Sprachlich zeigt die Formulierung „wir ham doch“ an, dass sie – mindestens bei ihrer primären Rezipientin – Wissen über diesen Sachverhalt annimmt.

Anschließend entsteht eine kurze Gesprächspause (Z. 055), in der J eine Bestätigung des angenommenen Wissens durch die Rezipientin(nen) abzuwarten scheint. Es ist zuerst C, die durch

<sup>99</sup> Es lässt sich nicht eindeutig rekonstruieren, ob es sich bei diesem Lachen um ein Nachverbrennen, also ein ‚reaktiviertes‘ Lachen handelt, das B’s Äußerung als Fortsetzung des Amusements über die ehemalige Lehrerin behandelt, oder ob diese Äußerung als Initiierung eines neuen Themas und damit als eigenständiger Lachanlass verstanden wird.

<sup>100</sup> Eine kurze Online-Recherche zu „Saugrobotern“ zeigt, dass die meisten dieser Geräte ein kreisrundes Design aufweisen: <https://tinyurl.com/y2rkhg8m>.

ein von Lachen gerahmtes „was?“ (Z. 056) ein Problem im Verstehen anzeigt. In Überlappung produziert die als primäre Rezipientin adressierte B mit „ja hat mein papa AUCH (.) ich find den voll putzig“ (Z. 057) eine Erwiderung, mit der sie (zumindest) bestätigt, Saugroboter zu kennen und diese ebenfalls „putzig“ zu finden. Bereits nach der Realisierung des bestätigenden „ja“ in turn-initialer Position beginnt J wieder, auf das Smartphonedisplay zu schauen und das Gerät nun vor ihr Gesicht zu halten. Durch das Timing der Wiederaufnahme des Smartphonegebrauchs zeigt J an, dass sie hier nicht die Produktion einer zweiten Bewertung als konditionell relevant behandelt, sondern eine Prä-Sequenz eröffnet hat, die der Aushandlung des Wissensstands der primären Rezipientin dient und ihr die rezipientenspezifische Ausgestaltung der projizierten, smartphone-assistierten Erzählung ermöglichen soll. Sobald dies geschehen ist, setzt sie – erkennbar durch eine *multimodale Gestalt*, in der die physische Handhabung des Smartphones bereits vor ihrer erneuten Übernahme des Rederechts aufgenommen wird – zur Elabotation der Erzählung an: Sie produziert mit *smile-voice* eine Äußerung, die durch „und dann (.)“ Kohärenzverhältnisse zwischen dem Besitz des Staubsaugerroboters und dem folgenden deiktischen Verweis auf die Geschehnisse auf dem Display ihres Smartphones („dann steht da“) herstellt.

Zu diesem Zeitpunkt ist noch nicht ersichtlich, welcher Medieninhalt von der Smartphoneinhaberin für die lokale Interaktion relevant gemacht wird. Erkennbar ist jedoch bereits hier, dass sie die *Affordanzen* des Smartphones nutzt, um eine während der F2F-Interaktion erhaltene Nachricht an einer sequenziellen Position in das Gespräch einzuführen, an dem das vorherige Thema zu einem erkennbaren Abschluss geführt wurde (vgl. DiDomenico et al., 2018, S. 678). Indem die Smartphoneinhaberin den Zeitpunkt der Integration einer empfangenen Textnachricht sensibel an die sequenzielle Organisation des Gesprächs anpasst, zeigt sie ihre „orientation to the interactional priority of the current, joint activity and to her own noticing as nonurgent, that is, as something that does not call for immediate attention but can wait until more important matters have been attended to“ (Helisten, 2019, S. 316): Erst nachdem ein geeigneter Slot an einer Übergangsposition nach Beendigung eines Themas gefunden ist, nutzt die Geräteinhaberin die grundsätzliche Flexibilität alltäglicher Gespräche, um schrittweise ein neues Thema in die Interaktion einzuführen.

Die Einführung der erhaltenen Textnachricht geschieht auch hier auf die bekannte Weise mittels einer *multimodalen Gestalt*, durch die die Verfügbarkeit einer smartphone-assistierten Erzählung projiziert wird. Erneut lässt sich für Interaktionsteilnehmerinnen nur durch das Zusammenspiel sprachlicher Äußerungen, verkörperter Handlungen und der Handhabung des mobilen Endgeräts inferieren, dass über das Smartphone abrufbare Medieninhalte für die kopräsen-

Interaktion relevant gemacht werden. In der so eröffneten Prä-Sequenz wird die projizierte Geschichtenerzählung methodisch vorbereitet. Dabei sind die interaktiv zu bearbeitenden Aufgaben – Verfügbarkeit einer Erzählung projizieren, *ticket* erhalten und Wissen der Rezipientinnen aushandeln, um ein rezipientenspezifisches Design zu ermöglichen – eindeutig durch den Einsatz körperlicher, visueller und materieller Ressourcen (vgl. König & Oloff, 2018a, S. 282) voneinander abgegrenzt: So richtet die Smartphoneinhaberin bspw. während der Wissensaushandlung ihren Blick auf die primäre Rezipientin und suspendiert die Bedienung des Smartphones. Diese wird erst fortgesetzt, als die Einschubsequenz zu einem projizierbaren Ende kommt und diese interaktionale Aufgabe als erfolgreich bearbeitet behandelt wird.

Indem die Smartphoneinhaberin wieder die typische Körperhaltung des smartphone-assistierten Erzählens einnimmt und eine sprachliche Äußerung produziert, die erkennbar die narrative Rekonstruktion vergangener Ereignisse einleitet, zeigt sie an, nun die Elaboration der Geschichte zu beginnen. Wie sich im weiteren Verlauf der Interaktion zeigt, ergeben sich jedoch zunächst Verstehensprobleme mit der verwendeten Referenzform für den abwesenden Kommunikationspartner, die vor der Elaboration der Erzählung bearbeitet werden müssen. Auch im folgenden Ausschnitt werden die *Affordanzen* des Smartphones als Ressource für die interaktiv relevant werdenden Reparaturversuche eingesetzt:

### Transkript 15: 181004\_„Saugroboter Worschter“\_Teil 3

058 J: [ <<lachend> und dann] (.) dann steht > da-  
059 (.) +also der heißt WORSCH <<lachend> ter+>; hhe  
--->+, , , ----Blick B-----+SP-->064  
060 (0.1)  
061 B: \$was?  
\$. . . .beugt sich vor-->075  
062 C: (wie, )  
063 J: WOR+SCHter. †  
-->+B--->064  
-----, , , , † RH senkt SP--->072  
064 (0.1)+(0.23)  
j: ---->+ zu C--->067  
065 C: <<:->>WORSCHter, >  
066 (0.17)  
067 J: <<f>WOR>--+hha  
----->+ nach unten, tlw. Augen geschlossen->072  
068 ((0.7 B und C lachen stumm))  
069 B: WOARSCHta?  
070 X: hha  
071 (0.14)  
072 J: <<:->, f>†!#WORSCH!+ #†te:r;%\$>

```

,,,,,----->+ zu C--->076
----->†SP vor Gesicht† dreht Display zu C--->075
c:                                     %Blick SP-Display->076
b: -----$....beugt sich
weiter vor, Blick auf SP-->075
#9 #10

```



073 (0.43)

074 J: (ich mein) (.) des\_\*is der spitzname von <<:-)>#von meinem Väter;>

```

a:                                     *beugt sich vor, Blick SP-->075
#11

```



075 <<lachend> und dann+ kommen immer so NA:CHrichten wie worsche \*erfordert ihre auf\$merksam+keit>; h†h

Mit ihrer Äußerung in Z. 058 leitet die Smartphoneinhaberin die Elaboration der Geschichten-erzählung ein. Sie bricht die Äußerungsproduktion jedoch ab, um in Z. 059 nach einer Mikro-Pause und eingeleitet durch den Reformulierungsindikator (Gülich, 1994, S. 85) „also“ zu erklären, welchen Namen der Saugroboter ihrer Familie trägt („WORSCHter“): Der Name wird hier erstmalig eingeführt, um den Rezipientinnen im weiteren Verlauf als erkennbare Referenzform („recognital“, Schegloff & Sacks, 1979, S. 17) für eine Figur der Geschichte dienen zu können. Die letzte Silbe des Namens wird von J lachend realisiert, womit sie sich an der potenziellen *laughability* des Namens zu orientieren scheint. Während der Äußerungsproduktion blickt sie zu B, die so erneut als primäre Rezipientin adressiert wird, und richtet ihren Blick erst mit der Realisierung der Lachpartikeln wieder auf das Smartphonedisplay. Nach einer kurzen

Stille (Z. 060) produzieren B (Z. 061: „was“?) und C (Z. 062: „wie,“) nacheinander jeweils unspezifische Reparatur-Initiierungen, die unbestimmt lassen, ob ein Problem des Hörens oder des Verstehens vorliegt (Egbert, 2009, S. 99). B beugt sich mit Beginn ihrer Äußerung über den Tisch und näher zu J – möglicherweise eine verkörperte Orientierung an einem Problem des Hörens.

Die beiden Reparatur-Initiierungen veranlassen J in Z. 063 zur Produktion eines Reparaturversuchs, indem sie den Namen des Saugroboters wiederholt – wodurch sie anzeigt, dass sie annimmt, dass die verwendete Referenzform die Problemquelle ihrer vorangegangenen Äußerung darstellt. Während der Turnproduktion schaut sie zu B, legt anschließend ihr Smartphone ab und blickt in der darauffolgenden Stille (Z. 064) zu C. Diese wiederholt in Z. 065 den Klang des Namens, wobei sie die Äußerung durch *smile-voice* und Frageintonation (Z. 065) als tentativen Verstehensversuch markiert. In Erwiderung setzt J erneut zu einer, nun in erhöhter Lautstärke realisierten<sup>101</sup> Wiederholung der Problemquelle an, bricht jedoch nach der ersten Silbe in Lachen aus (Z. 067), in das C in Z. 068 einstimmt<sup>102</sup>. Anschließend wiederholt auch B in Z. 069 mit Frageintonation ihren Verstehensversuch des Namens.

Indem beide Rezipientinnen ihre Erwiderungen mit Frageintonation als Verstehensversuche markieren, zeigen sie an, dass noch keine intersubjektiv verständliche Reparatur der Problemquelle erfolgt ist. Nach hörbaren Lachpartikeln (Z. 070) und einer kurzen Stille (Z. 071) setzt J daher zu einem erneuten Reparaturversuch an: Mit *smile voice*, erhöhter Lautstärke und starkem Akzent auf der ersten Silbe wiederholt sie den Namen. Dieses Mal fungiert die Äußerung dabei als Betrachtungsanweisung (Oloff, 2019a, S. 214) in einer smartphone-gestützten Zeigesequenz: Simultan führt J das Smartphone zunächst vor ihr Gesicht (#9) und dreht es mit der Realisierung der zweiten Silbe zu C (#10), sodass diese visuellen Zugriff zum Display erhält. Am Ende von J's Turn beginnt auch B, sich weiter über den Tisch zu beugen, um auf den Bildschirm schauen zu können. Anschließend entsteht erneut eine Gesprächspause, in der B und C das Display fokussieren und Zeit erhalten, sich den integrierten Medieninhalt betrachtend anzueignen (vgl. Oloff, 2019a, S. 215). Nach einem Moment der Betrachtung (Z. 073)

---

<sup>101</sup> Die Realisierung mit erhöhter Lautstärke kann als Anzeige verstanden werden, dass B annimmt, ein Problem des Hörens reparieren zu müssen. Aufgrund der insgesamt hohen Lautstärke im Café und der Vielzahl an Nebengeräuschen erscheint diese Interpretation durchaus plausibel. Dass nicht schon der erste Reparaturversuch mit erhöhter Lautstärke produziert wurde, zeigt jedoch, dass auch für B nicht eindeutig zu sein scheint, welche Dimension ihrer Äußerung für die Rezipientinnen problematisch ist.

<sup>102</sup> So wie B's Lachen in Z. 059 kann auch diese Lachen – nun als gemeinsame – implizite Orientierung an der Ungewöhnlichkeit des verwendeten Referenten gedeutet werden. Wenn diese Interpretation zutrifft, verschiebt sich auch das Verständnis für die Ursache der Verstehensprobleme: Diese lassen sich dem verwendeten Referenten zuschreiben und liegen weder in einer zu leisen Aussprache der Sprecherin noch in einer ungenügenden Zuhörerschaft der Rezipientinnen begründet.

produziert J in Z. 074 einen *Account*, der die Nicht-Erkennbarkeit der Referenzform damit erklärt, dass es sich um den Spitznamen ihres Vaters handelt<sup>103</sup>. Während der Äußerungsproduktion beugt sich nun auch A – die sich bis hier hin nicht beobachtbar an der Interaktionsepisode beteiligt hatte – nach vorne (#11), um auf das Display zu schauen. Dass keine der Rezipientinnen an der übergabe-relevante Stelle das Rederecht ergreift, wird von J als Anzeige von Verstehen und einer erfolgreichen Reparatur behandelt, sodass sie anschließend erneut die Elaboration der *smartphone-assistierten Geschichtenerzählung* aufnimmt.

Auch in diesem Ausschnitt werden die Progression einer Geschichtenerzählung und die simultane Bedienung des Smartphones suspendiert, um durch die Rezipientinnen angezeigte Probleme im Verstehen und Hören bearbeiten zu können (vgl. Kamunen, 2019, S. 422) – in diesem Fall um die Nicht-Erkennbarkeit einer verwendeten Referenzform für eine Figur der Erzählung zu reparieren. Die Bedienung des mobilen Endgeräts wird von der Erzählerin erst wieder aufgenommen (Z. 072), nachdem die Reparaturversuche durch sprachliche Äußerungen nicht erfolgreich waren und sie das Smartphone als Ressource in einem weiteren Reparaturversuch einsetzt. Indem sie den Rezipientinnen durch die Initiierung einer Zeigesequenz visuellen Zugriff zum Display gewährt, ermöglicht sie ihnen einen erweiterten epistemischen Zugang (Raclaw et al., 2016, S. 370) zur problematischen Referenzform. Durch die lesende Aneignung der Nachricht können die Teilnehmerinnen sich nun versichern, den Namen des Saugroboters richtig verstanden zu haben. Die anhaltende Uneindeutigkeit, ob bei den Rezipientinnen Probleme des Hörens oder Verstehens vorliegen, veranlasst die Smartphoneinhaberin darüber hinaus dazu, einen *Account* zu formulieren, der die Nicht-Erkennbarkeit der Referenzform damit erklärt, dass es sich um einen Spitznamen handelt. Das anschließende Ausbleiben weiterer fremd-initiiertter Reparaturen durch die Rezipientinnen wird – entsprechend der Präferenzorganisation in der Verwendung von personalen Referenzen – als Erkennbarkeit der Referenzform behandelt (vgl. Schegloff & Sacks, 1979, S. 19). Dabei macht sich die Smartphoneinhaberin den aktuellen Teilnehmerahmen, in dem alle Teilnehmerinnen epistemischen Zugang zur relevanten Textnachricht besitzen (vgl. Raclaw et al., 2016, S. 370), zu Nutze, um von der Zeigesequenz (die der Reparatur diene) zur Elaboration der Geschichtenerzählung überzugehen:

---

<sup>103</sup> Dass diese Äußerung als Erklärung geeignet ist, die Nicht-Erkennbarkeit des Namens zu reparieren, scheint damit zusammenzuhängen, dass Spitznahmen hochgradig idiosynkratische Arten der personalen Referenz sein ‚dürfen‘, die weder anderen bekannten Namen entsprechen, noch als Worte sinnhaft für andere Personen verständlich sein müssen. Insofern könnte die Erklärung, dass es sich um einen Spitznamen handelt, an dieser sequenziellen Position ein ‚agreeable‘ darstellen, der alle weiteren Erklärungen überflüssig macht und eine sequenzschließende Funktion übernimmt.

Transkript 16: 181004 „Saugroboter Worschter“\_Teil 4

074 J: (ich mein) (.) des\_\*is der spitzname von <<:-)> von meinem VAtter;>

a: \*beugt sich vor, Blick SP-->075

075 <<lachend># und dann+ kommen immer so NA:CHrichten wie worschter \*erfordert ihre auf\$merk#sam+keit;> h+h

-----,,-->+ nach links, auf SP-Display---  
-----+zu B->76  
-----  
-----,////////

->+ dreht Display zu sich

a: ,,,,,,--->\*sitzt gerade-->>

b: -----,////////,\$ setzt sich gerade auf, Blick zu C->>

#12

#13



076 B: das [is\_+n echter] WORSCHter al%ter;

j: ----->+ zu SP-Display

c: ----->%Blick zu B-->079

077 C: [hi hi hi ]

[Auslassung: Anhaltendes gemeinsames Lachen]

089 (0.84)

090 J: <<:-)> und dann musst du immer RUNter laufen zu dem (.) und den aus irgend >-

091 (0.19)

092 J: <<:-)> der steckt meistens irgendwo fest und dann musst du ihn so RAUS (.) holen >-

093 (0.08)

[Elaboration der Geschichtenerzählung]

In Z. 075 eröffnet J lachend die Ausgestaltung der Geschichte: Eingeleitet durch die Formulierung „und dann kommen immer so NA:CHrichten“ wird das Folgende als Illustration einer typischen Benachrichtigung des Saugroboters gerahmt. Erst an dieser Stelle des Gesprächs wird sprachlich expliziert, dass es sich bei dem integrierten Medieninhalt um eine Systembenachrichtigung handelt, die die Besitzerin darüber informiert, dass der Saugroboter „Worschter“ ihre „Aufmerksamkeit“ (Z. 075) erfordert. Mit der Produktion der einleitenden Formulierung

blickt die Smartphoneinhaberin auf das Display und bringt die DRS – auf bekannte Weise – als Vorlesen der Textnachricht hervor, während alle Rezipientinnen über den Tisch gebeugt das Display fokussieren (#12). Schon vor der Realisierung der finalen Silbe blickt J zu erneut zu B und dreht während der Produktion der Lachpartikel das Display wieder stärker zu sich. Indem sie den visuellen Zugriff auf das Display für die Rezipientinnen einschränkt und ihren Blick einer Interaktionspartnerin zuwendet, macht sie durch ihr verkörpertes Verhalten die Beendigung des Vorlesens erkennbar. Während C ihren Blick weiter auf das Smartphonedisplay gerichtet hält, lösen sich A und B im Verlauf der wörtlichen Zitation aus der ursprünglichen Teilnahmekonfiguration (#13) und reduzieren so ihr simultanes Engagement an der visuellen Fokussierung des Smartphonedisplays<sup>104</sup>.

B's anschließende Äußerung ratifiziert mit „das is\_n echter WORSCHter alter;“ nicht nur die *laughability* der vorgelesenen Nachricht. Die Verwendung der vorher noch reparaturbedürftigen Referenzform in einem neuen, ‚kreativen‘ Kontext zeigt an, dass auch die Verstehensprobleme erfolgreich bearbeitet wurden. Ihre amüsierte Charakterisierung des Saugroboters als „echter Worschter“<sup>105</sup> löst im Folgenden anhaltendes Gelächter und die Produktion weiterer Witze aus (nicht im Transkript enthalten). Erst als das gemeinsame Amüsement beendet ist und in der anschließenden Stille (Z. 089) keine der Anwesenden das Wort ergreift, setzt J die Ausgestaltung der Geschichte fort. Ab Z. 090 rekonstruiert sie nun sprachlich, welche Folgehandlungen derartige Systembenachrichtigungen normalerweise notwendig machen: Der Saugroboter muss im Haus gesucht und aus der Position befreit werden, in der er feststeckt. Die narrative Rekonstruktion erfolgt dabei ohne weitere Smartphoneintegration und wird durch zahlreiche verkörperte Enaktierungen animiert, die die notwendigen Handgriffe der Inhaberin sowie die Geräusche und Bewegungen des Saugroboters multimodal demonstrieren.

Die Analyse dieses ersten Falles zeigt, wie sich die Smartphoneinhaberin an ihrem mobilen Endgerät als synthetischer Komponente der Interaktion orientiert, gegenüber der sie eine gewisse Erwidungsverantwortung besitzt: Sie zeigt ihr *noticing* des Erhalts einer eingehenden

---

<sup>104</sup> Da es sich bei der vorgelesenen Nachricht aller Wahrscheinlichkeit nach um die gleiche Systembenachrichtigung handelt, die zur Bearbeitung des Verstehensproblems eingesetzt wurde, kann dieses ‚Disengagement‘ vor Vollendung des Vorlesens als Anzeichen gedeutet werden, dass der visuelle Zugriff – und damit die Möglichkeit des Mitlesens – von den Teilnehmerinnen nun als redundant behandelt wird. Alle Teilnehmerinnen hatten bereits (genügend) Zeit, sich die Nachricht (auch) lesend anzueignen und zeigen durch die Modifikation des Teilnehmerahmens an, dass sie den visuellen Zugriff nicht länger für ihr Verständnis des Vorlesens benötigen.

<sup>105</sup> Es lässt sich hier nicht eindeutig rekonstruieren, wie „echter Worschter“ von den anderen Teilnehmerinnen verstanden wird und warum diese Charakterisierung als *laughable* behandelt wird. Mit Rückgriff auf mein eigenes Common-Sense-Wissen lässt sich lediglich mutmaßen, dass der Name „Worschter“ an das Verb „wurschteln“ erinnert. Laut Duden wird dieses mit der Bedeutung „in einem gewissen Trott und ohne rechten Plan vor sich hinarbeiten“ verwendet, was im Kontext der Rekonstruktion des Erhalts einer Systemnachricht eines Saugroboters, der die Aufmerksamkeit seiner Besitzerin benötigt, eine durchaus amüsante – und passende – Koinzidenz darstellt.

Textnachricht über ihr verkörpertes Verhalten an und wendet sich der Handhabung ihres Smartphones simultan zum andauernden Gespräch der anderen Anwesenden zu. Ihr verkörpertes *noticing* entfaltet zunächst noch keine interaktionalen Konsequenzen – der divergierende Smartphonegebrauch bleibt unkommentiert. Die Smartphoneinhaberin nimmt anschließend die Teilnahme an der kopräsenten Unterhaltung wieder auf und wartet die beobachtbare Beendigung des aktuellen Themas ab, bevor sie in einer Gesprächspause den Smartphonegebrauch erneut aufnimmt und die Thematisierung der erhaltenen Nachricht durch die Initiierung eines *story preface* vorbereitet. Indem sie das *noticing* der erhaltenen Textnachricht erst an einer sequenziell passenden Position sprachlich thematisiert, priorisiert die Smartphoneinhaberin die aktuell im Vollzug befindlichen Gesprächsaktivität und behandelt den Erhalt der Nachricht als nicht-dringlich (vgl. Helisten, 2019, S. 316).

Die sequenzielle Entwicklung dieser *smartphone-assistierten Geschichtenerzählung* ist durch eine Einschubsequenz gekennzeichnet, in der Verstehensprobleme bearbeitet werden. Auf bereits bekannte Weise wird die Reparatursequenz eindeutig von der Geschichtenerzählung abgegrenzt: Diese ist durch die Einnahme der charakteristischen Körperhaltung des smartphone-assistierten Erzählens und die Alignierung der Rezipientinnen durch die visuelle Fokussierung der Erzählerin gekennzeichnet. Die Reparatursequenz zeichnet sich dagegen durch wechselseitige Blickverschränkungen und die (vorübergehende) Suspension des Smartphonegebrauchs aus. Nach erfolglosen Reparaturversuchen durch sprachliche Äußerungen werden auch in diesem Fall die Materialität der archivierten Textnachricht und die Mikro-Mobilität des Geräts als Ressourcen zur Bearbeitung eines Problems des Hörens eingesetzt: Indem den Teilnehmerinnen visueller Zugriff zum Display ermöglicht wird, können sie sich durch die lesende Aneignung versichern, die verwendete Referenzform akustisch richtig verstanden zu haben. Die Bearbeitung des fortbestehendes Problem des Verstehens – der Bedeutung der personalen Referenzform – erfolgt jedoch anschließend durch eine sprachliche Erklärung: ein weiterer empirischer Nachweis dafür, dass die soziale Bedeutung der integrierten Textnachrichten grundsätzlich durch die interaktiven Aushandlungsprozesse der Interaktionsteilnehmerinnen etabliert wird (vgl. Brown et al., 2018, S. 21; Keppler, 2019, S. 180).

Die Erzählwürdigkeit der empfangenen Nachricht wird im Verlauf der humoristischen Geschichtenerzählung zunächst dadurch hervorgebracht, dass die Erzählerin Textnachrichten eines nicht-menschlichen Agenten – ihres Staubsaugerroboters – und nicht, wie erwartbar wäre, von einer menschlichen KommunikationspartnerIn erhält. Besonders relevant ist jedoch, dass die Erzählerin rekonstruiert, inwiefern die erhaltene Nachricht unter anderen Umständen praktische Konsequenzen für anschließende Handlungsverläufe (Georgakopoulou, 2008, S. 601)

hätte – die Befreiung des Saugroboters: Da sich die Erzählerin jedoch räumlich entfernt von ihrem Saugroboter befindet, kann der Erhalt der Nachricht als Anlass und Ressource für die Entwicklung einer humoristischen Geschichtenerzählung genutzt werden, mit der die Anwesenden – durch das Vorlesen der Nachricht und lebhaftes, multimodales Enaktieren der üblichen Handlungsfolgen – unterhalten werden.

### 5.3.2 „Eilmeldungen“ – disjunkt platzierte Thematisierungen eingehender Textnachrichten

Anhand des ersten Falles wurde exemplarisch aufgezeigt, dass die SmartphoneinhaberInnen die *Affordanzen* des Smartphones nutzen können, um eingehende Textnachrichten an sequenziell passenden Stellen des Gesprächs zu thematisieren und zu *smartphone-assistierten Geschichtenerzählungen* zu entwickeln. Im Folgenden werde ich aufzeigen, dass das Bemerkte erhaltener Textnachrichten durchaus auch als Ereignis hervorgebracht werden kann, das eine plötzlichere Thematisierung an sequenziell disjunkten Positionen veranlasst (vgl. Helisten, 2019, S. 319): Sowohl durch die sequenzielle Positionierung als auch die methodische Hervorbringung des *noticings* wird dabei die Dringlichkeit der erhaltenen Nachricht herausgestellt. In und durch die anschließend hervorgebrachte Geschichtenerzählung wird die unmittelbare Handlungsrelevanz für die SmartphoneinhaberIn ausgehandelt und so *accountable* gemacht, warum andere Aktivitäten vorübergehend suspendiert werden (müssen), um die erhaltene Nachricht als Fokus der Interaktion zu behandeln.

Der folgende Ausschnitt stammt aus einem Gespräch zwischen sechs freundschaftlich Bekannten, die sich für ein „Running Dinner“<sup>106</sup> getroffen haben. Dabei handelt es sich um ein Event, bei dem drei Gänge eines Abendessens jeweils in einer anderen Wohnung eingenommen werden. Die Gerichte werden dabei meist von ‚Gastgeber‘-Teams aus jeweils zwei Personen zubereitet und bei jedem Gang treffen neue Konstellationen an TeilnehmerInnen aufeinander. In der folgenden Interaktionsepisode unterhält sich die Gruppe zunächst amüsiert über das ungewöhnliche Aussehen einer gemeinsamen Dozierenden von Veith (V) und Petra (P). Die Frage von Lisa (L), ob man die Dozierende googeln könne, veranlasst P dazu, nach ihrem Smartphone zu greifen und in diesem Zuge den Empfang einer neuen Textnachricht zu bemerken.

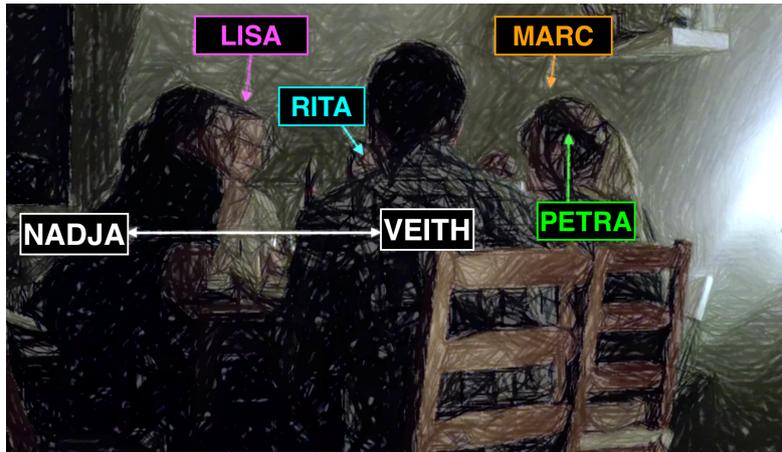
Der Fall unterscheidet sich damit insofern vom vorherigen, dass der Empfang der Nachricht von der GeräteinhaberIn nicht unmittelbar durch das *Monitoring* des Smartphonedisplays bzw. durch ein akustisches Signal, sondern erst im Zuge der Initiierung einer smartphone-gestützten

---

<sup>106</sup> [https://de.wikipedia.org/wiki/Running\\_Dinner](https://de.wikipedia.org/wiki/Running_Dinner)

Online-Suche bemerkt und anschließend sprachlich thematisiert wird<sup>107</sup>. Die Blicke der Smartphoneinhaberin Petra (P) werden mit „+“, ihre Gestik mit „±“ annotiert. Aufgrund der Kameraposition und den Lichtverhältnissen der Aufnahme ist es für einige TeilnehmerInnen und Passagen des Gesprächs nur schwer möglich, Details der Blickführung und des verkörperten Verhaltens zu transkribieren.

### Transkript 17: 180418 „Hass ich“\_Teil 1



001 L: kann man die GOOgeln?  
 002 (0.2)  
 003 P: ja;  
 004 M: he he he [<<:-)> kann man die (.) GU > ]-  
 005 P: [<<:-)> (xxx) ich hab\_s GRAD sch]on  
 gezeigt +bei der haupt ±(.) bei der #(-)  
 hauptspeise;>  
 +dreht sich um nach RH--->006  
 ±....greift nach SP-->006  
 #1

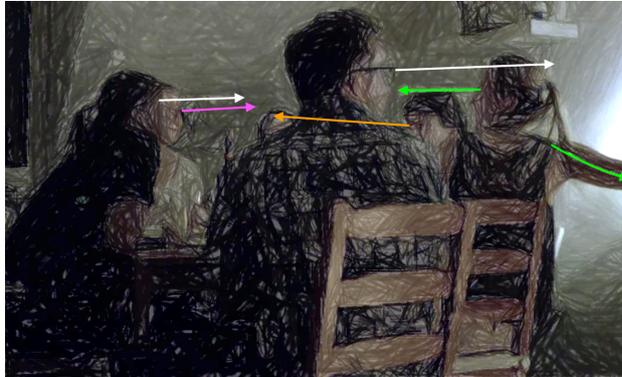


006 N: HAM wir jetzt ja wieder;+ hhe,

<sup>107</sup> Zeitnah zur Initiierung des Smartphonegebrauchs durch P kommt es zu einer Schismabildung (Egbert, 1997) in der Interaktion: Während Lisa (L), Rita (R), Marc (M) und Petra (P) über die erhaltene Textnachricht sprechen, entwickelt sich ab Z. 008 eine parallele Unterhaltung zwischen Nadja (N) und Veith (V) über eine „iPod-Station“, die auf der Kommode steht, auf der P ihr Handy abgelegt hatte. Zur besseren Lesbarkeit wurde die Interaktion zwischen N und L nicht im Transkript aufgenommen.

p: ----->+zu Gegenübern--->007  
 007 P: <<p, :-> SCH#EIße ja>;±+  
 ----->±  
 ----->+

#2



008 + (1.5)  
 p: +...auf SP--->019  
 009 P: <<f, all> HASS ich; >  
 010 oh SCHEI[ße ich hab ] ne böse nachricht  
 gekriegt;  
 011 R: [[<<prustend> kr >]  
 012 (1.3)  
 013 R: ne BÖ:se <<:-> [nach:richt > ]?

Der Ausschnitt beginnt damit, dass durch die sprachliche Interaktion ein ‚suchbares Objekt‘ hervorgebracht und ein sequenzieller Slot geschaffen wird, an dem die Initiierung einer smart-phone-gestützten Suchanfrage erfolgen kann (vgl. Brown et al., 2015, S. 511): L’s Frage in Z. 001, ob man die Dozierende googeln könne, wird nach einer kurzen Stille von P bejaht (Z. 003). In Überlappung zu M, der sich in Z. 004 über diese Frage amüsiert, erklärt P (Z. 005), dass sie bereits während der Hauptspeise anderen TeilnehmerInnen des Running Dinners Fotografien der Dozierenden gezeigt hat – und rahmt sie auf diese Weise als ‚zeigewürdig‘ (vgl. Licoppe, 2017, S. 71)<sup>108</sup>. Während der Äußerungsproduktion dreht sich P nach hinten rechts und greift nach ihrem Smartphone (Z. 005, #1), das auf der Kommode hinter ihr liegt. Mit ihrem verkörperten Verhalten behandelt sie L’s Frage also als Aufforderung, eine smartphone-gestützte Suchanfrage zu initiieren.

Während P nach ihrem Smartphone greift, produziert N in Z. 006 zunächst eine scherzhafte Äußerung, die auf einen früheren Zeitpunkt des Gesprächs Bezug nimmt, an dem schon einmal darüber gesprochen wurde, dass das Dessert (genau wie die Hauptspeise) sehr sättigend sei. In

<sup>108</sup> Ich werde in Kap. 6.1. ausführlicher auf die Hervorbringung der Zeigewürdigkeit von digitalen Fotografien eingehen.

Reaktion blickt P – ihren Oberkörper weiter Richtung Kommode gedreht – zu N und produziert einige Lachpartikeln und eine sprachliche Ratifizierung des scherzhaften Vergleichs, womit diese Nebensequenz (Jefferson, 1972) wieder geschlossen wird. In der anschließenden Gesprächspause kehrt P in ihre vorherige Sitzposition zurück und schaut auf ihr Smartphone, das sie nun vor sich hält. Anschließend produziert sie in Z. 009 mit „HASS ich;“ eine Äußerung, die durch ihre erhöhte Lautstärke und Sprechgeschwindigkeit den Charakter eines ‚fluchenden Herausplaters‘ (‚exclamatory imprecation‘, ebd. Goffman, 1978, S. 798)<sup>109</sup> erhält. Im unmittelbaren Anschluss produziert sie mit „oh SCHEIße ich hab ne böse nachricht gekriegt;“ ein „oh-prefaced environmental noticing“ (Drew & Chilton, 2000, S. 150), das mehrere Funktionen übernimmt: Durch das „oh“ in turn-initialer Position wird das Bemerkte als ‚spontan‘ konstruiert, gleichzeitig zeigt es als Erkenntnisprozessmarker (‚change-of-state token‘, Heritage, 1984b) an, dass sich das momentane Befinden der Sprecherin verändert hat. Es wird unmittelbar gefolgt von einer negativen Evaluation des anschließend explizierten Ereignisses – dem Erhalt einer „bösen“ Nachricht. Mit dieser Äußerung, die indexikalisch und zeitlich mit dem lokalen Wahrnehmungskontext der Sprecherin verknüpft ist (vgl. Helisten, 2019, S. 320), wird sprachlich für die Anwesenden *accountable* gemacht, was die Smartphoneinhaberin gerade auf dem nur für sie einsehbaren Display bemerkt hat.

R scheint sich mit ihrem prustenden Lachen (Z. 011), das kurz nach dem *oh-preface* in Überlappung einsetzt, implizit an der sequenziell disjunkten Platzierung des *noticings* und dem damit verbundenen abrupten Versuch eines Themenwechsels zu orientieren. Als die Smartphoneinhaberin sich in der entstehenden Gesprächspause (Z. 012) auf die Bedienung ihres mobilen Endgeräts fokussiert und diese damit als Hauptaktivität behandelt, produziert R mit Frageintonation („ne BÖ:se <<-> nach:richt?>“) eine Teilwiederholung (Egbert, 2009, S. 101), mit der sie ein Verstehensproblem anzeigt und eine Reparatur veranlasst.

Anhand des ersten Ausschnitts zeigt sich, dass die Thematisierung einer erhaltenen Textnachricht auf gänzlich andere Art und Weise als im ersten Fall erfolgt: Die Geräteinhaberin bemerkt eine Textnachricht, als sie den Smartphonegebrauch zum Zweck einer Suchanfrage initiiert. Es scheint also gerade die Platzierung des Smartphones außerhalb des Sichtbereichs und das Unterlassen einer regelmäßigen Beobachtung des Gerätedisplays, das nun während der Handhabung des Smartphones zur Produktion eines sprachlichen *noticings* führt: Dessen Dringlichkeit wird zum einen durch seine sequenziell disjunkte Positionierung, zum anderen durch die

---

<sup>109</sup> Flüche stellen nach Goffman (1978) eine spezifische Form von „blurtings“ dar. Als Herausplater stellen sie an alle Anwesenden den Anspruch, ihre Aufmerksamkeit auf die ‚Angelegenheiten‘ der SprecherIn zu richten (Goffman, 1978, S. 814).

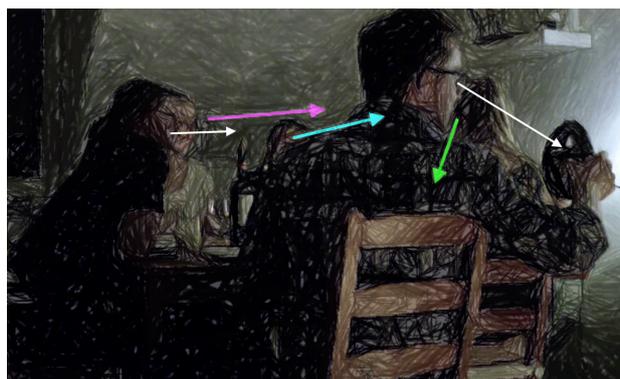
Kombination verschiedener sprachlicher Methoden – vor allem die Produktion eines Herausplatzers mit erhöhter Lautstärke und Sprechgeschwindigkeit sowie die Verwendung eines *change-of-state-tokens* – angezeigt. Mit ihrer anschließenden expliziten, sprachlichen Kategorisierung der Textnachricht macht sie für die Anwesenden ihr *noticing* auf dem für diese nicht einsehbaren Display *accountable*.

Durch die sequenzielle Positionierung und Art der Realisierung des *noticings* wird die aktuell im Vollzug befindliche Aktivität (der Online-Suche nach Bildern der Dozierenden) abrupt und ohne Deplatzierungsmarkierungen unterbrochen. Mir der anschließenden, exklusiven Fokussierung der Smartphoneinhaberin auf die Gerätebedienung wird die Dringlichkeit der erhaltenen Nachricht nun auch in ihrem verkörperten Verhalten angezeigt. Das initiale Lachen einer Anwesenden und die anschließende, mithilfe einer Teilwiederholung des *noticings* fremdinitiierten Reparatur zeigen eindeutig, dass dieses als erklärungsbedürftig behandelt wird. Auch wenn diese Erwiderung also (zunächst) nicht affiliativ ist, kooperieren die TeilnehmerInnen mit der Suspension der im Vollzug befindlichen Aktivität und dem abrupten Themenwechsel insofern, als dass die Rezipientin durch diese Frage ihre Aufmerksamkeit ebenfalls auf das *noticing* richtet und der Smartphoneinhaberin eine Gelegenheit eröffnet, ihre Kategorisierung der Textnachricht zu erklären: „Thus, they collaboratively prioritize dealing with the noticed trouble over the on-hold activity” (Helisten, 2019, S. 339). Die kollaborative Bearbeitung des problematischen *noticings* erfolgt nun, indem die Smartphoneinhaberin die Hervorbringung einer smartphone-gestützten Erzählung initiiert, mit der sie ihre Reaktion auf die erhaltene Nachricht erklärt:

#### Transkript 18: 180418 „Hass ich“\_Teil 2

013 R: ne BÖ:se <<:-)> nach:[richt? > ]  
 014 P: [so WEI::ßT] du schon wann  
 du kommst-  
 015 (0.5)  
 016 P: oh (.) vor EIner stunde;#

#3



017 (0.26)  
 018 R: ha [ha ha ha+ ha [ha ]  
 019 L: [ha ha ha+ ha #[ha ]  
 p: ---->+ blickt auf--->020  
 #4



020 N: [ ((hustet)) ] ((hustet)) °hh+ (.)  
 ((hustet))  
 p: ---->+ auf  
 SP--->024

021 (0.2)  
 022 P: [<<len> i:] [:ch KO:M ]me >-

In Überlappung zur finalen Silbe von R's Äußerung in Z. 013 beginnt P – ihren Blick auf das Smartphonedisplay gerichtet – eine Textnachricht vorzulesen (Z. 014): Das turn-initiale „so“ erscheint dabei als fragmentierte Version des Zitationsmarkers „und ich/er/sie so“ (Golato, 2000) und leitet die anschließende, wörtliche Wiedergabe der Nachricht ein, mit der der Autor<sup>110</sup> die Empfängerin fragt, ob sie wisse, „wann“ sie käme. Nach einer kurzen Stille (Z. 015) beschreibt P – eingeleitet vom Erkenntnisprozessmarker „oh“ – den Zeitpunkt des Erhalts der Nachricht. Durch den performativen Einsatz des *oh-preface* wird die Erzählwürdigkeit des Ereignisses – das Verpassen des Erhalts der Nachricht – hervorgebracht und die Aufmerksamkeit der RezipientInnen auf diesen Aspekt der rekonstruierten Ereignisse gelenkt.

Anschließend entsteht erneut eine kurze Gesprächspause (Z. 017), in der unter den Anwesenden zunächst Unsicherheit darüber zu herrschen scheint, ob die fragmentierte Erzählung ein Ende erreicht hat. Diese Uneindeutigkeit mag darin begründet sein, dass P in der typischen Körperhaltung des smartphone-assistierten Erzählens verharrt und nicht vom Display des mobilen Endgeräts zu den RezipientInnen schaut<sup>111</sup>. Als *accountable* wird, das P sich nicht weiter als

<sup>110</sup> Wie sich im weiteren Verlauf der Interaktion zeigt, handelt es sich um einen männlichen Autor – den Partner von P.

<sup>111</sup> Im Video lassen sich zumindest keine Veränderungen in der Körper- und Kopfpositionierung der Smartphoneinhaberin beobachten. Aufgrund der Kameraposition lässt sich jedoch nicht ausschließen, dass sie aus dieser Position zu den RezipientInnen hochschaut.

nächste Sprecherin selbst auswählt, beginnen zunächst R (Z. 018) und kurz darauf L (Z. 019) laut zu lachen, was P veranlasst, zu ihnen zu schauen. Das Lachen wird in Z. 020 zunächst durch das Husten von N übertönt und verebbt dann, woraufhin P sich wieder der Bedienung des mobilen Endgeräts zuwendet. Nach einer kurzen Stille (Z. 021) beginnt P nun in „typing voice“ (Komter, 2006, S. 208)<sup>112</sup> und simultan zur Bedienung des Smartphones mit „i:ch KO:Mme“ (Z. 022) die Produktion ihrer Antwort auf die erhaltene Nachricht zu verbalisieren. Mit dieser *multimodalen Gestalt* macht sie für die anderen Anwesenden *accountable*, dass sie die empfangene Nachricht als Teil eines nicht abgeschlossenen Ereignisses (vgl. Georgakopoulou, 2008, S. 601) behandelt und nun in situ die Produktion einer medial-vermittelten Erwiderung erfolgt.

Anhand des Lachens von R und L zeigt sich, dass bereits durch das Vorlesen der Textnachricht und die Kontextualisierung durch den Zeitpunkt des Empfangs das Ereignis auf eine Weise rekonstruiert wurde, dass es von ihnen als Lachanlass behandelt werden kann. Dass diese fragmentierte Rekonstruktion ein *laughable* konstituiert, kann damit erklärt werden, dass es sich bei der empfangenen Nachricht um einen Versuch der „micro-coordination“ (vgl. Ling & Yttri, 2002, S. 142 ff.) durch den Versender handelt: Die ständige Möglichkeit zur textbasierten Telekommunikation wird im Alltag dazu genutzt, Verabredungen und Zeitpläne kurzfristig miteinander zu koordinieren. Diese Art der Kommunikation ist für ihr Gelingen jedoch darauf angewiesen, dass die GeräteinhaberInnen ihre Erwiderungsverantwortung (Knorr Cetina, 2009, S. 68) wahrnehmen – denn nur durch zeitnahe Erwiderungen können Verabredungen spontan vereinbart werden. P scheint sich mit dem Erkenntnisprozessmarker „oh“ in Z. 016 implizit an ihrem Versäumnis der Aufrechterhaltung einer Erwiderungsverantwortung zu orientieren, so dass das Lachen von R und L als Amüsement über diesen Fauxpas der Smartphoneinhaberin erscheint.

Eine derartige Interpretation kann aufgrund der fragmentierten narrativen Rekonstruktion und einer fehlenden Explikation des TeilnehmerInnenverständnisses nur tentativ und mit Rückgriff auf ethnografisches Wissen formuliert werden. Eindeutig zeigt sich jedoch, dass die Erzählung

---

<sup>112</sup> Nach Komter (2006) können derartige Äußerungen, die oftmals in gesenkter Lautstärke und reduzierter Sprechgeschwindigkeit realisiert werden, als ‚Vorlesen des Geschriebenen‘ oder als ‚Selbstdiktat für das zu Schreibende‘ dienen. Sie werden als Methoden eingesetzt, um die in ‚Ausführung befindliche Arbeit‘ des Schreibens oder Tip-pens zu demonstrieren und das Rederecht zu behalten (vgl. Komter, 2006, S. 207–208). Während des Smartphonegebrauchs werden sie genutzt, um für die Anwesenden *accountable* zu machen, dass die GeräteinhaberInnen sich der Gerätebedienung zuwenden und um anzuzeigen, was auf dem Display eingegeben wird (vgl. Suderland, 2019, S. 104). Aus Perspektive der einzelnen Kamera lässt sich hier nicht beurteilen, ob P’s Äußerung den exakten Wortlaut dessen wiedergibt, was sie über das Display des Gerätes eingibt. Inwiefern Bediengesten und sprachliche Äußerungen miteinander synchronisiert werden, ließe sich lediglich durch eine Kombination von Video- oder Audioaufzeichnungen und Screen-Capture-Aufnahmen beantworten, wie sie bspw. von Brown et al. (2018) verwendet werden.

– ohne methodische Vorbereitung oder explizite Benennung der beteiligten AkteurInnen – für die zwei RezipientInnen bereits in dieser Form als Lachanlass verständlich ist. Dies verdeutlicht, dass Geschichtenerzählungen aufgrund von Kontextualisierungshinweisen, indexikalen Referenzen und geteiltem Hintergrundwissen der TeilnehmerInnen als solche *accountable* werden können (vgl. Spreckels, 2008, S. 410).

Während sich die Smartphoneinhaberin im Anschluss an die lachende Würdigung des *tellings* der Produktion einer Antwort an ihren abwesenden Kommunikationspartner zuwendet und dies verbalisiert, wird sie von M überlappt, der Probleme im Verstehen anzeigt. Erneut werden P's Blicke mit „+“ und ihre verkörperten Handlungen mit „±“ transkribiert. M's Blicke und Gesten werden mit „\*“ annotiert.

### Transkript 19: 180418 „Hass ich“ \_Teil 3:

022 P:           [<<len> i:][:ch KO:M       ]me >-  
 023 M:                           [wie HASS ich,]  
 024 P:   nur (-) ja+ dass ich natürlich nicht geANTwortet  
       <<:-> [hab > ]-+  
       ---->+...zu M-----  
       -----,,,,----->+  
 025 X:           [ha ha ] +ha ha  
       p:                           +auf SP--->030  
 026 P:   und ich KOMme in <<len>einer stunde;>>  
 027       (0.2)  
 028 M:   wie w? #(.) warum schreibt der (denn) HASS ich;  
       #5



029       ±(0.2)  
       p:   ±....beugt sich mit SP vor zu M--->049  
 030 P:   \*weiß du schon# wann du KOMMST zweiundzwanzig uhr  
       vierzig- (.)  
       m:   \*auf Smartphonedisplay--->037  
       #6



mit dem in Überlappung zu ihrer Äußerung begonnenen Lachen (Z. 025) schaut sie wieder auf das Display und nimmt in Z. 026 die Nachrichtenproduktion wieder auf. Dies macht sie erneut durch eine *multimodale Gestalt accountable*, in der sie den Nachrichtentext mit *typing voice* verbalisiert. An dieser Stelle zeigt sich, dass die lautsprachliche Realisierung der Displayeingebe, die nun für sich eine vollständige Turnkonstruktionseinheit darstellt, gleichzeitig mehrere Funktionen (vgl. Schegloff, 2007a, S. 169) für die lokale Interaktion übernimmt: Indem mit ihr *accountable* gemacht wird, welchem Zweck die Handhabung des Gerätes dient, bearbeitet sie das grundsätzliche Problem der *bystander inaccessibility* des Smartphonegebrauchs. Im spezifischen Kontext einer *im Entstehen befindlichen* Geschichtenerzählung (re-)konstruiert die Smartphoneinhaberin gleichzeitig den nächsten Zug in der Sequenz von Ereignissen ihrer Erzählung, indem sie vorliest, was sie ihrem Chatpartner antwortet. Auf diese Weise demonstriert sie auch die Konsequenzen der Nachricht für zukünftige Handlungsabläufe: Sie verkündet durch die Verbalisierung der Nachricht, wann sie nach Hause kommen wird<sup>113</sup>.

Nach einer kurzen Stille (Z. 027) ergreift M erneut das Rederecht und fordert eine weitere Erklärung für die empfangene Nachricht ein: Die wiederholten Neustarts am Anfang des Turns scheinen einen Versuch darzustellen, die Aufmerksamkeit der Smartphoneinhaberin zu rekrutieren (vgl. C. Goodwin, 1981, S. 61). Anhand des Videos lässt sich jedoch nicht eindeutig feststellen, ob diese Versuche erfolgreich sind. Nach einer Mikropause setzt M in jedem Fall die Äußerungsproduktion fort und zeigt durch die Formulierung der Frage dabei eindeutig an, dass er das Verhalten des abwesenden Chatpartner als beschwerdewürdig behandelt (Z. 028: „warum schreibt der (denn) HASS ich;“).

An dieser Stelle wird offenbar, dass die TeilnehmerInnen unterschiedliche Interpretationen der narrativ rekonstruierten Ereignisse entwickeln: Während P sich selbst die Verantwortung für die ungehaltene Nachricht aufgrund ihres Versäumnisses zuschreibt, behandelt M die Nachricht des Abwesenden selbst als beschwerdewürdig. In Reaktion auf seine Frage nach einer Erklärung für das Verhalten des Abwesenden beugt sich P mit ihrem Smartphone zu M (Z. 029), sodass dieser visuellen Zugriff zum Display erhält. Anschließend produziert sie nun eine adressaten- und kontextspezifisch modulierte Version ihrer Geschichtenerzählung, der sie weitere Informationen hinzufügt, die ihr die Verantwortung für die erhaltene Nachricht zuschreiben sollen (vgl. Mandelbaum, 1993, S. 263). Mit Beginn der sprachlichen Rekonstruktion in Z. 030 beugt sich auch M vor und schaut auf das Display, sodass er sich die Textnachrichten simultan

---

<sup>113</sup> Damit besitzt das Vorlesen der Erwiderung in diesem spezifischen Fall auch eine praktische Relevanz für die Anwesenden: Die Nachricht projiziert einen Zeitpunkt, an dem die Smartphoneinhaberin das gesellige Miteinander verlassen wird.

zur Produktion der Erzählung lesend aneignen kann. Die Smartphoneinhaberin liest erneut die erhaltenen Nachrichten vor, diesmal ergänzt um die Zeitmarken ihres Empfangs und in chronologischer Reihenfolge. Indem das Vorlesen der Zeitmarke der ersten Nachricht in der turn-finalen (Z. 030) und das der zweiten Nachricht in turn-initialer Position erfolgt (Z. 031), wird der zeitliche Abstand zwischen den Nachrichten in den Vordergrund der narrativen Rekonstruktion gerückt. Mit Realisierung des Wortlauts der erklärungsbedürftigen Nachricht (Z. 031: „HASS ich“) schaut die Erzählerin zu M und fügt in schnellem Anschluss einen *Account* hinzu (Z. 032), mit dem sie sich explizit die Verantwortung für die erhaltene Nachricht zuschreibt („=weil ich nicht geANTwortet hatte;=“). Während P in der ersten Version ihrer Erzählung den Zeitpunkt des Erhalts der fokalen Textnachricht als zeitliche Schätzung formulierte (Z. 016), nutzt sie in dieser anschließenden Version nun die *Affordanz* der IM-Applikation, dass „[t]he threaded chat . . . prominently displays the times of the message, proposing this as something that can be read aloud“ (Brown et al., 2018, S. 13): Das Vorlesen des exakten Empfangszeitpunkts der einzelnen Nachrichten wird von der Smartphoneinhaberin als Ressource verwendet, die ihre Version der Ereignisse stützen soll und ihr die Verantwortung für die Entwicklung der Ereignisse zuschreibt (vgl. Mandelbaum, 1993, S. 163).

Diese Version der narrativen Rekonstruktion veranlasst M dazu, im unmittelbaren Anschluss nach der sozialen Beziehung zwischen der Erzählerin und ihrem Telekommunikationspartner zu fragen (Z. 033: „=ist das dein BOYfriend?“). Mit dieser Frage zeigt er an, dass die Geschichte bis hierhin noch keinerlei explizite Informationen über den abwesenden Kommunikationspartner enthielt, diese aber für ihn zum Verständnis – und damit für eine angemessene Erwiderung – erforderlich sind. Außerdem macht die Äußerung die spezifische Mitgliedschaftskategorisierung (Sacks, 1972) eines ‚romantischen Partners‘ relevant und ruft somit implizit Fragen nach den spezifischen Rechten und Verantwortlichkeiten der Mitglieder dieser sozialen Kategorie auf. Nachdem P die Frage bejaht (Z. 034), entwickelt sich in der anschließenden *response sequence* eine angeregte Diskussion (nicht im Transkript enthalten) über den angemessenen Umgang mit Medien in romantischen Beziehungen sowie über die wechselseitigen Erwartungen und Verpflichtungen der PartnerInnen.

Die Analyse dieses zweiten Falles zeigt exemplarisch, dass eingehende Textnachrichten auch mittels *noticings* in der F2F-Interaktion thematisiert werden können, die durch ihre sequenzielle Platzierung und methodische Hervorbringung den Erhalt der Nachricht als dringliche Angelegenheit darstellen. Neben der disjunkten sequenziellen Positionierung werden derartige *noticings* auf der vokalen Ebene durch die methodische Verwendung von Herausplatzern, *oh-pre-faces* sowie erhöhte Lautstärke und Sprechgeschwindigkeit als dringlich dargestellt. Derartig

produzierte *noticings* unterbrechen die aktuelle, im Vollzug befindliche Aktivität (hier eine smartphone-gestützte Online-Suche). Indem die TeilnehmerInnen das *noticing* als *accountable* behandeln, Verstehensprobleme anzeigen oder Erklärungen einfordern, suspendieren sie die vorherige Aktivitäten und alignieren sich mit der durch das *noticing* eröffneten Sequenz, in der die Bearbeitung der mit dem Bemerkten verbundenen Störungen zum Fokus der lokalen Interaktion wird (vgl. Helisten, 2019, S. 339).

Die unterschiedlich ausgestalteten Versionen der durch das Vorlesen der Textnachrichten gestützten Geschichtenerzählung werden von der Smartphoneinhaberin hier methodisch eingesetzt, die abrupte Thematisierung und Interpretation der erhaltenen Nachricht jeweils rezipientenspezifisch *accountable* zu machen (vgl. Mandelbaum, 1993). Es handelt sich dabei um fragmentierte Geschichten von in Entwicklung befindlichen Ereignissen (Georgakopoulou, 2008, S. 601), in deren Kontext die Verbalisierung der in-situ produzierten, medial-vermittelten Erwiderung als weiteres Element einer Sequenz von Ereignissen fungiert. Diese Art der multimodalen Hervorbringung der narrativen Rekonstruktion macht außerdem für die Anwesenden nachvollziehbar, dass der Empfang der Nachrichten Konsequenzen für zukünftige Handlungsabläufe hat (Georgakopoulou, 2008, S. 601): Zum einen behandelt die Smartphoneinhaberin die erhaltene Nachricht als etwas, das im ‚Hier und Jetzt‘ der F2F-Interaktion beantwortet werden muss, zum anderen beeinflusst sie den Zeitpunkt, an dem sie das gemeinsame Treffen verlassen wird.

Anhand der Unterschiede in den zwei Versionen der Geschichtenerzählungen zeigt sich, dass die fragmentierte Variante, in der die Ereignisse nicht in chronologischer Reihenfolge rekonstruiert werden, für zwei Rezipientinnen aufgrund der Kontextualisierungshinweise und geteilten Wissens als Lachanlass verständlich wird. Für M ist diese Version jedoch nicht verständlich und er initiiert eine Fremdreparatur, die zunächst durch eine Erklärung bearbeitet wird. Nachdem M eine disaffiliative Interpretation der Ereignisse äußert, produziert die Smartphoneinhaberin eine modifizierte Version der Erzählung, in der sie dem Rezipienten visuellen Zugriff zum Display gewährt und die Ereignisse nun chronologisch geordnet rekonstruiert. Darüber hinaus wird die *Affordanz* der IM-Kommunikation, dass jede Nachricht mit einer Zeitmarke versehen wird, von der Erzählerin durch wörtliches Vorlesen (vgl. Brown et al., 2018, S. 13) als Ressource eingesetzt, um ihre Interpretation der rekonstruierten Ereignisse zu stützen. Aufgrund des heterogenen Wissens der Anwesenden und der unterschiedlichen Interpretationen der rekonstruierten Ereignisse zeigt sich an diesem Fall sehr eindrücklich, wie der gleiche Chatverlauf für unterschiedlichen AdressatInnen rezipientenspezifisch angepasst und für den jeweiligen kommunikativen Zweck, der mit der Erzählung verfolgt wird, ausgestaltet wird.

### 5.3.3 Fazit: In der Entstehung befindliche, smartphone-assistierte Geschichtenerzählungen

In diesem Kapitel wurde eine Variante *smartphone-assistierter Geschichtenerzählungen* diskutiert, die durch das beobachtbare Bemerkens einer auf dem mobilen Endgerät eingegangenen Textnachricht veranlasst wird. Damit rückten die *synthetischen* Komponenten mediatisierter F2F-Interaktionen in den Fokus der Analysen: Über Smartphones können zeit- und ortsunabhängig potenziell (interaktions-)relevante Informationen empfangen werden, die durch veränderte Displayanzeigen auf dem Gerät dargestellt werden (vgl. Knorr Cetina, 2009, S. 68–72). Diese technischen Möglichkeiten können dazu führen, dass die TeilnehmerInnen durch ein *Monitoring* des Displays anzeigen, den kommunikativen Affordanzen der mobilen Kommunikationstechnologien gegenüber eine (gewisse) Erwidungsverantwortung zu besitzen und auf eingehende Nachrichten reagieren zu müssen.

Das verkörperte Bemerkens einer erhaltenen Nachricht kann auf verschiedene Weisen in der F2F-Interaktion relevant gemacht werden<sup>114</sup>. Wie anhand der exemplarischen Analyse der Gesprächsepisoden gezeigt wurde, können die erhaltenen Textnachrichten an der nächsten, sequenziell passenden Position sprachlich in das Gespräch eingeführt und zu einer Geschichtenerzählung entwickelt werden: Durch die Positionierung des *noticings* nach der beobachtbaren Beendigung eines Themas wird angezeigt, dass die SmartphoneinhaberInnen die Abwicklung der aktuell im Vollzug befindlichen Aktivität priorisieren und den Erhalt der Nachricht nicht als dringliche Angelegenheit behandeln. Die Textnachrichten können aber auch durch disjunkt platzierte *noticings* in das Gespräch eingeführt werden: Diese sequenzielle Positionierung zeigt – im Zusammenspiel mit prosodischen und sprachlichen Ressourcen, mit denen ihre Dringlichkeit angezeigt wird – dass der Erhalt der Textnachricht als interaktionsrelevantes Ereignis priorisiert und damit eine momentan im Vollzug befindliche Aktivität der lokalen Interaktion unterbrochen wird. In allen Fällen muss die Aufmerksamkeit der Anwesenden rekrutiert werden, damit das *noticing* zum gemeinsamen Fokus der F2F-Interaktion werden kann. Dabei zeigt sich selbst bei disjunkt positionierten *noticings*, dass TeilnehmerInnen mit dem abrupten Themenwechsel kooperieren und aktuelle Aktivitäten unterbrechen, um die mit dem *noticing* verbundenen Störungen gemeinsamen zu bearbeiten (vgl. Helisten, 2019, S. 339).

---

<sup>114</sup> Dass es im gesamten Korpus nur zu wenigen Fällen durch den Erhalt von Textnachrichten veranlasster, *smartphone-assistierter Geschichtenerzählungen* kommt, liegt zu einem daran, dass den TeilnehmerInnen verschiedene Methoden der sprachlichen Thematisierung empfangener Nachrichten zur Verfügung stehen (vgl. Brown et al., 2018, S. 7–9, 15–19; DiDomenico et al., 2018, S. 19–22). Zum anderen begründet es sich darin, dass etliche GeräteinhaberInnen ihre Smartphones (über weite Teile der Interaktion) nicht als Technologie behandeln, denen gegenüber sie eine Erwidungsverantwortung besitzen – entsprechend werden die Geräte außerhalb des Hör- oder Sichtbereichs aufbewahrt und nur unregelmäßig auf aktualisierte Informationen ‚gecheckt‘.

Die Erzählwürdigkeit der anschließenden *smartphone-assistierten Geschichtenerzählungen* wird maßgeblich dadurch hervorgebracht, dass die GeräteinhaberInnen für die Anwesenden *accountable* machen, dass die empfangene Nachricht (potenzielle oder praktische) Folgen für zukünftige Handlungsabläufe hat (vgl. Georgakopoulou, 2008, S. 601): In nicht-dringlichen Erzählungen kann dies über die narrative Rekonstruktion von Handlungsabläufen erfolgen, die notwendig wären, wenn die EmpfängerIn sich in einer anderen Situation als der lokalen befände; in dringlichen Geschichten können dies zukünftige Handlungen mit praktischen Konsequenzen für die ErzählerIn (und die Anwesenden) sein. Allen diesen Erzählungen ist gemein, dass es sich nicht um die narrative Rekonstruktion von abgeschlossenen Geschehnissen, sondern um Geschichten von *im Entstehen befindlichen* Ereignissen handelt, deren weitere Entwicklung durch die Erzählung projiziert oder in situ durch das Verfassen einer Textnachrichten-Erweiterung realisiert wird.

Dass die RezipientInnen im Vollzug beider – sehr unterschiedlich in die Interaktion eingeführten – Geschichtenerzählungen Probleme des Verstehens anzeigen, sodass in Reparatursequenzen ein intersubjektives Verständnis (wieder-)hergestellt werden muss, mag mit den spezifischen Herausforderungen durch eingegangene Nachrichten ausgelöster *noticings* zusammenhängen: Im Vergleich zu anderen *noticings*, mit denen SprecherInnen anzeigen, dass sie etwas in der wahrnehmbaren Umwelt des Settings bemerkt haben und die Aufmerksamkeit der anderen Anwesenden auf dieses ‚etwas‘ lenken (vgl. Schegloff, 2007a, S. 219), handelt es sich beim Empfang von Textnachrichten um Ereignisse, die für andere Anwesende (meist) aufgrund der materiellen Eigenschaften des Gerätes (Größe und Platzierung des Displays) nicht wahrnehmbar sind. Entsprechend stehen die Bemerkenden auch in diesem Kontext vor der Aufgabe, die potenzielle *bystander inaccessibility* des Smartphonegebrauchs zu bearbeiten und für andere Anwesende *accountable* zu machen, was sie auf ihrem Smartphonedisplay bemerkt haben. Diese *Accountability* kann über die methodische Entwicklung einer Prä-Sequenz hergestellt werden oder ist für Anwesende aufgrund von Kontextualisierungshinweisen und geteiltem Wissen erkennbar. Dabei scheint durchaus ein Spannungsverhältnis dadurch zu entstehen, dass die SmartphoneinhaberInnen unvermittelt eine Nachricht einer bekannten KommunikationspartnerIn oder zu einem vertrauten Ereignis bemerken, in der lokalen Produktion der Geschichtenerzählung jedoch vor der Herausforderung stehen, die RezipientInnen über diese unvorhergesehenen Ereignisse auf eine Art und Weise zu informieren, dass diese für die Anwesenden sinnhaft verständlich werden.

Die Integration in situ empfangener Textnachrichten in *smartphone-assistierte Geschichtenerzählungen* stellt eine Methode der GeräteinhaberInnen dar, andere Anwesende – auf

humoristische oder problematisierende Weise – an Geschehnissen teilhaben zu lassen, die an anderen Orten stattfinden, über die sie medial-vermittelt erfahren und die eine potenzielle oder tatsächliche Relevanz für die Geräteinhaberin besitzen. Die grundsätzliche Flexibilität alltäglicher Gespräche ermöglicht es den SmartphoneinhaberInnen, die erhaltenen Textnachrichten als Anlass und – realisiert durch die multimodale Praxis des Vorlesens – materielle Ressource für die Realisierung von Geschichtenerzählungen einzusetzen und so zur Dynamik der thematischen Entwicklung der alltäglichen Gespräche zwischen Anwesenden beizutragen. So wie für die anderen Formen *smartphone-assistierter Geschichtenerzählungen* gilt dabei wiederum, dass direkte mündliche Gespräche sowohl eine Unterstützung für die Aushandlung des sozialen Sinns der archivierten oder empfangenen Textnachrichten bieten (McGregor, 2020, S. 158), als auch ein Forum darstellen, in dem Praktiken des Smartphonegebrauchs erprobt werden und ihre Angemessenheit ausgehandelt wird (vgl. Keppler, 2013, S. 90).

## 6 Die Integration digitaler Fotografien in smartphone-assistierten Geschichtenerzählungen

In den vorangegangenen Kapiteln standen *smartphone-assistierte Geschichtenerzählungen* im Fokus, die maßgeblich durch Praktiken des Vorlesens (und Lesens) von Textnachrichten hervorgebracht werden. Die *Affordanz* des Smartphones, auf Threads gespeicherter und gerade erhaltener Textnachrichten zugreifen zu können, wird von den TeilnehmerInnen als Ressource genutzt, die Beiträge der unterschiedlichen ChatteilnehmerInnen in ihrem Wortlaut zu zitieren und in ihrer chronologischen Abfolge so zu rekonstruieren, dass sie als Erzählungen von alltäglichen, aber bedeutungsvollen sozialen Geschehnissen für andere Anwesende *accountable* werden. Auch wenn neueste Instant-Messaging-Applikationen durch die Möglichkeiten zum Versand von Emojis, GIFs, Fotografien oder Videos Formen der multimedialen Kommunikation erlauben, stellt der überwiegend schriftsprachliche Charakter der Kurznachrichtenkommunikation die zentrale Ressource dieser *smartphone-assistierten Geschichtenerzählungen* dar und wurde entsprechend ins Zentrum der Analysen gestellt. Selbst wenn Bildinhalte als Ressource dieser Erzählungen eingesetzt wurden (vgl. Kap. 5.1.2), wurde die Visualität dieser Medieninhalte bis hierhin nicht als eigenständiges Thema der Analyse behandelt.

Im folgenden analytischen Teil dieser Arbeit werden nun jedoch eben diese visuellen Medieninhalte als Ressource für die Hervorbringung von Geschichtenerzählungen in den Fokus der Analysen gerückt. Grundsätzlich können über die mobilen Endgeräte Bilder, Fotografien und Videos zum einen durch Online-Suchanfragen gesucht, gefunden und in die F2F-Interaktion eingespielt werden (Brown et al., 2015, S. 508). Zum anderen verfügen moderne Smartphones über integrierte (Foto- und Video-)Kameras und Festplattenspeicher, die die Produktion und Archivierung von (audio-)visuellen Medieninhalten ermöglichen. Da in meinem Korpus jedoch kein einziger Fall der smartphone-gestützten Integration eines Videos in eine Geschichtenerzählung enthalten ist, werde ich im Folgenden ausschließlich die Integration digitaler Fotografien untersuchen. Den Analysen werden zunächst einige Überlegungen zu diesen visuellen Medieninhalten vorangestellt.

Die „Entwicklung des Smartphones zum aufzeichnungs- und internetfähigen Kommunikationsmedium erlaubt die mühelose Dokumentation des Selbst und seiner Erlebnisse und die Information Dritter in Form von Bildern“ (Meyer & Meier zu Verl, 2017, S. 79). Mit der Aneignung dieser technischen Möglichkeiten haben sich neue kulturelle Praktiken des alltäglichen Fotografierens ausgebildet, nicht zuletzt die ‚Selfie‘-Fotografie: Millionenfach fertigen die NutzerInnen mit ihren Smartphones Aufnahmen von sich selbst und ihren FreundInnen, Familienmitgliedern oder Haustieren in alltäglichen und außergewöhnlichen Situationen ihres Lebens

an und veröffentlichen sie über soziale Netzwerke. Allein auf der Plattform Instagram finden sich 437 Millionen Fotografien mit dem Hashtag „selfie“<sup>115</sup>.

Nicht nur in den Medien werden angeregte, oftmals kritische Diskussionen über die gesellschaftlichen und sozialen Konsequenzen dieser neuen Formen des Fotografierens geführt. Auch in der Wissenschaft herrscht ein reges Interesse an der Erforschung dieses Phänomens. Die vorhandenen Studien analysieren diese Fotografien jedoch zumeist ausschließlich als mediale Produkte in den ‚Online-Kontexten‘ ihrer Veröffentlichung (bspw. Fox & Rooney, 2015; Katz & Crocker, 2015; Tifentale & Manovich, 2015). Wesentlich seltener werden die Produktion sowie die Kommunikation mit und über derartige Fotografien als lokale Praxis untersucht (s. aber Gibbs et al., 2015; Weilenmann & Hillman, 2019). Um zur Schließung dieser Forschungslücke beizutragen, wird sich diese Arbeit deshalb einer spezifischen Form des lokalen Gebrauchs digitaler Fotografien widmen: Ihre *smartphone-gestützte Integration in Geschichtenerzählungen* in alltäglichen Unterhaltungen. Ich werde anhand der Analyse exemplarischer Fälle aufzeigen, wie digitale Fotografien als konstitutive Ressource in bildgestützten Geschichtenerzählungen (Kap. 6.1) und in der Post-Sequenz von Erzählungen als Erwiderung auf disaffiliatives Verhalten der RezipientInnen (Kap. 6.2) eingesetzt werden können.

Vorher ist es jedoch erforderlich, das spezifische Verständnis visueller – insb. fotografischer – Medieninhalte zu explizieren, das dieser Arbeit zugrunde liegt, und den relevanten Forschungsstand zum (smartphone-gestützten) Gebrauch von Bildern in alltäglichen Interaktionen darzustellen. Betrachtet man die gesellschaftliche Verbreitung der (Amateur-)Fotografie, so lässt sich zunächst festhalten, dass Bilder seit „dem Aufkommen technischer Apparate zu ihrer Produktion, Vervielfältigung und Diffusion, – insbesondere den Foto-, Video- und Fernsehgeräten“ (Burri, 2008, S. 342) – nicht mehr aus dem gesellschaftlichen Alltag wegzudenken sind. Sie stellen vielmehr bereits seit langer Zeit eine wichtige Form sozialer Kommunikation dar (vgl. Chalfen, 1987, S. 15). Doch trotz dieser gesellschaftlichen Verbreitung fotografischer Medien fand lange Zeit keine systematische und empirische Untersuchung dieser Bilder in der allgemeinen Soziologie statt<sup>116</sup>.

Im Unterschied dazu wird in den ethnomethodologisch fundierten Laborstudien (Knorr Cetina, 1995) und „studies of work“ (Bergmann, 2006) bereits seit den 1980er-Jahren die Rolle

---

<sup>115</sup> Stand: 01.12.2020

<sup>116</sup> Nennenswerte Ausnahmen stellen zum einen die frühen Untersuchungen von Pierre Bourdieu zur sozialen Nutzung von Fotografien im Alltag dar (1990, 1991), in denen dieser ihre Funktion als Mittel sozialer Integration rekonstruiert, das vor allem der Unterstützung von rituellen Ereignissen wie Geburtstagen oder Hochzeiten dient. Zum anderen finden sich auch bei Goffman (1979b) analytische Auseinandersetzungen mit Reklamebildern, anhand derer er das Verhältnis zwischen Bild(ern) und gesellschaftlichen Strukturen, insbesondere von geschlechtsspezifischen Verhaltensweisen und Darstellungspraktiken herausarbeitet.

visueller Repräsentationen in sozialen Prozessen der Wissensherstellung und -verbreitung – insbesondere in den Naturwissenschaften – empirisch untersucht (für eine Übersicht s. Burri & Dumit, 2008). Im Fokus steht dabei oftmals die Frage, wie Bilder aufgrund ihrer Materialität und Visualität zur Konstruktion wissenschaftlicher Fakten beitragen, indem sie Untersuchungsergebnisse stabilisieren und kommunizierbar machen (s. bspw. Knorr Cetina & Amann, 1988; Latour & Woolgar, 1986; Lynch, 1985). Ausgehend von diesen Untersuchungen des Einsatzes visueller Repräsentationen im Kontext der wissenschaftlichen Wissensproduktion entwickelte sich ein fruchtbares Forschungsfeld der ethnomethodologisch fundierten „visual analysis“ (C. Goodwin, 2000). Diese ist grundsätzlich von der Auffassung geprägt, dass visuelle Repräsentationen nicht losgelöst von den komplexen, sozialen Aktivitäten untersucht werden können, in die sie eingebettet sind.

Die vorliegende Arbeit schreibt sich dieser Perspektive<sup>117</sup> ein und versteht digitale Fotografien entsprechend als visuelle und materielle Objekte (vgl. Burri, 2012, S. 46), die „eine auf ihrer Fläche erkennbare Konfiguration von Objekten und Relationen [zeigen, Anm. D.S.], mit der zugleich eine durch die Art der Darstellung oder Aufnahme spezifische Sicht auf die im Bild sichtbaren Objekte und Motive eröffnet wird“ (Keppler, 2019, S. 181). Sie stellen zwei-dimensionale, statische Repräsentationen eines oder mehrerer Objekte dar, die von Ort zu Ort ‚transportiert‘ und reproduziert werden können, sowie unveränderbare Eigenschaften besitzen, die auch in den unterschiedlichen Kontexten ihres Gebrauchs konstant bleiben (vgl. „immutable mobiles“, Latour, 1986, 2009). Als visuelle Objekte lassen sich Fotografien dabei nicht unabhängig von den Interaktionsprozessen und Praktiken analysieren, in und durch die sie als relevant und bedeutungstragend hervorgebracht werden (vgl. C. Goodwin, 2000, S. 165): Fotografien erhalten ihre soziale Bedeutung im Kontext der spezifischen Praktiken ihres Gebrauchs, sodass die gleiche Fotografie in unterschiedlichen Settings auf unterschiedliche Art und Weise

---

<sup>117</sup> In der deutschsprachigen Medien- und Kommunikationssoziologie wurden einige Ansätze zur Analyse von Bildern und Fotografien entwickelt, in der diese als Medien kultureller Sinnvermittlung (Müller-Doohm & Neumann-Braun, 1995) verstanden werden, die in einem Wechselverhältnis mit Prozessen der sozialen und symbolischen Weltaneignung, -darbietung oder -erzeugung stehen (vgl. Keppler, 2000, S. 141). Bilder und Fotografien werden dabei als Untersuchungsgegenstand einer allgemeinen Kultursoziologie begriffen und verschiedene Methoden der sozialwissenschaftlichen Bildanalyse entwickelt. Beispielhaft zu nennen sind hier die struktural-hermeneutische Symbolanalyse (Müller-Doohm, 1993, 1997), die Segmentanalyse (Breckner, 2003; vgl. auch Waber, 1993) und insbesondere die dokumentarische Methode (Bohnsack, 2007a, 2007b). Dass diesen Ansätzen hier kein weiterer Raum eingeräumt wird, liegt zum einen daran, dass zu wenig auf die spezifische Bedeutung der Visualität der Bilder eingegangen wird (Burri, 2008, S. 344). Zum anderen handelt es sich um Methoden, die ausschließlich die Bilddokumente einer kultur- und sozialwissenschaftlichen Analyse unterziehen. Sie bieten – anders als die ethnomethodologisch fundierte *visual analysis* – weder einen theoretischen, noch methodischen Zugang zu den Prozessen, durch die die Bilder und Fotografien ihren sozialen Sinn entfalten. Auch wenn diese Ansätze nützliche Anleitungen für die Interpretation von Bildern bereitstellen, so lassen sie sich nicht produktiv für die vorliegende Arbeit nutzen, die davon ausgeht, dass die Bedeutung von Bildern sich grundsätzlich nicht unabhängig von ihren Nutzungskontexten analysieren lässt.

interpretiert werden kann. Um ihre soziale Sinnhaftigkeit rekonstruieren zu können, muss entsprechend empirisch untersucht werden, wie die Bedeutung einer Fotografie im dynamischen (und damit veränderbaren) Zusammenspiel mit anderen eingesetzten, multimodalen Ressourcen im Vollzug konkreter Interaktionsepisoden ausgehandelt und hervorgebracht wird (vgl. C. Goodwin, 2000, S. 166).

Aus der hier eingenommenen Perspektive wird zwar anerkannt, dass Fotografien als materiell-visuelle Objekte bestimmte, unveränderliche Eigenschaften aufweisen, sie ihre soziale Bedeutung jedoch nur im Kontext der spezifischen Praktiken ihres Gebrauchs erhalten. Diese Position lässt sich – analog zum *affordanz*theoretischen Verständnis des Technologiegebrauchs (vgl. Kap. 3.3) – in der Mitte von zwei Polen eines theoretischen Kontinuums verorten: Fotografien werden weder als „privilegierte Wiedergabe der Realität“ (Keppler, 2002, S. 89) verstanden, die eine kontextunabhängige, durch ihre materiellen Eigenschaften determinierte Bedeutung besitzen, noch als Medien, denen eine beliebige Bedeutung in sozialen Aushandlungsprozessen zugeschrieben werden kann (ebd., 2002, S. 89). Erneut wird hier also eine Position zwischen medien- und sozialdeterministischen Perspektiven eingenommen, die sowohl die materielle Visualität der untersuchten Medieninhalte als auch die grundsätzliche Kontextgebundenheit des sozialen Sinns jedweder Handlung anerkennt: Ein fotografisches Bild gewinnt seine Bedeutung demnach, „je nachdem, wie es kommunikativ verwendet wird“ (Keppler, 2002, S. 98).

Im Vergleich zu den zahlreichen Untersuchungen visueller Repräsentation in wissenschaftlichen und anderen professionellen Kontexten existieren kaum (konversationsanalytische) Studien, die den Gebrauch von Bildern und Fotografien in alltäglichen Interaktionen fokussieren. Dies liegt sicherlich nicht zuletzt daran, dass die Verbreitung mobiler Endgeräte, mit denen visuelle Medieninhalte produziert, gespeichert und über das Internet abgerufen werden können, ein verhältnismäßig junges Phänomen ist. Bedenkt man jedoch, wie lange Amateur- und Freizeitfotografie bereits eine verbreitete kulturelle Praxis darstellt, so mag die Nicht-Berücksichtigung damit verbundener Phänomene durchaus verwundern. Im Folgenden werde ich einige zentrale Ergebnisse dieser wenigen Untersuchungen vorstellen<sup>118</sup>, wobei ich neben Studien der smartphone-gestützten Bildintegration auch Analysen verwandter Phänomene – bspw. familiäre Dia-Abende oder die Betrachtung von Fotografien auf digitalen Kameras – diskutieren werde, um die anschließenden Analysen im bestehenden Forschungsstand zu verorten.

---

<sup>118</sup> In einigen Studien zum Smartphonegebrauch in F2F-Interaktion finden visuelle Medieninhalte en passant Erwähnung (vgl. bspw. Brown et al., 2018, S. 13; Robles et al., 2018, S. 158ff.; Weilenmann & Hillman, 2019, S. 12–15), ohne dass jedoch die visuelle Materialität dieser Medieninhalte systematisch berücksichtigt würde. Ich werde im Folgenden ausschließlich Ergebnisse derjenigen Studien vorstellen, die sich dezidiert mit der Visualität der integrierten Bilder und Fotografien auseinandersetzen.

In konversationsanalytischen Untersuchungen wurden bereits einige Erkenntnisse über die sequenzielle Organisation der smartphone-gestützten Bildintegration in alltägliche Gespräche generiert. Diese wird von Keppler (2019) basierend auf der Analyse von Audioaufnahmen in einem Verlaufsmodell der Smartphoneintegration beschrieben: In der Eröffnungsphase wird der Gebrauch des Smartphones initiiert und anschließend ein Suchprozess nach den relevanten Medieninhalten durchgeführt (Keppler, 2019, S. 179). Die anschließende Kernphase gliedert sich ebenfalls in zwei Phasen: eine Prä-Sequenz des Zeigens, in der der Medieninhalt kontextualisiert und ein „Verstehensfokus etabliert [wird, Anm. D.S.], der den gemeinsamen Vergegenwärtigungsprozess vorbereitend lenkt, indem er die Relevanz des jeweiligen . . . Bilds oder Videos für das laufende Gespräch hervorhebt“ (Keppler, 2019, S. 180). Im zweiten Schritt kommt es zum Zeigen des visuellen Medieninhalts und dessen gemeinsamer kommunikativer Aneignung. Die Smartphoneintegration wird in einer Beendigungsphase geschlossen, in der das Gezeigte im gemeinsamen Gespräch weiterverarbeitet wird.

Die interaktive Organisation des smartphone-gestützten Zeigens digitaler Fotografien wird von Oloff (2019a) aus multimodal-konversationsanalytischer Perspektive spezifisch fokussiert. Sie zeigt auf, dass diese (initialen) Zeigesequenzen in Adjazenzpaaren organisiert sind, die teilweise durch verkörperte Handlungen realisiert werden. In Paarsequenzen werden dabei folgende Aufgaben der Interaktionsorganisation bearbeitet: Zunächst muss die Aufmerksamkeit der (potenziellen) RezipientInnen erlangt werden, bevor eine „Fokussierung des Smartphone-displays“ stattfinden kann. Nach der Aufmerksamkeitsfokussierung und vor, simultan zu oder nach der Displayfokussierung formuliert die GeräteinhaberIn eine Betrachtungsanweisung (Oloff, 2019a, S. 214), mit der sie die Fotografie kontextualisiert und ihre Haltung anzeigt. Die RezipientInnen zeigen durch ihre ersten Erwidernungen ihre Alignment (und meist Affiliation) mit der Betrachtungsanweisung an. „[D]a der visuelle Zugang zum Display und die Möglichkeit, den Inhalt zu bewerten, einige Zeit in Anspruch nehmen können“ (Oloff, 2019a, S. 215), erfolgen explizite Bewertungen jedoch meist erst zu einem späteren Zeitpunkt und oftmals nach einer gemeinsamen, kommunikativen Aushandlung. Oloffs Fokussierung von Zeigesequenzen ergänzt also eine spezifische Phase des allgemeinen Verlaufsmodells der Smartphoneintegration um eine multimodale Analyse, die die Materialität des Smartphones und die Relevanz verkörperter Handlungen systematisch berücksichtigt.

Einig sind sich die verschiedenen AutorInnen, dass der kopräsenten Interaktion durch die smartphone-gestützte Integration digitaler Bildinhalte eine zusätzliche Ebene der Interpretation hinzugefügt wird, die beobachtbare interaktionale Konsequenzen für die TeilnehmerInnen hat (Raclaw et al., 2016, S. 377): Es bedarf der gemeinsamen, kommunikativen Aneignung der

Medieninhalte, um zu etablieren, „als was sie im jeweiligen Kontext gesehen und was in ihnen vor allem gesehen wird“ (Keppler, 2019, S. 181; vgl. a Oloff, 2019a, S. 215). Das Teilen von Fotografien in *mobile-supported sharing activities* wird also als gemeinsame Aktivität der Anwesenden hervorgebracht, die sich aus der Handhabung des mobilen Endgeräts, der auf dem Gerätedisplay angezeigten Fotografien sowie verkörperten Handlungen und sprachlichen Äußerungen zusammensetzt, die auf die Fotografie bezogen sind (vgl. Aaltonen et al., 2014, S. 129)<sup>119</sup>.

Das Zeigen und Betrachten der Fotografien verändert die sequenzielle Organisation der F2F-Interaktion dabei vor allem in zweierlei Hinsicht: Zum einen stattet es die SmartphoneinhaberIn – als InteraktionsteilnehmerIn, die mit der praktischen Aufgabe der Handhabung des mobilen Endgeräts betraut ist – mit exklusiven kommunikativen Rechten aus (Brown et al., 2015, S. 511; Keppler, 2013, S. 95). Die physische Bedienung des Geräts, die Möglichkeit zur Veränderung des Displayzugriffs und zur Weitergabe des Geräts können dabei als Ressource für die sequenzielle Organisation smartphone-gestützter Aktivitäten dienen. Zum anderen fügt die Zugänglichkeit der digitalen Fotografie der Interaktion eine zusätzliche semiotische Ressource hinzu, sodass die Turnkonstruktion nicht mehr nur über sprachliche Äußerungen (und verkörperte Handlungen), sondern über multimodale semiotische Einheiten stattfinden kann (vgl. Aaltonen et al., 2014, S. 139). Auf die Weise können auch lexiko-semantisch unvollständige Äußerungen durch ihren Bezug zu einer simultan betrachteten Fotografie für alle praktischen Zwecke verständlich werden. Im Vollzug der Interaktion verleihen sich sprachliche Äußerungen, verkörperte Handlungen und das semiotische Feld der Fotografie in einem dynamischen Zusammenspiel wechselseitig ihren Sinn (vgl. C. Goodwin, 2015b, S. 409).

Im Kontext von F2F-Interaktionen können visuelle Medieninhalte dabei verschiedene Funktionen übernehmen: Grundsätzlich können sie Impulse für die thematische Entwicklung der Gespräche liefern (bspw. Brown et al., 2018, S. 13; Robles et al., 2018, S. 158ff.). Konkret gewährt das Teilen digitaler Bildinhalte den RezipientInnen einen epistemischen Zugang zu einem ansonsten unzugänglichen Referenten, der es ihnen ermöglicht, zu einer (,eigenen‘) Einschätzung dieses Referenten zu gelangen (Raclaw et al., 2016, S. 362). Auf diese Weise fungieren die Fotografien als materiell-visuelle Ressource, die geeignet ist, die RezipientInnen zur Alignierung mit einer von der SmartphoneinhaberIn initiierten smartphone-gestützten Aktivität und

---

<sup>119</sup> Aaltonen et al. (2014) untersuchen den Gebrauch von auf einer digitalen Kamera gespeicherten Fotografien durch einen aphasischen Sprecher in Interaktionen mit seiner Ehefrau und einer ForscherIn. Digitale Kameras weisen aufgrund ihrer Größe, Tragbarkeit und Displayanzeige einige relevante materielle Gemeinsamkeiten zu Smartphones, insb. der hier fokussierten *Affordanz* zur Produktion und Speicherung digitaler Fotografien, auf, sodass die Erkenntnisse dieser Untersuchung unmittelbare Relevanz für die vorliegende Arbeit besitzen.

zur Affiliation mit der von ihr angezeigten Haltung zu leiten (Raclaw et al., 2016, S. 377). Dies gelingt nicht zuletzt, da ihnen oftmals von den TeilnehmerInnen als fotografische Repräsentation einer vorgängigen Realität (vgl. Keppler, 2002, S. 98) eine hohe Glaubwürdigkeit zugeschrieben wird und sie als ‚Belege‘ für Ereignisse vergangenen Ereignisse und Zustände anerkannt werden (Keppler, 2019, S. 185).

Die bis hierhin vorgestellten Studien rekonstruieren stärker kontext- bzw. aktivitätsunabhängige Sequenzmuster, Praktiken und Funktionspotenziale der Integration visueller Medieninhalte und die damit einhergehenden Veränderungen der Interaktionsorganisation. Der spezifische Einsatz visueller Medieninhalte in der interaktiven Hervorbringung von Geschichtenerzählungen ist im Gegensatz dazu aus konversationsanalytischer Sicht noch weitgehend unerforscht: Es finden sich zwar Erwähnungen davon, dass Bilder und Fotografien in Erzählungen integriert werden können (bspw. Brown et al., 2018, S. 13–15; Raclaw et al., 2016, S. 373–376), allerdings existieren kaum systematische Untersuchungen dieses Phänomens.

Am explizitesten wird die Idee durch visuelle Medieninhalte gestützter Geschichtenerzählungen von Isaacs et al. (2012) verfolgt: Die AutorInnen rekonstruieren Formen des ‚Teilens‘ visueller Medieninhalte, die – ebenso wie ‚kanonische‘ Geschichtenerzählungen – einer Drei-Phasen-Struktur aus Prä-Sequenz, Elaboration der Erzählung und Post-Sequenz folgen. Der Hauptteil der Erzählung wird dabei durch das Zeigen eines Videos, einer Fotografie oder anderen visuellen Medieninhalten auf dem Display des Smartphones konstituiert. Dabei argumentieren die AutorInnen, dass „[t]he technology can become the story because it is a physical representation of a prior experience and because it is readily available at a moment’s notice without any prior preparation“ (Isaacs et al., 2012, S. 621). Die Idee, Fotografien und Bilder als materiell-visuelle Repräsentationen vergangener Ereignisse zu verstehen, die eine visualisierte Form von Geschichtenerzählungen ermöglichen, erscheint vielversprechend. Der Aufsatz enthält jedoch keine systematische Analyse des Zusammenspiels der verschiedenen multimodalen Ressourcen im sequenziellen Verlauf der Interaktion, durch die die visuellen Medieninhalte überhaupt erst für die RezipientInnen als Repräsentation einer vergangenen Erfahrung innerhalb einer Geschichtenerzählung *accountable* werden.

Die wohl ausführlichste Auseinandersetzung mit dem Zusammenhang des Zeigens von Fotografien und der Erzählung von Geschichten findet sich in Keplers (1994, S. 186–209) Analysen familiärer Dia-Abende, in deren Kontext Erinnerungsbilder – Fotografien aus dem Familienleben, von Jubiläen, Hochzeiten oder gemeinsamen Reisen – gezeigt werden. Im Verlauf der Dia-Abende übernimmt diejenige TeilnehmerIn, die den Diaprojektor bedient, die Rolle der HauptsprecherIn und sorgt mit dem Bildwechsel für eine Strukturierung der sprachlichen

Interaktion. Die gezeigten Familienfotografien liefern narrative Impulse, die sprachliche – oftmals kollaborative – Geschichtenerzählungen veranlassen, die als kommunikative Akte der kollektiven Erinnerung fungieren, aus denen sich das Familiengedächtnis bildet (Keppler, 1994, S. 207). Dabei ist „das, was sich auf einem Foto relevantes zeigt, was es im Kontext der Familiengeschichte bedeutet, immer das Ergebnis einer kollektiven ‚Auslegung‘ der betreffenden Ereignisse“ (Keppler, 1994, S. 188).

Die hier vorgestellten Erkenntnisse über die kontextunabhängigen Organisationsprinzipien der Integration visueller Medieninhalte in F2F-Interaktion bilden den Ausgangspunkt für die folgenden Analysen. Die Darstellung zeigt jedoch auch, dass noch keine systematischen Analysen der methodischen smartphone-gestützten Integration digitaler Fotografien in alltägliche Geschichtenerzählungen vorliegen, die systematisch das Zusammenspiel multimodaler Ressourcen, der physischen Handhabung der mobilen Endgeräte und den visuellen Medieninhalten berücksichtigen. Aus dieser Perspektive werde ich in den folgenden zwei Kapiteln zunächst untersuchen, wie digitale Fotografien in den Hauptteil *smartphone-assistierter Geschichtenerzählungen* integriert werden, bevor anschließend rekonstruiert wird, wie Fotografien als Ressource in der Hervorbringung von *response sequences* von Erzählungen eingesetzt werden.

## **6.1 Bildgestützte, smartphone-assistierte Geschichtenerzählungen**

Als aufzeichnungsfähige Kommunikationsmedien ermöglichen Smartphones ihren NutzerInnen die spontane fotografische Dokumentation des Selbst und seiner Erlebnisse. Als internetfähige Technologien erlauben sie die medial-vermittelte Information abwesender Dritter mithilfe dieser Bilder (Meyer & Meier zu Verl, 2017, S. 79). Mit der Analyse von bildgestützten, *smartphone-assistierten Geschichtenerzählungen* widme ich mich in diesem Kapitel einer kommunikativen Praktik, durch die anwesende Dritte mithilfe der Integration digitaler Fotografien in Geschichtenerzählungen über das ‚Selbst der ErzählerIn und ihrer Erlebnisse informiert‘ werden: Derartige Geschichtenerzählungen zeichnen sich dadurch aus, dass eine (oder mehrere) digitale Fotografie(n) über das Smartphone abgerufen und in den Hauptteil der Erzählung integriert werden. Für die Hervorbringung dieser Variante spielt also die *Affordanz* des Smartphones, digitale Fotografien aufnehmen, archivieren und jederzeit wieder abrufen zu können, eine konstitutive Rolle. Bei den Fällen dieser Subkollektion handelt es sich – soweit es sich mithilfe der Videoaufnahmen rekonstruieren lässt – um mit der integrierten Kamera der mobilen Endgeräte aufgenommene Fotografien, mit denen die SmartphoneinhaberInnen eine ‚gute Zeit‘ dokumentierten (vgl. Chalfen, 1987, S. 138) – sei es durch die Aufnahme einer leckeren Mahlzeit, eines Naturschauspiels oder einer Freizeitaktivität.

Wie bei allen Formen des smartphone-assistierten Erzählens handelt es sich auch bei dieser Variante um eine multimodale Praktik, für die das Zusammenspiel sprachlicher Äußerungen mit materiellen, verkörperten und räumlichen Ressourcen eine konstitutive Rolle spielt (vgl. König & Oloff, 2018b, S. 213). Die durch das Vorlesen von Textnachrichten gekennzeichnete Variante *smartphone-assistierten Erzählens* ist maßgeblich von sprachlichen Äußerungen im Modus der DRS geprägt, mit denen die archivierten Nachrichten wiedergegeben werden. Dadurch weist sie an der Gesprächsoberfläche einige Gemeinsamkeiten zu mündlich-sprachlich verfertigten Geschichtenerzählungen auf, sodass sie bereits lediglich durch die Betrachtung der sprachlichen Äußerungen als solche identifizierbar sind. Die im Folgenden fokussierten bildgestützten Erzählungen werden konstitutiv durch die Integration und (kommunikative) Aneignung digitaler Fotografien hervorgebracht. Da hier die visuellen Medieninhalte von den TeilnehmerInnen als konstitutive Elemente der Erzählung behandelt werden, wird das konversationsanalytische Verständnis von Geschichtenerzählungen auf eine neue Weise herausgefordert: Geschichtenerzählungen werden meist als „in Form einer Diskurseinheit realisierte *verbale* Rekonstruktion eines Ablaufs realer oder fiktiver Handlungen oder Ereignisse, die im Verhältnis zum Zeitpunkt des Erzählens zurückliegen“ (Gülich & Hausendorf, 2000, S. 373, Herv. DS) verstanden. Wie zu zeigen sein wird, wird dieser Handlungs- oder Ereignisablauf in bildgestützten, *smartphone-assistierten Geschichtenerzählungen* jedoch nicht ausschließlich – und teilweise nicht einmal hauptsächlich – verbal rekonstruiert. Vielmehr wird durch das Zeigen der Fotografie(n) eine (oder mehrere) Handlung(en) oder Ereignis(se) des rekonstruierten Geschehens visuell repräsentiert.

Ziel der folgenden Analysen ist es, diejenigen Methoden der Hervorbringung bildgestützter, smartphone-assistierter Erzählungen zu rekonstruieren, die von den TeilnehmerInnen eingesetzt werden, um die Bedeutung der Fotografien im Kontext einer Geschichtenerzählung zu etablieren. Dabei gilt es aufzuzeigen, dass sich die InteraktionsteilnehmerInnen beobachtbar an diesen Fotografien als konstitutive Elemente der Rekonstruktion eines Ablaufs vergangener Handlungen oder Ereignisse orientieren (vgl. C. Goodwin, 1984, S. 243). Gelingt dieser Nachweis, fordert er die konversationsanalytische Narrationsforschung auf – auch entgegen hergebrachter Begriffsdefinitionen – bildgestützte, *smartphone-assistierte Geschichtenerzählungen* als multimodale und mediengestützte Variante des alltäglichen Erzählens in F2F-Interaktionen anzuerkennen.

Ich werde folgende forschungsleitende Fragestellungen mitführen, um die Spezifika bildgestützter, *smartphone-assistierter Geschichtenerzählungen* in alltäglichen Gesprächen rekonstruieren zu können:

- Auf welche Weise und an welchen sequenziellen Positionen wird die Integration des Smartphonegebrauchs vorbereitet?
- Wie werden die für den visuellen Zugriff der RezipientInnen erforderlichen Zeigesequenzen realisiert?
- Welche Teilnehmerahmen sind konstitutiv für die Hervorbringung bildgestützter, *smartphone-assistierter Geschichtenerzählungen*?
- Durch welche sprachlichen und verkörperten Praktiken werden die digitalen Fotografien so in die Interaktion eingeführt, dass sie erkennbar als konstitutive Bestandteile einer Geschichtenerzählung fungieren?

### **6.1.1 Die Eröffnung bildgestützter Geschichtenerzählungen durch informative Zeigesequenzen**

Ausgehend von diesen Fragestellungen möchte ich meine Analysen anhand eines Falls beginnen, in dem eine bildgestützte, *smartphone-assistierte Geschichtenerzählung* methodisch so eröffnet wird, dass eine digitale Fotografie als „showable“ (vgl. Licoppe, 2017, S. 64) – als ‚zeigbares Objekt‘ – hervorgebracht wird und anschließend als visuelle Ressource in der beobachtbaren und berichtbaren Produktion des Höhepunkts einer Erzählung fungieren kann. Der Ausschnitt stammt aus einem Gespräch zwischen den zwei Freundinnen Lea (L) und Marie (M), die sich nach einer universitären Lehrveranstaltung für eine zwanglose Unterhaltung auf einer Bank im Freien niedergelassen haben.

Die zu analysierende Geschichtenerzählung wird durch ein *oh-prefaced environmental noticing* (Drew & Chilton, 2000, S. 151) – dem beobachtbaren ‚Bemerkten‘ der Wetterverhältnisse – von L veranlasst, welches sie anschließend zur Initiierung einer thematisch verbundenen Erzählung nutzt, in die mehrere, auf ihrem Smartphone gespeicherte Fotografien integriert werden. Die Blicke von L werden mit „+“, die Gesten ihrer rechten Hand mit „±“ und die ihrer linken Hand durch „†“ annotiert. Die Blicke und Gesten von M werden durch „\*“ im Transkript kenntlich gemacht.

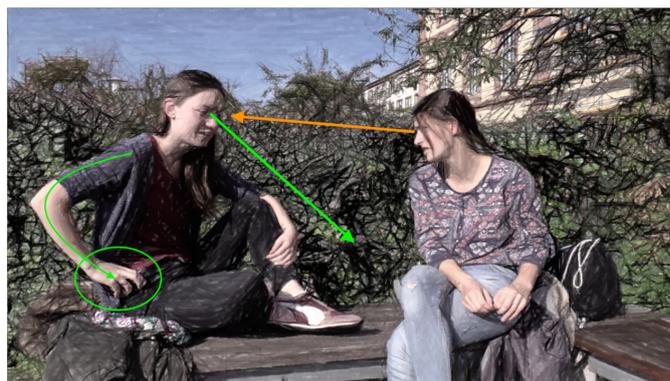
Transkript 20: 191022 „Abendrot“\_Teil 1



#

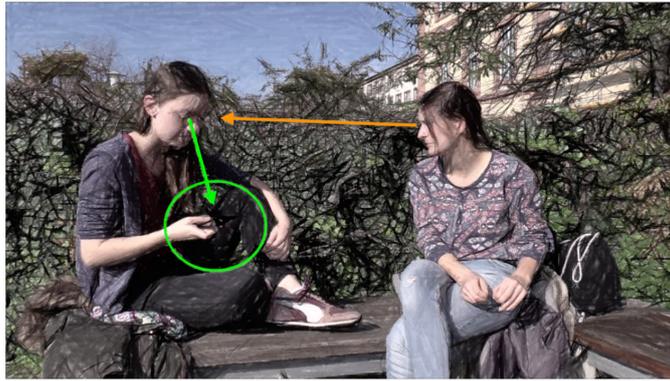
001 L: #+oh der\* Himmel ist so GEI:l heu:te;  
 + nach hinten R in Umgebung -->02  
 m: \* nach L in Umgebung -->02  
 002 M: das +is\*\_riCHTich schön.  
 ----->\*..... oben zu Himmel -->04  
 l: ,, ->+ zu M --->08  
 003 L: bo[ah hast du gestern das aBENDrot]  
 004 M: [\*setzen wir uns RAU::S? ]  
 -->\* zu L -->012  
 005 L: gesehn?  
 006 M: nee::;  
 007 ich war da [GRAD beim sport-]  
 008 L: [SOo+:: ±co]o#::l;  
 -----,,,,--->+ ... vorne R --->010  
 ± RH zu  
 Hosentasche -->010

#1



009 (0.6)  
 010 L: ±ich hab\_n ±FOto +ge#macht.  
 -----,,,,----->+ zu SP-Display-->020  
 ± nimmt SP-±....SP vor Gesicht, Daumen RH  
 bedient SP-->016

#2



011 (0.5)

Der Ausschnitt beginnt mit der Eröffnung eines Themas durch L, die in Z. 001 – während sie sich in der Umgebung umschaute – ein *oh-prefaced environmental noticing* produziert, mit dem sie anzeigt, ‚plötzlich‘ den ‚Himmel‘ zu bemerken. So thematisiert sie ein vorher nicht als relevant behandeltes Element der wahrnehmbaren Umgebung in der sprachlichen Interaktion. Mit ihrem *noticing*, das eine positive Bewertung des Himmels enthält, rekrutiert sie erfolgreich die Aufmerksamkeit ihrer Gesprächspartnerin M, die während der Äußerungsproduktion ebenfalls beginnt, sich in der Umgebung umzuschauen (vgl. Helisten, 2019, S. 339). Kurz darauf formuliert sie in Z. 002 – ihren Blick nach oben zum Himmel gerichtet – eine zweite, zustimmende Bewertung („richtig schön“). Kurz nach Beginn dieser Äußerung richtet L ihren Blick auf M und zeigt damit ihre Zuhörerschaft an.

In Z. 003–005 kommt es anschließend zu einer ausgedehnten Überlappung zwischen den zwei Sprecherinnen: Während L ihre Gesprächspartnerin fragt, ob sie „gestern das Abendrot“ gesehen habe und damit nach ihrem Wissen über ein thematisch verbundenes Ereignis fragt, produziert M in Z. 004 fast simultan ebenfalls eine, allerdings thematisch unverbundene Frage: Bezugnehmend auf das frühere Thema des gemeinsamen späteren Mensabesuchs (im Transkript nicht enthalten), fragt sie, ob sich die zwei „raus“ setzen werden. Indem L ihre Frage nach dem „Abendrot“ auch während dieser Überlappung ‚unbeirrt fortsetzt‘, erhält sie ihren Anspruch auf die Übernahme des Rederechts aufrecht (vgl. Lerner, 1989, S. 170) und beendet ihre Turnkonstruktionseinheit nach Beendigung der Überlappung. L’s Behauptung des Rederechts wird von M anerkannt, denn sie verneint in Z. 006 die Frage und formuliert in Z. 009 einen *Account* als Erklärung, warum sie das Abendrot nicht gesehen hat („ich war da GRAD beim sport“)<sup>120</sup>.

<sup>120</sup> Dass M ihr eigenes ‚Nicht-Wissen‘ als rechtfertigungsbedürftige Angelegenheit behandelt, lässt sich damit begründen, dass das „Abendrot“ – bzw. das Wetter am Ende eines vorangegangenen Tages – von ihr als Ereignis einer „commonly shared world“ (Pollner, 1974, S. 37) behandelt wird, das grundsätzlich für alle Mitglieder wahrnehmbar war und sie ebenfalls über Wissen von diesem Ereignis verfügen könnte. Dass das Wetter vergangener Tage als Thema behandelt wird, bei dem alle TeilnehmerInnen eigenes Erfahrungswissens beanspruchen können, wird auch die Analyse des zweiten Falls in diesem Kapitel zeigen.

Durch diese Sequenz in Z. 003–007 wird zwischen den Teilnehmerinnen eine Wissensasymmetrie (Drew, 1991) hinsichtlich des Wetters am gestrigen Abends etabliert. L beginnt die Produktion ihrer nächsten Äußerung (Z. 008) dabei mit deutlicher Überlappung zu M's *Account*. Indem L das Rederecht wieder übernimmt, bevor M's Äußerung einen Punkt möglicher Beendigung (vgl. Sacks et al., 1974, S. 709) erreicht hat, zeigt sie an, dass sie diese Sequenz als der Wissensabklärung dienend behandelt, die Erklärungen für diesen Wissensstand jedoch nicht als relevant erachtet.

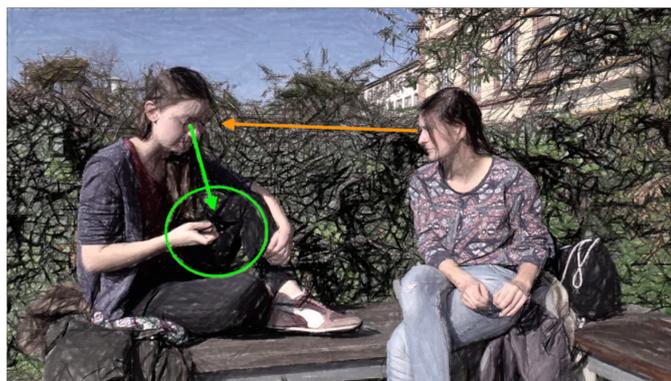
Während L mit deutlich gedehnter Aussprache („SO:: coo:l;“) eine weitere, positive Bewertung produziert, löst sie ihren Blick von ihrer Gesprächspartnerin, schaut nach rechts vorne und greift mit ihrer rechten Hand zu ihrer Hosentasche (#1). Anschließend entsteht zunächst eine kurze Gesprächspause (Z. 009), in der L versucht, etwas aus ihrer Hosentasche zu ziehen. Während sie in Z. 010 ihr Smartphone hervorholt, es vor ihr Gesicht hält und beginnt, es zu bedienen, berichtet sie, dass sie „ein Foto gemacht“ hat. Im Zusammenspiel mit ihrem verkörperten Verhalten wird mit diesem retrospektiv-prospektiv orientierten *Account* der selbst-initiierte Smartphonegebrauch – auf bekannte Art – durch eine *multimodale Gestalt* inferentiell für die Rezipientin als mit der Interaktion verbunden *accountable* gemacht. Auf diese Weise eröffnet L ein Setting *intrapersonaler Multiaktivität*, in dem die Bedienung des Smartphones in die andauernde sprachliche Interaktion eingebettet ist (vgl. Mondada, 2014a, S. 50–52).

In diesem ersten Ausschnitt wird zunächst durch das ‚Bemerken‘ des Wetters ein Themenwechsel eröffnet und anschließend die Integration einer thematisch verbundenen Fotografie von der Smartphoneinhaberin durch die Produktion einer *multimodalen Gestalt* projiziert. Auf diese Weise bereitet sie die Eröffnung einer *informativen Zeigesequenz* („informative showing“, Licoppe, 2017, S. 71) vor. Derartige Zeigesequenzen zeichnen sich dadurch aus, dass ein Wissensdefizit der potenziellen RezipientIn interaktiv etabliert wird oder von der Zeigenden zumindest angenommen werden kann. Durch ihr verkörpertes Verhalten projiziert die Zeigende dabei, dass die digitale Fotografie ein ‚zeigbares Objekt‘ darstellt, zu der die RezipientIn visuellen Zugriff erhalten kann. Die ‚Zeigewürdigkeit‘ („show-worthiness“, Licoppe, 2017, S. 71) der Fotografie wird dadurch hervorgebracht, dass sie methodisch als in Bezug zu etwas Neuem im Leben der Zeigenden stehend (vgl. ebd., 2017, S. 71) in die Interaktion eingeführt wird: Dies geschieht hier, indem zunächst das Nicht-Wissen der potenziellen Rezipientin über ein unlängst vergangenes Ereignis etabliert und anschließend die digitale Fotografie von der Smartphoneinhaberin sprachlich als neue visuelle Dokumentation dieses Ereignisses gerahmt wird.

Da die auf diese Weise im sequenziellen Verlauf der Prä-Sequenz relevant gemachte Fotografie anschließend jedoch zunächst im lokalen Speicher des mobilen Endgeräts gesucht und gefunden werden muss (vgl. Suderland, 2019, S. 91), fungiert die *multimodale Gestalt*, mit dem diese *informative Zeigesequenz* projiziert wird, als „Pre-pre“ (Schegloff, 1980)<sup>121</sup>, das die Bedienung des Smartphones als notwendige Vorarbeit für das eigentliche Zeigen rahmt.

### Transkript 21: 191022 „Abendrot“ \_Teil 2

010 L: **±ich hab\_n ±FOto +ge#macht.**  
 -----,, ,---->+ zu SP-Display-->020  
 ± nimmt SP-±....SP vor Gesicht, RH bedient SP-->016  
 #2



011 (0.5)  
 012 L: **aber des (-) also ich \*hab MEHRere gemacht;**  
 m: -----,, ,---->\*nach unten-->014  
 013 (--)  
 014 L: **aber das ist natürlich immer völllich  
 ver\*FÄ:LSCHT (-) weil dann: (.) ist ja von (-)  
 also bei kAmeras immer so mit  
 lichtverhältnissen ganz seltsam;**  
 m: -->\* zu L -->017  
 015 (0.5)  
 016 L: **aber ich fand\_s ±ME::ga: ±schö:n;**  
 ----->± LH SP ± SP quer-->

<sup>121</sup> Schegloff bezeichnet mit „preliminaries-to-preliminaries“ Handlungsprojektionen, die „serve to exempt what directly follows them from being treated as ‚produced in its own right‘. They make room for, and mark, what follows them as ‚preliminary‘“ (1980, S. 116). Ein zentraler Unterschied zu den von Schegloff untersuchten „Pre-Pres“ ist hier jedoch, dass die projizierten ‚Vorarbeiten‘ im Fall der Smartphoneintegration (zunächst) nicht in Form von sprachlicher Interaktion zwischen den GesprächsteilnehmerInnen geleistet werden. Denn auch wenn die Smartphoneintegration (oftmals) sprachlich angekündigt und interaktiv durch Reaktionen der Anwesenden ratifiziert wird, ist es anschließend die praktische Aufgabe der SmartphoneinhaberIn, die Bedienung des mobilen Endgeräts durchzuführen. Dass es im vorliegenden Datenmaterial (auch nach einer sprachlichen Ankündigung) selten zu solchen sprachlichen Bezugnahmen auf die smartphone-gestützten Suchprozesse kommt, zeigt zum einen, dass die Handhabung des Geräts von den Anwesenden als „preliminary“ verstanden wird, das nicht ‚in eigenem Recht‘ in der sprachlichen Interaktion behandelt werden muss. Zum anderen weist es darauf hin, dass die SmartphoneinhaberIn die intrapersonale Koordination (Deppermann, 2014) der für die Smartphonebedienung und die Teilnahme am Gespräch erforderlichen Ressourcen auf eine Art managt, die verhindert, dass die Integration des Smartphones zu kommunikativen Problemen führt (vgl. Keppler, 2013, S. 101).

Während L in Z. 010 berichtet, dass sie ein Foto gemacht hat, widmet sie sich, ihren Blick auf das Gerätedisplay fokussiert, nun der Bedienung ihres Smartphones. Nach einer kurzen Stille (Z. 011) beginnt sie in Z. 012 zunächst eine kontrastierende Äußerung („aber es“) bricht jedoch die Produktion ab, um – eingeleitet durch den Reformulierungsindikator „also“ (vgl. Gülich, 1994, S. 85) – ihren *Account* aus Z. 010 zu korrigieren, indem sie berichtet, dass sie „mehrere“ Fotos gemacht hat: Auf diese Weise macht sie nun die Integration mehrerer Fotografien erwartbar. Während dieser Äußerung blickt die Rezipientin zunächst zu Boden und zeigt so durch ihr reduziertes Engagement (vgl. C. Goodwin, 1981, S. 96) an, dass sie die anhaltende Bedienung des Smartphones als notwendige Vorarbeit für das projizierte Zeigen versteht, die nicht ‚in eigenem Recht‘ in der Interaktion behandelt werden muss. Während des Suchprozesses entsteht erneut eine kurze Stille (Z. 013), bevor L in Z. 014 nach einem recycelten Turn-Anfang (Schegloff, 1987b) („aber es“) einen elaborierten *Account* produziert, mit dem sie das Verhältnis zwischen der ‚Realität‘ und ihrer fotografischen Repräsentation thematisiert (vgl. Raclaw et al., 2016, S. 366). Nach einer erneuten kurzen Stille (Z. 015) wiederholt sie in einer reformulierten Variante und in gedehnter Aussprache ihre initiale, positive Bewertung des „Abendrots“<sup>122</sup>. Während dieser Äußerung greift sie nun auch mit ihrer linken Hand zum Smartphone und hält dieses anschließend quer vor sich. Durch diese *multimodale Gestalt* bestehend aus dem Zusammenspiel von sprachlicher ‚Rückkehr‘ zur Bewertung des ursprünglichen Referenten (des „Abendrots“) und der Veränderung in der Positionierung des mobilen Endgeräts zeigt sie ihrer Interaktionspartnerin an, dass der Suchprozess beendet ist und projiziert das eigentliche Zeigen der Fotografie(n).

Auffällig an der *suchbegleitenden Kommentierung* in diesem Ausschnitt ist, dass sich die Smartphoneinhaberin sprachlich explizit an einer Differenz zwischen den fotografischen Repräsentationen und ihrem Referenten in ‚der Welt da draußen‘ orientiert. Die Geräteinhaberin thematisiert hier also explizit das Verhältnis zwischen Wirklichkeit und fotografischer Repräsentation, das Gegenstand zahlreicher medienwissenschaftlicher Auseinandersetzungen mit Fotografien ist (vgl. bspw. Keppler, 2002, S. 98). Für die Analyse dieser Gesprächsepisode muss jedoch vor allem gefragt werden, welche praktische Aufgabe mit dieser einschränkenden Beschreibung der Fotografien bearbeitet wird, zu der die Rezipientin zu diesem Zeitpunkt noch keinen visuellen Zugriff hat. Zunächst lässt sich die Äußerung grundsätzlich als sprachliche Kontextualisierung in der Prä-Sequenz des eigentlichen Zeigens (vgl. Keppler, 2019, S. 181)

---

<sup>122</sup> Durch die Verwendung der Vergangenheitsform „fand“ und des deiktischen Verweises „\_S“ zeigt die Sprecherin an, dass sie sich auf ihre Wahrnehmung des Referenten in der bereits vergangenen Situation und nicht auf die Fotografie bezieht.

verstehen, die der Rezipientin prospektiv ein angemessenes Verständnis der visuellen Medieninhalte ermöglichen soll. Im spezifischen Kontext dieses Falls lässt sich außerdem argumentieren, dass die Smartphoneinhaberin durch ihre erste Bewertung im Anschluss zur projizierten Zeigesequenz eine zweite Bewertung durch die Rezipientin relevant macht, sobald diese über das notwendige Wissen verfügt, das „Abendrot“ bewerten zu können (vgl. Pomerantz, 1984, S. 58 f.). Indem sie die Fotografien als mediale Repräsentationen beschreibt, die die Realität lediglich verzerrt abbilden, produziert sie einen prospektiv orientierten *Account*, der eine (potenzielle) Erklärung für den Fall liefern kann, dass die Bewertung der Rezipientin nicht zustimmend ausfällt. Die Äußerungen erscheinen so als implizite Orientierung an der Präferenzorganisation von Bewertungen, in denen zustimmende zweite Bewertungen strukturell als präferierte Erwidierungen behandelt werden (Pomerantz, 1984) und die damit einen zentralen Ort für die Hervorbringung geteilter Verständnisse und (Geschmacks-)Urteile darstellen (vgl. M. H. Goodwin & Goodwin, 1992, S. 173). Mithilfe der sprachlichen ‚Vorarbeiten‘ der Smartphoneinhaberin ließe sich ein eventuell auftretender Widerspruch durch das ‚problematische‘ Verhältnis der Fotografie zur vorgängigen Realität erklären, sodass letztlich nicht aufzuklären wäre, zu welcher Bewertung des „Abendrots“ die Rezipientin käme<sup>123</sup>.

Nachdem der Suchprozess abgeschlossen und die Rezipientin durch die sprachlichen Betrachtungsanweisungen der Smartphoneinhaberin angehalten ist, die Fotografien als „verfälschte“ Repräsentationen eines schönen „Abendrots“ zu sehen, erfolgt nun das eigentliche Zeigen der Fotografien und ihre kommunikative Aneignung:

### Transkript 22: 191022\_„Abendrot“\_Teil 3

016 L: aber ich fand\_s ±ME::ga: ±schö:n;  
 ----->± LH SP + SP quer-->

017 (0.3)±\*(0.4)# ±#(1.1)

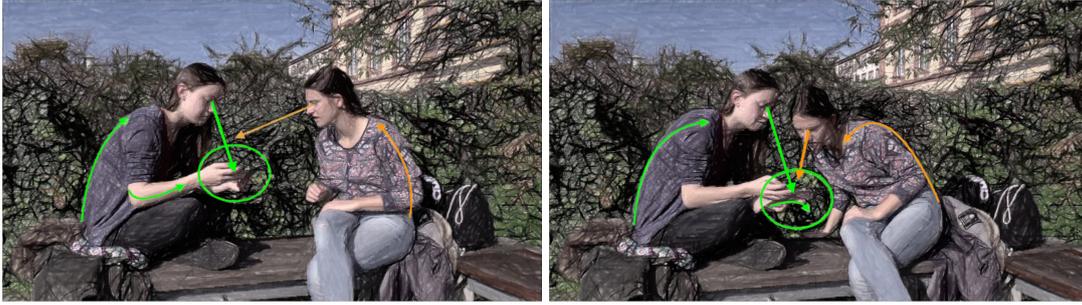
l: ---->±SP Richtung M± dreht Display zu M-->

m: ,,,-->\*.....-----zu SP-Display, beugt sich  
 vor--->021

#3            #4

---

<sup>123</sup> Wie sich im weiteren Verlauf der Interaktion zeigen wird, lässt sich diese Interpretation der ‚interaktiven Arbeit‘ dieses *Accounts* jedoch nicht mittels der *next-turn proof procedure* in den Daten verankern, da es nicht zur Produktion einer disaffiliativen Bewertung kommt. Evtl. zeigt sich daran, dass für die Rezipientin des Zeigens durch die sprachlichen Vorarbeiten eine ‚transparente‘ Sicht (vgl. C. Goodwin, 1996) durch die ‚verzerrte Fotografie‘ hindurch auf das visuell repräsentierte „Abendrot“ ermöglicht wird.



018 L: <<p>das GUCK;>± ±  
 ±swipe±  
 019 M: °hhh <<h>oh[h::\_WO::W-;> ]  
 020 L: [das war so VO::LL] vie±::l ±  
 -----±swipe±-->  
 021 und man hats halt perFEKT als ich ±gerAde  
 über #die brücke wollte \*ge[sehen+;]#  
 -----±swipe±-  
 ±----->±....SP in  
 rechte Hosentasche-->>  
 -----,,,,---->+zu M-->024  
 022 M: \* [oah wie ]  
 SCHÖ:N;  
 m: -----,,-->\*... Blick L, setzt  
 sich gerade hin-->024  
 #5 #6



023 L: [mega\* COO:L; + ]  
 024 M: [nee:\* ich glaub+ da] saß ich GRAD beim  
 spor?=  
 ---->\* zu L -->>  
 l: ----->+ ..... hinten R ,,,,---->025  
 025 =N+EE: da warn wir gerade bei laura (-) mit den  
 hAArn;  
 l: ->+ zu M -->>  
 026 L: ACHja;  
 027 M: ja;

Nachdem L simultan zu ihrer sprachlichen Äußerung in Z. 016 beginnt, das Smartphone mit beiden Händen quer zu halten, entsteht zunächst eine Stille von 1,8 Sekunden (Z. 017), in der die Displayfokussierung von den TeilnehmerInnen über verkörperte Handlungen organisiert

wird: L hält das Smartphone zunächst etwas weiter in M's Richtung, worauf sich diese nach vorne beugt und auf das Display des mobilen Endgeräts schaut (#3). Durch diese Veränderungen in der körperlichen Ausrichtung wird ein Teilnehmerahmen hervorgebracht, in dem sich das Smartphonedisplay im Blickfeld beider TeilnehmerInnen befindet. Anschließend produziert L in Z. 019 mit gesenkter Stimme einen deiktischen Referenten und das Imperativ eines visuellen Perzeptionsverbs („das GUCK;“), das die Aufmerksamkeit der Rezipientin auf die Fotografie lenkt (vgl. Oloff, 2019a, S. 214). Das Timing dieser Äußerung zielt dabei auf den Moment, an dem die Rezipientin das *showable* auf die in der Prä-Sequenz vorbereitete Weise sehen kann (vgl. Licoppe & Tuncer, 2019, S. 565). Anschließend führt die Smartphoneinhaberin eine ‚wischende‘ Bediengeste durch und scheint damit eine weitere Fotografie aufzurufen. Die Rezipientin reagiert darauf in Z. 019 zunächst mit hörbarem Einatmen, bevor sie in hoher Tonlage ein anerkennendes „oh WO::W“ produziert. Zum einen zeigt die Rezipientin damit sprachlich ihre Alignierung mit der *informativen Zeigesequenz* an. Zum anderen macht das „oh“ in turn-initialer Position als Erkenntnisprozessmarker eine Veränderung ihres Wissenstands kenntlich, bevor sie eine vage, aber eindeutig positive Evaluation produziert und so ihre Affiliation mit der Haltung der Zeigenden anzeigt (vgl. Oloff, 2019a, S. 214).

Bereits in Überlappung zu dieser für smartphone-assistierte Zeigesequenzen typischen, vagen ersten Evaluation setzt L die Kommentierung der Fotografien in Z. 020 fort: Zunächst produziert sie mit „das war so vo::ll viel“ eine deiktische Beschreibung, die aus der Beobachterperspektive als sprachliche Äußerung für sich unvollständig wirkt, aber für die Teilnehmerinnen – wie der weitere Verlauf der Interaktion zeigt – im Zusammenspiel mit der simultanen Betrachtung der Fotografischen pragmatisch verständlich ist (vgl. Aaltonen et al., 2014, S. 139). Am Ende dieser multimodalen semiotischen Einheit führt sie erneut eine wischende Bediengeste durch und produziert anschließend in Z. 021 eine fragmentierte narrative Rekonstruktion: „gerade“ als sie über eine Brücke gehen wollte, konnte sie das auf den Fotografien abgebildete Abendrot „perFEKT“ sehen, wobei sie mit der Realisierung von „gerade“ erneut eine wischende Bediengeste durchführt. Im Kontext dieser *informativen Zeigesequenz* wird schon durch diese sprachliche Äußerung eine multimodale Geschichtenerzählung hervorgebracht, in der chronologisch die Abfolge von zwei Ereignissen auf eine Weise rekonstruiert wird, die nachvollziehbar macht, ‚wie‘ es dazu kam, dass die Erzählerin das fotografisch repräsentierte Abendrot aus dieser Perspektive wahrnehmen und dokumentieren konnte. Schon vor der Beendigung dieser Äußerung führt sie das Smartphone bereits wieder in Richtung in Hosentasche und schaut vom Display zur Rezipientin (#6). Durch die Art und Weise der Hervorbringung dieser *multimodalen Gestalt* zeigt sie an, dass sie bereits das Ende der bildgestützten

Geschichtenerzählung erreicht hat. Dass dies für M *accountable* ist, zeigt sich daran, dass diese sich wieder gerade hinsetzt, ihren Blick vom Display löst und beginnt, in die Umgebung zu schauen (#6). Die Zeigesequenz wird hier also durch die verkörperte Auflösung des Teilnehmerahmens beobachtbar beendet. Gleichzeitig zeigt die Rezipientin durch ihre in minimaler Überlappung produzierte zweite Bewertung („so SCHÖ:N“) an, dass sie nun über das erforderliche Wissen zur Bewertung des Referenten verfügt und eine mit der Zeigenden übereinstimmende Haltung zum visuell repräsentierten „Abendrot“ einnimmt.

In der anschließenden *response sequence* produzieren beide Teilnehmerinnen mit einem Simultanstart voneinander unabhängige Versuche der Sequenzschließung: Während L in Z. 022 erneut eine Bewertung produziert, die gemeinsam mit der kurz darauf stattfindenden Abwendung ihres Blicks als „closing device“ (M. H. Goodwin & Goodwin, 1992, S. 171f.) erkennbar wird, kehrt die Rezipientin in Z. 023 zu ihrer Erklärung zurück, warum sie das Abendrot nicht gesehen hat. Dabei blickt sie wieder zu L, die ihren Blick jedoch zunächst nicht erwidert. Erst als M die Äußerungsproduktion abbricht und in Z. 024 eine Reformulierung als selbst-initiierte Selbstreparatur (vgl. Schegloff et al., 1977, S. 366) realisiert, schaut L zur Sprecherin. Anschließend wird in Z. 025–026 eine Schließungssequenz von den TeilnehmerInnen hervorgebracht und die Post-Sequenz der bildgestützten Geschichtenerzählung kollaborativ beendet.

Die Analyse dieses Falls zeigt, wie in der Prä-Sequenz eine Wissensasymmetrie etabliert und anschließend durch die sprachliche Erwähnung digitaler Fotografien eine smartphone-assistierte, *informative Zeigesequenz* vorbereitet wird. Nach der notwendigen Vorarbeit durch den Suchprozess, während dem die Smartphoneinhaberin *suchbegleitende Kommentierungen* produziert, wird das Zeigen der Fotografie durch eine *multimodale Gestalt* eröffnet. Während der anschließenden Phase des eigentlichen Zeigens wird eine fragmentierte, bildgestützte Geschichtenerzählung hervorgebracht, in der sich das Display des Smartphones im gemeinsamen Blickfeld beider Teilnehmerinnen befindet, sodass sich sprachliche Äußerungen und das semiotisch Feld der Fotografie(n) wechselseitig ihren Sinn verleihen (vgl. C. Goodwin, 2000, S. 166). Der sequenzielle Verlauf der Interaktion zeigt, wie die von der Smartphoneinhaberin initiierte *mobile-supported sharing activity* der Rezipientin einen epistemischen Zugang zur visuellen Repräsentation des vergangenen „Abendrots“ und damit die Produktion einer eigenen Bewertung dieses Referenten ermöglicht (vgl. Raclaw et al., 2016, S. 377). Während M zunächst eine vage, aber eindeutig affiliative Erwiderng produziert, formuliert sie nach eingehenderer Betrachtung der Fotografie(n) eine explizite zweite Bewertung, die ihr Verständnis der integrierten Fotografien als visuelle Repräsentationen eines ‚schönen Abendrots‘ eindeutig anzeigt (vgl. Oloff, 2019a, S. 214).

Als grundlegend multimodale Aktivität wird diese bildgestützte, *smartphone-assistierte Geschichtenerzählung* dabei nur durch das reflexive Verhältnis zwischen sprachlichen Äußerungen und digitalen Fotografien sinnhaft verständlich: Die sprachliche Erzählung kontextualisiert die Fotografien (Keppler, 2019, S. 180), indem sie deren Entstehungsgeschichte narrativ rekonstruiert – gleichzeitig dienen die Fotografien als visuell-materielle Repräsentationen der optischen Eindrücke der Erzählfigur (das ‚vergangene Ich‘ der Erzählerin) und damit als Höhepunkt der Erzählung. Die Sequenzorganisation dieser als *informativer Zeigesequenz* hervorgebrachten, bildgestützten Geschichtenerzählung entspricht zunächst grundsätzlich dem typischen Sequenzmuster von *storytellings* aus voneinander abgrenzbarer Prä-Sequenz, Hauptteil und Post-Sequenz (vgl. Licoppe & Tuncer, 2019, S. 566). Betrachtet man jedoch die sequenzielle Organisation des Hauptteils dieser Erzählung, so zeigt sich eine zentrale Differenz zu (überwiegend) sprachlich verfassten Geschichtenerzählungen: Aufgrund der linearen Sequenzialität sprachlicher Interaktion (vgl. Mondada, 2016a, S. 341) können die rekonstruierten Ereignisse immer nur nacheinander in das Gespräch eingeführt und in einer Erzählung miteinander verknüpft werden<sup>124</sup>. Für bildgestützte Geschichten zeigt sich hier jedoch, dass sich die Möglichkeiten der Rekonstruktion vergangener Ereignisse durch den simultanen Einsatz materiell-visueller Repräsentationen grundlegend ändern. Zum einen erfolgt die Hervorbringung des *big packages* der Erzählung nicht ausschließlich über sprachliche Äußerungen, sondern über multimodale, semiotische Einheiten (Aaltonen et al., 2014, S. 139)<sup>125</sup>, in denen sich sprachliche Äußerungen und Bildinhalte wechselseitig ihren Sinn verleihen. Zum anderen verändert sich die Zeitlichkeit der narrativen Rekonstruktion, denn die Rezipientin der Zeigesequenz hat bereits vor Beginn der sprachlichen Rekonstruktion visuellen Zugang zu mindestens einer Fotografie (Z. 018) und erhält im Verlauf der Interaktion Zugriff zu zwei weiteren (Z. 020, 021). Wenn die integrierten Fotografien von den Teilnehmerinnen als Momentaufnahmen (vgl. Larsen, 2005, S. 418) der optischen Eindrücke (des ‚Abendrots‘) der Erzählerin verstanden werden, heißt dies für die Abfolge der rekonstruierten Ereignisse folgendes: Die Erzählerin zeigt

---

<sup>124</sup> In den meisten Erzählungen erfolgt dies durch eine Darstellung der Ereignisse in ihrer chronologischen Abfolge. Insbesondere im Bereich der *small stories*-Forschung konnte jedoch gezeigt werden, dass die Abfolge der Ereignisse abhängig vom situativen Kontext und den praktischen Zielen der ErzählerIn neu geordnet werden kann (bspw. Georgakopoulou, 2007, Kap. 2, 2008, S. 601).

<sup>125</sup> Diese Argumentation soll nicht implizieren, dass in nicht-mediatisierten Geschichtenerzählung die narrative Rekonstruktion vergangener Ereignisse ausschließlich sprachlich erfolgt. Erzählungen in F2F-Interaktionen sind grundsätzlich als audio-visuelle Inszenierung zu verstehen, in der sich die TeilnehmerInnen aller ihnen zur Verfügung stehenden, multimodalen Ressourcen bedienen, um bspw. verkörperte Eleierungen hervorzubringen, mit denen sie vergangene Handlungen animieren (Sidnell, 2006; König & Oloff, 2018a, S. 302). Allerdings handelt es sich bei diesen gestisch-mimischen Verfahren der Hervorbringung narrativer Rekonstruktionen im Gegensatz zu visuellen Medieninhalten um flüchtige Phänomene und nicht um materiell-visuelle Objekte, die beliebig lange und wiederholt betrachtet werden können. Ich werde auf diesen Umstand zurückkommen.

und die Rezipientin betrachtet vor und während der Realisierung der sprachlichen Äußerungen, die die Erzählung sprachlich konstituieren, drei Momentaufnahmen eines Ereignisses, das erst im Vollzug der Interaktion als Höhepunkt der sequenziell verknüpften Handlungsfolge hervorgebracht wird. Da die Wahrnehmung der Fotografien visuell und der sprachlichen Äußerungen akustisch – also über unterschiedliche Kanäle – erfolgt und diese sich wechselseitig informieren, ist die Rekonstruktion vergangener Ereignisse in dieser multimodalen, bildgestützten Variante nicht mehr auf die gleiche Weise an das lineare Nacheinander einer sprachlichen Wiedergabe gebunden. Im Interaktionsverlauf wird hier durch die sprachlichen Äußerungen eine bestimmte Abfolge der Geschehnisse rekonstruiert, die – in diesem Fall – nicht mit dem Zeitpunkt der Integration der digitalen Fotografien als visuelle Repräsentationen eines Ereignisses identisch ist.

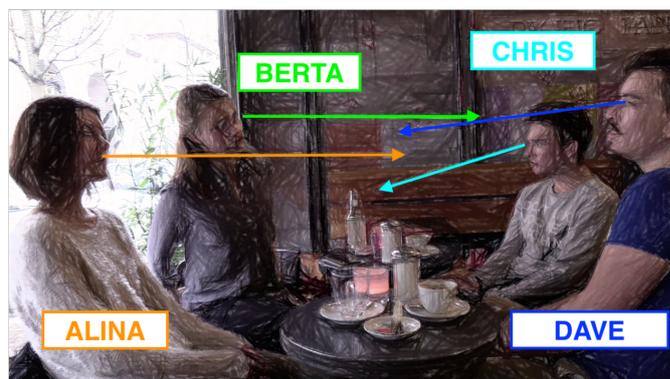
### **6.1.2 Die Eröffnung bildgestützter Geschichtenerzählungen durch evokative Zeigesequenzen**

Bereits durch die Analyse dieses ersten exemplarischen Falls konnten bereits einige der spezifischen Aufgaben der Interaktionsorganisation identifiziert werden, die typischerweise in der Hervorbringung bildgestützter, *smartphone-assistierter Geschichtenerzählungen* bearbeitet werden müssen: Neben den typischen Schritten der sequenziellen Organisation von Geschichtenerzählungen umfassen sie die Hervorbringung einer oder mehrerer Fotografien als zeigbare Objekte, die Koordination des als Vorarbeit erforderlichen Suchprozesses mit der F2F-Interaktion, sowie die rezipientenspezifische Kontextualisierung der Fotografien als konstitutives Element einer multimodalen, narrativen Rekonstruktion. Im Gegensatz zu mündlich-sprachlichen Erzählungen lässt sich die Abfolge der rekonstruierten Ereignisse dabei für bildgestützte Geschichtenerzählungen nicht mehr hinreichend als lineares Nacheinander beschreiben. Dies liegt daran, dass die simultane Betrachtung der Fotografien während der Produktion der sprachlichen Beiträge eine Gleichzeitigkeit der Rezeption der konstitutiven Elemente der Erzählung ermöglicht, die sowohl akustisch als auch visuell erfolgt.

Das zweite Exemplar einer bildgestützten, *smartphone-assistierten Erzählung* weist zunächst eine thematische Gemeinsamkeit zum ersten Fall auf: Auch in diesem Gesprächsausschnitt einer längeren Unterhaltung zwischen den vier Freunden Alina (A), Berta (B), Chris (C) und Dave (D) wird das Erleben eines ‚Naturereignisses‘ – hier des Schneefalls am Vortag – mithilfe der Integration einer digitalen Fotografie narrativ rekonstruiert. Der folgende Fall unterscheidet sich jedoch insofern, als dass es sich um ein Mehrpersonengespräch handelt, in dem alle Anwesenden Wissen über dieses Ereignis besitzen, und verschiedene TeilnehmerInnen

Geschichten ihres Erlebens des nächtlichen Schneefalls produzieren. Dies ermöglicht es der Smartphoneinhaberin, die Fotografien über eine ‚evokative Zeigesequenz‘ („evocative showing“, Licoppe, 2017, S. 72) in die Interaktion einzuführen: Derartige Zeigesequenzen werden auf eine Weise initiiert und realisiert, die (beobachtbar) voraussetzt, dass die RezipientInnen über hinreichendes Wissen verfügen, das zeigbare Objekt auf angemessene Art sehen zu können (ebd., 2017, S. 78). Darüber hinaus wird hier eine digitale Fotografie in die Geschichtenerzählung integriert, die nicht die optischen Eindrücke der Erzählerin, sondern eine Momentaufnahme ihrer verkörperten Handlungen in einer spezifischen, materiellen Umgebung repräsentiert (vgl. Larsen, 2005, S. 425). Wie zu zeigen sein wird, ermöglicht die Integration derartiger Fotografien die Hervorbringung von Geschichtenerzählungen, deren *Accountability* noch stärker durch die Visualität der Bildinhalte gestützt wird. Im folgenden Transkript werden die Blicke von B mit „+“, die Gesten ihrer rechten Hand mit „±“ und die ihrer linken Hand mit „†“ annotiert. Die Blicke und Gesten von A werden durch „\*“, die von C durch „•“ und von D durch „\$“ kenntlich gemacht.

### Transkript 23: 181217\_„Es hat geschneit“\_Teil 1



b: +>>-- zu D --->003  
001 D: das kann schon ganz SCHÖN sein so; (.)  
002 ich mag den GERUCH von regen [zum beispiel;  
] ]  
003 A: [gestern hat\_s  
<<:-)>ge:+SCHN]EI:T;>  
b: --, , , , -->+ . . . . --->006  
004 (0.58)  
005 D: wa[s?]  
006 B: [+j]a;  
, , -->+ schräg rechts--->009  
007 (0.11)  
008 A: <<:-)>GEstern hat\_s geschne[it; >]  
009 D: [ja ] AL+ter;  
b: ----- , , , , -->+ zu D-->012  
010 (.)

011 B: auf m[einer TErrasse lag schnee; ]  
 012 D: [ich bin ges? (.) ich bin irg]endwie am (-)  
 samstag ± (.)†NACHT#(.)±+um eins oder zwei±±#  
 nach hause gekommen;  
 b: -----,, , , , --->+ SP-Display -->019  
 b: ±..RH zeigt auf± ±bedient  
 SP-->032  
 b: †...LH nimmt SP †bedient  
 SP-->032



013 da war einfach (-) NICHTS drau[ßen ]  
 014 C: [ich B]IN [(.) ja  
 ]  
 015 D: [un  
 mor]gens aufgewacht (.) KOMplett weiß [irgendwie.]

D's Äußerungen in den Z. 001–002 sind noch Teil der *response sequence* zu B's Erzählung, in denen er eine positive Bewertung abgibt und erklärt, dass er den „Geruch“ von Regen mag. Dies veranlasst A nun in Z. 003 zu einer in Überlappung produzierten Erwähnung eines anderen ‚Wetterereignisses‘: Mit *smile voice* berichtet sie, dass es „gestern“ geschneit hat und macht so ein neues Themenangebot (Schegloff, 2007a, S. 169). Nach einer kurzen Stille (Z. 004) produzieren D und B fast zeitgleich eine Erwiderung: D initiiert mit „was?“ (Z. 005) eine Reparatur und zeigt so ein unspezifisches Problem des Verstehens oder Hörens an, während B mit „ja“ (Z. 006) ihre Zustimmung anzeigt. Indem A in Z. 008 ihre Äußerung wiederholt, realisiert sie einen Reparaturversuch und priorisiert zunächst die Wiederherstellung eines intersubjektiven Verständnisses. Die erfolgreiche Reparatur der Problemquelle wird von D in Z. 009 bestätigt, und die ‚Erwähnungswürdigkeit‘ („mentionability“, Schegloff & Sacks, 1973, S. 301) des ‚Schneefalls‘ durch sein „alter“ in turn-finaler Position verstärkend ratifiziert.

Nachdem das Verständnisproblem erfolgreich bearbeitet wurde, ergreift B in Z. 011 das Wort und berichtet, dass auf ihrer „Terrasse“ Schnee lag. Kurz nach Beginn dieser Äußerung beginnt jedoch auch D in Überlappung zu sprechen (Z. 012) und initiiert – nach einer Selbstreparatur im gleichen Turn (vgl. Schegloff et al., 1977, S. 366) – den Beginn einer Erzählung. Obwohl D sich selbst das Rederecht nimmt, bevor B's Äußerung eine übergaberelevante Stelle erreicht,

und er die Turnproduktion unbeirrt fortsetzt, wird dies von der unterbrochenen Sprecherin B akzeptiert: Anstatt um das Rederecht ‚zu kämpfen‘, hebt sie nach der Überlappung und noch während D spricht ihre rechte Hand in einer Meldegeste<sup>126</sup>, greift mit der linken Hand zu ihrem Smartphone auf dem Tisch (#1) und beginnt dann, den Blick auf das Display gerichtet, das Gerät mit beiden Händen zu bedienen (#2). Die SmartphoneinhaberIn übernimmt hier also die Rolle einer ZuhörerIn und nutzt die mit dieser diskursiven Identität in Mehrpersonengesprächen verbundenen Handlungsmöglichkeiten (vgl. DiDomenico et al., 2018), um den Gebrauch ihres mobilen Endgeräts zu initiieren. Das Timing der verkörperten Initiierung des Smartphonegebrauchs weist diesen als etwas aus, das durch die sprachliche Interaktion veranlasst wird. Der weitere Gebrauch des Smartphones wird von den InteraktionsteilnehmerInnen jedoch als divergente Aktivität gerahmt (vgl. Oloff, 2019b Z. 32–33): B zeigt durch ihr verkörpertes Verhalten an, dass sie die Alignierung mit der sprachlichen Interaktion vorübergehend suspendiert und sich der Gerätebedienung als Hauptaktivität zuwendet. Dies wird von den anderen Anwesenden weder durch Kommentare, noch visuelles *Monitoring* als ‚bemerkenswert‘ behandelt, vielmehr fährt D in Z. 013 mit der narrativen Rekonstruktion seines Erlebens des nächtlichen Schneefalls fort. Dies führt im weiteren Gesprächsverlauf zur Hervorbringung einer Serie von Geschichtenerzählungen, in der anschließend auch C sein individuelles Erleben des gestrigen Schneefalls rekonstruiert (nicht im Transkript enthalten). In beiden Geschichten betonen die Erzähler, dass sie am Wochenende bis spät abends unterwegs waren und den Schneefall am nächsten Morgen als plötzliches und unerwartetes Ereignis erlebt haben, wobei sich A in beiden Fällen als amüsierte ZuhörerIn positioniert.

Währenddessen bleibt B durchgehend mit der Bedienung ihres Smartphones beschäftigt, bis sie diese während der *response sequence* zu C’s Geschichte beendet und im nun folgenden Ausschnitt eine *evokative Zeigesequenz* initiiert:

---

<sup>126</sup> Aus BeobachterInnenperspektive lässt sich hier argumentieren, dass die Meldegeste als verkörperte Variante einer Ankündigung dient, mit der die SmartphoneinhaberIn gleichzeitig anzeigt, dass sie sich nun der Gerätebedienung zuwendet, ihr Engagement an der F2F-Interaktion reduziert und das Rederecht prospektiv für einen Zeitpunkt nach Beendigung des Suchprozesses ‚reserviert‘ (vgl. Keppler, 2019, S. 179). Da sich jedoch keine beobachtbare Orientierung anderer Anwesender an dieser *multimodalen Gestalt* feststellen lässt, kann diese Interpretationsmöglichkeit nicht in den Daten verankert werden. Die verkörperte Ankündigung durch eine Meldegeste, die unkommentierte Zuwendung zum Smartphonegebrauch sowie der Umgang anderer Anwesenden mit diesem entspricht damit dem Interaktionsverlauf des Falls „Gnumma?“ (vgl. Kap. 5.1.2).

Transkript 24: 181217\_„Es hat geschneit“\_Teil 2

032 D: **abter um †\*sechs# nach house kommen und um elf aufstehen das ist au+ch ne LE[ISTung •joa;]**  
 b: -----  
 -----,,-->+ Blick A --->033  
 b: ----->+  
 b: ->†.....---- dreht SP zu A -->036  
 a: \*.....Blick SP-Display----->037  
 c: •... Blick zu  
 B & A -->

#3



033 B: **bei [mir] aus;** [so \*sah\_•s +GES]#tern  
 -----,,,,,,,,,---->+Blick zu  
 SP-Display-->035  
 a: \*,,,beugt sich vor  
 ----->036  
 c: •... Blick zu  
 B & A -->Z.036

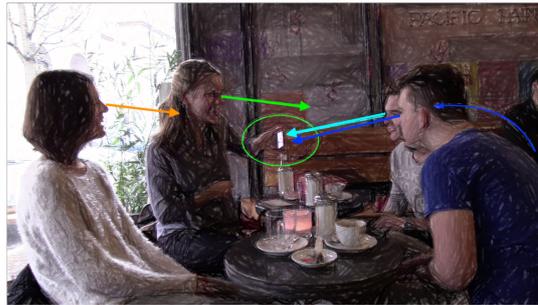
#4

034 C: **[ja-]**



035 **(0.15)+(0.25)**  
 b: --,-->+ Blick zu C--->036  
 036 B: **hh hi hi+† •\*\$hi#**  
 b: ----->+ zu D -->042  
 b: ----->† dreht SP zu C&D-->  
 a: ---,,----->\* lehnt sich zurück, lacht-->>  
 c: ----->• Blick zu Display --> 040  
 d: \$ Blick zu SP-Display, beugt sich vor-->044

#5



037 (0.2) \* (0.41)  
a: ,,-->\* Blick zu C & D-->045  
038 C: hhe  
039 B: hi hi ha  
040 (1.6) •±(0.25)  
b: ,---->† dreht SP zu sich -->>  
c: ---->• Blick nach vorn-->044  
041 C: al[TER;]  
042 D: [\$läu+]ft;  
-->\$ Blick geradeaus, sitzt gerade-->>  
b: ----->+ Blick SP-Display -->044  
043 D: he he  
044 B: es •+hat GESCHNE+IT;†  
--->+, zu D-----+ zu SP-Display -->>  
† tippt 3x Display-->046  
c: -->• Blick B--->045  
045 A: he •he †he \*h°  
----->\* Blick rechts-->>  
c: -->• Blick unten-->>  
b: ----->†  
046 (0.81)

Der Ausschnitt beginnt in Z. 032 mit D's Äußerung, die noch zur *reponse sequence* auf C's Geschichte gehört und mit der er das Aufstehen nach wenigen Stunden Schlaf als „Leistung“ anerkennt. Während der Realisierung dieser Äußerung beendet B die Bedienung des Smartphones und dreht das Gerätedisplay nun zu ihrer Sitznachbarin A (#3). So gelingt es ihr, die Aufmerksamkeit von A zu rekrutieren und die Displayfokussierung zu veranlassen, denn A richtet ihren Blick nun auf das Smartphone. Nach einem kurzen Moment richtet B ihren Blick auf die Rezipientin und scheint sich durch dieses visuelle *Monitoring* zu versichern, dass A das Display fokussiert. Anschließend produziert B in Z. 033 in teilweiser Überlappung zu D's Äußerung mit „so sah\_s GESTern bei mir aus“ eine Betrachtungsanweisung, die den visuellen Medieninhalt als fotografische Repräsentation des gestrigen ‚Schneefalls bei ihr zuhause‘ ausweist und während der sie wieder auf das Gerätedisplay schaut. Simultan beugt sich A nach vorne (#4)

und zeigt so ihre Alignierung mit der smartphone-gestützten Zeigesequenz an. B's kontextualisierende Äußerung wird dabei erst realisiert, als sich die Zeigende versichert hat, dass die Rezipientin in der Lage ist, die digitale Fotografie auf ‚angemessene‘ Art (als visuelle Repräsentation des „Schneefalls“) zu sehen (vgl. Licoppe & Tuncer, 2019, S. 565).

Anhand der Eröffnung dieser Zeigesequenz zeigen sich zentrale Differenzen zum vorangegangenen Fall: Vor und während der verkörperten Eröffnung des Zeigens wird die auf dem Gerätedisplay angezeigte Fotografie sprachlich nicht erwähnt: „However, the showing sequence is initiated in a fluid way, without any explicit cues and as a kind of incipient action” (Licoppe & Tuncer, 2019, S. 558). Die Abwesenheit einer (sprachlichen) Prä-Sequenz ist typisch für *evokative Zeigesequenzen*. Durch diese Art der Eröffnung wird die Rezipientin zum einen als kompetent angesehen, die Eröffnung des Zeigens durch den visuellen Zugriff auf das Smartphonedisplay zu erkennen. Zum anderen wird ihr zugeschrieben, die Bedeutung der eingebundenen Fotografie in diesem Moment der Interaktion selbstständig erkennen und sie aus einer angemessenen Perspektive zu betrachten (vgl. Licoppe, 2017, S. 80)<sup>127</sup>. Durch die ausschließlich verkörperte Eröffnung der Zeigesequenz signalisiert die Smartphoneinhaberin, dass sie annimmt, die Fotografie aufgrund ihrer thematischen Verbundenheit auf eine derart reduzierte Weise als zeigbares und zeigwürdiges Objekt in die Interaktion einführen zu können<sup>128</sup>.

Mit der erfolgreichen Initiierung dieser *evokativen Zeigesequenz* geht die Hervorbringung eines neuen Teilnehmerrahmens einher, in dem A und B nun beide das Smartphonedisplay fokussieren. Dieses körperliche Arrangement projiziert bereits non-verbal eine Schismabildung (vgl. Egbert, 1997) im Mehrpersonengespräch, die durch B's Äußerung in Z. 033 nun auch hörbar wird: Während C und D in die *response sequence* einer Geschichtenerzählung involviert sind

---

<sup>127</sup> Wie Licoppe (2017) argumentiert, setzen derartige *evokative Zeigesequenzen* durch ihren (fast) vollständigen Verzicht auf eine Prä-Sequenz voraus, dass 1.) Zeigende und RezipientInnen über geteiltes Wissen verfügen, 2.) das auf dieses Wissen durch das ‚gezeigte Objekt‘ auf erkennbare Weise verwiesen wird, sowie 3.) dass durch das Design der Zeigesequenz für die RezipientInnen erkennbare Erwartungen an eine ‚angemessene‘ Erwiderung und Kriterien der Beurteilung des ‚gezeigten Objekts‘ projiziert werden. Derartig gestaltet, macht eine *evokative Zeigesequenz* eine Erwiderung konditionell relevant, die anzeigt, dass „the recipient has got the gist of the implicit reference, and just doing this may be enough“ (Licoppe, 2017, S. 80).

<sup>128</sup> In ihren Analysen von Zeigesequenzen in Videotelefonaten betonen Licoppe und Tuncer, dass „part of what provides for the possibility of a showing sequence to flow swiftly and efficiently from this kind of topical talk, is that the relevant item is recognizable as ready-at-hand with respect to the action of showing it“ (Licoppe & Tuncer, 2019, S. 558–559). Im Unterschied zu Objekten in der unmittelbaren Umgebung der TeilnehmerInnen sind digitale Fotografien gerade nicht unmittelbar zur Hand, sondern müssen zunächst im lokalen Speicher des Smartphones oder über eine Online-Suche gesucht und gefunden werden, bevor das eigentliche Zeigen erfolgen kann (vgl. Keppler, 2019, S. 180; Suderland, 2019, S. 91). Dass die Eröffnung dieser smartphone-gestützten Zeigesequenz trotzdem (erfolgreich) auf diese Weise erfolgen kann, mag zum einen an den gelockerten Teilnahmeverpflichtungen von Mehrpersonengesprächen liegen, die es der RezipientIn ermöglichen, sich einer anderen Aktivität zuzuwenden, während gleichzeitig andere Anwesende die aktuell im Vollzug befindliche Gesprächsaktivität (der *response sequence*) aufrechterhalten können. Die Zeigbarkeit der Fotografie wird darüber hinaus dadurch verstärkt, dass sie in Bezug zu einem Thema der Interaktion steht, das von der Rezipientin der Zeigesequenz eröffnet wurde (Z. 003).

(#3), fokussieren sich A und B auf die smartphone-gestützte Integration einer digitalen Fotografie. Die sprachliche Betrachtungsanweisung von B in Z. 033 veranlasst jedoch auch C dazu, zu den zwei auf das Display fokussierten Freundinnen herüberzuschauen (#4), während er ein bestätigendes Rezeptionssignal als Erwiderung auf D's Äußerung produziert (Z. 034). In der anschließenden Stille (Z. 035) betrachtet er aufmerksam das Display, während die Smartphoneinhaberin vom Gerät auf und zu C schaut, sodass die beiden sich nun gegenseitig anschauen. Anschließend beginnt B zu Lachen und zeigt so ihre Haltung zu den integrierten Medieninhalten an. Währenddessen blickt sie zu D und beginnt das Smartphone in Richtung von C und D zu drehen, die unmittelbar die Displayfokussierung initiieren, wobei sich der am weitesten entfernt sitzende D nach vorne über den Tisch beugt (#5).

Auf diese Weise eröffnet B nun eine weitere *evokative Zeigesequenz*, mit der C und D als Rezipienten adressiert werden. Während der Eröffnung des Zeigens lehnt sich A zurück und erleichtert beiden neuen Rezipienten so die Sicht auf das Display. Gleichzeitig beginnt sie lautlos zu lachen und zeigt damit ihre Affiliation mit der Haltung der Zeigenden an. Das Timing, mit dem die Smartphoneinhaberin diese Zeigesequenz eröffnet und damit die Schismabildung der Interaktion wieder beendet, ist dabei auf den sequenziellen Verlauf der sprachlichen Interaktion von C und D abgestimmt (vgl. Licoppe & Tuncer, 2019, S. 547): Zum einen erfolgt sie in der Gesprächspause (Z. 035), die auf das zustimmende Rezeptionssignal von C folgt und deswegen als Beendigung der *response sequence* zu seiner Geschichtenerzählung interpretiert werden kann. Zum anderen wird die Zeigesequenz erst initiiert, nachdem C durch seine Beobachtung des Smartphonegebrauchs<sup>129</sup> (#4) eine reduzierte Alignierung mit der sprachlichen Interaktion mit D anzeigt.

In der anschließenden Stille (Z. 037) betrachten C und D aufmerksam das Display, während B und A sie beobachten. Dann beginnen beide Rezipienten nacheinander zu Lachen (Z. 038, 039) und produzieren damit – auch in diesem Fall – zunächst vage, jedoch eindeutig affiliative Erwiderungen (vgl. Oloff, 2019a, S. 215). Anschließend entsteht erneut eine fast zweisekündige Stille, in der zunächst C die Betrachtung der Fotografie beobachtbar beendet, woraufhin B die smartphone-assistierte Zeigesequenz durch eine Drehung des Displays, die den visuellen Zugriff für die RezipientInnen einschränkt, schließt. Erst nach dieser verkörperten Beendigung des Zeigens produzieren beide Rezipienten in Überlappung zunächst anerkennende, sprachliche

---

<sup>129</sup> Der divergierende Smartphonegebrauch wird in Mehrpersonengesprächen oftmals erst (sprachlich) *accountable* gemacht, wenn die NutzerInnen bemerken, dass sie von anderen Anwesenden beobachtet werden (Oloff, 2019b Z. 809–810). Analog dazu lässt sich fragen, ob das Bemerken des Beobachtet-Werdens die Smartphoneinhaberin hier dazu veranlasst, den für C und D ‚opaken‘ Smartphonegebrauch dadurch *accountable* zu machen, dass sie ihnen (ebenfalls) visuellen Zugriff zum Display des Smartphones gewährt.

Bewertungen (Z. 041: „alTER“; Z. 042: „läuft“). D beginnt zu Lachen (Z. 043), wodurch er sich mit der Haltung der Smartphoneinhaberin affiliert. In Z. 044 formuliert B – den Blick zunächst auf D, dann auf das Gerätedisplay gerichtet – die Aussage „es hat geSCHNEIT;“. Damit produziert sie einen *return home* zur themen-eröffnenden Äußerung von A (Z. 003) und rahmt diese retrospektiv als Anlass für die Initiierung dieser bildgestützten, *smartphone-assistierte* *Geschichtenerzählung*. Die Sequenz wird durch das Lachen von A in Z. 045 und die Auflösung eines gemeinsamen visuellen Fokus beobachtbar beendet.

Die Analyse dieser Gesprächsepisode zeigt, wie B einen länger andauernden Suchprozess beendet und anschließend durch ihr verkörpertes Verhalten eine smartphone-gestützte, *evokative Zeigesequenz* eröffnet. Nach der erfolgreichen Aufmerksamkeits- und Displayfokussierung wird die digitale Fotografie durch eine sprachliche Äußerung (Z. 033: „so sah\_s GESTern bei mir aus;“) kontextualisiert und als visuelle Repräsentation des ‚Schneefalls‘ gerahmt. Im Anschluss an die affiliative Erwidern der Rezipientin wird eine weitere Zeigesequenz eröffnet, die die beiden anderen Gesprächsteilnehmer adressiert. Im Gegensatz zum ersten Fall lässt sich hier nicht argumentieren, dass bereits durch sprachliche Äußerungen eine narrative Rekonstruktion eines Handlungs- oder Ereignisverlaufs produziert wird, in der dann eine digitale Fotografie als visuelle Repräsentation optischer Eindrücke fungiert. Die Argumentation, dass auch in diesem Fall eine multimodale Geschichtenerzählung hervorgebracht wird, lässt sich nur durch eine Betrachtung der Bildinhalte stützen. Aus der eingenommenen Kameraperspektive lässt sich zwischenzeitlich das Smartphonedisplay einsehen, sodass durch ein Hereinzoomen in den entsprechenden Bildausschnitt das Motiv der Fotografie zu erkennen ist:



Abb. 4: 181217\_Es hat geschneit\_Teil 2 – #4, Z. 034

Im Zentrum des Motivs steht eine menschliche Gestalt, die einen Kopfstand macht und dabei ihre Beine im Spagat spreizt. Auf der rechten Seite des Bilds sieht man eine Hauswand, der Boden stellt sich als weiße Fläche dar. Diese spezifische Konfiguration der in der Fotografie

abgebildeten Objekte eröffnet im Kontext dieser sprachlichen Interaktion eine Sicht darauf (vgl. Keppler, 2019, S. 181), zu welchen Handlungen der ‚nächtliche Schneefall‘ die Zeigende veranlasste: Es handelt sich um eine Momentaufnahme einer körperlichen Handlungen der Smartphoneinhaberin – eines Kopfstands mit Spagat –, die auf einer ungewöhnlichen ‚materiellen Bühne‘ – ihrer schneebedeckten Terrasse – stattfindet (vgl. Larsen, 2005, S. 425).

Beide *evokativen Zeigesequenzen* werden von der Smartphoneinhaberin durch Lachen begleitet (Z. 036; 039). Durch dieses zeigt sie zum einen ihre Haltung zu der Fotografie an und rahmt sie als etwas Amüsantes. Darüber hinaus scheint sich B mit ihrem zeigebegleitenden Lachen implizit daran zu orientieren, dass die Fotografie mehr zeigt, als sprachlich angekündigt wird (Z. 033: „so sah\_s GESTern bei mir aus;“). Damit macht sie es zur Aufgabe der RezipientInnen, während der Betrachtung zu entdecken, was die Fotografie zu einer unterhaltsamen visuellen Repräsentation des Schneefalls macht. Diese Art der Rahmung trägt zur Hervorbringung einer Erzählung bei, in dessen Vollzug durch das Zusammenspiel von sprachlicher Äußerung und digitaler Fotografie rekonstruiert wird, wie es auf der Terrasse der Smartphoneinhaberin aussah und zu welchen Handlungen sie der nächtliche Schneefall veranlasste.

Dass diese *evokative* smartphone-gestützte *Zeigesequenz* von den RezipientInnen als multimodale Rekonstruktion einer Sequenz vergangener und kausal verbundener Ereignisse verstanden wird, in der die Erzählerin durch ihre visuelle Repräsentation in der Fotografie als Figur der Erzählwelt animiert wird (vgl. Couper-Kuhlen, 2007, S. 81), zeigt sich implizit bereits an ihrem affiliativen Lachen (Z. 038, 043, 045). Expliziter wird es in den sprachlichen Erwidern von C und D (Z. 041–042): Sowohl das „alter“ von C, als auch das „läuft“<sup>130</sup> von D reagieren anerkennend auf etwas, ohne dass der Referent dieser Äußerungen explizit benannt würde. Diese werden im Zuge dieser ‚durch und durch‘ multimodalen Geschichtenerzählung nicht selbst sprachlich rekonstruiert, sondern durch die fotografische Momentaufnahme verkörperter Handlungen visuell repräsentiert und so kontextualisiert, dass die Fotografie als solche *accountable* wird. Die RezipientInnen zeigen durch ihre ‚angemessenen‘ Erwidern an, die impliziten Bedeutungen dieser bildgestützten Geschichtenerzählung verstanden zu haben (vgl. Licoppe, 2017, S. 80), woraufhin die SmartphoneinhaberIn eine ausleitende sprachliche Äußerung produziert, mit der die Erzählung retrospektiv an die Themeneröffnung von A (Z. 03) rückgebunden wird<sup>131</sup>.

---

<sup>130</sup> D’s „läuft“ scheint eine fragmentierte Version von „läuft bei dir/bei dir läuft“ (Büttner & Frank, 2016) darzustellen, die als alternative, jugendsprachliche Redewendung für ‚Du hast es drauf!‘ gebraucht wird (ebd., 2016, S. 234). Wenn diese Interpretationsmöglichkeit zutrifft, stützt dies die Argumentation weiter, dass sich die in Z. 042 geäußerte Anerkennung auf die visuell repräsentierten Handlungen oder Leistungen der Erzählerin bezieht.

<sup>131</sup> Dies erscheint insbesondere hinsichtlich der Positionierung der *smartphone-assistierten Geschichtenerzählung* nach einer Serie von Geschichten (Ryave, 1978) relevant: Die beiden vorangegangenen Erzählungen fokussierten

### 6.1.3 Fazit: Bildgestützte, smartphone-assistierte Geschichtenerzählungen

Die Analyse dieser zwei exemplarischen Fälle bildgestützter, *smartphone-assistierter Geschichtenerzählungen* zeigt, wie unterschiedlich diese Form multimodaler *tellings* eröffnet und digitale Fotografien in die F2F-Interaktion integriert werden können: Die meisten bildgestützten Erzählungen im vorliegenden Datenkorpus werden durch *informative Zeigesequenzen* realisiert. In dieser Variante wird methodisch eine Prä-Sequenz hervorgebracht, mit der vor allem zwei Aufgaben der Interaktionsorganisation bearbeitet werden: Zum einen dient sie dazu, die Verfügbarkeit einer Erzählung anzuzeigen, deren Verbindung zum vorangehenden Gespräch darzustellen und so einen sequenziellen ‚Slot‘ für die RezipientInnen zu schaffen, sich mit der projizierten Aktivität zu (dis-)alignieren. Darüber hinaus gilt es, den Gebrauch des Smartphones zu initiieren und eine smartphone-gestützte Zeigesequenz zu projizieren, mit der eine digitale Fotografie als zeigbares Objekt in die Interaktion eingeführt werden kann.

Neben der Konstruktion von ‚Erzählwürdigkeit‘ muss in Prä-Sequenzen bildgestützter Geschichtenerzählungen also auch die ‚Zeigewürdigkeit‘ (vgl. Licoppe, 2017, S. 71) einer digitalen Fotografie etabliert werden: In den meisten Fällen geschieht dies, indem die Fotografien als visuelle Dokumentation von etwas Neuem im Leben der Zeigenden gerahmt werden: Bilder neuer Objekte (bspw. eines neuen Pullovers), Momentaufnahmen kürzlich vergangener Erlebnisse (bspw. eines Sonnenuntergangs) oder Aktivitäten (bspw. ein Ausflug zu einer Sehenswürdigkeit). Gemeinsam ist all diesen *informativen Zeigesequenzen*, dass interaktiv ein Wissensdefizit auf Seiten der RezipientInnen etabliert wird oder angenommen werden kann.

Da die digitalen Fotografien, deren Integration projiziert wird, jedoch nicht immer unmittelbar ‚zur Hand‘ sind, müssen sie vor dem eigentlichen Zeigen im lokalen Speicher des Smartphones (oder über eine Online-Suche) gesucht und gefunden werden. Die prospektiven RezipientInnen von Zeigesequenzen sind für die Produktion relevanter Erwidern auf einen visuellen Zugriff zu den digitalen Fotografien angewiesen, sodass durch einen (länger andauernden) Suchprozess eine Phase in der Interaktion entsteht, in der die projizierte Aktivität der bildgestützten Erzählung nicht fortschreiten kann: Der Suchprozess wird in den meisten Fällen als notwendige Vorarbeit des projizierten Zeigens und nicht selbst als Gegenstand der Interaktion in eigenem Recht (vgl. Schegloff, 1980, S. 116) von den TeilnehmerInnen behandelt. Die Phasen des Suchens nach lokal gespeicherten Fotografien werden dabei oftmals mit *suchbegleitenden Kommentierungen* gefüllt (vgl. Suderland, 2019, S. 111 ff.): Diese dienen grundsätzlich dazu, das

---

das plötzliche Wahrnehmen des Schneefalls nach dem Aufwachen. Durch die Rückbindung von B's Erzählung an die Themeneröffnung von A macht die Smartphoneinhaberin hier *accountable*, dass ihre Geschichtenerzählung nicht als weitere Erzählung in dieser Serie zu verstehen ist, sondern durch die initiale, themeneröffnende Äußerung veranlasst wurde.

„Problem“ der *bystander inaccessibility* des mobilen Endgeräts zu bearbeiten und dessen Gebrauch *accountable* zu machen. Im Kontext einer im Vollzug befindlichen, bildgestützten Erzählung können Fotografien durch diese Kommentierung kontextualisiert und so eine spezifische Sichtweise vorbereitet werden, die erforderlich ist, um die Bilder anschließend als konstitutives Element der multimodalen, narrativen Rekonstruktion erkennen zu können<sup>132</sup>.

Anhand des zweiten Falles zeigt sich jedoch, dass – in selteneren Fällen – digitale Fotografien auch durch *evokative Zeigesequenzen* in die Interaktion eingeführt werden können. Diese zeichnen sich durch einen (fast) vollständigen Verzicht auf die methodische Entwicklung einer Präsequenz aus. Sie werden auf eine Weise eröffnet, mit der die Zeigende ihre Annahme kenntlich macht, dass die RezipientIn über genügend Wissen verfügt, die verkörperte Eröffnung des Zeigens zu erkennen und die Fotografie auf angemessene Art und Weise zu sehen (vgl. Licoppe, 2017, S. 72). *Evokative Zeigesequenzen* erfordern, dass Zeigende und RezipientInnen über geteiltes Wissen verfügen, auf das durch das gezeigte Objekt verwiesen wird. Außerdem muss das Design der Zeigesequenz für die RezipientInnen deutlich machen, welche Erwartungen an eine „angemessene“ Erwiderung und die Beurteilung des gezeigten Objekts gestellt werden (vgl. Licoppe, 2017, S. 80).

Unabhängig davon, ob die Integration der digitalen Fotografien über *informative* oder *evokative Zeigesequenzen* erfolgt, weist die Phase des eigentlichen Zeigens in der Kernsequenz bildgestützter Geschichtenerzählungen zentrale Gemeinsamkeiten auf: Nach Beendigung eines eventuell notwendigen Suchprozesses kann die eigentliche Zeigesequenz eröffnet werden, die grundsätzlich in Paarsequenzen organisiert ist, „von denen einzelne Teile auch rein verkörperter Natur sein können“ (Oloff, 2019a, S. 214). Durch diese Paarsequenzen werden Aufmerksamkeits- und Displayfokussierung koordiniert, sodass ein Teilnehmerahmen hervorgebracht wird, der sowohl der Zeigenden als auch der RezipientIn visuellen Zugriff zur digitalen Fotografie ermöglicht. Dieser Teilnehmerahmen ist konstitutiv für die Hervorbringung bildgestützter, *smartphone-assistierter Geschichtenerzählungen*, denn durch ihn wird sichergestellt, dass die vorbereitenden und während des Zeigens produzierten sprachlichen Äußerungen der SmartphoneinhaberIn ihren Sinn durch ihren Bezug zur visuell-materiellen Oberfläche der digitalen Fotografien erhalten können. Er ermöglicht es, dass auch lexiko-semantisch unvollständige Äußerung durch ihren Bezug zur visuellen Repräsentation pragmatisch verständlich werden (vgl.

---

<sup>132</sup> Wie sich anhand des untersuchten Mehrpersonengesprächs gezeigt hat, kann die Notwendigkeit eines Suchprozesses auch ein Anlass für die SmartphoneinhaberIn sein, von einer Behauptung des Rederechts abzusehen und dieses anderen Anwesenden überlassen. Dies kann dazu führen, dass andere SprecherInnen ein aktuelles Thema des Gesprächs weiter fortsetzen und die SmartphoneinhaberIn nach Beendigung der Suche zum einen erneut eine Möglichkeit finden muss, dass Rederecht zu übernehmen, und zum anderen die Produktion der bildgestützten Erzählung an den neuen sequenziellen Kontext anpassen muss.

Aaltonen et al., 2014, S. 139). Die Äußerungen der SmartphoneinhaberIn lassen sich grundsätzlich als Betrachtungsanweisungen beschreiben, mit denen die Fotografie kontextualisiert (vgl. Keppler, 2019, S. 180) und der RezipientIn eine bestimmte Sichtweise nahegelegt wird. Typisch für smartphone-gestützte Zeigesequenzen ist dabei, dass die RezipientInnen im Anschluss an die Displayfokussierung zunächst mit vagen, evaluativen Erwidern auf die Betrachtungsanweisungen reagieren, mit denen sie ihre (Dis-)Alignierung mit der Zeigesequenz und ihre (Dis-)Affiliation mit der von der ErzählerIn eingenommenen Haltung sprachlich anzeigen (vgl. Oloff, 2019a, S. 214–215).

Im Kontext bildgestützter, *smartphone-assistierter Geschichtenerzählungen* übernehmen die kontextualisierenden Äußerungen spezifisch die Funktion, die Fotografie(n) erkennbar als Element(e) in einer Sequenz vergangener und kausal miteinander verbundener Ereignisse oder Handlungen hervorzubringen. Unterschiedliche Fotografien können dabei – abhängig von der relationalen Konfiguration von Objekten auf ihrer Bildoberfläche – als visuelle Repräsentation der optischen Eindrücke der ErzählerIn oder als Momentaufnahmen verkörperter Handlungen fungieren, durch die das ‚vergangene Ich‘ der Erzählerin als Figur der Erzählwelt animiert wird (vgl. Couper-Kuhlen, 2007, S. 81). Gelingt es, dass die Fotografien von den RezipientInnen als Elemente einer Geschichtenerzählungen gesehen werden, wird eine durch und durch multimodale, narrative Rekonstruktion hervorgebracht, in denen sich sprachliche Äußerungen und digitale Fotografien wechselseitig ihren Sinn verleihen. Wie auch in anderen Formen *smartphone-assistierter Erzählungen* wird ihre Beendigung beobachtbar durch Veränderungen in der Blickführung und der Positionierung des mobilen Endgeräts angezeigt. Im Fall von bildgestützten Geschichtenerzählung geht dies damit einher, den visuellen Zugriff zum Display für die RezipientInnen einzuschränken und somit den für diese Form des Erzählens typischen Teilnehmerahmen aufzulösen.

Im Anschluss an eine eingehende Betrachtung der digitalen Fotografien, dem simultanen Hören der Äußerungen der ErzählerIn sowie einer eventuell notwendigen kommunikativen Aneignung produzieren die RezipientInnen nun konkrete Erwidern (vgl. Oloff, 2019a, S. 214–215), mit denen sie ihr Verständnis der bildgestützten Geschichtenerzählungen anzeigen. In der Post-Sequenz wird außerdem die Bedeutung der Erzählung für die Anwesenden ausgehandelt, sodass das Gezeigte weiter im gemeinsamen Gespräch verarbeitet werden kann.

Insbesondere hinsichtlich der Hervorbringung des Hauptteils bildgestützter, *smartphone-assistierter Geschichtenerzählungen* lassen sich dabei zentrale Differenzen zu überwiegend sprachlich konstruierten Geschichten identifizieren. Geschichten werden in der KA üblicherweise als *verbale* Rekonstruktion (vgl. Gülich & Hausendorf, 2000, S. 373) verstanden, die von den

ErzählerInnen durch die Produktion eines *big packages* sprachlicher Äußerungen hervorgebracht wird<sup>133</sup>. Die Analysen bildgestützter, *smartphone-assistierter Geschichtenerzählungen* fordern dieses Verständnis nun in zweierlei Hinsicht hinaus: Zum einen handelt es sich um eine grundsätzlich multimodale Praktik, da ihre erkennbare Hervorbringung auf die Verbindung zwischen sprachlichen Äußerungen und dem semiotischen Feld der Fotografie(n) angewiesen ist. Die Rekonstruktion des Handlungsablaufs erfolgt nicht nur – und teilweise nicht einmal überwiegend – verbal, da die Fotografien als visuelle Repräsentationen von Handlungen oder Ereignissen der Geschichte fungieren<sup>134</sup>. Zum anderen sind derartige bildgestützte Erzählungen nicht an die lineare Sequenzialität sprachlicher Interaktion gebunden (vgl. Mondada, 2016a, S. 341). Aufgrund der unterschiedlichen Zeitlichkeiten des Erzählens und Zeigens, sowie des Hörens und Sehens können die Elemente des rekonstruierten Handlungsablaufs nicht mehr nur nacheinander, sondern auch zeitgleich und in Überlappung zueinander in die Erzählung eingeführt werden. Auf diese Weise werden sie erst im Vollzug der Interaktion und durch die Produktion multimodaler semiotischer Einheiten in eine sinnhafte Ordnung und chronologische Abfolge gebracht. Die Analysen dieses Kapitels haben gezeigt, dass sich die InteraktionsteilnehmerInnen beobachtbar an den integrierten Fotografien als konstitutiven Elemente der Rekonstruktion eines Ablaufs vergangener Handlungen oder Ereignisse orientieren (vgl. C. Goodwin, 1984, S. 243), sodass aus Perspektive einer multimodalen KA auch derartige bildgestützte Varianten des Geschichtenerzählens als solche anerkannt werden sollten.

Die hier diskutierten Fälle der Integration digitaler Fotografien in den Hauptteil *smartphone-assistierter Geschichtenerzählungen* waren dabei von einer grundsätzlichen Ausrichtung der RezipientInnen an den Zeigesequenzen und ihrer inhaltlich gleichgerichteten Haltung mit den

---

<sup>133</sup> Licoppe und Tuncer (2019) betonen im Fazit ihrer Analysen von Zeigesequenzen, dass diese eine ähnliche, sequenzielle Grundstruktur wie Geschichtenerzählungen aufweisen. Ausgehend von eben solchen kanonischen Definitionen von *storytellings* als vorwiegend sprachlich verfassten Aktivitäten identifizieren die AutorInnen die konstitutive Multimodalität von Zeigesequenzen als zentrales Differenzkriterium zu Geschichtenerzählungen: „where story-telling can be accomplished and recognized through, and as, talk-in-interaction, showing sequences are multi-modal through and through. They are . . . sequences which attend to, and operate on, the lived visual ecologies in relation to talk as a constitutive feature of their accomplishment as such” (Licoppe & Tuncer, 2019, S. 566). Eben diese Unterscheidung trifft nun für bildgestützte, *smartphone-assistierte Geschichtenerzählung* gerade nicht zu, denn es handelt es sich um Erzählungen, die durch Zeigesequenzen gestützt werden und als multimodales Phänomen zu begreifen sind, das konstitutiv durch die Verknüpfung sprachlicher Äußerungen mit dem semiotischen Feld der Fotografien hervorgebracht wird.

<sup>134</sup> Isaacs et al. (2012) argumentieren, dass ein visueller Medieninhalt zur Geschichte werden kann, „because it is a physical representation of a prior experience and because it is readily available at a moment’s notice without any prior preparation“ (Isaacs et al., 2012, S. 621). Ausgehend von den Analysen dieses Kapitels muss diese Schlussfolgerung in zweifacher Hinsicht relativiert werden: Zum einen zeigt sich bereits an den beiden exemplarischen Fällen, dass digitale Fotografien als Elemente von Geschichtenerzählungen fungieren können, auch wenn sie nicht unmittelbar zur Verfügung stehen, sondern zunächst im lokalen Speicher des Geräts gesucht und gefunden werden müssen. Zum anderen – und wesentlich zentraler – lässt sich festhalten, dass digitale Fotografien erst in und durch das Zusammenspiel multimodaler Ressourcen im Vollzug von Zeigesequenzen als visuell-materielle Repräsentation einer vergangenen Erfahrung oder Handlung hervorgebracht werden.

ErzählerInnen gekennzeichnet. Wie ich im abschließenden Analysekapitel zeigen werde, kann die Integration digitaler Fotografien durch eine ErzählerIn auch in Erwiderung auf ‚störende‘ Verhaltensweisen der RezipientInnen erfolgen – in diesen von Disalignierung und Disaffiliation gekennzeichneten, sequenziellen Kontexten der Bildintegration lassen sich sowohl die Prozesse der kommunikativen Aneignung der Fotografien beobachten, als auch das Potenzial von *mobile-supported sharing activities* untersuchen, andere Anwesende zu einer inhaltlich gleichgerichteten Haltung zu leiten.

## 6.2 Bildgestützte, smartphone-assistierte Erwidern der ErzählerIn auf spaßhafte Provokationen

Bis hierhin wurde ausschließlich die smartphone-assistierte Integration von Medieninhalten in den Hauptteil von Geschichtenerzählungen untersucht. Dabei wurde rekonstruiert, wie gespeicherte Nachrichten-Threads mithilfe von Praktiken des Vorlesens und digitale Fotografien durch Praktiken des Zeigens als konstitutive Bestandteile narrativer Rekonstruktionen hervorgebracht werden. Die diskutierten Fälle sind dabei von einer grundsätzlichen Alignierung der RezipientInnen mit der *smartphone-assistierten Geschichtenerzählungen* gekennzeichnet. Diese ermöglicht es den ErzählerInnen die multimodale, narrative Rekonstruktion einer Sequenz vergangener Ereignisse zu produzieren. Doch nicht alle Geschichtenerzählungen verlaufen (aus Sicht der ErzählerInnen) derart ‚reibunglos‘: RezipientInnen können durch ihr Verhalten sowohl eine unzureichende oder fehlende Alignierung – verstanden als formale Ausrichtung an der Aktivität des Erzählens (vgl. Kotthoff, 2017, S. 12) – anzeigen, als auch eine inkongruente oder widersprechende Haltung zu den rekonstruierten Ereignissen einnehmen und sich so mit der ErzählerIn disaffilieren (vgl. Stivers, 2008, S. 32). Durch disalignierende und disaffilierende Verhaltensweisen nehmen die RezipientInnen als ‚Ko-AutorInnen‘ maßgeblichen Einfluss auf die Geschichtenerzählungen (vgl. Mandelbaum, 2013, S. 501).

Dieses Kapitel stellt Geschichtenerzählungen in den analytischen Fokus, in deren Verlauf es zu einer spezifischen Form disaffiliativer Erwidern kommt: Eine der RezipientInnen produziert einen ‚tease‘<sup>135</sup> (Drew, 1987) – eine spaßhaft formulierte Provokation – durch die die narrative Rekonstruktion, die Art und Weise des Erzählens oder die von der ErzählerIn eingenommene Haltung als ‚abweichend‘ – d. h. nicht den Normalitätserwartungen der ProvokateurIn entsprechend – interpretiert wird (vgl. Drew, 1987, S. 243). In den im Folgenden analysierten Fällen veranlasst die durch eine spaßhafte Provokation eröffnete *Teasing*-Sequenz die ursprüngliche ErzählerIn zur Initiierung einer smartphone-gestützten Zeigesequenz, mit der sie eine digitale Fotografie in die F2F-Interaktion integriert, um sich zur Provokation zu

---

<sup>135</sup> Ich verwende in dieser Arbeit bewusst den englischen Begriff des ‚Teasings‘ für eine Bandbreite an Äußerungen, die als ‚mocking but playful jibes‘ (Drew, 1987, S. 219) verstanden werden. Übersetzungen für den Begriff Teasing sind u. a. ‚Neckerei‘, ‚Stichelei‘ oder ‚Frotzelei‘. Diese sind im Deutschen jedoch nicht nur alltags-sprachlich mit unterschiedlichen Konnotationen belegt, sie werden auch in der gesprächs- und gattungsanalytischen Forschung als Begriffe ‚zweiter Ordnung‘ (Schütz, 1971) verwendet. Wie Günthner (vgl. 2006, S. 92) aufzeigt, werden Unterscheidungen zwischen ‚Necken‘ als vorwiegend ‚harmlos‘ und ‚scherzhaft‘, ‚Sticheln‘ als wiederholte ‚weniger harmlose‘ Provokation und ‚Frotzelei‘ als gekennzeichnet durch eine Doppelstruktur aus Provokation und Spielmodalität vorgenommen – auch wenn diese nicht vollkommen trennscharf sind. Nicht zuletzt aufgrund des beschränkten Umfangs der Kollektion, die diesem Kapitel zugrunde liegt, erscheint es mir sinnvoll, mit dem vageren Überbegriff des ‚Teasings‘ zu arbeiten und über die feinen Unterscheide der verschiedenen Spielarten von (mehr oder minder) spaßhaften Provokationen hinwegzusehen.

positionieren. Bevor ich die Analysen derartiger bildgestützter, *smartphone-assistierter Erwid-  
derungen auf spaßhafte Provokationen* beginne, möchte ich zunächst die typischen Verlaufs-  
strukturen und methodischen Verfahrensweisen von *Teasing*-Sequenzen darstellen, um einen  
analytischen Rahmen für die anschließende Ergebnisdarstellung zu schaffen.

Spaßhafte Provokationen besitzen – unabhängig von den konkreten Details ihrer Realisierung  
– einige grundlegende, gemeinsame Eigenschaften: Sie sind keine themen-initiiierenden Äuße-  
rungen, sondern reagieren auf einen, meist unmittelbar vorangegangenen Redezug einer Spre-  
cherIn, die dadurch zum Ziel des *Teasings* wird (vgl. Drew, 1987, S. 233). *Teasings* nehmen  
meist Bezug auf Äußerungen wie Übertreibungen, Selbst-Überhöhungen oder Beschwerden,  
die von den ProduzentInnen der Provokation aufgrund ihrer Formulierungsdynamik oder se-  
quenziellen Eigenschaften (bspw. ‚unnötige‘ Wiederholungen) als *abweichend* interpretiert  
werden. Auf diese Weise rahmen Provokationen die Zieläußerung als rechtfertigungsbedürftig.  
Gleichzeitig werden sie jedoch methodisch auf eine Weise produziert, die offenkundig macht,  
dass sie nicht vollständig ernst gemeint sind, bspw. indem lexikalische Übertreibungen verwen-  
det oder ein hoher Grad an Gegensätzlichkeit zur Zieläußerung hergestellt wird (vgl. Drew,  
1987, S. 231–232)<sup>136</sup>. *Teasings* sind als spaßhafte Provokationen also grundsätzlich durch eine  
Doppelstruktur gekennzeichnet, in der Kritik am Verhalten des Gegenübers mit einer spaßhaf-  
ten Modalität gepaart wird.

Durch die Realisierung spaßhafter Provokationen wird ein sequenzieller Kontext geschaffen,  
in dem sich die RezipientIn zu diesem *Teasing* positionieren muss. Die Erwidderungen der Re-  
zipientIn lassen sich dabei auf einem Kontinuum verorten, an dessen Polen die ausschließliche  
Orientierung an entweder der spaßhaften oder der ernsthaften Dimension der Provokation steht.  
Häufig reagieren die RezipientInnen mit ernsthaften Erwidderungen (‚po-faced receipts‘, Drew,  
1987) oder Lachen in Kombination mit einer Zurückweisung der *abweichenden* Kategorisie-  
rung (Günthner, 2006, S. 92). Auf diese Weise wird ein Sequenzmuster eröffnet, das dem Ver-  
lauf von ‚Vorwurfs-Rechtfertigungs-Sequenzen‘ gleicht und in dem die RezipientIn an der  
Glaubwürdigkeit oder Angemessenheit ihrer ursprünglichen Darstellung festhält (Drew, 1987,  
S. 243). Durch die Orientierung am sachlichen Inhalt der Kritik behandeln die RezipientInnen  
die Provokation als Versuch der Ausübung von sozialer Kontrolle über ihre Darstellungen (vgl.  
Drew, 1987, S. 250) und zeigen ihren Widerstand gegen diese an. In alternativen

---

<sup>136</sup> Die wohl detailliertesten Analysen der methodischen Hervorbringung des erkennbar ‚unernsten‘ Charakters  
spaßhafter Provokationen werden von Günthner (2006, 2012b) für ‚Frotzeleien‘ durchgeführt: Sie zeigt, dass in  
‚Frotzeläußerungen‘ verschiedene Verfahren wie Aufladung mit Lachpartikeln, expressive Prosodie, der Wechsel  
der Interaktionsrolle durch ‚fremde Stimme‘ und hyperbolische Überzeichnungen miteinander kombiniert werden,  
um sie als ‚spielerische‘ Provokationen erkennbar zu machen (vgl. ebd., 2012b, S. 174).

Verlaufsformen von *Teasing*-Sequenzen findet eine stärkere Orientierung an der spaßhaften Dimension der Provokation statt: Dies geschieht entweder dadurch, dass sie ignoriert und somit als ‚nicht-ernstzunehmende‘ Äußerung behandelt wird, oder indem die Spaßmodalität aufgegriffen und humoristisch erweitert wird – bspw. durch die Produktion einer ‚schlagfertigen‘ ‚Retourkutsche‘ (Günthner, 2012b, S. 180), mit der eine Provokation an die ProvokateurIn zurückgesendet wird.

Diese Kurzdarstellung der methodischen Hervorbringung von *Teasing*-Sequenzen dient als analytischer Rahmen für die Untersuchung bildgestützter, *smartphone-assistierter Erwidern* der ErzählerIn auf spaßhafte Provokationen. Ich werde in der Analyse dieser Praktik folgende Forschungsfragen mitführen:

- Welcher Verlauf von *Teasing*-Sequenzen veranlasst die ErzählerIn zur Produktion bildgestützter, *smartphone-assistierter Erwidern* und wie werden digitale Fotografien methodisch in derartige Sequenzen eingeführt?
- Welches Funktionspotential besitzen digitale Fotografien als Ressource für Erwidern auf Provokationen, mit denen Elemente der vorangegangenen Erzählung als *abweichend* gerahmt wurden?
- Welche Relevanz besitzt die notwendige kommunikative Aneignung der Fotografien für die Auswirkungen der bildgestützten Erwidern auf den weiteren Interaktionsverlauf?

Da es sich bei spaßhaften Provokationen um doppelbödige Interpretationen vorangegangener Äußerungen einer ErzählerIn handelt, mit denen auf spaßhafte Weise die Angemessenheit ihrer Darstellungen in Frage gestellt wird, kommt es in *Teasing*-Sequenzen beobachtbar zur Disalignierung mit der Geschichtenerzählung und zur Disaffiliation mit der ErzählerIn. Die Analyse bildgestützter, *smartphone-assistierter Erwidern* auf derartige Provokationen ist deshalb nicht nur geeignet, um eine weitere Praktik der Integration visueller Medieninhalte empirisch zu untersuchen. Sie kann methodologisch als Heuristik eingesetzt werden, um zu rekonstruieren, inwiefern die Integration digitaler Fotografien geeignet ist, RezipientInnen zur Alignierung und Affiliation zu leiten (Raclaw et al., 2016, S. 377).

### **6.2.1 An der Spaßmodalität orientierte, bildgestützte Erwidern**

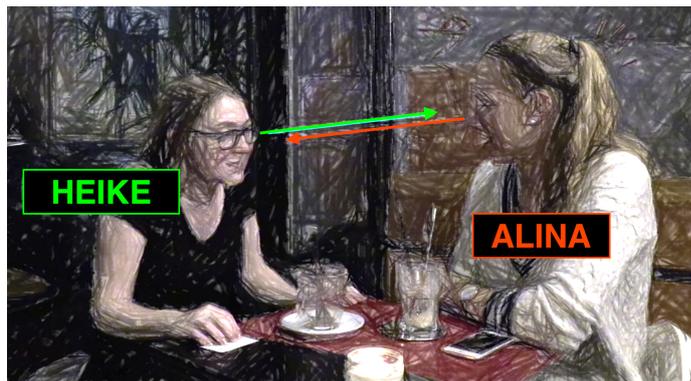
Ich möchte mit der Analyse eines Falles beginnen, in dem die ErzählerIn eine bildgestützte, *smartphone-assistierte Erwidern* einsetzt, um die spaßhafte Modalität eines *Teasings* aufzugreifen. Die folgende Gesprächsepisode stammt aus einer Unterhaltung zwischen den zwei FreundInnen Heike (H) und Alina (A), die sich in einer Serie von Geschichten (vgl. Ryave,

1978) über aktuelle Entwicklungen mit ihren ‚Männerbekanntschaften‘ austauschen. Um rekonstruieren zu können, wie der Verlauf einer *Teasing*-Sequenz die Produktion einer *bildgestützten, smartphone-assistierten Erwidern* veranlasst, werde ich nachzeichnen, welche Elemente der Geschichtenerzählung zum Ziel der spaßhaften Provokation werden. Aus Platzgründen wird die – zunächst ohne Smartphoneintegration realisierte – Geschichtenerzählung zusammenfassend dargestellt, um anschließend die *Teasing*-Sequenz anhand konkreter Transkriptausschnitte zu analysieren: Die folgende Episode ist Teil einer längeren Beschwerdeerzählung über das kommunikative Verhalten eines Abwesenden (vgl. Günthner, 1999a). In ihr erzählt H von einem ‚Typ‘, der sie über WhatsApp angeschrieben hat. Sie beklagt sich, dass der ‚Typ‘ sowohl ihre Telefonnummer und als auch ihren Studiengang kennt, obwohl sie sich nicht erklären kann, wie er an diese Informationen gelangt ist. Der Hauptteil der Erzählung wird durch Rekonstruktion der ausgetauschten Chatnachrichten im Modus der DRS ausgestaltet. Zunächst ist die Erzählung durch die Alignierung und Affiliation der Rezipientin A gekennzeichnet: Sie hört aufmerksam zu, produziert Zuhörersignale und stimmt der Einschätzung zu, dass das Verhalten des Abwesenden als ‚gruselig‘ zu bewerten ist.

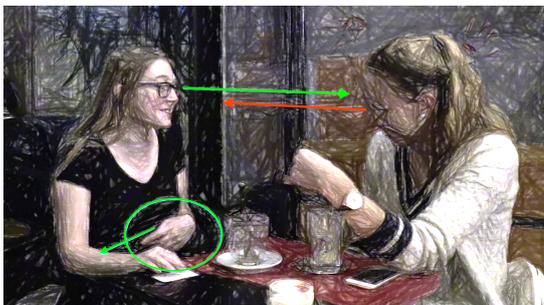
Dieses Verhältnis zwischen den Gesprächspartnerinnen ändert sich erst, als H beginnt, von Komplimenten zu erzählen, die ihr der ‚Typ‘ geschickt hat. Dabei scheint es vor allem die Art der Wiedergabe der Komplimente zu sein, die von der Rezipientin als Form der Selbstüberhöhung interpretiert und zum Anlass genommen wird, wiederholt spaßhafte Provokationen zu produzieren: Sie fragt die Erzählerin bspw. mit *smile voice*, „wie viele hundert Typen“ sie „am Start“ hätte (Z. 180, nicht im Transkript enthalten). Durch diese hyperbolische Überzeichnung rahmt die Rezipientin nicht länger die Nachrichten des Abwesenden, sondern die ‚Anzahl der romantischen Kontakte‘ der Erzählerin als *abweichendes* Verhalten. H behandelt diese Provokationen zuerst als „byplay“ (M. H. Goodwin, 1997), das ein nicht-fokales Element der Geschichte aufgreift, und versucht zu verhindern, dass dieses zum Fokus der Interaktion wird (vgl. ebd., 1997, S. 84). Auf diese Weise gelingt es ihr zunächst, das Rederecht wieder zu übernehmen und die Geschichtenerzählung fortzusetzen.

Der folgende Ausschnitt setzt mit H’s Ausleitung aus ihrer Erzählung über die Chatkommunikation mit dem fremden Mann ein und zeigt, wie die Rezipientin weitere spaßhafte Provokationen produziert. Die beiden FreundInnen sitzen in einem Café in einem *L-Arrangement* und schauen sich gegenseitig an. Entsprechend werden zur besseren Lesbarkeit des Transkripts nur Veränderungen der Blickführung annotiert. Heikes (H) Blicke werden mit „+“, ihre Gesten mit „\*“ transkribiert, Blicke und Gesten von Alina (A) sind durch „%“ kenntlich gemacht.

Transkript 25: 181205 „Creep“\_Teil 1



184 H: ich dachte so ja: (.) WAS ist das hier?  
 185 was? (.) hä:,  
 186 <<lachend> du [bist richtig GRU:selich;>]  
 187 A: [<<:-=)warum ge: >]  
 188 <<lachend> warum gehen die alle so auf dich AB?>  
 189 [<<lachend> ich komm GAR > ]  
 190 H: [<<lachend, pp >ich weiß au-> ]  
 191 A: <<lachend> nicht  
 drauf kla(h)a(h)r;>  
 192 H: !DAN!:ke.+  
 +Blick geradeaus-->194  
 193 A: ((lacht, 0.8 Sek.))  
 194 A: !EY! [gib mir ma+ was ab von]  
 h: -----,,-->+  
 195 H: [nhehe KEIne ahnung; > ]  
 196 A: <<:-)>deiner BEU:te;>  
 197 <<lachend>OHne witz.+>  
 h: ----->+ auf Tisch-->199  
 198 A: ((lacht))  
 199 H: [<<lachend> du +\*kan#nst den +\*HA:#ben;>=]  
 \*LH zeigt zu rechter Seite  
 \*RH Griff Tasche->201  
 ----->+..--zu A,->+Blick in Raum-->200  
 #1 #2



200 A: [((lacht)) > ]  
 201 H: [=<<lachend> ich \*zeig# dir h den +hh WAR#\*te.> hhh]  
 -----\*RH SP aus Tasche \*.....-----

#3

#4



202 A: [ ((lacht)) ]

Der Ausschnitt beginnt damit, dass die Rezipientin erneut durch eine spaßhafte Provokation den Fokus auf die ‚romantischen Kontakte‘ der Erzählerin lenkt: Bereits in Überlappung zu H’s Schlussfolgerung aus der rekonstruierten Handlungsfolge (Z. 186), mit der sie den Abwesenden erneut als „gruselig“ bewertet, setzt A zur Produktion einer Äußerung an (Z. 187: „warum ge:“). Sie unterbricht diese jedoch zunächst und passt die Beendigung von H’s Redezug in Z. 189 ab. Dann äußert sie mit einem recycelten Turn-Anfang ihr Unverständnis für das Interesse „aller“ ‚Typen‘ an der Erzählerin. Typisch für eine spaßhafte Provokation ist dabei die lachende Realisierung sowie die Verwendung einer „extreme-case“-Formulierung (Pomerantz, 1986), die in Kombination mit einer hyperbolischen Überzeichnung (Günthner, 2012b, S. 178) das Interesse der Verehrer spaßhaft als *accountable* behandelt („warum gehen die alle so auf dich AB?“).

Im Anschluss beginnen beide TeilnehmerInnen mit einem Simultanstart zu sprechen: Lachend und mit gesenkter Lautstärke setzt H zu einer Erwiderung an (Z. 190: „ich weiß au-“), bricht die Äußerungsproduktion jedoch ab und überlässt A das Rederecht. Diese ergänzt lachend: „ich komm GAR nicht drauf kla(h)a(h)r;“ (Z. 189, 191), wodurch sie ihr Unverständnis über diesen Umstand betont. Während H vorher versuchte, die Fokussierung dieses Aspekts der Erzählung durch lachende Zurückweisungen der abweichenden Kategorisierung (Günthner, 2006, S. 92) zu verhindern, formuliert sie in Z. 192 nun eine andere Art der Erwiderung: Mit ernstem Gesichtsausdruck ‚bedankt‘ sie sich ‚entrüstet‘ und wendet ihren Blick von A ab. Indem sie ironisch das zweite Teil des Adjazenzpaars ‚Kompliment-Dank‘ produziert (vgl. Pomerantz, 1978), orientiert sich H beobachtbar an der Herabwürdigung, die in As demonstrativ geäußertem Unverständnis enthalten ist. Während dieser Äußerung und der anschließend Gesprächspause amüsiert sich A lachend auf Kosten der Erzählerin (Z. 194–195). Im Kontrast zu H’s Gesichtsausdruck sowie ihrem empörten ‚Bedanken‘ entsteht dadurch zunehmend der Eindruck, dass die *Teasing*-Sequenz droht, in einen Konflikt überzugehen.

Bevor es jedoch dazu kommt, produziert A mit *smile voice* eine Aufforderung an H, ihr etwas von der „Beute“ ‚abzugeben‘ (Z. 194, 196). In Überlappung realisiert H zunächst noch eine von Lachpartikeln initiierte ‚ernsthafte Erwiderung‘ mit der sie anzeigt, keine Erklärung für das Interesse ihrer Verehrer geben zu können (Z. 195: ‚n(hehe) keine Ahnung‘). Diese Äußerung wird von A jedoch nicht erwidert, stattdessen produziert sie in Z. 197 lachend mit ‚ohne Witz‘ eine äußerungsmodalisierende Kommentarphrase (vgl. Torres Cajo, 2017, S. 230), mit der sie der spaßhaften Aufforderung Nachdruck verleiht. Anschließend bricht sie in anhaltendes Lachen aus (Z. 198, 200, 202). Während dieses Lachens richtet H ihren Blick zunächst auf den Tisch, bevor sie A wieder anschaut und lachend der Aufforderung nachkommt: Sie bietet A an, dass sie den Abwesenden ‚haben‘ könne (Z. 199), während sie mit ihrer linken Hand zu ihrer rechten Körperseite zeigt (#1) und mit der rechten zu ihrer Gesäßtasche greift (#2). Anschließend kündigt sie sprachlich an, ihr den Abwesenden zu ‚zeigen‘ (Z. 201), wobei sie der Äußerung in der turn-finalen Position eine *Warte-Aufforderung* anfügt. Währenddessen zieht sie ihr Smartphone aus der Tasche (#3) und initiiert die Bedienung des mobilen Endgeräts (#4). Durch die Hervorbringung dieser *multimodalen Gestalt* projiziert H nun die smartphone-gestützte Integration einer digitalen Fotografie. Dies erfolgt erst, nachdem die Rezipientin sich wiederholt mit der im Vollzug befindlichen Erzählung disaligniert und durch die Fokusverschiebung mit der Story-Produzentin disaffiliert. Die *Teasing*-Rezipientin versucht zunächst wiederholt, die Provokationen durch Ignorieren als ‚byplay‘ (M. H. Goodwin, 1997) abzutun, und widersetzt sich anschließend durch ernste Erwiderungen der Interpretation ihrer Erzählung als *abweichend*. Erst auf die im Modus der Fiktionalisierung (vgl. Bergmann, 1998, S. 114)<sup>137</sup> produzierte Aufforderung von A, die ‚Beute‘ mit ihr zu teilen, reagiert H mit einer ‚schlagfertigen‘ Erwiderung und projiziert die Integration einer digitalen Fotografie mithilfe sprachlicher und verkörperter Methoden. Durch die Hervorbringung dieser *multimodalen Gestalt* wird für die Gesprächspartnerin *accountable* gemacht, dass das Smartphone als Ressource mobilisiert wird, um die ‚punktuelle Fiktionalisierung‘ (Bergmann, 1998, S. 116) des ‚Teilens der Beute‘ auszubauen. Mit der *Warte-Aufforderung* in turn-finaler Position wird sowohl die Fokussierung auf die Bedienung des mobilen Endgeräts erwartbar gemacht (vgl. Keppler, 2019, S. 179–180), als auch die Rezipientin zur vorübergehenden Suspension der aktuell im Vollzug befindlichen (*Teasing*-)Aktivität angehalten.

---

<sup>137</sup> Dass es sich hier um eine durch die Rezipientin eingeleitete ‚punktuelle Fiktionalisierung‘ (Bergmann, 1998, S. 116) handelt, lässt sich nur im Rückgriff auf das alltagsweltliche Verständnis romantischer Sozialbeziehungen argumentieren: (Männliche) Verehrer als ‚Beute‘ mit FreundInnen zu teilen ist eine mögliche, aber unwahrscheinliche Folgehandlung, die zwar ‚lapidar formuliert‘ (ebd., S. 166), aber von den TeilnehmerInnen ohne Schwierigkeiten als fiktionale Alternativpunkte der Erzählung identifiziert wird.

Nach einem kurzen Suchprozess eröffnet die Smartphoneinhaberin eine *informative Zeigesequenz*, mit der sie eine Fotografie des Abwesenden in die Interaktion einführt:

**Transkript 26: 181205\_„Creep“\_Teil 2**

200 A: [ ((lacht)) ]  
 201 H: [=<<lachend> ich \*zeig# dir h den +hh WAR#\*te.> hhh]  
 -----\*RH SP aus Tasche \*.....-----  
 Bedienung SP-->Z.210  
 ----->+SP-Display-->214  
 #3 #4



202 A: [ ((lacht)) ]  
 203 H: °h der [is also der is SECHSunzwanzig ]  
 204 A: [ <<:->ich glaub ich will den nich HABen;> ]  
 205 H: der is SECHSunzwanzig;  
 206 und der sieht echt GUT aus.  
 207 und der is auch wirklich !NETT!;  
 208 (0.8)  
 209 is aber halt (.) n (.) CREE:P.  
 210 (4.1)\*%#(3.2)%#(0.6)%  
 h: -->\*dreht Display zu A-->214  
 a: %Blick Display-->220  
 a: %swipe %

#5 #6



211 H: <<pp>(hat) keine;)>  
 212 (2.4)  
 213 A: ach:+NJA:..  
 h: --->+Blick A-->215  
 214 (0.5)\*(1.3)  
 h: ,,-->\*

215 H: <<all>joa HÄSSlich ist der +nich;>=  
----->+Display-->220

216 A: =NEE (das) nich;  
217 (0.6)

218 A: a[ber ]  
219 H: [nja:]

220 A: er is auch nich \*so das %HIGH+light;  
-----, , , , , , , , %Blick zu H-->222  
h: -----+zu A -->222  
\*.....SP auf Tisch-->>

221 (3.1)

222 H: auf jeden fall+ war% das so rIchtich SELTsam;%  
-----, , , >+in Raum-->>  
a: --->%... Blick Uhr LH , , , , %

223 A: hmhm;  
224 (1.1)

225 H: wieviel UHR is\_es?

Während H die angekündigte Fotografie auf ihrem Smartphone sucht, initiiert sie die Produktion einer *suchbegleitenden Kommentierung*. Mit „also“ beginnt sie in Z. 203 die erste Eigenschaft des abwesenden Verehrers zu beschreiben, während A in Überlappung zunächst eine Ablehnung des fiktionalen Angebots produziert: Sie will den Abwesenden „nicht haben“ (Z. 204): Trotz des spaßhaften Charakters der Offerte zeigt A sowohl durch die deutliche Verzögerung (das ‚Angebot‘ erfolgte in Z. 199) als auch das modalisierende „ich glaube“ in turn-initialer Position an, dass sie ihre Ablehnung als dispräferiert behandelt (vgl. Schegloff, 2007a, S. 14). H reagiert auf diese Zurückweisung jedoch nur insofern, als dass sie in Z. 205 – durch eine wiederholte Nennung des Alters des Abwesenden – erneut beginnt, eine dreiteilige Liste (vgl. Jefferson, 1990, S. 63) von positiven Eigenschaften (‚Alter‘, ‚Aussehen‘ und ‚Charakter‘, Z. 205-207) zu produzieren. Diese Äußerungen fungieren als Kontextualisierungen der Fotografie und können im Zusammenhang der spielerischen Fortsetzung der *Teasing*-Sequenz als ‚Anpreisen‘ ihrer ‚Beute‘ interpretiert werden. Als in der Stille in Z. 208 keine Erwiderung durch A erfolgt, fügt H lachend hinzu: „is aber halt (.) n (.) CREE:P“ und erzeugt so eine Kontraststruktur, durch die die positiven Eigenschaften des Verehrers einer schwerwiegenden negativen Eigenschaft gegenübergestellt werden. Aufgrund der performativen Pause erhält die Nennung der negativen Eigenschaft als abschließendes Element der Liste den Charakter der Pointe eines Witzes (vgl. Sacks, 1978, S. 258), wird jedoch von A nicht erwidert.

Es entsteht zunächst eine Gesprächspause, in der H nach ca. vier Sekunden die Suche nach der Fotografie des Abwesenden beobachtbar beendet, indem sie das Display zu A dreht. Die veränderte Positionierung des mobilen Endgeräts veranlasst A zur Displayfokussierung (#5),

wobei A nach einer kurzen Zeit der Betrachtung das Gerät durch eine ‚wischende‘ Bediengeste manipuliert<sup>138</sup> (#6). Auch wenn aus der eingenommenen Beobachterperspektive nicht erkennbar ist, welchem Zweck diese Bediengeste dient, scheint sie für H als sinnhafte Handlung erkennbar (vgl. Brown et al., 2013, S. 1036), denn sie entgegnet im unmittelbaren Anschluss: „(hat) keine“ (Z. 212)<sup>139</sup>. Nach einer eingehenden Betrachtung der Fotografie produziert A in gedehnter Aussprache mit „ach NJA::“ (Z. 213) zunächst eine vage, jedoch eindeutig disaffiliative Evaluation. Wie Oloff (2019) zeigen konnte, produzieren die RezipientInnen smartphone-gestützter Zeigesequenzen zunächst meist unspezifische Evaluationen „im Sinne der projizierten Positionierung“ (ebd., 2019a, S. 214) der Zeigenden, bevor sie nach eingehenderer Betrachtung konkrete Einschätzungen formulieren. Dass hier keine derartige initiale Erwiderung erfolgt, macht für die Zeigende die Realisierung einer dispräferierten Antwort erwartbar – die dann auch in Z. 213 erfolgt.

In der anschließenden Stille (Z. 214) dreht H das Display stärker zu sich, führt einige Bediengesten durch und produziert dann in Z. 215 mit schneller Sprechgeschwindigkeit eine eingeschränkte Zustimmung („joa“) gefolgt von einer modifizierten Bewertung des Aussehens des „Creeps“ („HÄSSlich ist der nich“): Durch die Verwendung einer Litotes tritt „mit der Negation [des Gegenteils, Anm. DS] etwas Unbestimmtes an die Stelle, an der eine eigentlich notwendige Bestimmungsleistung“ (Bergmann, 1999, S. 182) zu erwarten wäre. Indem die Bewertung so konstruiert wird, dass sie anzeigt, wie der Referent nicht beurteilt wird, eröffnet sie prospektiv einen größeren Spielraum für mögliche übereinstimmende Bewertungen der Rezipientin und fungiert als Methode, das Ausmaß des aufgetretenen Widerspruchs zu entschärfen. A stimmt dieser modifizierten Einschätzung im schnellen Anschluss zunächst zu (Z. 216), formuliert jedoch in den Z. 218 und 220 anschließend erneut eine als dispräferiert verfertigte, aber eindeutig widersprechende Bewertung, in der sie ebenfalls eine Litotes verwendet („nicht so das HIGHlight“). Während der Äußerungsproduktion übergeht sie dabei H’s Versuch der Sequenzschließung (Z. 219: „nja:“).

Im Anschluss an das Auftreten dieses offenkundigen Widerspruchs in der Haltung zum visuell repräsentierten ‚Verehrer‘ entsteht zunächst eine mehrsekündige Pause (Z. 221), nach der H in

---

<sup>138</sup> In dieser Passage zeigt sich für smartphone-gestützte Zeigesequenzen erneut, dass „ein Teilen visueller Inhalte auch a priori ein Teilen . . . des Gerätes beinhaltet“ (Oloff, 2019a, S. 213). Wird eine Zeigesequenz auf eine Weise eröffnet, in der der visuelle Zugriff zu einem Medieninhalt eine Bedingung für die Produktion der konditionell relevant gewordenen Erwiderung der RezipientIn darstellt, kann neben einer eingehenden Betrachtung auch die Manipulation der Displayanzeige erforderlich sein, um diese Erwiderung produzieren zu können. Den RezipientInnen wird – zumindest in Interaktionen zwischen FreundInnen – die Berechtigung zugestanden, derartige Manipulationen selbstinitiiert durchzuführen.

<sup>139</sup> Im Kontext dieser Zeigesequenz erscheint es plausibel, dass A durch ihre Bediengeste versucht, weitere Fotografien aufzurufen und H sie informiert, dass ihr Verehrer keine weiteren Fotografien in seinem Profil hat.

Z. 222 mit „auf jeden fall war da so richtig SELTsam;“ einen *return home* produziert, mit dem sie zum Fazit ihrer Erzählung zurückkehrt, während A auf ihre Armbanduhr schaut. Die Sequenz wird von A lediglich durch die Produktion eines zustimmenden Rezeptionssignals geschlossen (Z. 223) und die Gesprächsbeendigung nach einer kurzen Stille durch eine Frage nach der Uhrzeit (Z. 225) initiiert (nicht im Transkript enthalten).

Die Analyse dieses ersten Falles zeigt, wie die Smartphoneinhaberin zunächst ernsthafte auf die wiederholten, spaßhaften, aber zunehmend herabwürdigenden Provokationen reagiert. Die Integration einer digitalen Fotografie wird erst in Reaktion auf eine scherzhafte Aufforderung der Provokateurin initiiert, mit der sie die Erzählerin dazu anhält, ihren Verehrer mit ihr zu ‚teilen‘. In ihrer ‚schlagfertigen‘, sprachlichen Erwiderung orientiert sich die Erzählerin an der spaßhaften Dimension des *Teasings* und projiziert mithilfe einer *multimodalen Gestalt* die Eröffnung einer *informativen Zeigesequenz*. In dieser wird eine digitale Fotografie des abwesenden, der Rezipientin nicht bekannten ‚Verehrers‘ in die Interaktion integriert. Die Zeigesequenz wird auf bekannte Weise hervorgebracht und sequenziell organisiert – die sprachlichen Äußerungen, mit denen die Fotografie kontextualisiert und eine bestimmte Betrachtungsweise vorbereitet wird, werden spezifisch auf die Fortsetzung der *punktuellen Fiktionalisierung* angepasst: Durch sie wird die ‚Beute angepriesen‘ und der Rezipientin nahegelegt, die zu zeigende Fotografie als Bild eines ‚gutaussehenden‘, ‚netten‘ Mannes im ‚besten Alter‘ zu sehen – der aber ‚gruselig‘ ist.

Vergleicht man nun den Verlauf der *mobile-supported sharing activity* mit dem von Disalignierung und Disaffiliation gekennzeichneten Verlauf der Geschichtenerzählung, so zeigt sich, dass die Integration der digitalen Fotografie geeignet ist, die Rezipientin zur Alignierung zu bewegen (Raclaw et al., 2016, S. 364): A kommt der *Warte-Aufforderung* nach und suspendiert die Progression der *Teasing*-Sequenz, ermöglicht der Smartphoneinhaberin die Produktion einer *suchbegleitenden Kommentierung* und aligniert sich mit der verkörperten Eröffnung der Zeigesequenz, indem sie das Smartphonedisplay fokussiert. Durch die Fokussierung ihrer visuellen Aufmerksamkeit auf den gezeigten Medieninhalt, ihre Bediengesten sowie ihre sprachliche Erwiderung nimmt sie die für die Zeigesequenz erforderliche instrumentelle und epistemische Haltung (vgl. C. Goodwin, 2007a, S. 61) ein, sodass die Begutachtung des visuell repräsentierten Mannes als gemeinsame Aufgabe der F2F-Interaktion hervorgebracht wird. Die eindeutige Alignierung mit der Zeigesequenz führt jedoch nicht zu einer Affiliation mit den Bewertungen des abgebildeten Mannes: Auch wenn die Smartphoneinhaberin die Fotografie auf spezifische Weise als fotografische Repräsentation eines ‚gutaussehenden Mannes‘ rahmt, zeigt sich im Zuge der kommunikativen Aneignung des Bildes, dass die Rezipientin zu einer

anderen Beurteilung kommt. Dies wird von ihr auch durch eine widersprechende Bewertung expliziert – trotz des methodischen Versuchs der Zeigenden, das Auftreten eines Widerspruchs zu verhindern.

Im Zuge der kommunikativen Aneignung stellt sich also heraus, dass sich in der Fotografie für die Interaktionsteilnehmerinnen etwas Unterschiedliches zeigt (vgl. Keppler, 2019, S. 181) und die Rezipientin nicht zur Affiliation mit der Haltung der Zeigenden geleitet werden kann. Trotzdem wird durch die Analyse dieser Gesprächsepisode offenkundig, dass die *Affordanz* des Smartphones, spontan digitale Fotografie aufrufen zu können, von den TeilnehmerInnen kreativ genutzt werden kann, um die Spaßmodalität von *Teasings* aufzugreifen und in multimodalen, punktuellen Fiktionalisierungen fortzuführen. Auf diese Weise kann die Rezipientin zur Alignierung geführt und das (potenzielle) Auftreten einer konflikthafteren Situation vermieden werden.

### **6.2.2 An der sachlichen Kritik orientierte, bildgestützte Er widerungen**

Anhand der Analyse eines zweiten Falles werde ich aufzeigen, dass digitale Fotografien auch als Ressource für die Produktion ‚ernster‘, *smartphone-assistierter Er widerungen* auf spaßhafte Provokationen eingesetzt werden und dabei durchaus das Potenzial besitzen, die ProvokateurInnen zu einer inhaltlichen Gleichgerichtetheit mit der Haltung der ErzählerIn zu führen (vgl. Raclaw et al., 2016, S. 370). Darüber hinaus werden die Herausforderungen für die Smartphoneintegration näher beleuchtet, die dadurch entstehen, dass die Suche nach Medieninhalten mit dem sequenziellen Fortschreiten der sprachlichen Interaktion koordiniert werden muss. Im Unterschied zum vorangegangenen Fall stammt der folgende Ausschnitt aus einem Mehrpersönengespräch zwischen den Studienfreunden Fritz (F) und Moritz (M) sowie Gregor (G) und Andi (A), die später hinzugekommen sind. Er ist Teil einer längeren Erzählung von F, der seinen Freunden berichtet, dass er mit dem letzten Gehalt seines Studentenjobs zwei Parfüms gekauft hat. Das Transkript setzt ein, nachdem die Anwesenden die Auswahl der gekauften Parfüms positiv bewertet haben (Z. 022) und F seine Geschichte dadurch weiterentwickelt, dass er sich noch einen Universitäts-Pullover mit „schwarzer Schrift“ kaufen möchte.

Die Blicke von F werden mit „+“ und seine verkörperten Handlungen mit „\*“ transkribiert. Die Blicke und Gesten von A werden mit „%“, von M mit „§“ und von G mit „&“ annotiert. Zur besseren Lesbarkeit des Transkripts werden nicht alle Veränderungen der Blickführung jedes Teilnehmers festgehalten, sondern vor allem diejenigen verkörperten Handlungen annotiert, an denen sich die TeilnehmerInnen als relevant für die *Teasing*-Sequenz und die Hervorbringung der smartphone-gestützten Integration der Fotografie orientieren. Sind die Blicke einzelner

Teilnehmer nicht transkribiert, sind sie (in der Regel) auf den jeweils primären Sprecher gerichtet.

Transkript 27: 180503\_„Understatement“\_Teil 1



022 G: [JA:] das is\_n guter duft.  
 023 F: dann äh dann wollte ich mir so n (.) uni würzburg  
 024 hoodie kaufen mit SCHWARZER schrift; (.)  
 024 (so) komplett underSTATEment; weiß\_te wie ich  
 mein,  
 025 G: !KOM!plEtt (.) under<<:-)statement>;  
 026 F: +\*#k[ann ich dir]  
 +...Blick SP-->028  
 \*griff RH zu SP-->>  
 027 A: [<<:-)>VO:LL] untersta#tEMENT;> %chch  
 %Blick zu A-->

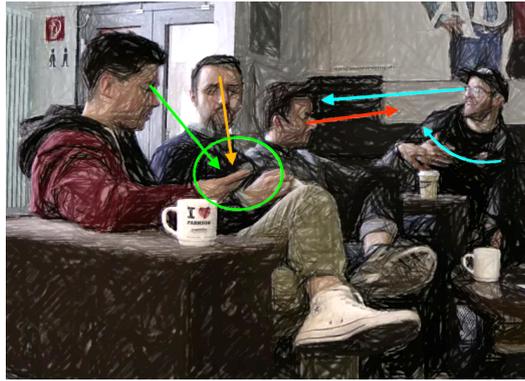
#1

#2



028 F: <<pp>( +kann\_s\_+zeig) ich [(no\_)ma;>]  
 +Blick A+Blick SP-->029  
 029 A: [ ch ch ]<<lachend>  
 fEtt drauf (.)%uni #W(h)Ü:RZ(h)+bu(h)rg;>% [hähä h h  
 ]  
 %RH fährt quer über Brust %  
 f: -----+,,Blick zu A-->031

#3



030 F: [ <<all, f>nee] das siehst\_e aber nich %weil der SCHWARZ ist.>  
 a: -----, , , ->%Blick zu F-->032  
 031 auf\_e schwar\_schwarze schrift auf +SCHWARZem pulli.  
 -----, , , , , , , , , , -->+-->Blick SP-->033  
 032 (.)  
 033 A: top.%+=  
 --->% Blick zu Boden-->034  
 f: ---->+Blick zu A-->  
 034 G: =ja +ich hab <<:-) \$hier auch (.) äh \*in +GRAU uni wÜrzburg %draufstehen.>  
 \$LH zeigt auf Brust-->036  
 f: -->+-----Blick SP----- + zu G-->  
 f: \*LH zu SP-->>  
 a: ----->% Blick zu F-->>  
 035 (0.3)\* (1.0) \*(0.2)  
 f: -->+Blick SP-->038  
 f: \*Grimasse zu G\*  
 036 G: s(hh)\_\$(h)AU(h)CH\_n(h) Und(h)ersta(hh)tement.  
 , , , -->\$  
 037 (1.2)  
 038 F: UND +noch des äh: (.) pOrte(.)monnaie +vom gräggers;  
 -->+...-----Blick A-----, , , +SP-->040  
 039 (1.8)  
 040 A: kommste da mit ZWEIhundert euro +aus?  
 f: -->+ zu A-->>

Im Anschluss an G's positive Bewertung des gekauften Parfüms (Z. 022) fährt F – mit seinem Blick A als primären Rezipienten (vgl. Kidwell, 2013, S. 109) adressierend<sup>140</sup> – mit der narrativen Rekonstruktion seiner Kaufabsichten fort: Er erzählt, sich einen Universitätspullover mit „schwarzer Schrift“ kaufen zu wollen (Z. 023) und bewertet das Design des Pullovers als „komplett underSTATEment“ (Z. 024), wobei er diese Beurteilung mit einem verständnisichernden Frageanhängsel („weiß\_te wie ich mein,“) versieht. Es ist diese Bewertung, die nun zum Ziel eines *Teasings* wird, als G sich in Z. 025 als nächsten Sprecher selbst auswählt und die Bewertung wörtlich wiederholt. Diese Wiederholung wird durch die übersteigerte Akzentuierung des Steigerungspartikel („!KOM!plEtt“) und die Realisierung des Endes der Äußerung in *smile voice* als spaßhafte Provokation erkennbar, die die Bewertung des Erzählers als unangemessen interpretiert<sup>141</sup>.

F reagiert in Z. 026 unmittelbar auf diese spöttische Bemerkung: Er greift mit der rechten Hand zu seinem Smartphone (#1) und beginnt mit „kann ich dir“ eine ernste, sprachliche Erwiderung. Er bricht die Äußerungsproduktion jedoch ab, da A – seinen Blick auf G gerichtet – in Überlappung und *smile voice* in Z. 027 eine Variation der unangemessenen Bewertung („VOLL understatement“) produziert. Während A auf diese Weise mit dem Produzenten des *Teasings* koalitiert und die spaßhafte Provokation erweitert (vgl. Günthner, 2012b, S. 185), setzt F simultan seine Zuwendung zum Smartphonegebrauch fort (#2). In Z. 028 kündigt er erneut, nun jedoch in deutlich gesenkter Lautstärke an, etwas „(no)\_ma“ zeigen zu können. Dabei scheint er kurz die Bedienung des Smartphones zu unterbrechen, denn er schaut zu A, bevor er seinen Blick wieder auf das Display richtet. Aber auch dieses Mal erfolgt keine beobachtbare Orientierung der anderen Anwesenden an der durch diese *multimodale Gestalt* projizierten Smartphoneintegration. Stattdessen wird die *Teasing*-Sequenz in Z. 029, erneut in Überlappung zu F's Äußerung, durch A erweitert, indem dieser lachend und seinen Blick auf G gerichtet eine sprachlich-verkörpernte Enaktierung (vgl. Sidnell, 2006) des „fetten“ Aufdrucks des Universitätsnamens auf der Brust produziert (#3). Während dieser Überlappung ist F visuell auf den Gebrauch

---

<sup>140</sup> F's Erzählungen über die Investition seines Gehalts wurde durch eine Frage von A veranlasst. Bevor A und G zur Gruppe hinzugestoßen sind, hatte F bereits dem neben ihm sitzenden M von seinen Einkäufen erzählt. Bei M handelt es sich also um einen ‚wissenden‘, bei A und G um ‚nicht-wissende‘ Rezipienten (vgl. C. Goodwin, 1979, S. 100): Nur für sie stellen die rekonstruierten Ereignisse Neuigkeiten dar.

<sup>141</sup> Es lässt sich argumentieren, dass die Ironisierung der Wiederholung durch ein ‚Spiel‘ mit der Präferenzorganisation von Bewertungssequenzen *accountable* gemacht wird: Indem G im unmittelbaren Anschluss eine zweite Bewertung produziert, die den ganzen Turn einnimmt und durch die Wiederholung lexikalischer Elemente als ‚gleichwertige Beurteilung‘ konstruiert wird (vgl. Pomerantz, 1984, S. 66–67), orientiert er sich an den ‚Regeln‘ der Hervorbringung einer präferierten Erwiderung. Die Spaßhaftigkeit seiner Äußerung wird durch die Kombination aus überspitzter Intonation und Akzentuierung, *smile voice* und Mikropause zwischen Steigerungspartikel und evaluativem Adjektiv erkennbar. Darüber hinaus besitzt G noch kein eigenes Wissen über den Pullover, das ihm eine zweite Bewertung ermöglicht – sodass für alle anderen Teilnehmer, die selbst keine zweite Bewertung verfertigen, eindeutig erkennbar ist, dass hier eine spaßhafte Provokation formuliert wird.

des mobilen Endgeräts fokussiert. Gegen Ende dieser multimodalen Darstellung schaut er zu A (Z. 029) und unterbricht erneut die Smartphonebedienung, um simultan zu A's Lachen eine weitere ernsthafte Erwiderung auf die spaßhafte Provokation – diesmal mit erhöhter Sprechgeschwindigkeit – zu produzieren (Z. 030): Er erklärt, warum „understatement“ eine ‚angemessene‘ Beurteilung darstellt („das siehste aber nich weil“). Während der Äußerungsproduktion beginnt A, zu ihm zu schauen. Anschließend fügt F eine explikative Beschreibung des Designs des Pullovers an (Z. 031: „schwarze schrift auf SCHWARZem pulli.“). Dieser ernsthaft formulierte Widerstand gegen die Interpretation seiner Bewertung als unangemessen führt – nach einer Mikropause in Z. 032 – zu einer sprachlich positiven („top“), aber verhalten produzierten Ratifizierung der Erklärung durch A, der anschließend seinen Blick abwendet und auf den Boden schaut. A scheint damit von der Produktion fortgesetzter Provokationen abzusehen, die *Teasing*-Sequenz wird jedoch von G erweitert, indem er eine punktuelle Fiktionalisierung produziert: Er behauptet mit *smile voice*, dass er auf seinem Pullover ebenfalls einen gleichfarbigen Aufdruck des Universitätsnamens trägt, während er mit der linken Hand auf seine Brust zeigt (Z. 034: „ich hab <<:-) hier auch (.) äh in GRAU uni wÜrzburg draufstehen.>>“)<sup>142</sup>.

Anders als zuvor wird diese Erweiterung der *Teasing*-Sequenz nun jedoch nicht vom Publikum goutiert. Lediglich F, der Adressat der Provokation, zeigt eine beobachtbare Reaktion: Er schneidet G in der anschließenden Stille (Z. 035) eine Grimasse<sup>143</sup>, die einem gekünstelten Lachen gleicht, und mit der er anzeigt, die erneute Provokation nicht lustig zu finden. Auch G's lachend formulierte Fortsetzung in Z. 036 („das ist auch understatement“) wird von den anderen Anwesenden ignoriert und so als misslungen interpretiert (vgl. Günthner, 2006, S. 105). F nutzt die anschließende Stille (Z. 037), um die narrative Rekonstruktion seiner Kaufabsichten erneut aufzunehmen (Z. 038), woraufhin A – nach einer kurzen Stille (Z. 039) – durch seine Frage eine Alignierung mit der Wiederaufnahme der Geschichtenerzählung anzeigt (Z. 040). Dies erlaubt F, seine Geschichte zu Ende zu führen. Anschließend entwickelt sich zunächst ein Gespräch über die Studieninhalte der Freunde, während dem F zur Bedienung seines Smartphones zurückkehrt (nicht im Transkript enthalten) und etliche Minuten später eine *evokative Zeigesequenz* eröffnet, mit der er eine Fotografie eines Pullovers zeigt.

---

<sup>142</sup> Auch hier lässt sich nur durch Rückgriff auf ‚unser‘ Alltagswissen explizieren, warum es sich hierbei um eine für alle TeilnehmerInnen als solche erkennbare ‚Fiktionalisierung‘ handelt: Ob auf G's grauem Pullover wirklich in ‚grauer Schrift‘ der Universitätsname aufgedruckt ist oder nicht, lässt sich zum einen von allen Anwesenden optisch überprüfen. Zum anderen stünde es im Widerspruch zu seiner vorangegangenen ‚Frotzelei‘, in der er sich über F's positive Beurteilung eines derartigen Designs lustig gemacht hat. Eine ‚ernst gemeinte‘ Information, einen vergleichbaren Pullover zu tragen, wäre somit sequenziell ‚falsch‘ platziert.

<sup>143</sup> Die Grimasse lässt sich zwar im Bewegtbild erkennen, aufgrund der Kameraperspektive jedoch leider nicht intersubjektiv nachvollziehbar in einem Standbild festhalten.

Bevor ich im Folgenden analysieren werden, welche Auswirkungen diese (verzögerte) bildgestützte, *smartphone-assistierte Erwidernng* auf die spaßhaften Provokationen hat, möchte ich zunächst den Verlauf dieser *Teasing*-Sequenz reflektieren und herausstellen, welche spezifischen Herausforderungen der Smartphoneintegration der Rezipient der Provokation zu bewältigen hat: Im Verlauf dieser Gesprächsepisode wird ein Element der narrativen Rekonstruktion – hier die Bewertung des Designs eines Pullovers als „understatement“ – zum Ziel spaßhafter Provokationen, mit denen die Rezipienten diese als unangemessen interpretieren. Die Disalignierung mit der im Vollzug befindlichen Erzählung und die Disaffilierung mit der von F angezeigten Haltung veranlassen den Erzähler bereits in Reaktion auf die erste Provokation zur Produktion einer ersten sprachlichen Erwidernng – einer ‚Handlungsankündigung‘ (Schegloff, 1980), den „Pullover“ zeigen zu können. In Kombination mit der simultanen Zuwendung zur Bedienung seines Smartphones wird so eine *multimodalen Gestalt* hervorgebracht, die die Integration des mobilen Endgeräts als (potenziell) relevante nächste Handlung projiziert (vgl. Brown et al., 2015, S. 511).

Seitens seiner Interaktionspartner findet jedoch keine Alignierung mit der projizierten Smartphoneintegration statt: Das Publikum koaliert vielmehr in Überlappung zu dieser Ankündigung mit dem Provokateur und erweitert die *Teasing*-Sequenz. Auch F’s wiederholte Projektion der Smartphoneintegration wird von den Anwesenden nicht ratifiziert, stattdessen amüsieren sie sich weiter auf Kosten des ursprünglichen Erzählers (vgl. Günthner, 2006, S. 94) und erweitern die Provokationen kollaborativ. Diese Disalignierung veranlasst F wiederholt dazu, die simultan zur sprachlichen Interaktion durchgeführte Manipulation seines mobilen Endgeräts zu suspendieren und stattdessen expliziten Widerspruch gegen die Interpretation seiner Bewertung als *abweichend* zu äußern. Erst nach einem nachdrücklichen und erklärenden Widerspruch beginnen die Anwesenden mit F zu koalieren, in dem sie den fortgesetzten Versuch der Erweiterung der *Teasing*-Sequenz durch G ignorieren und sich wieder mit der Hervorbringung der Geschichtenerzählung alignieren.

Die wiederholte Suspension der Smartphonebedienung zeigt, dass der Geräteinhaber in diesem Setting *intrapersonaler Multiaktivität* die Teilnahme an der F2F-Interaktion und den Mediengebrauch als sich wechselseitig ausschließende Aktivitäten behandelt und diese in einer *exclusive ordner*<sup>144</sup> organisiert. Daran, dass F wiederholt die Produktion sprachlicher Erwidernng priorisiert, um Widerstand gegen die spaßhaften Provokationen zu leisten, zeigt sich eine

---

<sup>144</sup> Zur Erinnerung: Mit der „exclusive order“ von Multiaktivitätssettings bezeichnet Modanda, dass „one or more courses of action are abandoned in favour of a single and exclusive other course of action“ (Modanda, 2014a, S. 70).

spezifische Herausforderung der smartphone-gestützten Integration von Medieninhalten in als *turn-by-talk* organisierte Aktivitäten: Für die Schaffung eines Slots in der Interaktion, in dem die smartphone-gestützte Integration eines Medieninhalts vorbereitet und realisiert werden kann (vgl. Brown et al., 2015, S. 511), ist die interaktive Zusammenarbeit aller Anwesenden erforderlich. In Situationen, in denen sich diese nicht mit der multimodal projizierten Smartphoneintegration alignieren, sondern die aktuell im Vollzug befindliche Aktivität durch ihre Gesprächsbeiträge dynamisch fortsetzen, zeigt sich anschaulich die Ambivalenz aus Funktions- und Problempotential des Smartphones als Ressource für die Hervorbringung sozialer Handlungen (hier einer ernsthaften Erwiderung auf eine spaßhafte Provokation): Die GeräteinhaberInnen orientieren sich beobachtbar an der *Affordanz* des Smartphones, spontan Medieninhalte aufrufen zu können – die Herausforderung besteht jedoch darin, die Zeitlichkeit des für die Integration erforderlichen Suchprozesses mit dem Fortschreiten der sprachlichen Interaktion zu koordinieren. Abhängig von der Progression der Gesprächsaktivität kann es geschehen, dass der sequenzielle Kontext, in dem das angekündigte Zeigen eines Medieninhalts eine relevante nächste Handlung gewesen wäre, nicht mehr existiert, oder andere Zugzwänge der Interaktion – bspw. das fortgesetzte Amüsement der Interaktionspartner auf Kosten des Smartphoneinhabers – die Realisierung einer alternativen Erwiderung erforderlich machen.

Anhand derartiger Dynamiken zeigt sich, dass die über *multimodale Gestalten* projizierte Smartphoneintegration von den InteraktionspartnerInnen als solche erkannt und behandelt werden muss, damit ein Medieninhalt als Ressource für einen konditionell relevanten und sequenziell passenden Beitrag eingesetzt werden kann. Der Verlauf dieser Gesprächsepisode verdeutlicht einmal mehr, dass der Einsatz der bereits diskutierten *Warte-Aufforderungen* (meist) der Bearbeitung eben dieses Problems der Interaktionsorganisation dient: Mit ihnen werden die TeilnehmerInnen zur Suspension der aktuell im Vollzug befindlichen Aktivität aufgefordert. Werden sie ratifiziert, wird es den GeräteinhaberInnen so ermöglicht, den Suchprozess durchzuführen und anschließend an einer passenden Stelle der Interaktion eine smartphone-assistierte Handlung zu realisieren.

Da die Teilnehmer in diesem Fall die sprachliche Interaktion fortsetzen, wird die *Teasing*-Sequenz und anschließend die Geschichtenerzählung von F beendet, bevor der Suchprozess nach der Fotografie abgeschlossen ist. Anschließend beginnen die anderen Anwesenden eine Unterhaltung über ein gänzlich anderes Thema. Wie der weitere Verlauf der Interaktion zeigt, können spaßhafte Provokationen durchaus einen irritierenden Charakter haben, der den Rezipienten eine längere Zeit beschäftigt (vgl. Günthner, 2012b, S. 192) – denn F fokussiert sich während des neuen Themas für einen längeren Zeitraum auf den bereits eröffneten Suchprozess. Das

folgende Transkript setzt ein, als F nach dessen Beendigung eine *evokative Zeigesequenz* eröffnet und M die Fotografie eines Pullovers zeigt. Dies führt zunächst zu einer Schismabildung in der Mehrparteieninteraktion: Während M und F das Smartphonedisplay fokussieren, unterhalten sich A und G weiter über Studieninhalte. Aus Gründen der besseren Lesbarkeit und des analytischen Fokus dieses Kapitels wurde das Transkript in zweierlei Hinsicht bearbeitet: Zum einen enthält es lediglich die ersten Äußerungen der parallel zueinander laufenden Gespräche – ab Z. 065 wird nur die Interaktion zwischen M und F wiedergegeben. Zum anderen wurde eine längere, smartphone-assistierte Einschubsequenz, in der sich F und M eine zweite Fotografie kommunikativ aneignen, entfernt, da anhand des folgenden Ausschnitts vor allem rekonstruiert werden soll, wie mit der verspätet eröffneten Zeigesequenz zunächst ein Anwesender adressiert wird und die so entstehende Schismabildung anschließend dadurch aufgelöst wird, dass auch G und A – die vorherigen Provokateure der *Teasing*-Sequenz – involviert werden:

#### Transkript 28: 180503 „Understatement“\_Teil 2<sup>145</sup>

f: -->> Bedient SP, Blick Display -->070  
 061 A: [ . un\_dann äh ] (-) [muss] ich heute [noch ]  
 062 F: [<<pp>DER hier;] (.)\$# (.) [häh?] (-- ) [achso;]  
 m: §...Blick zu SP -->070  
 #4



063 A: mir äh n bisschen BRANDstiftung an[gucken; ]  
 064 F: [(hier) al]so  
 \*#UNgefähr so::;  
 f: \*dreht Display zu M-->065  
 #5

<sup>145</sup> Zur Erinnerung: Die Blicke von F werden mit „+“, seine verkörpert Handlungen mit „\*“ transkribiert. Die Blicke und Gesten von A werden mit „%“, von M mit „\$“ und von G mit „&“ annotiert.



065 M: mhm=  
 066 F: =bloß HALT (.) mit hoo[die;]  
 067 M: [nen ] HOODie,\*  
 f: -----,/////////\*  
 068 F: jaja (.) die gibt\_s grad HIER nich.  
 [069-081: Smartphone-assistierte Einschubsequenz entfernt]  
 082 aber DIE sin\_auch qualitativ \*(.) gut  
 \*Display zu M-->087  
 083 [guck;]  
 084 M: [aber ] ham die auch so\_n GU:ten schnitt?  
 085 (-)  
 086 F: ja &das muss# ich ja erst\_ma TES:ten;=  
 g: &reckt Kopf, Blick zu SP von F-->096  
 #6



087 M: =die ham doch \*n SCHEISS schnitt;#  
 f: -->\*dreht Display nach vorn, streckt  
 Arm aus-->Z.095

#7



088 F: ne%e:=  
 a: %Blick zu SP-Display-->Z.095  
 089 =ich TESTe den erst ma.  
 090 M: das sind bestimmt solche BEU:tel.  
 091 (0.4)  
 092 F: ja %muss ich erst ma TESTen #halt; nee,  
 a: %lehnt sich nach vorn, Richtung SP-->105  
 #8



093 (.)

In Z. 062 zeigt F, seinen Blick auf das Display des Smartphones fokussiert, durch eine in deutlich gesenkter Lautstärke produzierte Äußerung („DER hier“) die Beendigung des Suchprozesses an und eröffnet eine *evokative Zeigesequenz*. Bereits mit dieser deiktisch formulierten Äußerung gelingt es ihm, die Aufmerksamkeit von M zu rekrutieren, der seinen Kopf dreht und auf das Display des Smartphones schaut (#4). Durch die Art der Initiierung der Zeigesequenz setzt F voraus, dass M – obwohl gerade ein anderes Thema behandelt wird und er den Pullover nicht erneut ausdrücklich erwähnt – in der Lage ist, die konditionellen Relevanzen der Aufmerksamkeits- und Displayfokussierung zu erkennen. Zudem nimmt er an, dass M über das notwendige Wissen verfügt, die integrierte Fotografie auf angemessene Art ‚sehen‘ zu können (vgl. Licoppe, 2017, S. 78).

Nach der erfolgreichen Eröffnung der Zeigesequenz produziert F zunächst ein fragendes „häh“, das kurz darauf von einem Erkenntnisprozessmarker gefolgt wird: Diese Äußerungsfragmente machen als *suchbegleitende Kommentierung* für den Rezipienten des Zeigens *accountable*, dass vorübergehend ein Problem in der Gerätebedienung aufgetreten und bewältigt worden ist. Anschließend produziert F in Z. 064 mit „(hier) also UNgefähr so::;“ eine Betrachtungsanweisung, die M dazu anhält, die Fotografie als visuelle Repräsentation eines ähnlichen Pullovers zu sehen. Während dieser Äußerung lehnt sich F zu M und hält das Smartphone zwischen beide, woraufhin M seinen Kopf näher zum Gerät bewegt (#5). Seinen Blick auf das Display gerichtet, produziert M einen *continuer*, mit dem er anzeigt, diese Betrachtungsanweisung als Vorbereitung für weitere Erläuterungen zu verstehen. Im schnellen Anschluss konkretisiert F, was den

abgebildeten Pullover von dem zu kaufenden Modell unterscheidet (Z. 066: „bloß HALT (.) mit hoodie;“). M's mit minimaler Überlappung produzierte, verständnisanzeigende Komplettierung der Erklärung (Z. 067: „nen HOODie;“) wird von F bestätigt und anschließend erklärend gerechtfertigt, dass es dieses Modell „hier“ nicht gäbe (Z. 068).

In den aus dem Transkript entfernten Z. 069–081 kommt es anschließend zu einer smartphonegestützten Einschubsequenz, die dadurch eröffnet wird, dass M eine Präferenz für einen anderen Bekleidungshersteller äußert. Dies veranlasst F dazu, dessen Homepage aufzurufen und ein alternatives Pullovermodell gemeinsam mit M zu betrachten und zu diskutieren. Die Einschubsequenz wird beendet, indem F die Einschätzung äußert, dass die Pullover der Marke, dessen Kauf er plant, „auch qualitativ“ gut sind (Z. 082), und er nach einer Drehung des Smartphones M – diesmal explizit mit dem Imperativ eines Perzeptionsverbs („guck“) – zur erneuten Fokussierung des Displays auffordert (Z. 083). M's Frage, ob der gezeigte Pullover „auch so\_n GU:Ten schnitt?“ hat, wird durch das turn-initiale „aber“ als Widerspruch positioniert (Z. 084) und eröffnet im weiteren Verlauf eine von Dissens gekennzeichnete kommunikative Aneignung (Z. 085–092) dieser Eigenschaft des Pullovers: In Frage steht, ob die visuelle Repräsentation einen epistemischen Zugang zu diesem Aspekt bietet und dessen Evaluation ermöglicht (Raclaw et al., 2016, S. 370) oder ob, wie M wiederholt insistiert (Z. 086, 088–089, 092), der Pullover dafür ‚getestet‘ (d. h. anprobiert) werden muss.

Im Verlauf dieser Phase des eigentlichen Zeigens handeln die zwei Freunde im Dissens miteinander aus, was für sie in der Fotografie zu sehen ist. Für den analytischen Fokus dieses Kapitels besonders relevant ist jedoch, dass die andauernde, visuelle Fokussierung von M und F auf das Smartphonedisplay und die von Widerspruch gekennzeichnete kommunikative Aneignung die Aufmerksamkeit von G zu erregen scheint. Seine Unterhaltung mit A über Studieninhalte ist in der Zwischenzeit zum Erliegen gekommen und beide Teilnehmer schauen in den Raum. Mit Einsetzen von F's Äußerung in Z. 086 reckt G sich (#6) und versucht ebenfalls auf das Smartphonedisplay zu schauen. Diese verkörperte Anzeige von Interesse an der gezeigten Fotografie wird von F als Anlass behandelt, eine weitere *evokative Zeigesequenz* zu eröffnen: Während M in Z. 087 eine widersprechende Einschätzung produziert, dreht F das Smartphone schräg nach vorn und streckt seinen rechten Arm aus (#7), sodass nun auch G und A das Display einsehen können. Mit der Realisierung von F's wiederholtem Widerspruch in Z. 088 beginnt auch A in Richtung Smartphone zu schauen, wobei er sich nach einiger Zeit der Betrachtung in Z. 092 nach vorne beugt, um die Fotografie besser sehen zu können (#8). Durch diese komplexe Verschränkung der kommunikativen Aneignung der Fotografie mit der simultanen Eröffnung einer weiteren Zeigesequenz wird die vorher entstandene Schismabildung der Interaktion

aufgelöst und die gemeinsame Aufmerksamkeit aller Interaktionsteilnehmer nun auf die digitale Fotografie gerichtet.

Anhand dieses Ausschnitts zeigt sich, dass die spaßhaften Provokationen des ersten Ausschnitts F noch länger beschäftigt haben (vgl. Günthner, 2012b, S. 192) und er die projizierte Integration einer Fotografie – obwohl bereits seit einiger Zeit ein anderes Thema behandelt wird – weiterhin als relevante Aktivität behandelt. Die Art der Eröffnung dieser Zeigesequenz und ihre Auswirkungen auf die Mehrpersoneninteraktion ähneln dabei stark der bereits in Kap. 6.1.2 diskutierten Variante (Fall: „es hat geschneit“): Der Smartphoneinhaber eröffnet nach Abschluss eines länger andauernden Suchprozesses eine *evokative Zeigesequenz*, indem er seinen nächsten Sitznachbarn als wissenden Rezipienten adressiert. Dieser ist in der Lage, die Zeigesequenz als solche zu erkennen und die Fotografie als ‚zeigbares Objekt‘ auf angemessene Art zu sehen (vgl. Licoppe, 2017, S. 78). Durch die unmittelbar erfolgende, verkörperte Alignierung des Rezipienten mit der Zeigesequenz kommt es auch in diesem Fall zu einer Schismabildung. Nachdem G und A ihr aktuelles Gesprächsthema zu einem beobachtbaren Ende geführt haben und als der Zeigende das visuelle *Monitoring* der im Vollzug befindlichen kommunikativen Aneignung durch A bemerkt, eröffnet der Smartphoneinhaber eine weitere *evokative Zeigesequenz*, die nun A und G als – ebenfalls wissende – Rezipienten adressiert.

Das Smartphone wird teilweise als *sticky media device* (Mantere & Raudaskoski, 2017) bezeichnet, da seine NutzerInnen Schwierigkeiten haben können, ihre Aufmerksamkeit vom Gerätedisplay zu lösen. Aufgrund der Analysen von „Es hat geschneit“ und dem vorliegenden Fall lässt sich wiederum argumentieren, dass der konvergente Smartphonegebrauch eine gewisse ‚Anziehungskraft‘ auf andere Anwesende ausübt: Sie beginnen die smartphone-assistierte Zeigesequenz zu beobachten, woraufhin die GeräteinhaberInnen dies zum Anlass nehmen, auch anderen Anwesenden visuellen Zugriff zu ermöglichen. Dies führt zur Auflösung des Schismas und der gemeinsamen Fokussierung der Aufmerksamkeit auf die integrierte Fotografie: Auf diese Weise zeigt sich eindrücklich, wie sehr die visuelle Materialität derartiger Medieninhalte – die eine genaue Betrachtung erforderlich machen, um ihre Relevanz für die aktuelle Interaktion zu erkennen – geeignet ist, andere Anwesende zur Alignierung mit einer *mobile-supported sharing activity* zu leiten (Raclaw et al., 2016, S. 377). Zum anderen lässt sich auch hier argumentieren, dass die Eröffnung einer smartphone-gestützten Zeigesequenz als Ressource genutzt wird, um einen aufgetretenen Dissens zu entschärfen: Indem simultan zur von Widerspruch geprägten Aneignung der Fotografie die beiden anderen Anwesenden mit einer neuen Zeigesequenz adressiert werden, werden nun ihre Erwidernungen konditionell relevant gemacht, wie der folgende Ausschnitt zeigt:

Transkript 29: 180503 „Understatement“\_Teil 3<sup>146</sup>

094 A: welche GRÖße hast du bestellt?  
 095 F: <<all>ich\*& hab% noch NIX bestellt.>=  
 ,,,,,,-->\* gestikuliert mit RH, SP in Hand-->097  
 g: ---,,,,,-->&  
 a: ----->% Blick F-->098  
 096 =ich geh: zum campus shop und teste den ma wie der  
 SITZT un\_so; #nee,  
 #9



097 \*aber #ich find \*halt (-)#s\_n (.) gu\*[%ter pulli;#]  
 \*Display zu A \*Displ. mehr zu sich\*z zu A-->098  
 a: ----->% zu SP-  
 Display-->101  
 #10 #11 #12



098 A: [ich find\_n ]  
 eigentlich ganz CO:OL;\*  
 f: ,,,,,, \*  
 099 G: ja der \*[IS auch cool; ]  
 f: \*... Display zu M-->100  
 100 M: [ja ich find den] auch NIce;=\*  
 f: -----,,-->\*  
 101 +=nur ich be%FÜRCHte halt dass es so n+\*beu:tel  
 is. \*  
 f: + dreht Display mehr zu sich ,,-->+\* zu SP-  
 Display\*  
 a: ----->%..----- zu M-----  
 ----->105

<sup>146</sup> Zur Erinnerung: Die Blicke von F werden mit „+“, seine verkörpertten Handlungen mit „\*“ transkribiert. Die Blicke und Gesten von A werden mit „%“, von M mit „\$“ und von G mit „&“ annotiert.

102 F: +ja das PRO:bier ich dann und dann sag ich dir  
 beschEId;+ okay?=  
 +.....-----SP runter,,,,,,,,,  
 ,,,,,,,,, +Display näher zu sich-->>

102 A: =ich schätze halt das is so n fruit of the LOOM (.)  
 [pullover is; ]

103 M: [hhhehehehhhhh] [hihih]hi

104 G: [hh hh]

105 A: <<:-)>kennt% ihr DIE?>  
 --,,,,,,,,-->% zu G -->>

Nach einer Weile der Betrachtung veranlasst A durch seine Frage in Z. 094 („welche GRÖÙe hast du bestellt?“) die Eröffnung einer Einschubsequenz, in der F Hintergrundinformationen zu seinen Kaufabsichten produziert. In seinen Er widerungen orientiert sich F dabei beobachtbar an den von M geäußerten Zweifeln hinsichtlich der Passform des Pullovers (Z. 096: „NIX bestellt“; Z. 097: „teste den ma wie der sitzt“). Während der Produktion dieser Äußerungen gestikuliert F mit beiden Händen, sodass das Zeigen der Fotografien vorübergehend unterbrochen wird (#9). Die Zeigesequenz wird mit der nächsten Äußerung wieder eröffnet und durch das turn-initiale „aber“ als kontrastiv gerahmt (#10): F produziert in Z. 097 erneut eine positive Bewertung des Pullovers, die durch die Verwendung von „ich find halt“ als subjektive Evaluation (vgl. Wiggins & Potter, 2003, S. 520)<sup>147</sup> verfertigt wird. Während der Formulierungsarbeit („halt (-) s\_n (.)“) dreht er das Display noch einmal zu sich (#11), bevor er es mit der Produktion des positiv evaluierenden Adjektivs („guter pulli“) wieder zu A dreht (#12).

Dieser beginnt, bevor F’s Äußerung ein erkennbares Ende erreicht hat, ebenfalls, eine (verhalten) positive Bewertung zu produzieren („eigentlich ganz CO:OL). Anhand der erweiterten Überlappung zeigt sich zum einen, dass A bereits aufgrund der emergenten Struktur von F’s Turn („ich find halt (-) s\_n (.)“) erkannt hat, dass eine bewertende Äußerung realisiert wird (vgl. M. H. Goodwin & Goodwin, 1992, S. 162). Zum anderen zeigt A so, dass er seinen epistemischen Zugang zur Fotografie nutzt, um eine eigenständige Bewertung zu verfertigen, die nicht lediglich aufgrund des von F verwendeten ‚Bewertungsadjektivs‘ positiv ausfällt (vgl. M. H. Goodwin & Goodwin, 1992, S. 165). Anschließend ist es G – der vorherige

<sup>147</sup> Wiggins und Potter (2003) unterscheiden zwischen „subjective“ und „objective evaluations“: Die Begriffe ‚subjektiv‘ und ‚objektiv‘ verweisen dabei nicht auf die (Un-)Genauigkeit oder (In-) Korrektheit der Bewertungen, vielmehr soll damit folgende Differenz zwischen verschiedenen, evaluierenden Gesprächspraktiken betont werden: „Subjective terms . . . foreground the subject; objective terms . . . foreground the object“ (ebd., 2003, S. 516). Die AutorInnen zeigen, dass natürliche Sprache Ressourcen zur Verfügung stellt, Äußerungen zu produzieren, die entweder die Eigenschaften eines bewerteten Referenten oder die Einstellungen des Bewertenden in den Vordergrund stellen. Durch ihre Verwendung lassen sich unterschiedliche Funktionspotentiale evaluierender Gesprächspraktiken realisieren (ebd., 2003, S. 516f.).

„Provokationsführer“ – der eine zweite Bewertung produziert: Diese wird als ‚objektive‘ Bewertung konstruiert: „ja der IS auch cool;“. Indem er den Fokusakzent auf „IS“ legt, bezieht G sich auf eine Eigenschaft des Pullovers und formuliert nicht nur ein subjektives Geschmacksurteil (vgl. Wiggins & Potter, 2003, S. 520).

Auch M äußert nun eine positive Bewertung (Z. 100), die ebenfalls durch die Realisierung in erweiterter Überlappung als ‚eigenständige‘ Beurteilung konstruiert wird. Dieses subjektive Geschmacksurteil räumt jedoch nicht seine Zweifel über die Passform aus, wie er in schnellem Anschluss in Z. 101 hinzufügt. Währenddessen beginnt F, die Zeigesequenz zu schließen, indem er das Smartphone zunächst zu sich dreht (was mindestens für G und A aufgrund ihrer Sitzposition den visuellen Zugriff beendet), es während seiner Erwiderung in Z. 102 absenkt und die Bedienung beendet. Seine Äußerung, die die Beurteilung der Passform erneut in die Zukunft verschiebt, schließt mit einem (ungehalten wirkenden) Frageanhängsel („okay?“), das einen Versuch der Sequenzschließung darstellt. Die von M geäußerten Zweifel werden jedoch unmittelbar von A aufgegriffen, der die Vermutung äußert, dass es sich um einen Pullover der Marke „Fruit of the LOOM“ handelt (Z. 103). Die kategoriale Identifikation des gezeigten Pullovers wird zunächst mit Lachen von M und G erwidert (Z. 103–104) und entwickelt sich im Anschluss zu einer angeregt-amüsierten ‚Kategorie-Evaluation‘ (Wiggins & Potter, 2003, S. 526) der Kleidungsstücke dieser Marke. An dieser Entwicklung zeigt sich, wie in der Beendigungsphase der Smartphoneintegration das „Gezeigte . . . in Erzählungen und Konversationen aufgegriffen“ wird und „der Belebung des Kommunikationsgeschehens“ (Keppler, 2019, S. 180) dient.

Die Analyse dieser beiden miteinander verbundenen Zeigesequenzen und der kommunikativen Aneignung der Fotografie ist – gerade aufgrund der Komplexität des Falles – geeignet, um insbesondere zwei Argumente zu verdeutlichen: Zunächst zeigt sich anhand der Art der Eröffnungen und der Verläufe der beiden Zeigesequenzen, dass die Integration visueller Medieninhalte geeignet ist, andere Anwesende zur Alignierung mit der durch das Zeigen eröffneten Aktivität zu leiten (vgl. Raclaw et al., 2016, S. 370) – selbst dann, wenn aufgrund der Dauer des Suchprozesses mittlerweile im Gespräch ein ganz anderes Thema behandelt wird. Darüber hinaus wird deutlich, dass die Integration eines visuellen Medieninhalts den Anwesenden einen anders garteten epistemischen Zugang zu einem abwesenden Referenten verschafft, der geeignet ist, sie zur Affiliation mit der Haltung des Zeigenden zu leiten (vgl. Raclaw et al., 2016, S. 377): Während G und A die sprachlichen Beschreibungen des Referenten noch veranlassten, spaßhafte Provokationen zu produzieren, mit denen die Bewertung des Erzählers als unangemessen dargestellt wurden, führt die Betrachtung und kommunikative Aneignung der visuellen

Repräsentation des Pullovers dazu, dass die Anwesenden sich auf eine (grundsätzlich) positive Bewertung des Pullovers einigen können. Im Zusammenspiel mit den sprachlichen Äußerungen des Zeigenden fungiert die digitale Fotografie hier also als Ressource, die Anwesenden zur Affiliation zu leiten.

Gleichzeitig wird an diesem Fall jedoch auch erkennbar, dass diese Affiliation kein Effekt der Fotografie ist, sondern dass der visuelle Medieninhalt der Interaktion eine zusätzliche Interpretationsebene hinzufügt (vgl. Raclaw et al., 2016, S. 377). Im Prozess der kommunikativen Aneignung wird ausgehandelt, was sich für die Anwesenden in einer Fotografie zeigt und welche Relevanz sie im Kontext einer spezifischen Aktivität besitzt: Die Verständigung auf eine inhaltlich gleichgerichtete Bewertung des Aussehens des Pullovers ist ein Resultat der bildgestützten, *smartphone-assistierten Erwiderung*. Im Zuge der kommunikativen Aneignung zeigt sich jedoch auch, dass die Rezipienten etwas in der Fotografie sehen, das so vom Zeigenden nicht antizipiert wurde: Die Passform wird – als relevante Eigenschaft eines Pullovers – nun zum Fokus der Interaktion und auf unterschiedliche Weise von den Anwesenden beurteilt, so dass es zur Disaffiliation hinsichtlich der Bewertung eines neuen Aspekts des visuell repräsentierten Referenten kommt.

### **6.2.3 Fazit: Bildgestützte Erwiderungen auf spaßhafte Provokationen**

Den beiden hier diskutierten Beispielen und allen weiteren Fällen dieser Subkollektion ist gemein, dass die RezipientInnen einen Aspekt der Geschichtenerzählung durch spaßhafte Provokationen als *abweichend* interpretieren. Diese eröffnen eine *Teasing*-Sequenz, in der sich die ProvokationsrezipientIn zur Provokation positionieren und die Sequenz dadurch weiterentwickeln, dass sie sich eher an der sachlichen Kritik oder der spaßhaften Modalität des *Teasings* orientieren. Eine Möglichkeit der Positionierung ist für die Provozierten dabei die Hervorbringung einer *multimodalen* Gestalt, mit der eine *bildgestützte, smartphone-assistierte Erwiderung* projiziert wird. Insbesondere die sprachlichen Äußerungen, mit denen die Smartphoneintegration erwartbar gemacht wird, machen für die Anwesenden dabei *accountable*, ob sich die GeräteinhaberIn eher an der spaßhaften oder der ernsten Dimension des *Teasings* orientiert.

Betrachtet man die sequenziellen Positionen, an denen bildgestützte, *smartphone-assistierte Erwiderungen* in *Teasing*-Sequenzen projiziert werden, so zeigt sich, dass dies in den meisten Fällen erst geschieht, nachdem die ProvokationsrezipientIn mehrere ernste sprachliche Antworten realisiert hat. Mit diesen versucht sie zum einen, die Provokationen nicht zum Fokus der Interaktion werden zu lassen und die ProvokateurInnen zur erneuten Alignierung mit der Geschichtenerzählung zu bewegen. Zum anderen verteidigt sie ihre Darstellung der Ereignisse

gegen die Interpretation als *abweichend* (vgl. Drew, 1987, S. 250) und versucht die ProvokateurInnen so zur Affiliation zu leiten. Anhand des abweichenden Falls der Kollektion („understatement“) konnte deutlich gemacht werden, dass die InteraktionsteilnehmerInnen eine *smartphone-assistierte Erwiderung*, wenn sie als erste Antwort projiziert wird, zunächst nicht ratifizieren, sondern die *Teasing*-Sequenzen mit ihren Beiträgen fortsetzen. Es scheint also eine Präferenz für Progressivität zu geben (vgl. Schegloff, 2007a, S. 54), bei der von allen TeilnehmerInnen unmittelbar realisierbare (spaßhafte oder ernste) Erwiderungen der RezipientInnen oder Erweiterungen des *Teasings* durch die ProvokateurIn(nen) priorisiert werden. Sie werden Handlungen vorgezogen, die einer ‚Vorarbeit‘ bedürfen – so wie bildgestützte, *smartphone-assistierte Erwiderungen*, die zunächst die Suche nach einer Fotografie erforderlich machen. Diese Präferenz für im nächsten Turn realisierbare Fortsetzungen der *Teasing*-Sequenz lässt sich damit erklären, dass spaßhafte Provokationen nicht nur dazu dienen, eine Äußerung der ursprünglichen SprecherIn als *abweichend* zu interpretieren. Aufgrund ihres spaßhaften Charakters tragen sie auch zu einer ‚geselligen Atmosphäre‘ bei (vgl. Günthner, 2012b, S. 198), in der sich die ProvokateurInnen gemeinsam auf Kosten der RezipientIn amüsieren. Von dieser wird durchaus erwartet, nicht als ‚SpielverderberIn‘ aufzutreten (vgl. Günthner, 2012b, S. 183), sondern die Entwicklung der *Teasing*-Sequenz zuzulassen oder sogar selbst zu erweitern.

Die TeilnehmerInnen an *Teasing*-Sequenzen orientieren sich an der Eröffnung bildgestützter, *smartphone-assistierter Erwiderungen* also als „nth-and-alternative-device“ (Jefferson, 1980, S. 86): Sie werden als multimodale Praktik behandelt, die erst nach dem Einsatz anderer sprachlicher Methoden, bspw. ernsten Erwiderungen, dem Ignorieren der Provokation, schlagfertigen Retourkutschen oder spaßhaften Erweiterungen, von der SmartphoneinhaberIn initiiert und den anderen Anwesenden als ‚angemessene Erwiderung‘ ermöglicht wird. Dies verdeutlicht zum einen, dass die Smartphoneintegration in als *turn-by-turn-talk* organisierten Sequenzen voraussetzt, dass von allen Beteiligten kollaborativ ein Slot geschaffen wird, in dem die Integration des Smartphones – inklusive des notwendigen Suchprozesses – vorbereitet und durchgeführt werden kann. Aufgrund der (potenziell) unterschiedlichen Zeitlichkeiten von Suchprozess und sprachlicher Interaktion muss die Projektion der Smartphoneintegration von anderen Anwesenden als solche erkannt und die aktuell im Vollzug befindliche Aktivität eventuell suspendiert werden, damit die Integration des relevant gewordenen Medieninhalts an einer ‚passenden‘ sequenziellen Position erfolgen kann. Zum anderen zeigt sich spezifisch für *Teasing*-Sequenzen das Funktionspotenzial bildgestützter Erwiderungen: Sie werden nach anhaltender Disaffiliation und wiederholten widersprüchlichen Einschätzungen realisiert. Unabhängig davon, ob sich die Zeigende dabei eher an der spaßhaften oder der ernsten Dimension der Provokationen

orientiert, wird die Integration visueller Medieninhalte methodisch eingesetzt, um der Interaktion einen neuen ‚Impuls‘ zu geben, mit dem anhaltende Widersprüche aufgelöst oder das Auftreten möglicher Konflikte vermieden werden kann.

Werden die bilgestützten, *smartphone-assistierten Erwidernngen* – nach Abschluss des erforderlichen Suchprozesses – durch *evokative* oder *informative Zeigesequenzen* von den RezipientInnen der Provokation eröffnet, zeigt sich, dass *mobile-supported sharing activities* eine effektive Methode darstellen, andere Anwesende zur Alignierung zu bewegen (vgl. Raclaw et al., 2016, S. 364): In allen Fällen führt die Eröffnung einer smartphone-gestützten Zeigesequenz dazu, dass die RezipientInnen zur erfolgreichen Hervorbringung der Aufmerksamkeits- und Displayfokussierung beitragen und sich zunächst mit der durch die Betrachtungsanweisung der Zeigenden angelegten Sichtweise auf die Fotografie alignieren (Oloff, 2019a, S. 214–215). Auf die eröffnete Zeigesequenz wird von den RezipientInnen nicht unmittelbar mit der Produktion weiterer spaßhafter Provokationen reagiert – vielmehr erhält die SmartphoneinhaberIn zunächst eine Chance, die *Teasing*-Sequenz auf die von ihr angelegte Art und Weise zu entwickeln.

Die Alignierung der RezipientInnen mit der Zeigesequenz und die Hervorbringung des für die kommunikative Aneignung der Fotografie erforderlichen Teilnahmerahmens (C. Goodwin, 2007a, S. 55) ist keinesfalls gleichbedeutend damit, dass es im Verlauf der Zeigesequenzen auch zur Affiliation – zur Gleichgerichtetheit in der Haltung gegenüber des relevanten Aspekts der integrierten Fotografie – kommt (vgl. Raclaw et al., 2016, S. 377): Im Zusammenspiel mit den sprachlichen Äußerungen der Zeigenden besitzen digitale Fotografien durchaus das Potenzial, die RezipientInnen davon zu überzeugen, dass die ursprüngliche Darstellung der ErzählerIn nicht unangemessen war. Aber auch wenn von der SmartphoneinhaberIn mit der Integration der Fotografie ein bestimmter Zweck verfolgt wird, „ist die Deutungshoheit der bildgebenden Akteure grundsätzlich dadurch gefährdet, dass andere anderes in einem Bild sehen können.“ (Keppler, 2019, S. 188). Die RezipientInnen können auf Grundlage ihres epistemischen Zugangs zur digitalen Fotografie zu widersprechenden Bewertungen kommen, oder in der Fotografie gar neue Aspekte relevant machen, die nun ihrerseits zum Fokus der kommunikativen Aneignung werden. Anhand der untersuchten Fälle zeigt sich somit vor allem, dass der F2F-Interaktion durch die Bildintegration eine zusätzliche Interpretationsebene hinzugefügt wird (Raclaw et al., 2016, S. 377). Der Ausgang der erforderlichen kommunikativen Aneignung ist grundsätzlich offen, da „das, was für die Gesprächsteilnehmer in einem Foto zu sehen ist, stets auch davon abhängt, . . . wie sie es sich – sei es im Konsens, sei es im Dissens – im Gang der Konversation aneignen“ (Keppler, 2019, S. 185).

## **7 Fazit: Smartphone-assistierte Geschichtenerzählungen als multimodale Praktik**

In dieser Studie wurden *smartphone-assistierte Geschichtenerzählungen* in alltäglichen Gesprächen aus Perspektive der multimodalen KA untersucht. In den Analysen wurde dabei fokussiert, wie über das Smartphone abrufbare Medieninhalte – Kurznachrichten und digitale Fotografien – gemeinsam mit sprachlichen Äußerungen und verkörperten Handlungen als Ressource für die Hervorbringung von Geschichtenerzählungen eingesetzt werden. Darüber hinaus wurde rekonstruiert, wie sich die beobachtbare Orientierung der TeilnehmerInnen den *Affordanzen* der mobilen Endgeräte auf die sequenzielle Organisation und methodische Hervorbringung der *smartphone-assistierten* Erzählungen auswirkt. Im Folgenden werde ich zunächst die zentralen Ergebnisse dieser Arbeit zusammenfassen (7.1). Anschließend werden die mit der Smartphoneintegration einhergehenden Veränderungen der Formen des alltäglichen Erzählens diskutiert (7.2) und abschließend ein kurzes Resümee gezogen (7.3).

### **7.1 Zusammenfassung der zentralen Ergebnisse**

Bei *smartphone-assistierten Geschichtenerzählungen* in F2F-Interaktionen handelt es sich um eine multimodale Praktik, die kollaborativ durch wechselseitig aufeinander bezogene sprachliche Äußerungen, verkörperte Handlungen aller TeilnehmerInnen und den Gebrauch der mobilen Endgeräte hervorgebracht wird. In Kap. 5 wurden unterschiedliche Varianten *smartphone-assistierter Geschichtenerzählungen* fokussiert, die konstitutiv durch Praktiken des Vorlesens von Textnachrichten hervorgebracht werden. Derartige Erzählungen zeichnen sich dadurch aus, dass die TeilnehmerInnen sich beobachtbar an der *Affordanz* des Smartphones als *Archiv sozialer Originale* orientieren: In Instant-Messaging-Applikationen und vergleichbaren Anwendungen werden asynchrone, computer-vermittelte Interaktionsverläufe in all ihren Details archiviert und lassen sich mittels weniger Bediengesten aufrufen. Die aufgerufenen Nachrichtenthreads können wiederum durch wörtliche Wiedergaben als Ressource in der Hervorbringung narrativer Rekonstruktionen eingesetzt werden.

In Kap. 5.1 und 5.2 standen *smartphone-assistierte Geschichtenerzählungen* im Fokus, in denen die Nachrichtenthreads als abgeschlossene Interaktionen behandelt und als Ressource für die Entwicklung des Hauptteils von Erzählungen eingesetzt werden. Diese Erzählungen weisen grundsätzlich die typische dreigliedrige Struktur von Geschichten auf. Diese besteht aus einer Eröffnung, in der die Kernaktivität eingeleitet wird, dem Hauptteil, in dem Rekonstruktions- und Bewertungsarbeit geleistet wird, und einem Abschluss, in dem das Verständnis des *tellings*

etabliert wird. In den ersten zwei Kapiteln wurde aufgezeigt, dass die *smartphone-assistierten Geschichtenerzählungen* sowohl durch die methodische Hervorbringung einer Prä-Sequenz eröffnet (Kap. 5.1), als auch durch einen regulären Redebeitrag initiiert und erst im weiteren Verlauf zu einer Erzählung ausgearbeitet werden können (Kap. 5.2).

Die in Kap. 5.1 analysierten methodisch vorbereiteten, *smartphone-assistierten Geschichtenerzählungen* werden von den GeräteinhaberInnen aus ihrer diskursiven Identität als SprecherInnen heraus initiiert und durch die Hervorbringung einer Prä-Sequenz eingeleitet. In dieser wird grundsätzlich die Verfügbarkeit einer Geschichte projiziert, ihre Erzählwürdigkeit dargestellt und ihre Verbindung zum vorangegangenen *turn-by-turn-talk* kenntlich gemacht. Die spezifische Verfügbarkeit einer *smartphone-assistierten* Erzählung wird erkennbar durch die Hervorbringung *multimodaler Gestalten* (Mondada, 2014b, S. 140) projiziert: Sprachliche Äußerungen, die simultane – oftmals methodisch synchronisierte – visuelle Fokussierung sowie die physische Handhabung des mobilen Endgeräts verleihen sich wechselseitig ihren Sinn (Goodwin, 2007a, S. 70), sodass für die RezipientInnen *accountable* wird, dass der Gebrauch des mobilen Endgeräts eine praktische Relevanz für die projizierte Aktivität besitzt. Insbesondere in thematisch verbundenen Geschichtenerzählungen erfolgt die Initiierung des Smartphonegebrauchs jedoch häufig ohne sprachliche Ankündigung (vgl. Keppler, 2019, S. 179) – seine Sinnhaftigkeit ist für die RezipientInnen lediglich anhand der *multimodalen Gestalten* inferentiell erschließbar.

In der Prä-Sequenz orientieren sich die RezipientInnen oftmals durch ihr verkörpertes Verhalten an der potenziellen *bystander inaccessibility* (Mantere, 2020) des Smartphonegebrauchs: Da der praktische Zweck der Gerätebedienung aufgrund der materiell-technischen Eigenschaften der mobilen Endgeräte nicht ohne Weiteres ersichtlich ist, verschaffen sie sich selbst-initiiert visuellen Zugang zum Smartphonedisplay und damit Informationen über den Zweck der Bedienung. Das visuelle *Monitoring* des Displays durch die RezipientInnen wird von den GeräteinhaberInnen meist – ohne weitere Kommentierung – ermöglicht und stellt eine verkörperte Praktik dar, den Smartphonegebrauch *accountable* zu machen. Die Hervorbringung dieses spezifischen Teilnehmerahmens wird von den SmartphoneinhaberInnen als verkörperte Alignierung mit der projizierten Erzählung und – wenn die RezipientInnen nicht das Rederecht ergreifen – als *ticket* für die Produktion des Hauptteils der Erzählung behandelt.

In Kap. 5.2 wurde gezeigt, dass nicht alle *smartphone-assistierten* Geschichtenerzählung durch eine klassische Prä-Sequenz vorbereitet werden. Stattdessen kann der Gebrauch des Smartphones von den GeräteinhaberInnen aus ihrer diskursiven Identität als ZuhörerIn initiiert und eine *smartphone-assistierte* Aktivität anschließend durch einen regulären Redebeitrag eröffnet

werden. Dieser wird erst im weiteren Interaktionsverlauf zu einer Erzählung entwickelt. Da alltägliche Gespräche von einer grundsätzlichen Konkurrenz um das Rederecht gekennzeichnet sind, muss dabei zunächst ein sequenzieller Slot geschaffen werden, in dem der selbst-initiierte Smartphonegebrauch thematisiert werden kann. Die GeräteinhaberInnen bedienen sich dazu unterschiedlicher sprachlicher Methoden (bspw. „acknowledgement tokens“ und anschließendem Themenwechsel, Jefferson, 1984, S. 199) und verkörperter Verfahren (bspw. Meldegesten). Als spezifische Methode, einen Slot für die Smartphoneintegration in derartigen Gesprächskontexten zu schaffen, fungieren *Warte-Aufforderungen*: Diese werden von den GeräteinhaberInnen grundsätzlich dafür eingesetzt, für andere Anwesende *accountable* zu machen, dass sie ihre Aufmerksamkeit auf die Bedienung des Smartphones fokussieren und ihre Beteiligung an der sprachlichen Interaktion vorübergehend reduzieren (Keppler, 2019, S. 179). In den hier untersuchten sequenziellen Kontexten dienen sie insbesondere dazu, die InteraktionspartnerInnen zur Suspension der gerade im Vollzug befindlichen Aktivität aufzufordern.

Wird diese Aufforderung von den InteraktionspartnerInnen anerkannt und die aktuelle Aktivität unterbrochen, ergibt sich anschließend die Möglichkeit, einen eventuell erforderlichen smartphone-gestützten Suchprozess als notwendige Vorarbeit für die Integration abzuschließen (Suderland, 2019, S. 101) und den aufgerufenen Medieninhalt anschließend in der Interaktion zu thematisieren. Vergleichbar zum visuellen *Monitoring* wird auch das ‚Warten‘ der RezipientInnen – d. h. die Suspension einer im Vollzug befindlichen Aktivität und der Verzicht auf die Initiierung eines neuen Themas – von den GeräteinhaberInnen als Alignierung mit der projizierten, smartphone-assistierten Aktivität behandelt und als Möglichkeit genutzt, anschließend eine Erzählung zu entwickeln.

Unabhängig davon, wie die smartphone-assistierte Erzählung initiiert wurde, findet in ihrer Kernsequenz vor allem Rekonstruktions- und Bewertungsarbeit statt. Durch diese wird die narrative Rekonstruktion einer Ereignisfolge adressaten- und aktivitätsspezifisch ausgestaltet, so dass die Haltung der ErzählerInnen angezeigt und den RezipientInnen die Produktion angemessener Erwidierungen ermöglicht wird. Der Hauptteil dieser Art *smartphone-assistierter Geschichtenerzählungen* wird konstitutiv durch das Vorlesen von Textnachrichten hervorgebracht. Diese multimodale Praktik ist durch Einnahme einer spezifischen Körperhaltung der Story-ProduzentInnen gekennzeichnet: Indem die ErzählerInnen während der DRS der Textnachrichten durchgehend auf das Display des Smartphones schauen, wird die wörtliche Wiedergabe als Vorlesen des Wortlauts *accountable* gemacht. Die Teilnahmerahmen dieser durch das Vorlesen von Textnachrichten gekennzeichneten Erzählungen sind dabei zu einem gewissen Grad variabel: Die RezipientInnen müssen eine instrumentelle Haltung einnehmen, die es

ihnen ermöglicht, für die im Vollzug befindliche smartphone-assistierte Erzählung relevante Handlungen und Erwidern zu produzieren (vgl. Goodwin, 2007, S. 61). Dies setzt voraus, dass sie mindestens die ErzählerIn anschauen und ihren sprachlichen Äußerungen zuhören. Oftmals beobachten die RezipientInnen jedoch auch das Displays des Smartphones. Dieses visuelle *Monitoring* dient dabei als verständnis- und intersubjektivitätssicherndes Verfahren, da es den RezipientInnen das simultane (Mit-)Lesen der Textnachrichten ermöglicht.

Die Ausgestaltung der Geschichte wird von den GeräteinhaberInnen insbesondere durch das Vorlesen der Textnachrichten realisiert. Entsprechend handelt es sich um sprachliche Äußerungen im Modus der DRS, in denen die persönliche, räumliche und zeitliche Deixis der Nachrichten aus Sicht der ursprünglichen AutorIn wiedergegeben wird, um die Form und den Inhalt der originalen Nachrichten zu bewahren (vgl. Holt & Clift, 2007, S. 5). Die vorgelesenen Nachrichten werden durch den Einsatz sprachlicher (und seltener verkörperter) Methoden als Handlungen in einer sequenziell und kausal verknüpften Abfolge von Ereignissen konstruiert und somit als narrative Rekonstruktion hervorgebracht. Grundsätzlich geschieht dies über Zitationsmarker, insbesondere bestehend aus Personalpronomen und *verbum dicendi* oder alternativen Formen wie „und er/sie so“ (Golato, 2000). Teilweise werden diese Zitationsmarker indexikalisch an die Wiedergabe von IM-Kommunikation angepasst (bspw. „und dann kommt“), um kenntlich zu machen, dass der Erhalt einer Textnachricht rekonstruiert wird – in vielen Fällen kann dies jedoch nur anhand der *multimodalen Gestalten* inferiert werden (bspw. „und dann hatte die mich letzte Woche gefragt“). Darüber hinaus werden von den ErzählerInnen auch Modulationen der Stimmqualität und Prosodie eingesetzt, um anzuzeigen, wessen Nachrichten vorgelesen werden und um einen Wechsel zwischen unterschiedlichen AutorInnen kenntlich zu machen.

Auch wenn die Materialität der archivierten Textnachrichten die ErzählerInnen in ihrer Wiedergabe auf deren Wortlaut verpflichtet, so handelt es sich keinesfalls um neutrale narrative Rekonstruktionen: Die ErzählerInnen leisten explizite Bewertungsarbeit durch ihre sprachlichen Äußerungen in der Einleitung der Erzählung und während der Elaboration. Darüber hinaus setzten sie para-sprachliche Verfahren (bspw. Lachpartikeln) und prosodische Modulationen ein, um während des Vorlesens implizit ihre Haltung anzuzeigen und den RezipientInnen ein bestimmtes Verständnis der Geschichte naheulegen (vgl. Brown et al., 2018, S. 20; Raclaw et al., 2016, S. 369).

Die Beendigung der narrativen Rekonstruktion wird nicht nur sprachlich, sondern auch durch verkörperte Handlungen von den ErzählerInnen *accountable* gemacht: An dieser sequenziellen Position lösen sie ihren Blick vom Display und schauen zu den RezipientInnen. Außerdem wird

die Bedienung des Smartphones beendet oder zumindest vorübergehend unterbrochen – bspw. indem der TeilnehmerInnen, in dem die RezipientInnen visuellen Zugriff zum Display haben, aufgelöst wird. Die anschließende Post-Sequenz ist durch eine Rückkehr zum *turn-by-turn-talk* gekennzeichnet. In ihr zeigen die RezipientInnen durch ihre Erwidernungen ihr Verständnis der smartphone-assistierten Erzählung an und die Bedeutung der Geschichte für die lokale Interaktion wird gemeinsam ausgehandelt. Auch wenn die Bedienung des Smartphones für diese Sequenz nicht mehr unmittelbar relevant ist, werden oftmals noch einzelne Bediengesten durchgeführt, bevor die Handhabung endgültig beendet (oder eine weitere *mobile-supported sharing activity* eröffnet) wird. Dies führt dazu, dass es in der Post-Sequenz zwar vermehrt zur wechselseitigen Blickverschränkungen zwischen den TeilnehmerInnen kommt, der Blick der SmartphoneinhaberIn aber immer wieder zwischen Gerätedisplay und ihren GesprächspartnerInnen wechselt.

In den Kap. 5.1 und 5.2 wurden der typische sequenzielle Verlauf und die konstitutiven Praktiken *smartphone-assistierter Geschichtenerzählungen* rekonstruiert, in denen die archivierten Textnachrichten als abgeschlossene Interaktionen behandelt werden. Von dieser Variante wurden in Kap. 5.3 noch einmal durch den Empfang von Textnachrichten veranlasste Erzählungen unterschieden, bei denen die *Affordanz* des Smartphones zur orts- und zeitunabhängigen Telekommunikation in den Fokus des Interesses rückt: Anhand dieser Variante zeigt sich anschaulich, dass alltägliche F2F-Interaktionen durch die Präsenz mobiler Endgeräte zu synthetischen Situationen (Knorr Cetina, 2009, S. 61) werden können, wenn die TeilnehmerInnen ihre Smartphones so handhaben, dass sie auf akustische oder visuelle Signale eingehender Benachrichtigungen (oder Telefonate) reagieren können. Durch das Beobachten des Displays orientieren sich die TeilnehmerInnen daran, dass die telekommunikativen *Affordanzen* der mobilen Endgeräte für sie eine gewisse Erwidernungsverantwortung erzeugen. Das Bemerkens des Empfangs einer neuen Nachricht kann von den GeräteinhaberInnen durch ein *noticing* interaktionsrelevant gemacht und anschließend durch das Vorlesen der Textnachricht(en) zu einer *smartphone-assistierten Geschichtenerzählung* ausgebaut werden.

Diese Erzählungen weisen grundsätzlich die gleiche sequenzielle Struktur wie die Varianten der vorangegangenen Kapitel auf und werden durch dieselben konstitutiven Verfahrensweisen hervorgebracht. Sie unterscheiden sich jedoch zum einen dadurch, dass sie durch ein sequenziell passend oder disjunkt positioniertes *noticing* initiiert werden, mit dem die SmartphoneinhaberIn versucht, die Aufmerksamkeit der Anwesenden auf den Erhalt der empfangenen Nachricht zu lenken. Durch die sequenzielle Positionierung und Art der Gestaltung des *noticings* zeigen die GeräteinhaberInnen dabei an, ob sie die erhaltene Nachricht als dringliche

Angelegenheit behandeln und dadurch aktuell im Vollzug befindliche Aktivitäten unterbrechen; oder ob sie die Benachrichtigung als nicht-dringlich auffassen und als Ressource für die Initiierung eines neuen Themas nach Beendigung des vorangegangenen einsetzen. Zum anderen handelt es sich um narrative Rekonstruktionen, mit denen die GeräteinhaberInnen die Anwesenden über Ereignisse an einem anderen Ort informieren. Mit ihnen kann *accountable* gemacht werden, dass eine empfangene Nachricht (potenzielle oder praktische) Folgen für zukünftige Handlungsabläufe hat (vgl. Georgakopoulou, 2008, S. 601). Derartige Geschichten handeln somit von *im Entstehen befindlichen* Ereignissen, deren weitere Entwicklung durch die Erzählung vorhergesagt oder in situ durch das Verfassen einer Antwortnachricht realisiert wird. Wie vorangegangene Studien gezeigt haben, kann die Produktion medial-vermittelter Erwidierungen im Kontext kopräserter Interaktionen dabei auch als kollaborative Aufgabe der Anwesenden gestaltet werden (Brown et al., 2018, S. 15–19; McGregor, 2020, S. 108–113).

Insgesamt stellen die durch das Vorlesen von Textnachrichten gestützten Geschichtenerzählungen ein Verfahren dar, durch das die gespeicherten, (überwiegend) schriftsprachlichen Interaktionen als bedeutungsvolle und erzählwürdige Ereignisse hervorgebracht und ihre soziale Bedeutung gemeinsam ausgehandelt wird. Wie sich anhand der analysierten Einzelfälle gezeigt hat, können – unabhängig von der konkreten Variante der *smartphone-assistierter Geschichtenerzählung* – im Zuge des Gebrauchs der mobilen Endgeräte im Kontext in F2F-Interaktionen immer wieder ‚Probleme‘ auftreten. Zentral für die Aufrechterhaltung eines sinnhaft geordneten Smartphonegebrauchs sind deswegen Praktiken, die es den TeilnehmerInnen erlauben, diese ‚Probleme‘ zu bearbeiten: „Successful collaboration around mobile phones does not mean that there are no problems, but rather that ‚troubles‘ get dealt with as they arise. During co-located phone use, the local order can be maintained with repair“ (McGregor, 2020, S. 149). Wie die TeilnehmerInnen mit dem Auftreten von interaktionalen Problemen umgehen und welcher Methoden sie sich in ihrer Reparatur bedienen, wurde insbesondere anhand der unvorbereitet eröffneten *smartphone-assistierten Geschichtenerzählungen* aufgezeigt (Kap. 5.2). Diese Variante erscheint aufgrund des Fehlens einer Prä-Sequenz anfällig für das Auftreten von Problemen des Verstehens und Missverständnissen in der sequenziellen Abfolge, die interaktiv bearbeitet werden müssen. Werden derartige Probleme interaktionsöffentlich relevant gemacht, wird ihre Reparatur – und damit die (Wieder-)Herstellung eines intersubjektiven Verständnisses – von den TeilnehmerInnen vor der Progression der Geschichtenerzählung priorisiert. Reparaturen von Verstehensproblemen werden von den RezipientInnen meist in zweiter Position durch sprachliche Äußerungen (bspw. Frageworte, Teilwiederholungen oder Verstehensversuche, Egbert, 2009, S. 99–101) initiiert. Auf diese Weise werden Einschubsequenzen

eröffnet, in denen die Reparatur der Problemquelle von den ErzählerInnen durchgeführt wird. Reparaturen von Missverständnissen in der sequenziellen Abfolge werden dagegen von den Story-ProduzentInnen in dritter Position selbst-initiiert realisiert. Sie behandeln eine vorangegangene Äußerung der RezipientInnen als unpassende Erwidern auf die smartphone-assistierte Erzählung und zeigen so an, dass ihre Handlungsabsicht von den RezipientInnen nicht ‚richtig‘ erkannt wurde (vgl. Egbert, 2009, S. 130). In den Analysen der Bearbeitung dieser interaktionalen Probleme konnte in Erweiterung zu McGregor (2020, S. 149) aufgezeigt werden, dass Sprache als eine zentrale, jedoch nicht immer als primäre Ressource für die Reparaturen fungiert. Die materiell verfestigten Handlungsmöglichkeiten der mobilen Endgeräte und integrierten Medieninhalte werden von den TeilnehmerInnen ebenfalls als Ressourcen für Reparaturen eingesetzt: Die gleichen Medieninhalte, die zur Hervorbringung der Geschichtenerzählung eingesetzt werden, können – im Zusammenspiel mit sprachlichen Äußerungen und verkörperten Handlungen – für die Bearbeitung der Problemquellen genutzt werden, indem die SmartphoneinhaberInnen visuellen Zugriff zum Display und damit eine lesende oder betrachtende Aneignung der Medieninhalte ermöglichen.

Die Analysen von Reparaturen sind insbesondere geeignet, die spezifischen Verantwortlichkeiten, die mit der diskursiven Identität der SmartphoneinhaberIn verbunden sind, zu beleuchten. Im Prozess der kollaborativen (Wieder-)Herstellung eines intersubjektiven Verständnisses werden diejenigen normativen Erwartungen der TeilnehmerInnen explizit, an denen sie sich sonst nur implizit orientieren. Die TeilnehmerInnen orientieren sich beobachtbar an der Verantwortung der SmartphoneinhaberInnen, den Gebrauch des Smartphones für andere Anwesende als interaktionsrelevant *accountable* zu machen. Damit die integrierten Medieninhalte als konstitutive Elemente einer Geschichtenerzählung und hinsichtlich ihrer sequenziellen Implikationen erkennbar sind, müssen sie von den GeräteinhaberInnen adressaten- und aktivitätsspezifisch kontextualisiert werden (Keppler, 2019, S. 180). Gelingt dies nicht, werden von den TeilnehmerInnen Reparaturen initiiert und der visuelle Zugriff zu den relevanten Medieninhalten eingefordert.

Das Auftreten von Problemen und ihrer anschließenden interaktivem Bearbeitung ist zunächst ein ‚normales‘ Phänomen von *talk-in-interaction*, in dem Reparaturen als zentrale, intersubjektivitätssichernde Verfahren fungieren (Schegloff, 1992). Im Kontext des Smartphonegebrauchs in F2F-Interaktionen stellt die materiell bedingte *bystander inaccessibility* der mobilen Endgeräte eine spezifische Herausforderung dar, da die Verwendung der Geräte für andere Anwesende durch spezifische Verfahren beobachtbar-und-berichtbar gemacht werden muss. Darüber hinaus scheint außerdem ein Zusammenhang zwischen dem Auftreten interaktionaler Probleme

und den kulturell verbreiteten, persönlichen Nutzungsmustern mobiler Endgeräte zu bestehen: Sowohl in den Analysen der *unvorbereitet eröffneten* als auch der *durch den Empfang einer Textnachricht veranlassten* Geschichtenerzählungen zeigt sich, dass die GeräteinhaberInnen ihre Smartphones als hochgradig vertraute und persönliche Geräte behandeln. Zum einen stellen sie ein *Archiv sozialer Originale* dar, dessen Inhalte Teil der fortlaufenden Soziabilität der InhaberInnen mit relevanten Anderen sind und die ihre Relevanz im Zusammenhang mit anderen (medial-vermittelten und kopräsenten) Interaktionen erhalten (vgl. Brown et al., 2018, S. 2). Dies kann dazu führen dazu, dass die SmartphoneinhaberInnen gespeicherte Textnachrichten im Kontext eines laufenden Gesprächs als bedeutungsvoll erachten und der Gesprächsverlauf sie dazu veranlasst, einen ‚passenden‘ Nachrichtenthread zu thematisieren. Zum anderen ermöglichen die telekommunikativen *Affordanzen* der mobilen Kommunikationsgeräte es, in situ neue Nachrichten (oder Anrufe) eben dieser relevanten Anderen zu empfangen, die – wie gezeigt wurde – ebenfalls die Produktion *smartphone-assistierter Geschichtenerzählungen* veranlassen können.

In beiden Konstellationen ergibt sich für die GeräteinhaberInnen die Herausforderung, die (zuschriebene) Relevanz der archivierten oder in situ empfangenen Textnachrichten auf eine Art und Weise *accountable* zu machen, die das (Un-)Wissen der RezipientInnen über die abwesenden KommunikationspartnerInnen und die vorangegangene Interaktionsgeschichte berücksichtigt. Insbesondere ohne Prä-Sequenz eröffnete Thematisierungen von Textnachrichten erzeugen für andere Anwesende dabei immer wieder Schwierigkeiten, die für die GeräteinhaberInnen scheinbar so ‚offensichtliche‘ Relevanz der Medieninhalte für die lokale Interaktion zu erkennen. Diese Probleme werden von den TeilnehmerInnen mithilfe von Reparaturen bearbeitet, in denen wiederum die materiell-verfestigten Handlungsmöglichkeiten des Smartphones als Ressource eingesetzt werden.

In Kap. 5 wurden die typischen sequenziellen Verlaufsstrukturen, Teilnehmerahmen und methodischen Verfahrensweisen von *smartphone-assistierten Geschichtenerzählungen* herausgearbeitet, die durch das Vorlesen von Textnachrichten gekennzeichnet sind. Dabei wurde nicht systematisch berücksichtigt, dass moderne Kurznachrichtenkommunikation bereits selbst oftmals visuelle Medieninhalte beinhaltet (z. B. GIFs, Videos oder Fotografien). Die Integration visueller Medieninhalte – insbesondere von digitalen Fotografien – wurde deshalb in den Fokus des zweiten analytischen Teils dieser Arbeit gestellt. Digitale Fotografien werden in dieser Arbeit als visuell-materielle Objekte verstanden, die „eine auf ihrer Fläche erkennbare Konfiguration von Objekten und Relationen [zeigen, Anm. D.S.], mit der zugleich eine durch die Art der Darstellung oder Aufnahme spezifische Sicht auf die im Bild sichtbaren Objekte und

Motive eröffnet wird“ (Keppler, 2019, S. 181). Ausgehend von der grundsätzlichen Perspektive, dass Fotografien ihre soziale Bedeutung immer nur im Kontext ihres konkreten, kommunikativen Gebrauchs gewinnen, wurden in Kap. 6 bildgestützte, *smartphone-assistierte Geschichtenerzählungen* (Kap. 6.1) und *smartphone-assistierte Erwidern* auf spaßhafte Provokationen (Kap. 6.2) analysiert.

Für die Hervorbringung *bildgestützter, smartphone-assistierter Geschichtenerzählungen* spielt die *Affordanz* des Smartphones, digitale Fotografien aufnehmen, archivieren und jederzeit wieder abrufen zu können, eine konstitutive Rolle. Die bildgestützten Erzählungen gleichen in ihrer sequenziellen Grundstruktur anderen Formen des smartphone-assistierten Erzählens. Sie weisen im Vergleich zu durch das Vorlesen von Textnachrichten gekennzeichneten Geschichtenerzählungen jedoch auch einige essenzielle Unterschiede auf: Bildgestützte Geschichtenerzählungen können ebenfalls auf unterschiedliche Weise initiiert werden, in den meisten Fällen erfolgt dies durch die methodische Vorbereitung einer *informativen Zeigesequenz*. In der Prä-Sequenz *informativer Zeigesequenzen* wird die Verfügbarkeit einer Erzählung angezeigt und deren Verbindung zum vorangehenden Gespräch verdeutlicht. Dadurch wird für die RezipientInnen ein Slot geschaffen, der es ihnen ermöglicht, sich mit der projizierten Aktivität zu (dis-)alignieren. Es gilt außerdem, den Gebrauch des Smartphones einzuleiten und eine *smartphone-gestützte Zeigesequenz* zu projizieren, mit der eine digitale Fotografie als zeigbares Objekt in die Interaktion eingeführt werden kann. Neben der Konstruktion von ‚Erzählwürdigkeit‘ muss in Prä-Sequenzen bildgestützter Geschichtenerzählungen also auch die ‚Zeigewürdigkeit‘ (vgl. Licoppe, 2017, S. 71) einer digitalen Fotografie etabliert werden. Dies geschieht in den meisten Fällen, indem die Fotografie als visuelle Dokumentation von etwas Neuem im Leben der SmartphoneinhaberInnen gerahmt wird, über das die RezipientInnen informiert werden können. Gemeinsam ist *informativen Zeigesequenzen* somit, dass interaktiv ein Wissensdefizit auf Seiten der RezipientInnen etabliert oder zumindest von der Zeigenden angenommen wird.

In selteneren Fällen erfolgt die Initiierung bildgestützter Erzählungen durch *evokative Zeigesequenzen*, die sich durch einen (fast) vollständigen Verzicht auf die methodische Entwicklung einer Prä-Sequenz auszeichnen. Mit dieser Art der Eröffnung orientieren sich die GeräteinhaberInnen beobachtbar an der Annahme, dass die RezipientInnen über genügend Wissen verfügen, die verkörperte Eröffnung des Zeigens zu erkennen und die Fotografie auf angemessene Weise zu sehen (vgl. Licoppe, 2017, S. 72). Die GeräteinhaberInnen setzen voraus, dass das Design der Zeigesequenz für die RezipientInnen erkennbare Erwartungen über angemessene Erwidern und Kriterien der Beurteilung der gezeigten Fotografie projiziert.

Findet seitens der RezipientInnen eine Alignierung mit der projizierten Zeigesequenz statt, wird die Kernsequenz bildgestützter Geschichtenerzählungen in der Phase des eigentlichen Zeigens entwickelt. Diese weist – unabhängig von der Art der Eröffnung der bildgestützten Erzählung – immer einige typische Merkmale auf. Sie wird durch smartphone-gestützte Zeigesequenzen eröffnet, die grundsätzlich in Paarsequenzen organisiert sind, „von denen einzelne Teile auch rein verkörperter Natur sein können“ (Oloff, 2019, S. 214). Durch diese Paarsequenzen werden die Aufmerksamkeits- und Displayfokussierung koordiniert, sodass der für bildgestützte Geschichtenerzählungen typische Teilnehmerahmen hervorgebracht wird: In diesem wird das Smartphonedisplay so zwischen den TeilnehmerInnen positioniert, dass Zeigende und RezipientInnen visuellen Zugriff zur Fotografie erhalten. Auf diese Weise wird sichergestellt, dass die während des Zeigens produzierten sprachlichen Äußerungen der SmartphoneinhaberInnen durch ihren Bezug zur digitalen Fotografie für die RezipientInnen sinnhaft verständlich werden können.

Die sprachlichen Äußerungen der SmartphoneinhaberIn dienen dazu, die Fotografie für die RezipientInnen in den Kontext der Erzählung einzuordnen. Dabei werden die visuellen Medieninhalte als Elemente einer Sequenz von Ereignissen oder Handlungen dargestellt, die zeitlich und kausal miteinander verknüpft sind. Abhängig von der relationalen Konfiguration von Objekten auf ihrer Bildeoberfläche können unterschiedliche Fotografien dabei als visuelle Repräsentationen optischer Eindrücke oder als Momentaufnahmen verkörperter Handlungen fungieren, durch die das ‚vergangene Ich‘ der ErzählerIn als Figur der Erzählwelt animiert wird (vgl. Couper-Kuhlen, 2007, S. 81). In bildgestützten, *smartphone-assistierten Geschichtenerzählungen* erfolgt die narrative Rekonstruktion eines Handlungs- oder Ereignisablaufs also nicht ausschließlich – und teilweise nicht einmal hauptsächlich – verbal, denn „Bilder lassen sich . . . nicht ‚einfach‘ in Sprache ‚übersetzen‘: sie eröffnen . . . eine spezifische Sicht auf das visuell repräsentierte Motiv“ (vgl. Keppler, 2019, S. 181). Die sprachlichen Äußerungen der ErzählerInnen dienen dazu, den RezipientInnen eine gewisse Sichtweise nahezu legen, in der die Fotografien als konstitutive Elemente einer Geschichtenerzählungen *accountable* werden. Gelingt dies, wird eine durch und durch multimodale, narrative Rekonstruktion hervorgebracht, in denen sich sprachliche Äußerungen und digitale Fotografie wechselseitig ihren Sinn verleihen. Auch in bildgestützten Varianten findet die Würdigung der Erzählung und die Aushandlung ihres sozialen Sinns anschließend in der Post-Sequenz statt. Diese dient insbesondere der kommunikativen Aneignung dessen, was sich für die TeilnehmerInnen in den visuell repräsentierten Motiven der Fotografien zeigt (vgl. Keppler, 2019, S. 181).

Die Integration visueller Medieninhalte fügt F2F-Interaktionen eine zusätzliche Interpretationsebene hinzu, die eindeutige Konsequenzen für ihren Verlauf hat (vgl. Raclaw et al., 2016, S. 377). Die Relevanz der kommunikativen Aushandlung, welche Bedeutung die visuellen Medieninhalte für die TeilnehmerInnen besitzen, zeigt sich besonders explizit, wenn die Integration digitaler Fotografien in sequenziellen Kontexten untersucht wird, die von Disalignierung und Disaffiliation gekennzeichnet sind. Wie in Kap. 6.2 gezeigt wurde, kommt die *Affordanz* des Smartphones, spontan archivierte oder online-abrufbare Fotografien in die Interaktion einzuspielen, nicht nur im Hauptteil von Geschichtenerzählungen zum Einsatz. In diesem Kapitel wurden abschließend bildgestützte, *smartphone-assistierte Erwidernngen auf Teasings* untersucht. Bei spaßhaften Provokationen handelt es sich um Äußerungen, mit denen die RezipientInnen ein Element der Erzählung als *abweichend*, d. h. nicht ihren Normalitätserwartungen entsprechend, interpretieren. Durch die Produktion eines *Teasings* zeigt die RezipientIn sowohl ihre Disalignierung mit der im Vollzug befindlichen Geschichtenerzählung als auch ihre Disaffiliation mit der Haltung der ErzählerIn an. Auf diese Weise wird diese (und evtl. anwesende Andere) dazu veranlasst, sich zu der Provokation zu positionieren und es kommt zur Hervorbringung einer *Teasing*-Sequenz. Mit der Analyse bildgestützter Erwidernngen in diesem sequenziellen Kontext wurde zum einen die grundsätzliche Herausforderung der Smartphoneintegration in als *turn-by-turn-talk* organisierten Sequenzen beleuchtet. Zum anderen konnte so rekonstruiert werden, inwiefern die RezipientInnen durch die Integration digitaler Fotografien zur Alignierung und Affiliation mit der SmartphoneinhaberIn geleitet werden können (vgl. Raclaw et al., 2016, S. 377).

In *Teasing*-Sequenzen operiert grundsätzlich eine Präferenz für Progressivität. Das bedeutet, dass die TeilnehmerInnen im nächsten *turn* realisierbare Erwidernngen solchen Handlungen vorziehen, die erst einer ‚Vorarbeit‘ bedürfen – so wie *smartphone-assistierte, bild-gestützte Erwidernngen*, die zunächst die Suche nach einer digitalen Fotografie erforderlich machen. Diese Präferenz lässt sich damit erklären, dass *Teasing*-Sequenzen aufgrund ihres spaßhaften Charakters zu einer ‚geselligen Atmosphäre‘ beitragen (vgl. Günthner, 2012, S. 198), in der sich die ProvokateurInnen auf Kosten der RezipientIn lustig machen und von dieser erwartet wird, die Entwicklung der Sequenz zuzulassen oder selbst fortzusetzen. Deshalb wird Eröffnung bildgestützter, *smartphone-assistierter Erwidernngen* von der GeräteinhaberIn erst nach anderen Beiträgen zu dieser Aktivität initiiert und von anderen Anwesenden als ‚angemessene Erwidernngen‘ anerkannt. Dies verdeutlicht grundsätzlich, dass aufgrund der (potenziell) unterschiedlichen Zeitlichkeiten von Suchprozess und sprachlicher Interaktion von allen Beteiligten kollaborativ ein sequenzieller Kontext geschaffen werden muss, damit die

Smartphoneintegration in F2F-Interaktionen an einer ‚passenden‘ sequenziellen Position erfolgen kann (vgl. Brown et al., 2015, S. 511).

Die bildgestützten, *smartphone-assistierten Erwidernngen* in *Teasing*-Sequenzen werden ebenfalls durch *evokative* oder *informative Zeigesequenzen* realisiert. Anhand des anschließenden Gesprächsverlaufs zeigt sich, dass *smartphone-gestützte Zeigesequenzen* eine etablierte Methode darstellen, andere Anwesende zur Alignierung zu bewegen (vgl. Raclaw et al., 2016, S. 364): Die RezipientInnen tragen in allen Fällen zur Hervorbringung der Aufmerksamkeits- und Displayfokussierung bei und richten sich zunächst an der durch die Betrachtungsanweisung der Zeigenden angelegten Sichtweise auf die Fotografie aus (Oloff, 2019, S. 214–215). Dies lässt sich damit erklären, dass das Zeigen einer Fotografie anschließend eine Erwiderung durch die RezipientInnen konditionell relevant macht, die konstitutiv auf ihre Betrachtung angewiesen ist. „[D]a der visuelle Zugang zum Display und die Möglichkeit, den Inhalt zu bewerten, einige Zeit in Anspruch nehmen können“ (Oloff, 2019, S. 215), führen *smartphone-assistierte Erwidernngen* auf *spaßhafte Provokationen* dazu, dass die Zeigenden zunächst eine Chance erhalten, die *Teasing*-Sequenz auf die von ihnen angelegte Weise weiterzuentwickeln.

Die formale Ausrichtung der RezipientInnen an der Zeigesequenz ist jedoch nicht gleichbedeutend damit, dass es im weiteren Verlauf auch zur Affiliation – zur Gleichgerichtetheit in der Haltung zur integrierten Fotografie – kommt. Bildgestützte, *smartphone-assistierte Erwidernngen* besitzen durchaus das Potenzial, die RezipientInnen zu Entwicklung einer übereinstimmenden Haltung mit der ErzählerIn zu bewegen. Die SmartphoneinhaberInnen verfügen jedoch keinesfalls über eine Deutungshoheit über die integrierten Bilder: Anhand der untersuchten Fälle zeigt sich, dass „das, was für die Gesprächsteilnehmer in einem Foto zu sehen ist, stets auch davon abhängt, was sie in ihm sehen, d. h. wie sie es sich – sei es im Konsens, sei es im Dissens – im Gang der Konversation aneignen“ (Keppler, 2019, S. 185). Im Anschluss an die Integration kommt es zur kommunikativen Aneignung der digitalen Fotografie, in der die TeilnehmerInnen aushandeln, was für sie in dem Bild zu sehen und wie der fotografisch repräsentierte Referent zu bewerten ist.

## 7.2 Veränderungen des alltäglichen Erzählens durch die Integration mobiler Endgeräte

Aufbauend auf dieser zusammenfassenden Darstellung der multimodalen Praktiken der Hervorbringung *smartphone-assistierter Geschichtenerzählungen* möchte ich nun zur eingangs gestellten Frage zurückkehren, wie die Handlungsangebote mobiler Kommunikationstechnologien „die *Formen* der alltäglichen Kommunikation *modifizieren* oder gar grundlegend *verändern*“ (Keppler, 2013, S. 92, Herv. i. O.). Für die vorliegende Arbeit bedeutet dies, anhand der empirischen Analysen zu reflektieren, welche „procedural consequentiality of mediation“ (Arminen et al., 2016, S. 305) sich für alltägliche Geschichtenerzählung aufgrund der beobachtbaren Orientierung an den *Affordanzen* der mobilen Endgeräte ergibt: Im Folgenden werden deshalb die Konsequenzen des Mediengebrauchs auf die sequenzielle und zeitliche Entwicklung der interaktiven Hervorbringung der Erzählungen rekonstruiert. Diese abschließende Reflektion wird dabei zunächst 1.) die spezifische Multimodalität *smartphone-assistierter Geschichtenerzählungen* fokussieren, bevor 2.) der Zusammenhang zwischen Smartphonegebrauch und der Progressivität der Erzählungen sowie 3.) abschließend die visuelle Materialität konstitutiver Elemente von *smartphone-assistierten, narrativen Rekonstruktionen* diskutiert werden.

1.) Das „Erzählen in zeitlich-räumlicher Ko-Präsenz ist eine grundlegend körperlich und räumlich situierte Aktivität“ (König & Oloff, 2018a, S. 221). In *smartphone-assistierten Geschichtenerzählungen* erfolgt die Hervorbringung dieser Aktivität nun nicht nur über sprachliche Äußerungen und verkörperte Handlungen der TeilnehmerInnen, sondern auch konstitutiv durch die Handhabung des mobilen Endgeräts und die Integration von – (überwiegend) schriftsprachlichen oder (audio-)visuellen – Medieninhalten. Das bedeutet, dass sie durch komplexe *multimodale Gestalten* realisiert werden, in denen sich die unterschiedlichen, von den TeilnehmerInnen mobilisierten Ressourcen wechselseitig ihren Sinn verleihen. Damit das Zusammenspiel der unterschiedlichen Modalitäten intersubjektiv verständlich werden kann, müssen die für *smartphone-assistierte Geschichtenerzählungen* spezifischen Teilnahmerahmen hervorgebracht werden. Diese unterscheiden sich voneinander, je nachdem, ob Textnachrichten ausschließlich vorgelesen oder simultan lesend angeeignet werden bzw. den RezipientInnen ein visueller Medieninhalt gezeigt wird. In jedem Fall sind es mindestens die SmartphoneinhaberInnen, oftmals aber auch die RezipientInnen, die im Vollzug der Erzählung visuell das Smartphonedisplay fokussieren.

Aus diesen konstitutiven Teilnahmerahmen ergibt sich eine zentrale Differenz zu ‚nicht-meditatisierten‘ Geschichtenerzählungen: Die SprecherInnen in F2F-Interaktionen orientieren sich

normalerweise an der *Regel* „*When a speaker gazes at a recipient he should make eye contact with that recipient*“ (Goodwin, 1979, S. 106, Herv. i. O.). Wenn SprecherInnen eine RezipientIn adressieren, die sie nicht anschaut, kommt es oftmals zur Produktion von Stockungen, Pausen oder Neustarts, bis die visuelle Aufmerksamkeit und damit die Zuhörerschaft der RezipientIn gesichert ist (vgl. Goodwin, 1981, S. 60–77). In den Prä-Sequenzen und Hauptteilen *smartphone-assistierter Geschichtenerzählungen* zeigt sich jedoch, dass diese Regel während des simultanen Smartphonegebrauchs ausgesetzt oder zumindest erheblich gelockert wird. Es kommt wesentlich seltener zu wechselseitigen Blickverschränkungen, da die visuelle Fokussierung des Smartphonedisplays eine Bedingung für die konstitutiven Praktiken *smartphone-assistierten Geschichtenerzählungen* darstellt – diese umfassen u. a. das Vorlesen von Textnachrichten, ihre lesende Aneignung oder die Betrachtung (audio-)visueller Medieninhalte. Die RezipientInnen werden von der ErzählerIn jedoch immer wieder an sequenziellen Schlüsselpositionen angeschaut, bspw. um ihr Verständnis der impliziten Bedeutung einer Äußerung zu ‚checken‘ oder das Timing einer Betrachtungsanweisung mit der Displayfokussierung koordinieren zu können.

Mit der Hervorbringung von Teilnehmerahmen, die eine visuelle Fokussierung des Smartphonedisplays ermöglichen, gehen also hinsichtlich der Blickführung der TeilnehmerInnen eindeutige Differenzen zu nicht-mediatisierten Geschichtenerzählungen einher. Darüber hinaus hat die notwendige physische Manipulation des mobilen Endgeräts in Rahmen *smartphone-assistierter Geschichtenerzählungen* Konsequenzen für den Einsatz von Gesten durch die ErzählerInnen, denn „[g]leichzeitige manuelle Tätigkeiten erschweren die gestische Ausgestaltung der Erzählung . . . sowie die Aufrechterhaltung eines festen Teilnehmerrahmens“ (König & Oloff, 2018b, S. 303): In den meisten Fällen verzichten die ErzählerInnen auf eine gestische Ausgestaltung oder nutzen lediglich Gesten, die mit einer Hand realisiert werden können, sodass der visuelle Zugriff zum Smartphone aufrechterhalten werden kann. Anhand des Falles „understatement“ (vgl. Kap. 6.2), in dem der Geräteinhaber während einer Zeigesequenz mit beiden Händen gestikuliert, zeigt sich, dass dies zu einer Unterbrechung des Zeigens und einer dynamischen Veränderung des Teilnehmerahmens führt.

Mit der spezifischen Multimodalität *smartphone-assistierter Geschichtenerzählungen* und ihren konstitutiven Teilnehmerahmen sind darüber hinaus grundlegende Veränderungen in den Beteiligungsrollen der TeilnehmerInnen verbunden. Diese können mit den ‚kanonischen‘ Bezeichnungen ErzählerIn und ZuhörerIn (vgl. bspw. Quasthoff, 2001, S. 1300) nur unzureichend erfasst werden. Wenn Erzählen in Interaktion grundlegend als körperlich und räumlich situierte Aktivität verstanden wird, bedeutet dies ganz grundsätzlich, dass die ZuhörerInnen einer

Erzählung immer „gleichzeitig auch ZuschauerInnen und Mitakteure einer audio-visuellen . . . Inszenierung“ (König & Oloff, 2018b, S. 303) der narrativen Rekonstruktion sind. Für *smartphone-assistierte Geschichtenerzählungen* werden diese Beteiligungsrollen nun noch einmal komplexer: Die RezipientInnen tragen konstitutiv zur Hervorbringung einer intersubjektiv verständlichen, smartphone-assistierten Erzählung bei, indem sie sich als ZuhörerInnen, BeobachterInnen des Smartphonegebrauchs und – in einigen Fällen – als LeserInnen der archivierten Textnachrichten oder als BetrachterInnen der Fotografien positionieren. Die Story-ProduzentInnen wiederum agieren in und durch ihre Äußerungen, verkörperten Handlungen und die Manipulation der mobilen Endgeräte als ErzählerInnen, SmartphonebedienerInnen und – insbesondere im Hauptteil der Geschichte – als VorleserInnen der archivierten Textnachrichten oder als Zeigende visueller Medieninhalte.

2.) Die Handhabung des Mobiltelefons während *smartphone-assistierter* Geschichtenerzählungen wirkt sich jedoch nicht nur auf die Beteiligungsrollen der TeilnehmerInnen, ihre Blickführung oder den Einsatz von Gestik aus. Eine zentrale Konsequenz ihrer spezifischen Multimodalität ist auch, dass es zu einer komplexen Verschränkung des Smartphonegebrauchs mit der Progression der Geschichtenerzählung kommt. Sowohl auf Ebene der Turnproduktion als auch hinsichtlich der sequenziellen Entwicklung der einzelnen Phasen der Erzählung ist der Fortschritt der narrativen Rekonstruktion mit der physischen Manipulation der mobilen Endgeräte verknüpft. Hinsichtlich der Realisierung sprachlicher Äußerungen der GeräteinhaberInnen während der simultanen Bedienung des Smartphones zeigt sich, dass diese an der linguistischen Gesprächsoberfläche immer wieder durch Stockungen, Mikropausen, Selbstabbrüche und -korrekturen gekennzeichnet sind. Mithilfe einer multimodalen Analyse lassen sich diese vermeintlichen Formulierungsschwierigkeiten funktional als Synchronisierung mit der Bedienung des Smartphones interpretieren. Dabei kann insbesondere der Neustart sprachlicher Äußerungen als methodisches Verfahren identifiziert werden, mit dem die unterschiedlichen Zeitlichkeiten von Äußerungen und physischer Manipulation des Smartphones auf eine Weise miteinander synchronisiert werden, dass sich sprachliche Äußerungen, Bediengesten und die visuelle Oberfläche des Displays wechselseitig ihren Sinn verleihen können (vgl. Goodwin, 2015, S. 418–419) und als *multimodale Gestalten accountable* werden.

Neben der Turnproduktion muss auch die sequenzielle Entwicklung der gesamten Geschichtenerzählung mit dem Gebrauch des mobilen Endgeräts koordiniert werden. Die Realisierung von Handlungen, die für die Progression *smartphone-assistierter Geschichtenerzählungen* konstitutiv sind, sind oftmals auf den visuellen Zugriff zum relevant gewordenen Medieninhalt angewiesen – sei es, damit die ErzählerIn eine Ereignisabfolge durch Praktiken des Vorlesens

rekonstruieren kann oder damit die RezipientInnen durch die Betrachtung einer digitalen Fotografie eine angemessene Erwiderung produzieren können. Dies setzt voraus, dass die entsprechenden Medieninhalte im lokalen Speicher des Smartphones oder über eine Online-Suche zunächst gesucht, gefunden, und anschließend zugänglich gemacht wurden (Brown et al., 2015, S. 510; Suderland, 2019, S. 91). Die für die Integration notwendigen Suchprozesse können sowohl der sprachlichen Initiierung der smartphone-assistierten Erzählung vorgeschaltet werden als auch im Rahmen der Prä-Sequenz oder zwischen Prä-Sequenz und Hauptteil der Geschichte stattfinden.

Die GeräteinhaberInnen stehen dabei vor der Aufgabe, die Teilnahme an der sprachlichen Interaktion und die smartphone-gestützten Suchprozesse als *intrapersonale Multiaktivität* zu organisieren. Es gelingt ihnen dabei immer wieder, beide Aktivitäten so miteinander zu koordinieren, dass sie parallel zueinander und ohne beobachtbare Störungen ablaufen („parallel order“, Mondada, 2014a, S. 69) oder die Bedienung des Smartphones so in die sprachliche Interaktion einzubetten („embedded order“, ebd., S. 69), dass die Progression beider Aktivitäten wechselseitig aneinander angepasst wird. Wie sehr die Entwicklung der Geschichtenerzählungen mit der Manipulation der mobilen Endgeräte zusammenhängt, zeigt sich am offensichtlichsten, wenn die beiden Aktivitäten in einer sich wechselseitig ausschließenden Ordnung („exclusive order“, ebd. S. 69) organisiert werden und es zur Hervorbringung von Einschubsequenzen kommt, die eigens der Durchführung eines smartphone-gestützten Suchprozesses dienen. Die Weiterentwicklung der Geschichtenerzählung wird dabei so lange pausiert, bis der Suchprozess abgeschlossen und relevante Medieninhalte gefunden wurden.

In derartigen Phasen der Suche wird der Abschluss des Suchprozesses also als Voraussetzung für die Fortsetzung der Geschichtenerzählung behandelt. Die GeräteinhaberInnen bedienen sich unterschiedlicher multimodaler Methoden, um den im Vollzug befindlichen Suchprozess *accountable* zu machen und die Teilnahmemöglichkeiten der RezipientInnen zu managen. Dabei lassen sich *suchbegleitende Kommentierungen* von Praktiken der *kollaborativen Suche* unterscheiden. Die Praktiken der *suchbegleitenden Kommentierung* dienen dazu, den Verlauf der physischen Manipulation des Smartphones (und die Geschehnisse auf dem oftmals nicht einsehbaren Gerätedisplay) für andere Anwesende durch Beschreibungen und Einschätzungen sinnhaft verständlich zu machen. Sie ermöglichen für die RezipientInnen jedoch keine Teilnahmemöglichkeiten an der Suche. Dies wird durch Praktiken der *kollaborativen Suche* ermöglicht, in denen die SmartphoneinhaberInnen den Anwesenden Fragen stellen und ihre Antworten bei der Formulierung von Suchbegriffen oder Auswahl von Suchergebnissen berücksichtigen (Brown et al., 2015, S. 512–513; Suderland, 2019, S. 102–111). Auch in nicht-

mediatisierten Interaktionen kann es durchaus zu Wortsuchsequenzen (Schegloff et al., 1977, S. 363) kommen, die von der aktuellen SprecherIn oder kollaborativ von den InteraktionsteilnehmerInnen (vgl. Goodwin, 1987) bearbeitet werden müssen, bevor eine Aktivität fortgesetzt werden kann. Im Fall von *smartphone-assistierten Geschichtenerzählungen* handelt es sich dabei jedoch um ein Multiaktivitätssetting, indem vor allem die GeräteinhaberIn vor der praktischen Aufgabe steht, die physische Manipulation des Gerätes zum Zweck der Suche durchzuführen, bevor die multimodale Erzählung fortgesetzt werden kann.

3.) Bei der spezifischen Multimodalität und der Verschränkung des Smartphonegebrauchs mit der Progression einer gemeinsamen Aktivität handelt es sich also um Differenzen, die smartphone-assistierte von nicht-mediatisierten Aktivitäten in F2F-Interaktion unterscheiden. Diese Unterschiede sind jedoch zunächst nicht spezifisch für narrative Rekonstruktionen als Lösungen für das kommunikative Problem der Flüchtigkeit sozialer Ereignisse. Eine spezifische Veränderung für das alltägliche Erzählen ergibt sich jedoch durch die materielle Visualität der integrierten Medieninhalte. Dieses Argument werde ich im Folgenden zunächst für Textnachrichten und anschließend für digitale Fotografien explizieren.

Werden in mündlich-sprachlichen Geschichtenerzählungen die vergangenen Äußerungen einer Interaktion im Modus der DRS rekonstruiert, handelt es sich immer „um Stilisierungen von Redewiedergaben . . . , deren Authentizität kaum ausmachbar ist“ (Kotthoff, 2018, S. 237, FN 11). Durch die Verwendung von DRS wird zwar beansprucht, die vergangene Äußerung in ihrem Inhalt und ihrer Realisierungsform wiederzugeben, da die ursprüngliche Situation vergangen ist, lässt sich dieser Authentizitätsanspruch jedoch nicht überprüfen. Werden dagegen archivierte Textnachrichten – d.h. Ausschnitte persistenter, nicht-flüchtiger Interaktionen – durch Praktiken des Vorlesens integriert, fügt dies der smartphone-assistierten Erzählung eine zusätzliche Dimension epistemischer Autorität hinzu (vgl. Raclaw et al., 2016, S. 377). Dies liegt daran, dass die Verankerung des Wortlauts in der Materialität der Textnachrichten für die Authentizität der Wiedergabe verbürgt und diese potenziell für alle Anwesenden überprüfbar macht. Die *Affordanzen* des Smartphones werden oftmals auch genutzt, um den RezipientInnen visuellen Zugriff zum Display zu gewähren, sodass sie die Textnachrichten simultan zur Wiedergabe (mit-)lesen können. Dieser erweiterte epistemischen Zugriff zu einem bewertbaren Referenten erlaubt den RezipientInnen (vermeintlich), zu einer ‚eigenständigen‘ Interpretation der archivierten Interaktion zu gelangen (Holt & Clift, 2007, S. 12), und kann deshalb eingesetzt werden, um sie zur Alignierung und Affiliation zu leiten (Raclaw et al., 2016, S. 377). Die durch die Materialität der archivierten Threads ermöglichte lesende Aneignung verweist somit auf zwei zentrale Differenzen zu narrativen Rekonstruktionen flüchtiger sozialer Situationen:

Während in den Nachrichtenthreads vergangene Interaktionen in all ihren Details bewahrt werden, können mündlich-sprachliche Interaktionen nicht wiederholt betrachtet bzw. erlebt werden. Deshalb können sie (normalerweise) auch nur durch den Einsatz sprachlicher und verkörperter Ressourcen von den ErzählerInnen rekonstruiert werden – im Gegensatz dazu ermöglichen die archivierten, schriftsprachlichen Interaktionen auch eine narrative Rekonstruktion, die durch Praktiken des Vorlesens und (Mit-)Lesens gestützt werden kann.

Geschichtenerzählungen werden in der KA üblicherweise als verbale Rekonstruktion (vgl. Gülich & Hausendorf, 2000, S. 373) verstanden, die von den ErzählerInnen durch die Produktion eines *big packages* sprachlicher Äußerungen hervorgebracht werden. Vergleicht man den Hauptteil der durch das *Vorlesen von Textnachrichten* gestützten Geschichtenerzählungen mit dem von mündlich-sprachlichen Erzählungen, so finden sich an der linguistischen Gesprächsoberfläche aufgrund der sprachlichen Äußerungen im Modus der DRS einige Gemeinsamkeiten. Betrachtet man jedoch die Kernsequenz bildgestützter, *smartphone-assistierter Geschichtenerzählungen*, so zeigen sich grundlegendere Differenzen zu überwiegend sprachlich konstruierten Geschichten. Zum einen handelt es sich um eine multimodale Praktik, die konstitutiv auf die Verbindung sprachlicher Äußerungen und dem semiotischen Feld der Fotografie(n) angewiesen ist. Die Rekonstruktion des Handlungsablaufs erfolgt nicht nur – und teilweise nicht einmal überwiegend – verbal, da die Fotografie(n) als visuelle Repräsentation(en) von Handlungen oder Ereignissen der Geschichte fungieren. Zum anderen sind diese bildgestützten Erzählungen nicht an die lineare Sequenzialität sprachlicher Interaktion gebunden (vgl. Mondada, 2016, S. 341). Aufgrund der unterschiedlichen Zeitlichkeiten des Erzählens und Zeigens sowie des Hörens und Sehens können die Elemente des rekonstruierten Handlungsablaufs nicht mehr nur nacheinander, sondern auch zeitgleich und in Überlappung zueinander in die Erzählung eingeführt werden. Auf diese Weise werden sie erst im Vollzug der Interaktion und durch die Produktion multimodaler semiotischer Einheiten (Aaltonen et al., 2014, S. 139) in eine sinnhafte Ordnung und chronologische Abfolge gebracht – wobei die Fotografien als visuelle Repräsentationen von Elementen der Erzählungen einer (potenziell) beliebig langen oder erneuten Betrachtung unterzogen werden können.

Trotz ihrer unterschiedlichen Medialität ist archivierten Textnachrichten und digitalen Fotografien eine Eigenschaft gemein: Sie besitzen – unabhängig vom konkreten Kontext ihres Gebrauchs – stabile, unveränderliche Eigenschaften (vgl. C. Goodwin, 2000, S. 165). Für beide lässt sich außerdem argumentieren, dass sie Formen der registrierenden Konservierung (Bergmann, 1985) vergangener Interaktionen oder Ereignisse darstellen: In Nachrichtenthreads wird vergangene IM-Kommunikation in allen für die InteraktionsteilnehmerInnen verfügbaren

Details gespeichert. (Digitale) Fotografien können dagegen als Momentaufnahmen eines vergangenen Ereignisses aus einer spezifischen Kameraperspektive verstanden werden (vgl. Larsen, 2005, S. 418). Es handelt sich also um unterschiedliche Formen einer technischen Aufzeichnung – und damit selbst bereits um Lösungen des kommunikativen Problems der Flüchtigkeit sozialer Situationen, die mithilfe der *Affordanzen* des Smartphones produziert, gespeichert und rezipiert werden können. Obwohl das Smartphone damit als Archiv (unterschiedlichster Formen) registrierender Konservierungen sozialer Situationen verstanden werden kann, zeigt sich anhand der Analysen dieser Arbeit, dass dadurch das alltägliche Geschichtenerzählen keinesfalls obsolet wird: Die Integration von Textnachrichten oder digitalen Fotografien in smartphone-assistierten Erzählungen ist nicht nur eine Aktivität ‚in eigenem Recht‘, sondern eine zentrale kommunikative Praxis, durch die der soziale Sinn dieser dokumentierten Ereignisse etabliert und ausgehandelt wird (vgl. Brown et al., 2018, S. 21). Durch ihre Einbettung in narrative Rekonstruktionen werden die gespeicherten Threads und digitalen Fotografien als konstitutive Elemente erzählwürdiger Ereignisse hervorgebracht und unterschiedlichste soziale Handlungen in F2F-Interaktionen realisiert. Dies zeigt, dass das Geschichtenerzählen in Interaktion keinesfalls seine Bedeutung für die gemeinsame Ausdeutung des sozialen – und heutzutage oftmals medial-vermittelten – Erlebens verloren hat. Im Gegenteil hat sich mit *smartphone-assistierten Geschichtenerzählungen* eine modifizierte Form dieser multimodalen Praxis entwickelt, in und durch die sowohl die medialen Erfahrungen unseres Alltags thematisiert werden, als auch unser zunehmend vom Gebrauch mobiler Kommunikationstechnologien durchdrungener Alltag konstituiert wird.

### 7.3 Resümee

Anhand des Datenkorpus dieser Arbeit und den Befunden der stetig wachsenden Anzahl an Studien der empirischen „Handyforschung“ wird zunehmend offensichtlich, dass sich Smartphones zu mobilen Kommunikationstechnologien entwickelt haben, die alle Aspekte unseres alltäglichen Lebens durchdringen, sodass uns die Präsenz und der Gebrauch der mobilen Endgeräte in sozialen Situationen vertraut, fast schon selbstverständlich geworden ist. Entsprechend kann argumentiert werden, dass „practices of mobile phone use have become constitutive to local order in everyday life and that the study of these practices is key to understanding what social order is now like“ (McGregor, 2020, S. 160). Dieser Status von mobilen Kommunikationstechnologien fordert die wissenschaftliche Forschung auf, Perspektiven zu entwickeln, die dieses Phänomen befremden, um uns für die Details des Smartphonegebrauchs und dessen Auswirkungen auf alltägliche F2F-Interaktionen zu sensibilisieren. Die

multimodale KA bietet einen entsprechenden methodologischen Zugang. Aufbauend auf Jahrzehnten der Erforschung der Organisationsprinzipien von (sprachlichen) Interaktionen, lassen sich mithilfe der multimodalen KA die Konsequenzen des Mediengebrauchs auf die sequenzielle und zeitliche Entwicklung der interaktiven Hervorbringung zentraler Aktivitäten alltäglicher F2F-Interaktionen identifizieren (vgl. Arminen et al., 2016, S. 305). Diese Herangehensweise besitzt das Potenzial, die Prozesse der Mediatisierung auf der Mikro-Ebene alltäglicher Unterhaltungen empirisch greifbar zu machen.

Bereits jetzt lässt sich aufbauend auf den Ergebnisse konversationsanalytischer Arbeiten den Befunden, dass die Präsenz und Nutzung von Smartphones die TeilnehmerInnen hauptsächlich von der Interaktion mit Anwesenden ablenken und das Gesprächsgeschehen stören (Chotpitayasunondh & Douglas, 2016, 2018; Roberts & David, 2016), zwei Argumente entgegensetzen: Zum einen lässt sich eindeutig aufzeigen, dass die TeilnehmerInnen bereits über etablierte Ethno-Methoden verfügen, den Smartphonegebrauch sinnhaft und intersubjektiv verständlich zu machen. Der Einsatz der mobilen Endgeräte wird dabei sensibel auf den sequenziellen Kontext der Interaktionen abgestimmt, sowie adressaten- und aktivitätsspezifisch kontextualisiert. Zum anderen wird durch die Ergebnisse dieser Arbeit und anderer konversationsanalytischer Studien offenkundig, dass Smartphones in sozialen Situationen längst nicht nur für interpersonale Fernkommunikation genutzt werden. Durch die kreative Aneignung der Handlungsmöglichkeiten der mobilen Endgeräte haben sich Nutzungspraktiken entwickelt, in denen Smartphones als Ressource für gemeinsame, kollaborative Aktivitäten der anwesenden InteraktionspartnerInnen genutzt werden, die das Gesprächsgeschehen beleben und erweitern.

Die Analyse *smartphone-assistierter Geschichtenerzählungen* zeigt dabei exemplarisch, dass die Praktiken des kollaborativen Smartphonegebrauchs und alltägliche Unterhaltungen in einem wechselseitig verschränkten Verhältnis zueinanderstehen: Kollaborative Formen des Gebrauchs mobiler Endgeräte können das Interaktionsgeschehens beleben und den Gesprächen zusätzliche Ebenen der Interpretation hinzufügen. Die direkten mündlichen Gespräche bieten wiederum eine Unterstützung für die Aushandlung des sozialen Sinns der eingespielten Medieninhalte und ein Forum, in dem angemessene Praktiken des Smartphonegebrauchs ausgehandelt werden. Diese Dynamiken zeigen, dass sich „[d]er Wandel der Verständigungsverhältnisse, . . . nach wie vor wesentlich innerhalb der alltäglichen Face-to-Face-Kommunikation [vollzieht], in der diese zwar ein zunehmend verändertes Gesicht gewinnt, ohne aber damit ihre Bedeutung für die Erzeugung individueller wie gemeinschaftlicher Orientierungen zu verlieren“ (Keppler, 2013, S. 101). Zukünftig gilt es, diese Erforschung konkreter Situationen des Mediengebrauchs fortzusetzen, um den kulturpessimistischen Generalisierungen, die in der

zunehmenden Verbreitung mobiler Kommunikationstechnologien eine Gefahr für unser soziales Miteinander vermuten, ein empirisch fundiertes Verständnis der vielfältigen, oftmals unvorhergesehen und sich – nicht zuletzt mit dem technischen Fortschritt – stets weiterentwickelnden Praktiken des Smartphonegebrauchs in alltäglichen Gesprächen entgegenzusetzen.

## Literaturverzeichnis

- Aagaard, J. (2020). Digital Akrasia: A Qualitative Study of Phubbing. *AI & Society*, 35, 237–244. <https://doi.org/10.1007/s00146-019-00876-0>
- Aaltonen, T., Arminen, I., & Raudaskoski, S. (2014). Photo Sharing as a Joint Activity Between an Aphasic Speaker and Others. In M. Nevile, P. Haddington, T. Heinemann, & M. Rauniomaa (Hrsg.), *Interacting with Objects* (S. 125–144). John Benjamins.
- Abels, H. (2013). Ethnomethodologie. In G. Kneer & M. Schroer (Hrsg.), *Handbuch Soziologische Theorien* (S. 87–110). VS Verlag für Sozialwissenschaften.
- Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying Mediated Interaction. *Research on Language and Social Interaction*, 49(4), 290–309. <https://doi.org/10.1080/08351813.2016.1234614>
- Arndt, O., Plant, S., Burkart, G., Johnson, D., & Kenichi, F. (2010). *Welt in der Hand: Zur globalen Alltagskultur des Mobiltelefons*. Spector Books.
- Ayaß, R. (2014). Using Media as Involvement Shields. *Journal of Pragmatics*, 72, 5–17. <https://doi.org/10.1016/j.pragma.2014.02.003>
- Ayaß, R. (2015). Doing Data: The Status of Transcripts in Conversation Analysis. *Discourse Studies*, 17(5), 505–528. <https://doi.org/10.1177/1461445615590717>
- Bærentsen, K. B., & Trettvik, J. (2002). An Activity Theory Approach to Affordance. *Proceedings of the Second Nordic Conference on Human-computer Interaction*, 51–60. <https://doi.org/10.1145/572020.572028>
- Bamberg, M., & Georgakopoulou, A. (2008). Small Stories as a New Perspective in Narrative and Identity Analysis. *Text & Talk*, 28(3), 377–396. <https://doi.org/10.1515/TEXT.2008.018>
- Bergmann, J. (1993). Alarmiertes Verstehen: Kommunikation in Feuerwehrnotrufen. In T. Jung & S. Müller-Doohm (Hrsg.), „Wirklichkeit“ im Deutungsprozess: Verstehen und Methoden in den Kultur- und Sozialwissenschaften (S. 283–328). Suhrkamp.
- Bergmann, J. R. (1981). Ethnomethodologische Konversationsanalyse. In P. Schröder & H. Steger (Hrsg.), *Dialogforschung* (S. 9–52). Pädagogischer Verlag Schwann.
- Bergmann, J. R. (1985). Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit: Aufzeichnungen als Daten der interpretativen Soziologie. In W. Bonß & H. Hartmann (Hrsg.), *Entzauberte Wissenschaft: Zur Relativität und Geltung soziologischer Forschung* (Bd. 3, S. 299–320). Schwarz.
- Bergmann, J. R. (1988a). *Der Untersuchungsansatz der ethnomethodologischen Konversationsanalyse (Kurseinheit 2)*. FernUniversität GHS Hagen.
- Bergmann, J. R. (1988b). *Ethnomethodologie: Untersuchungen zur methodischen Erzeugung von sozialer Wirklichkeit im alltäglichen Handeln (Kurseinheit 1)*. FernUniversität GHS Hagen.
- Bergmann, J. R. (1988c). *Organisationsprinzipien der sozialen Interaktion: Objekte der Konversationsanalyse (Kurseinheit 3)*. FernUniversität GHS Hagen.
- Bergmann, J. R. (1990). On the Local Sensitivity of Conversation. In I. Markova & K. Foppa (Hrsg.), *The Dynamics of Dialogue* (S. 201–226). Harvester Wheatsleaf.
- Bergmann, J. R. (1998). Authentisierung und Fiktionalisierung in Alltagsgesprächen. In H. Willems & M. Jurga (Hrsg.), *Inszenierungsgesellschaft: Ein einführendes Handbuch* (S. 107–123). Westdeutscher Verlag.

- Bergmann, J. R. (1999). Diskrete Exploration: Über die moralische Sinnstruktur eines psychiatrischen Frageformats. In T. Luckmann & J. R. Bergmann (Hrsg.), *Kommunikative Konstruktion von Moral* (Bd. 2, S. 169–190). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-663-12193-0\\_6](https://doi.org/10.1007/978-3-663-12193-0_6)
- Bergmann, J. R. (2000). Konversationsanalyse. In E. von Kardorff, U. Flick, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (S. 524–537). Rowohlt.
- Bergmann, J. R. (2006). Studies of Work. In F. Rauner (Hrsg.), *Handbuch Berufsbildungsforschung* (S. 639–646). W. Bertelsmann Verlag.
- Bergmann, J. R. (2011). Qualitative Methoden der Medienforschung: Einleitung und Rahmung. In J. R. Bergmann & R. Ayaß (Hrsg.), *Qualitative Methoden der Medienforschung* (S. 13–41). Verlag für Gesprächsforschung.
- Blyth, C., Recktenwald, S., & Wang, J. (1990). I'm like, „Say What?!“: A New Quotative in American Oral Narrative. *American Speech*, 65(3), 215–227. <https://doi.org/10.2307/455910>
- Bohnsack, R. (2007a). Die dokumentarische Methode in der Bild- und Fotointerpretation. In R. Bohnsack, I. Nentwig-Gesemann, & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung* (S. 69–91). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-90741-3\\_4](https://doi.org/10.1007/978-3-531-90741-3_4)
- Bohnsack, R. (2007b). „Heidi“: Eine exemplarische Bildinterpretation auf der Basis der dokumentarischen Methode. In R. Bohnsack, I. Nentwig-Gesemann, & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung* (S. 325–340). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-90741-3\\_16](https://doi.org/10.1007/978-3-531-90741-3_16)
- Bourdieu, P. (1990). *Photography: A Middle-Brow Art*. Stanford UP.
- Bourdieu, P. (1991). Towards a Sociology of Photography. *Visual Anthropology Review*, 7(1), 129–133. <https://doi.org/10.1525/var.1991.7.1.129>
- Breckner, R. (2003). Körper im Bild: Annäherungen an eine methodische Analyse am Beispiel einer Fotografie von Helmut Newton. *Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung*, 1, 33–60.
- Brown, B., McGregor, M., & Laurier, E. (2013). iPhone in Vivo: Video Analysis of Mobile Device Use. *CHI'13. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1031–1040. <https://doi.org/10.1145/2470654.2466132>
- Brown, B., McGregor, M., & McMillan, D. (2014). 100 Days of iPhone Use: Understanding the Details of Mobile Device Use. *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, 223–232. <https://doi.org/10.1145/2628363.2628377>
- Brown, B., McGregor, M., & McMillan, D. (2015). Searchable Objects: Search in Everyday Conversation. *CSCW '15. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 508–517. <https://doi.org/10.1145/2675133.2675206>
- Brown, B., O'Hara, K., McGregor, M., & McMillan, D. (2018). Text in Talk: Lightweight Messages in Co-Present Interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(6), 1–25. <https://doi.org/10.1145/3152419>
- Bucholtz, M. (2000). The Politics of Transcription. *Journal of Pragmatics*, 32(10), 1439–1465. [https://doi.org/10.1016/S0378-2166\(99\)00094-6](https://doi.org/10.1016/S0378-2166(99)00094-6)

- Burkart, G. (2000a). Das Mobiltelefon und die Veränderung der Kommunikation im sozialen Raum. *Soziale Welt*, 51(2), 209–227.
- Burkart, G. (2000b). Mobile Kommunikation: Zur Kulturbedeutung des „Handy“. *Soziale Welt*, 51(2), 209–231.
- Burkart, G. (2007). *Handymaniamania. Wie das Mobiltelefon unser Leben verändert hat*. Campus Verlag.
- Burri, R. V. (2008). Bilder als soziale Praxis: Grundlegungen einer Soziologie des Visuellen. *Zeitschrift für Soziologie*, 37(4), 342–358.
- Burri, R. V. (2012). Visual Rationalities: Towards a Sociology of Images. *Current Sociology*, 60(1), 45–60. <https://doi.org/10.1177/0011392111426647>
- Burri, R. V., & Dumit, J. (2008). Social Studies of Scientific Imaging and Visualization. In U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Hrsg.), *The Handbook of Science and Technology Studies* (2. Aufl., S. 297–317). MIT Press.
- Büttner, D., & Frank, M. (2016). „Bei dir läuft“. In E. Arslan & K. Bozay (Hrsg.), *Symbolische Ordnung und Bildungsungleichheit in der Migrationsgesellschaft* (S. 215–237). Springer VS. [https://doi.org/10.1007/978-3-658-13703-8\\_12](https://doi.org/10.1007/978-3-658-13703-8_12)
- Caron, A. H., & Caronia, L. (2007). *Moving Cultures: Mobile Communication in Everyday Life*. McGill-Queen's Press.
- Chalfen, R. (1987). *Snapshot Versions of Life*. University of Wisconsin Press.
- Chotpitayasunondh, V., & Douglas, K. M. (2016). How „Phubbing“ Becomes the Norm: The Antecedents and Consequences of Snubbing via Smartphone. *Computers in Human Behavior*, 63, 9–18. <https://doi.org/10.1016/j.chb.2016.05.018>
- Chotpitayasunondh, V., & Douglas, K. M. (2018). The Effects of „Phubbing“ on Social Interaction. *Journal of Applied Social Psychology*, 48(6), 304–316. <https://doi.org/10.1111/jasp.12506>
- Churchill, L. (1971). Ethnomethodology and Measurement. *Social Forces*, 50(2), 182–191. <https://doi.org/10.1093/sf/50.2.182>
- Clift, R. (2006). Indexing Stance: Reported Speech as an Interactional Evidential. *Journal of Sociolinguistics*, 10(5), 569–595. <https://doi.org/10.1111/j.1467-9841.2006.00296.x>
- Clift, R. (2007). Getting There First: Non-Narrative Reported Speech in Interaction. In R. Clift & E. Holt (Hrsg.), *Reporting Talk. Reported Speech in Interaction* (S. 120–149). Cambridge UP.
- Costall, A. (1995). Socializing Affordances. *Theory & Psychology*, 5(4), 467–481. <https://doi.org/10.1177/0959354395054001>
- Couldry, N., & Hepp, A. (2016). *The Mediated Construction of Reality* (1. Aufl.). Polity Press.
- Couper-Kuhlen, E. (1999). Coherent Voicing: On Prosody in Conversational Reported Speech. In W. Bublitz, U. Lenk, & E. Ventola (Hrsg.), *Coherence in Spoken and Written Discourse* (S. 11–34). John Benjamins. <https://doi.org/10.1075/pbns.63.05cou>
- Couper-Kuhlen, E. (2007). Assessing and Accounting. In E. Holt & R. Clift (Hrsg.), *Reporting Talk: Reported Speech in Interaction* (S. 81–119). Cambridge UP.
- Davidson, C. (2009). Transcription: Imperatives for Qualitative Research. *International Journal of Qualitative Methods*, 8(2), 36–52. <https://doi.org/10.1177/160940690900800206>
- De Fina, A., & Georgakopoulou, A. (2012). *Analyzing Narrative: Discourse and Sociolinguistic Perspectives*. Cambridge UP.
- de Ruiter, J. P., & Albert, S. (2017). An Appeal for a Methodological Fusion of Conversation

- Analysis and Experimental Psychology. *Research on Language and Social Interaction*, 50(1), 90–107. <https://doi.org/10.1080/08351813.2017.1262050>
- Deppermann, A. (2008). *Gespräche analysieren: Eine Einführung* (4. Aufl.). VS Verlag für Sozialwissenschaften.
- Deppermann, A. (2013). Multimodal Interaction from a Conversation Analytic Perspective. *Journal of Pragmatics*, 46(1), 1–7. <https://doi.org/10.1016/j.pragma.2012.11.014>
- Deppermann, A. (2014). Multimodal Participation in Simultaneous Joint Projects: Interpersonal and Intrapersonal Coordination in Paramedic Emergency Drills. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Hrsg.), *Multiactivity in Social Interaction: Beyond Multitasking* (S. 247–281). John Benjamins.
- Deppermann, A., & Helmer, H. (2013). *Zur Grammatik des Verstehens im Gespräch: Inferenzen anzeigen und Handlungskonsequenzen ziehen mit „also“ und „dann“*. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/1304>
- DiDomenico, S. M., & Boase, J. (2013). Bringing Mobiles into the Conversation: Applying a Conversation Analytic Approach to the Study of Mobiles in Co-Present Interaction. In D. Tannen & A. M. Trester (Hrsg.), *Discourse 2.0: Language and New Media* (S. 119–132). Georgetown UP.
- DiDomenico, S. M., Raclaw, J., & Robles, J. S. (2018). Attending to the Mobile Text Summons: Managing Multiple Communicative Activities Across Physically Copresent and Technologically Mediated Interpersonal Interactions. *Communication Research*, 669–700. <https://doi.org/10.1177/0093650218803537>
- Döbler, T. (2014). Das Ende der Verbindlichkeit? Veränderungen sozialer Beziehungen durch mobiles Kommunikationsverhalten. In J. Wimmer & M. Hartmann (Hrsg.), *Medienkommunikation in Bewegung: Mobilisierung – Mobile Medien – Kommunikative Mobilität* (S. 139–154). Springer VS.
- Drew, P. (1984). Speakers' Reportings in Invitation Sequences. In J. M. Atkinson & J. Heritage (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (S. 129–151). Cambridge UP.
- Drew, P. (1987). Po-Faced Receipts of Teases. *Linguistics*, 25(1), 219–253. <https://doi.org/10.1515/ling.1987.25.1.219>
- Drew, P. (1989). Recalling Someone From the Past. In D. Roger & P. Bull (Hrsg.), *Conversation: An Interdisciplinary Perspective* (S. 96–115). Multilingual Matters.
- Drew, P. (1991). Asymmetries of Knowledge in Conversational Interactions. In K. Foppa (Hrsg.), *Asymmetries in Dialogue* (S. 29–48). Harvester Wheatsheaf.
- Drew, P. (1998). Complaints About Transgressions and Misconduct. *Research on Language and Social Interaction*, 31(3–4), 295–325. <https://doi.org/10.1080/08351813.1998.9683595>
- Drew, P., & Chilton, K. (2000). Calling Just to Keep in Touch: Regular and Habitualised Telephone Calls as an Environment of Small Talk. In J. Coupland (Hrsg.), *Small talk* (S. 137–162). Pearson Education.
- Duranti, A. (2007). Transcripts, Like Shadows on a Wall. *Mind, Culture, and Activity*, 13(4), 301–310. [https://doi.org/10.1207/s15327884mca1304\\_3](https://doi.org/10.1207/s15327884mca1304_3)
- Dürscheid, C. (2016). Neue Dialoge – alte Konzepte? Die schriftliche Kommunikation via Smartphone. *Zeitschrift für germanistische Linguistik*, 44(3), 437–468. <https://doi.org/10.1515/zgl-2016-0023>

- Dwyer, R. J., Kushlev, K., & Dunn, E. W. (2018). Smartphone Use Undermines Enjoyment of Face-to-Face Social Interactions. *Journal of Experimental Social Psychology*, 78, 233–239. <https://doi.org/10.1016/j.jesp.2017.10.007>
- Eberle, T. S. (2009). Ethnomethodologie. In R. Buber & H. H. Holzmüller (Hrsg.), *Qualitative Marktforschung: Konzepte – Methoden – Analysen* (S. 93–109). Gabler. [https://doi.org/10.1007/978-3-8349-9441-7\\_6](https://doi.org/10.1007/978-3-8349-9441-7_6)
- Egbert, M. M. (1997). Schisming: The Collaborative Transformation from a Single Conversation to Multiple Conversations. *Research on Language and Social Interaction*, 30(1), 1–51. [https://doi.org/10.1207/s15327973rlsi3001\\_1](https://doi.org/10.1207/s15327973rlsi3001_1)
- Egbert, M. M. (2009). *Der Reparatur-Mechanismus in deutschen Gesprächen*. Verlag für Gesprächsforschung.
- Ehmer, O. (2011). *Imagination und Animation: Die Herstellung mentaler Räume durch animierte Rede*. De Gruyter.
- Eisentraut, S. (2016). *Mobile Interaktionsordnungen im Jugendalter: Zur Soziologie des Handygebrauchs*. Beltz Juventa.
- Esbjörnsson, M., Juhlin, O., & Weilenmann, A. (2007). Drivers Using Mobile Phones in Traffic: An Ethnographic Study of Interactional Adaptation. *International Journal of Human-Computer Interaction*, 22(1–2), 37–58. <https://doi.org/10.1080/10447310709336954>
- Figeac, J., & Chaulet, J. (2018). Video-Ethnography of Social Media Apps' Connection Cues in Public Settings. *Mobile Media & Communication*, 6(3), 407–427. <https://doi.org/10.1177/2050157917747642>
- Fitzgerald, R. (2012). Membership Categorization Analysis: Wild and Promiscuous or Simply the Joy of Sacks? *Discourse Studies*, 14(3), 305–311. <https://doi.org/10.1177/1461445612440776>
- Ford, C. E., & Fox, B. A. (2010). Multiple Practices for Constructing Laughables. In D. Barth-Weingarten, E. Reber, & M. Selting (Hrsg.), *Prosody in Interaction* (S. 339–368). John Benjamins.
- Forrester, M. (1999). Conversation and Instruction with Apprenticeship: Affordances for Learning. In P. Ainley & H. Rainbird (Hrsg.), *Apprenticeship: Towards a New Paradigm of Learning* (S. 86–98). Routledge.
- Fortunati, L. (2002). The Mobile Phone: Towards New Categories and Social Relations. *Information, Communication & Society*, 5(4), 513–528. <https://doi.org/10.1080/13691180208538803>
- Fox, J., & Rooney, M. C. (2015). The Dark Triad and Trait Self-Objectification as Predictors of Men's Use and Self-Presentation Behaviors on Social Networking Sites. *Personality and Individual Differences*, 76, 161–165. <https://doi.org/10.1016/j.paid.2014.12.017>
- Garfinkel, H. (1963). A Conception of, and Experiments with „Trust“ as a Condition of Stable Concerted Action. In O. J. Harvey (Hrsg.), *Motivation and Social Interaction: Cognitive Determinants* (S. 187–236). Ronald Press Co.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice Hall.
- Garfinkel, H. (2020). *Studien zur Ethnomethodologie*. Campus Verlag.
- Garfinkel, H. (1959). Aspects of the Problem of Common Sense Knowledge of Social Structures. *Transactions of the IVth World Congress of Sociology*, 51–65.
- Garfinkel, H., & Sacks, H. L. (1986). On Formal Structures of Practical Action. In H. Garfinkel

- (Hrsg.), *Ethnomethodological Studies of Work* (S. 160–193). Routledge.
- Garfinkel, H., & Wieder, L. D. (1992). Two Incommensurable Asymmetrically Alternate Technologies of Social Analysis. In G. Watson & R. M. Seiler (Hrsg.), *Text in Context: Contributions to Ethnomethodology* (S. 175–206). SAGE.
- Gaver, W. W. (1991). Technology Affordances. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 79–84. <https://doi.org/10.1145/108844.108856>
- Gehrau, V. (2019). Gespräche über Medien in Zeiten von Mobilkommunikation und sozialen Onlinenetzen. *Publizistik*, 64(1), 37–58. <https://doi.org/10.1007/s11616-018-00465-z>
- Gehrau, V., & Goertz, L. (2010). Gespräche über Medien unter veränderten medialen Bedingungen. *Publizistik*, 55, 153–172.
- Georgakopoulou, A. (2007). *Small Stories, Interaction and Identities*. John Benjamins. <https://doi.org/10.1075/sin.8>
- Georgakopoulou, A. (2008). ‚On MSN with Buff Boys‘: Self- and Other-Identity Claims in the Context of Small Stories. *Journal of Sociolinguistics*, 12(5), 597–626. <https://doi.org/10.1111/j.1467-9841.2008.00384.x>
- Georgakopoulou, A. (2015). Small Stories Research. In A. De Fina & A. Georgakopoulou (Hrsg.), *The Handbook of Narrative Analysis* (S. 255–271). John Wiley & Sons. <https://doi.org/10.1002/9781118458204.ch13>
- Gergen, K. J. (2002). The Challenge of the Absent Presence. In J. E. Katz & M. Aakhus (Hrsg.), *Perpetual Contact: Mobile Communication, Private Talk, Public Performance* (S. 227–241). Cambridge UP.
- Geser, H. (2004). *Towards a Sociological Theory of the Mobile Phone*.
- Geser, H. (2005). Is the Cell Phone Undermining the Social Order? Understanding Mobile Technology from a Sociological Perspective. In P. Glotz, S. Bertschi, & C. Locke (Hrsg.), *Thumb Culture: The Meaning of Mobile Phones for Society* (S. 23–36). transcript Verlag.
- Gibbs, M., Meese, J., Arnold, M., Nansen, B., & Carter, M. (2015). #Funeral and Instagram: Death, Social Media, and Platform Vernacular. *Information, Communication & Society*, 18(3), 255–268. <https://doi.org/10.1080/1369118X.2014.987152>
- Gibson, J. J. (1977). The Theory of Affordances. In R. Shaw & J. Bransford (Hrsg.), *Perceiving, Acting and Knowing* (S. 67–82). John Wiley & Sons.
- Gibson, J. J. (1982). *Wahrnehmung und Umwelt: Der ökologische Ansatz in der visuellen Wahrnehmung*. U&S Psychologie.
- Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum.
- Glotz, P., Locke, C., & Bertschi, S. (Hrsg.). (2005). *Daumenkultur: Das Mobiltelefon in der Gesellschaft*. transcript Verlag.
- Goffman, E. (1959). *The Presentation of Self in Everyday Life* (1. Aufl.). Doubleday.
- Goffman, E. (1963). *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Free Press.
- Goffman, E. (1964). The Neglected Situation. *American Anthropologist*, 66(6), 133–136. [https://doi.org/10.1525/aa.1964.66.suppl\\_3.02a00090](https://doi.org/10.1525/aa.1964.66.suppl_3.02a00090)
- Goffman, E. (1967). *Interaction Ritual: Essays on Face-to-Face Behavior*. Doubleday.
- Goffman, E. (1974). *Das Individuum im öffentlichen Austausch: Mikrostudien zur öffentlichen Ordnung*. Suhrkamp.
- Goffman, E. (1978). Response Cries. *Language*, 54(4), 787–815.

- <https://doi.org/10.2307/413235>
- Goffman, E. (1979a). Footing. *Semiotica*, 25(1–2), 1–30.  
<https://doi.org/10.1515/semi.1979.25.1-2.1>
- Goffman, E. (1979b). *Gender Advertisements*. Harper & Row.
- Goffman, E. (1981). *Forms of Talk*. University of Pennsylvania Press.
- Goffman, E. (2003). *Wir alle spielen Theater: Die Selbstdarstellung im Alltag*. Piper.
- Goffman, E. (2018). *Interaktion im öffentlichen Raum* (Neuausgabe). Campus Verlag.
- Goggin, G. (2006). *Cell Phone Culture: Mobile Technology in Everyday Life*. Routledge.
- Goggin, G. (2013). *Mobile Phone Cultures*. Routledge.
- Golato, A. (2000). An Innovative German Quotative for Reporting on Embodied Actions: Und Ich So/und Er So ‘and I’m Like/and He’s Like’. *Journal of Pragmatics*, 32(1), 29–54.  
[https://doi.org/10.1016/S0378-2166\(99\)00030-2](https://doi.org/10.1016/S0378-2166(99)00030-2)
- Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural Conversation. In G. Psathas (Hrsg.), *Everyday Language: Studies in Ethnomethodology* (S. 97–121). Irvington.
- Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, 50(3–4), 272–302. <https://doi.org/10.1111/j.1475-682X.1980.tb00023.x>
- Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*. Academic Press.
- Goodwin, C. (1984). Notes on Story Structure and the Organization of Participation. In J. M. Atkinson & J. Heritage (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (S. 225–246). Cambridge UP.
- Goodwin, C. (1986). Between and Within: Alternative Sequential Treatments of Continuers and Assessments. *Human Studies*, 9(2/3), 205–217.
- Goodwin, C. (1987). Forgetfulness as an Interactive Resource. *Social Psychology Quarterly*, 50(2), 115–130. <https://doi.org/10.2307/2786746>
- Goodwin, C. (1996). Transparent Vision. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Hrsg.), *Interaction and Grammar* (S. 370–404). Cambridge UP.
- Goodwin, C. (2000). Practices of Seeing: Visual Analysis: An Ethnomethodological Approach. In T. van Leeuwen & C. Jewitt (Hrsg.), *Handbook of Visual Analysis* (S. 157–182). SAGE.
- Goodwin, C. (2007a). Participation, Stance and Affect in the Organization of Activities. *Discourse & Society*, 18(1), 53–73. <https://doi.org/10.1177/0957926507069457>
- Goodwin, C. (2007b). Environmentally Coupled Gestures. In S. D. Duncan, J. Cassell, & E. T. Levy (Hrsg.), *Gesture and the Dynamic Dimension of Language: Essays in Honor of David McNeill* (Bd. 1, S. 195–212). John Benjamins.
- Goodwin, C. (2015a). Narrative as Talk-in-Interaction. In A. Georgakopoulou & A. De Fina (Hrsg.), *The Handbook of Narrative Analysis* (S. 197–218). Wiley-Blackwell.
- Goodwin, C. (2015b). Professional Vision. In S. Reh, K. Berdelmann, & J. Dinkelaker (Hrsg.), *Aufmerksamkeit: Geschichte—Theorie—Empirie* (S. 387–425). Springer VS.  
[https://doi.org/10.1007/978-3-531-19381-6\\_20](https://doi.org/10.1007/978-3-531-19381-6_20)
- Goodwin, C., & Goodwin, M. H. (1992). Context, Activity and Participation. In P. Auer & A. D. Luzio (Hrsg.), *The Contextualization of Language* (S. 77–99). John Benjamins.
- Goodwin, C., & Goodwin, M. H. (2004). Participation. In A. Duranti (Hrsg.), *A Companion to*

- Linguistic Anthropology* (S. 222–244). Wiley-Blackwell.
- Goodwin, C., & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, 19(1), 283–307. <https://doi.org/10.1146/annurev.an.19.100190.001435>
- Goodwin, M. H. (1980). Processes of Mutual Monitoring Implicated in the Production of Description Sequences. *Sociological Inquiry*, 50(3–4), 303–317. <https://doi.org/10.1111/j.1475-682X.1980.tb00024.x>
- Goodwin, M. H. (1997). Byplay: Negotiating Evaluation in Storytelling. In G. R. Guy, C. Feagin, D. Schiffin, & J. Baugh (Hrsg.), *Towards a Social Science of Language: Papers in Honor of William Labov* (Bd. 2, S. 77–102). John Benjamins.
- Goodwin, M. H., & Goodwin, C. (1992). Assessments and the Construction of Context. In A. Duranti (Hrsg.), *Rethinking Context: Language as an Interactive Phenomenon* (S. 147–189). Cambridge UP.
- Gordon, C. (2012). Beyond the Observer’s Paradox: The Audio-Recorder as a Resource for the Display of Identity. *Qualitative Research*, 13(3), 299–317. <https://doi.org/10.1177/1468794112442771>
- Graffe, J. (2014). Begrüßungen 2.0: Eine kontrastive Darstellung der Einstiegssequenzen in SMS und privaten Nachrichten im sozialen Netzwerk Facebook. In N. Bahlo & K. König (Hrsg.), *SMS, WhatsApp & Co.: Gattungsanalytische, kontrastive und variationslinguistische Perspektiven zur Analyse mobiler Kommunikation* (S. 41–58). Westfälische Wilhelms-Universität.
- Greer, T. (2016). Multiple Involvements in Interactional Repair: Using Smartphones in Peer Culture to Augment Lingua Franca English. In M. Theobald (Hrsg.), *Friendship and Peer Culture in Multilingual Settings* (S. 197–229). Emerald Group Publishing Limited. <https://doi.org/10.1108/S1537-466120160000021010>
- Greschke, H. (2014). „Mein Smartphone ist mein Schatz“: Intimität in transnationalen Familien. In K. Hahn (Hrsg.), *E<3Motion* (S. 151–167). Springer VS. [https://doi.org/10.1007/978-3-658-02732-2\\_9](https://doi.org/10.1007/978-3-658-02732-2_9)
- Greschke, H., Dreßler, D., & Hierasimowicz, K. (2017). Die Mediatisierung von Eltern-Kind-Beziehungen im Kontext grenzüberschreitender Migration. In F. Krotz, C. Despotović, & M.-M. Kruse (Hrsg.), *Mediatisierung als Metaprozess: Transformationen, Formen der Entwicklung und die Generierung von Neuem* (S. 59–80). Springer VS. [https://doi.org/10.1007/978-3-658-16084-5\\_4](https://doi.org/10.1007/978-3-658-16084-5_4)
- Gülich, E. (1994). Formulierungsarbeit im Gespräch. In S. Cmejrkova (Hrsg.), *Writing vs. Speaking: Language, Text, Discourse, Communication* (S. 77–95).
- Gülich, E., & Hausendorf, H. (2000). Vertextungsmuster: Narration. In K. Brinker, G. Antos, W. Heinemann, & S. Sager (Hrsg.), *Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung: Bd. 1. Halbband* (S. 369–385). De Gruyter.
- Gülich, E., & Mondada, L. (2008). *Konversationsanalyse: Eine Einführung am Beispiel des Französischen*. Niemeyer.
- Günthner, S. (1997). Direkte und indirekte Rede in Alltagsgesprächen: Zur Interaktion von Syntax und Prosodie in der Redewiedergabe. In P. Schlobinski (Hrsg.), *Syntax des gesprochenen Deutsch* (S. 227–262). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-322-88924-9\\_9](https://doi.org/10.1007/978-3-322-88924-9_9)
- Günthner, S. (1999a). Beschwerdeerzählungen als narrative Hyperbeln. In J. R. Bergmann & T. Luckmann (Hrsg.), *Kommunikative Konstruktion von Moral: Struktur und Dynamik*

- der Formen moralischer Kommunikation* (Bd. 1, S. 174–205). Opladen.
- Günthner, S. (1999b). Polyphony and the ‚Layering of Voices‘ in Reported Dialogues: An Analysis of the Use of Prosodic Devices in Everyday Reported Speech. *Journal of Pragmatics*, 31(5), 685–708. [https://doi.org/10.1016/S0378-2166\(98\)00093-9](https://doi.org/10.1016/S0378-2166(98)00093-9)
- Günthner, S. (2006). Zwischen Scherz und Schmerz: Frotzelaktivitäten in Alltagsinteraktionen. In H. Kotthoff (Hrsg.), *Scherzkommunikation: Beiträge aus der empirischen Gesprächsforschung* (S. 81–108). Verlag für Gesprächsforschung.
- Günthner, S. (2012a). Kleine interaktionale Erzählungen als Ressourcen der Fremd- und Selbststilisierung. In F. Kern, M. Morek, & S. Ohlhus (Hrsg.), *Erzählen als Form—Formen des Erzählens* (S. 65–83). De Gruyter. <https://doi.org/10.1515/9783110280579.65>
- Günthner, S. (2012b). *Vorwurfsaktivitäten in der Alltagsinteraktion: Grammatische, prosodische, rhetorisch-stilistische und interaktive Verfahren bei der Konstitution kommunikativer Muster und Gattungen* (Reprint). De Gruyter. <https://doi.org/10.1515/9783110919974>
- Guski, R. (1996). *Wahrnehmen: Ein Lehrbuch*. Kohlhammer.
- Haakana, M. (2007). Reported Thought in Complaint Stories. In E. Holt & R. Clift (Hrsg.), *Reporting Talk: Reported Speech in Interaction* (S. 150–178). Cambridge UP.
- Haddington, P., Keisanen, T., Mondada, L., & Nevile, M. (Hrsg.). (2014). *Multiactivity in Social Interaction: Beyond Multitasking*. John Benjamins.
- Haddington, P., & Rauniomaa, M. (2011). Technologies, Multitasking, and Driving: Attending to and Preparing for a Mobile Phone Conversation in a Car. *Human Communication Research*, 37(2), 223–254. <https://doi.org/10.1111/j.1468-2958.2010.01400.x>
- Harper, R. (2001). The Mobile Interface: Old Technologies and New Arguments. In R. Harper, B. Brown, & N. Green (Hrsg.), *Wireless World. Social and Interactional Aspects of the Mobile Age* (S. 36–49). Springer VS.
- Hartmann, M. (2010). Mediatisierung als Mediation: Vom Normativen und Diskursiven. In M. Hartmann & A. Hepp (Hrsg.), *Die Mediatisierung der Alltagswelt* (S. 35–47). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-92014-6\\_3](https://doi.org/10.1007/978-3-531-92014-6_3)
- Hartmann, M. (2013). *Domestizierung*. Nomos.
- Haviland, J. B. (2000). Pointing, Gesture Spaces, and Mental Maps. In D. McNeill (Hrsg.), *Language and Gesture* (S. 13–46). Cambridge UP. <https://doi.org/10.1017/CBO9780511620850.003>
- Hazel, S. (2016). The Paradox from Within: Research Participants Doing-Being-Observed. *Qualitative Research*, 16(4), 446–467. <https://doi.org/10.1177/1468794115596216>
- Heath, C. (1982). The Display of Reciprocity: An instance of a Sequential Relationship in Speech and Body movement. *Semiotica*, 42(2–4), 147–168. <https://doi.org/10.1515/semi.1982.42.2-4.147>
- Heath, C. (1986). *Body Movement and Speech in Medical Interaction* (1. Aufl.). Cambridge UP.
- Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in Qualitative Research*. SAGE.
- Heath, C., & Luff, P. (2000). *Technology in Action*. Cambridge UP.
- Hector, T. M. (2017). Nutzungskontexte und Dialogizität von WhatsApp-Sprachnachrichten. *studInteraktionSprache (studentische Arbeitspapierreihe)*, 27. <http://arbeitspapiere.sprache-interaktion.de/stud/arbeitspapiere/arbeitspapier27.pdf>
- Helisten, M. (2019). Disjunctively Positioned Problem-Noticings in Managing Multiactivity.

- Research on Language and Social Interaction*, 52(4), 318–342.  
<https://doi.org/10.1080/08351813.2019.1657274>
- Hepburn, A. (2004). Crying: Notes on Description, Transcription, and Interaction. *Research on Language and Social Interaction*, 37(3), 251–290.  
[https://doi.org/10.1207/s15327973rlsi3703\\_1](https://doi.org/10.1207/s15327973rlsi3703_1)
- Hepp, A., Berg, M., & Roitsch, C. (2014). *Mediatisierte Welten der Vergemeinschaftung: Kommunikative Vernetzung und das Gemeinschaftsleben junger Menschen*. Springer VS.
- Hepp, A., & Hartmann, M. (2010). Mediatisierung als Metaprozess: Der analytische Zugang von Friedrich Krotz zur Mediatisierung der Alltagswelt. In M. Hartmann & A. Hepp (Hrsg.), *Die Mediatisierung der Alltagswelt* (S. 9–20). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-92014-6\\_1](https://doi.org/10.1007/978-3-531-92014-6_1)
- Heritage, J. (1984a). *Garfinkel and Ethnomethodology*. Polity Press.
- Heritage, J. (1984b). A Change-of-State Token and Aspects of its Sequential Placement. In J. M. Atkinson & J. Heritage (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (S. 299–345). Cambridge UP.
- Heritage, J. (1988). Explanations as Accounts: A Conversation Analytic Perspective. In C. Antaki (Hrsg.), *Analyzing Lay Explanation: A Case Book of Methods* (S. 127–144). SAGE.
- Heritage, J. (2011). Conversation Analysis: Practices and Methods. In D. Silverman (Hrsg.), *Qualitative Research: Issues of Theory, Method and Practice* (3. Aufl., S. 159–183). SAGE.
- Heritage, J. (2012). The Epistemic Engine: Sequence Organization and Territories of Knowledge. *Research on Language and Social Interaction*, 45(1), 30–52.  
<https://doi.org/10.1080/08351813.2012.646685>
- Heritage, J., & Raymond, G. (2005). The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction. *Social Psychology Quarterly*, 68(1), 15–38.
- Heritage, J., & Stivers, T. (1999). Online Commentary in Acute Medical Visits: A Method of Shaping Patient Expectations. *Social Science & Medicine*, 49(11), 1501–1517.  
[https://doi.org/10.1016/S0277-9536\(99\)00219-1](https://doi.org/10.1016/S0277-9536(99)00219-1)
- Hillman, T., Weilenmann, A., Jungselius, B., & Lindell, T. L. (2016). Traces of Engagement: Narrative-Making Practices with Smartphones on a Museum Field Trip. *Learning, Media and Technology*, 41(2), 351–370. <https://doi.org/10.1080/17439884.2015.1064443>
- Hjarvard, S. (2008). The Mediatization of Society: A Theory of the Media as Agents of Social and Cultural Change. *NORDICOM Review*, 29, 105–134.
- Hoffmann, D., & Winter, R. (Hrsg.). (2018). *Mediensoziologie: Handbuch für Wissenschaft und Studium*. Nomos.
- Höflich, J. R. (1996). *Technisch vermittelte interpersonale Kommunikation: Grundlagen, organisatorische Medienverwendung, Konstitution „elektronischer Gemeinschaften“*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-99336-6>
- Höflich, J. R. (2006). Das Mobiltelefon im Spannungsfeld zwischen privater und öffentlicher Kommunikation: Ergebnisse einer internationalen explorativen Studie. In P. Glotz, C. Locke, & S. Bertschi (Hrsg.), *Daumenkultur: Das Mobiltelefon in der Gesellschaft* (S. 143–158). transcript Verlag.
- Höflich, J. R. (2011). *Mobile Kommunikation im Kontext: Studien zur Nutzung des Mobiltelefons im öffentlichen Raum*. Peter Lang.
- Höflich, J. R. (2016). *Der Mensch und seine Medien*. Springer VS.

- Höflich, J. R., & Hartmann, M. (2006). *Mobile Communication in Everyday Life: Ethnographic Views, Observations and Reflections*. Frank & Timme GmbH.
- Höflich, J. R., Kircher, G. F., Linke, C., & Schlote, I. (Hrsg.). (2010). *Mobile Media and the Change of Everyday Life* (1st Edition). Peter Lang.
- Höflich, J. R., Wimmer, J., & Hartmann, M. (2014). Doing Mobility: Menschen in Bewegung. Aktivitätsmuster, Zwischenräume und mobile Kommunikation. In *Medienkommunikation in Bewegung: Mobilisierung – Mobile Medien – Kommunikative Mobilität* (S. 31–45). VS Verlag für Sozialwissenschaften.
- Holfelder, U., & Ritter, C. (2013). Filmen im Alltag: Handyfilme in der Perspektive einer medienweltlichen Ethnografie. *kommunikation@gesellschaft*, 14, 27. <https://doi.org/10.5167/uzh-92083>
- Holler, J., & Kendrick, K. H. (2015). Unaddressed Participants' Gaze in Multi-Person Interaction: Optimizing Reciprocity. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00098>
- Holt, E. (1996). Reporting on Talk: The Use of Direct Reported Speech in Conversation. *Research on Language and Social Interaction*, 29(3), 219–245. [https://doi.org/10.1207/s15327973rlsi2903\\_2](https://doi.org/10.1207/s15327973rlsi2903_2)
- Holt, E. (2000). Reporting and Reacting: Concurrent Responses to Reported Speech. *Research on Language and Social Interaction*, 33(4), 425–454. [https://doi.org/10.1207/S15327973RLSI3304\\_04](https://doi.org/10.1207/S15327973RLSI3304_04)
- Holt, E. (2007). ‚I’m Eyeing Your Chop Up Mind’: Reporting and Enacting. In E. Holt & R. Clift (Hrsg.), *Reporting Talk: Reported Speech in Interaction* (S. 47–80). Cambridge UP.
- Holt, E. (2010). The Last Laugh: Shared Laughter and Topic Termination. *Journal of Pragmatics*, 42(6), 1513–1525. <https://doi.org/10.1016/j.pragma.2010.01.011>
- Holt, E. (2017). Indirect Reported Speech in Storytelling: Its Position, Design, and Uses. *Research on Language and Social Interaction*, 50(2), 171–187. <https://doi.org/10.1080/08351813.2017.1301302>
- Holt, E., & Clift, R. (2007). Introduction. In E. Holt & R. Clift (Hrsg.), *Reporting Talk: Reported Speech in Interaction* (S. 1–15). Cambridge UP.
- Hömke, P., Holler, J., & Levinson, S. C. (2017). Eye Blinking as Addressee Feedback in Face-To-Face Conversation. *Research on Language and Social Interaction*, 50(1), 54–70. <https://doi.org/10.1080/08351813.2017.1262143>
- Humphreys, L. (2005). Cellphones in Public: Social Interactions in a Wireless Era. *New Media & Society*, 7(6), 810–833. <https://doi.org/10.1177/1461444805058164>
- Hutchby, I. (2001a). *Conversation and Technology: From the Telephone to the Internet*. Blackwell Publishers.
- Hutchby, I. (2001b). Technologies, Texts and Affordances. *Sociology*, 35(2), 441–456. <https://doi.org/10.1177/S0038038501000219>
- Hutchby, I. (2003). Affordances and the Analysis of Technologically Mediated Interaction: A Response to Brian Rappert. *Sociology*, 37(3), 581–589. <https://doi.org/10.1177/00380385030373011>
- Hutchby, I., & Wooffitt, R. (1998). *Conversation Analysis: Principles, Practices and Applications*. Wiley-Blackwell.
- İkizoğlu, D. (2019). „What did it say?“. Mobile Phone Translation App as Participant and

- Object in Family Discourse. *Journal of Pragmatics*, 147, 1–16. <https://doi.org/10.1016/j.pragma.2019.05.001>
- Innis, H. A. (2008). *The Bias of Communication*. University of Toronto Press. <http://site.ebrary.com/id/10382206>
- Isaacs, E., Szymanski, M., Yamauchi, Y., Glasnapp, J., & Iwamoto, K. (2012). Integrating Local and Remote Worlds Through Channel Blending. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 617–626. <https://doi.org/10.1145/2145204.2145299>
- Ito, M., & Okabe, D. (2005). Technosocial Situations: Emergent Structurings of Mobile Email Use. In M. Ito, D. Okabe, & M. Matsuda (Hrsg.), *Personal, Portable, Pedestrian: Mobile Phones in Japanese Life* (S. 257–273). MIT Press.
- Jäckel, M. (Hrsg.). (2005). *Mediensoziologie: Grundfragen und Forschungsfelder*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-80675-8>
- Jefferson, G. (1972). Side Sequences. In D. N. Sudnow (Hrsg.), *Studies in Social Interaction* (S. 294–33). Free Press.
- Jefferson, G. (1978). Sequential Aspects of Storytelling in Conversation. In J. Schenkein (Hrsg.), *Studies in the Organization of Conversational Interaction* (S. 219–248). Academic Press.
- Jefferson, G. (1980). The Abominable Ne? An Exploration of Post-Response Pursuit of Response. In H. Steger & P. Schröder (Hrsg.), *Dialogforschung: Jahrbuch 1980 des Instituts für deutsche Sprache* (S. 53–88). Pädagogischer Verlag Schwann.
- Jefferson, G. (1983). Two Explorations of the Organization of Overlapping Talk in Conversation. *Tilburg Papers in Language and Literature*, 28.
- Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens „Yeah“; and „Mm Hm“. *Papers in Linguistics*, 17(2), 197–216. <https://doi.org/10.1080/08351818409389201>
- Jefferson, G. (1985). An Exercise in the Transcription and Analysis of Laughter. In T. Van Dijk (Hrsg.), *Handbook of Discourse Analysis: Discourse and Dialogue* (Bd. 3, S. 25–34). Academic Press.
- Jefferson, G. (1988). On the Sequential Organization of Troubles Talk in Ordinary Conversation. *Social Problems*, 35(4), 418–442.
- Jefferson, G. (1990). List Construction as a Task and Resource. In G. Psathas (Hrsg.), *Interactional Competence* (S. 63–92). University Press of America.
- Jefferson, G. (1996). A Case of Transcriptional Stereotyping. *Journal of Pragmatics*, 26(2), 159–170. [https://doi.org/10.1016/0378-2166\(96\)00010-0](https://doi.org/10.1016/0378-2166(96)00010-0)
- Jefferson, G. (2015). On the Interactional Unpacking of a „Gloss“. In P. Drew, J. Heritage, G. H. Lerner, & A. Pomerantz (Hrsg.), *Talking About Troubles in Conversation* (S. 127–164). Oxford UP.
- Jefferson, G., & Lee, J. R. E. (1981). The rejection of advice: Managing the problematic convergence of a ‘troubles-telling’ and a ‘service encounter’. *Journal of Pragmatics*, 5(5), 399–422. [https://doi.org/10.1016/0378-2166\(81\)90026-6](https://doi.org/10.1016/0378-2166(81)90026-6)
- Kääntä, L. (2014). From Noticing to Initiating Correction: Students’ Epistemic Displays in Instructional Interaction. *Journal of Pragmatics*, 66, 86–105. <https://doi.org/10.1016/j.pragma.2014.02.010>
- Kaerlein, T. (2018). *Smartphones als digitale Nahkörpertechnologien: Zur Kybernetisierung*

- des Alltags*. transcript Verlag.
- Kamunen, A. (2019). How to Disengage: Suspension, Body Torque, and Repair. *Research on Language and Social Interaction*, 52(4), 406–426. <https://doi.org/10.1080/08351813.2019.1657287>
- Katz, J. E. (2011). *Magic in the Air: Mobile Communication and the Transformation of Social Life*. Transaction Publishers.
- Katz, J. E., & Aakhus, M. (Hrsg.). (2002). *Perpetual Contact: Mobile Communication, Private Talk, Public Performance*. Cambridge UP.
- Katz, J. E., & Crocker, E. T. (2015). Selfies and Photo Messaging as Visual Conversation: Reports from the United States, United Kingdom and China. *International Journal of Communication*, 9, 1861–1872.
- Kaukomaa, T., Peräkylä, A., & Ruusuvuori, J. E. (2013). Turn-Opening Smiles: Facial Expression Constructing Emotional Transition in Conversation. *Journal of Pragmatics*, 55, 21–42. <https://doi.org/10.1016/j.pragma.2013.05.006>
- Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge UP.
- Kendrick, K. H. (2017). Using Conversation Analysis in the Lab. *Research on Language and Social Interaction*, 50(1), 1–11. <https://doi.org/10.1080/08351813.2017.1267911>
- Kendrick, K. H., & Holler, J. (2017). Gaze Direction Signals Response Preference in Conversation. *Research on Language and Social Interaction*, 50(1), 12–32. <https://doi.org/10.1080/08351813.2017.1262120>
- Keppler, A. (1994). *Tischgespräche: Über Formen kommunikativer Vergemeinschaftung am Beispiel der Konversation in Familien*. Suhrkamp.
- Keppler, A. (2000). Medien- und Kommunikationssoziologie: Verschränkte Gegenwart. Die Untersuchung kultureller Transformationen. *Soziologische Revue*, 23, 140–153. <https://doi.org/10.1524/srsr.2000.23.sonderheft5.140>
- Keppler, A. (2002). Visuelle Gestalt und kultureller Gehalt: Über den sozialen Gebrauch fotografischer Bilder. In K. Neumann-Braun (Hrsg.), *Medienkultur und Kulturkritik* (S. 89–99). Opladen.
- Keppler, A. (2013). Reichweiten alltäglicher Gespräche: Über den kommunikativen Gebrauch alter und neuer Medien. In A. Bellebaum & R. Hettlage (Hrsg.), *Unser Alltag ist voll von Gesellschaft: Sozialwissenschaftliche Beiträge* (S. 85–104). Springer VS.
- Keppler, A. (2018). Medien, Lebenswelt und Alltagshandeln. In D. Hoffmann & R. Winter (Hrsg.), *Mediensoziologie: Handbuch für Wissenschaft und Studium* (S. 71–85). Nomos. <https://doi.org/10.5771/9783845264196-70>
- Keppler, A. (2019). „Zeig mal“: Smartphones im Gespräch. In C. Marx & A. Schmidt (Hrsg.), *Interaktion und Medien: Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation* (S. 177–190). Universitätsverlag Winter.
- Kidwell, M. (2013). Framing, Grounding and Coordinating Conversational Interaction: Posture, Gaze, Facial Expression, and Movement in Space. In C. Müller, A. J. Cienki, E. Fricke, S. L. Ladewig, D. McNeill, & S. Tessendorf (Hrsg.), *Body—Language—Communication: International Handbook on Multimodality in Human Interaction* (S. 100–112). De Gruyter.
- Knorr Cetina, K. (1989). Spielarten des Konstruktivismus: Einige Notizen und Anmerkungen. *Soziale Welt*, 40(1/2), 86–96.

- Knorr Cetina, K. (1995). Laboratory Studies: The Cultural Approach to the Study of Science. In S. Jasanoff, G. Markle, J. Peterson, & T. Pinch (Hrsg.), *Handbook of Science and Technology Studies* (S. 140–166). SAGE. <https://doi.org/10.4135/9781412990127.n7>
- Knorr Cetina, K. (2009). The Synthetic Situation: Interactionism for a Global World. *Symbolic Interaction*, 32(1), 61–87.
- Knorr Cetina, K., & Amann, K. (1988). The Fixation of (Visual) Evidence. *Human Studies*, 11(2–3), 133–169.
- Koike, C. (2001). An Analysis of Shifts in Participation Roles in Japanese Storytelling in Terms of Prosody, Gaze, and Body Movements. *Annual Meeting of the Berkeley Linguistics Society*, 27(1), 381–392. <https://doi.org/10.3765/bls.v27i1.3418>
- Komter, M. L. (2006). From Talk to Text: The Interactional Construction of a Police Record. *Research on Language and Social Interaction*, 39(3), 201–228. [https://doi.org/10.1207/s15327973rlsi3903\\_2](https://doi.org/10.1207/s15327973rlsi3903_2)
- König, K. (2015). „Muss leider absagen. Muss noch nen referat fertig machen“: Zur Dialogizität von Absagen und Verabredungsablehnungen in der SMS-Kommunikation. *Linguistik Online*, 70(1).
- König, K. (2019). Sequential Patterns in SMS and WhatsApp Dialogues: Practices for Coordinating Actions and Managing Topics. *Discourse & Communication*, 13(6), 612–629. <https://doi.org/10.1177/1750481319868853>
- König, K., & Bahlo, N. (2014). SMS, WhatsApp & Co.: Forschungsstand und Analyseperspektiven. In N. Bahlo & K. König (Hrsg.), *SMS, WhatsApp & Co.: Gattungsanalytische, kontrastive und variationslinguistische Perspektiven zur Analyse mobiler Kommunikation* (S. 1–16). Westfälische Wilhelms-Universität.
- König, K., & Hector, T. M. (2017). Zur Theatralität von WhatsApp-Sprachnachrichten: Nutzungskontexte von Audio-Postings in der mobilen Messenger-Kommunikation. *Networks*, 79.
- König, K., & Oloff, F. (2018a). Die Multimodalität alltagspraktischen Erzählens. *Zeitschrift für Literaturwissenschaft und Linguistik*, 48, 277–307. <https://doi.org/10.1007/s41244-018-0093-7>
- König, K., & Oloff, F. (2018b). Ansätze zu einer multimodalen Erzählanalyse: Einführung in das Themenheft. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 19 (Themenheft Erzählen multimodal), 207–241.
- Kopomaa, T. (2000). *The City in Your Pocket: Birth of the Mobile Information Society*. Gaudeamus.
- Koskinen, I. (2004). *Seeing with Mobile Images: Towards Perpetual Visual Contact*. Communications in the 21st Century: The Mobile Information Society, Budapest.
- Koskinen, I., Kurvinen, E., & Lehtonen, T.-K. (2002). *Mobile Image*. Edita.
- Kotthoff, H. (2017). Erzählen in Gesprächen: Eine Einführung in die konversationsanalytische Erzählforschung mit Übungsaufgaben. *Freiburger Arbeitspapiere zur Germanistischen Linguistik*, 38, 1–77.
- Kotthoff, H. (2018). Nicht ausgebaute Erzählungen: Befunde aus Kontexten des Spracherwerbs und der institutionellen Kommunikation. *Zeitschrift für Literaturwissenschaft und Linguistik*, 48, 225–248. <https://doi.org/10.1007/s41244-018-0091-9>
- Kress, G. R., & Leeuwen, T. V. (Hrsg.). (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Oxford UP.

- Krotz, F. (1995). *Die Mediatisierung kommunikativen Handelns: Wie sich Alltag und soziale Beziehungen, Kultur und Gesellschaft durch die Medien wandeln*. Westdeutscher Verlag.
- Krotz, F. (2007). *Mediatisierung: Fallstudien zum Wandel von Kommunikation*. VS Verlag für Sozialwissenschaften.
- Krotz, F. (2008). Kultureller und gesellschaftlicher Wandel im Kontext des Wandels von Medien und Kommunikation. In T. Thomas (Hrsg.), *Medienkultur und soziales Handeln* (S. 43–62). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-90898-4\\_2](https://doi.org/10.1007/978-3-531-90898-4_2)
- Krotz, F., Despotovic, C., & Kruse, M.-M. (Hrsg.). (2014). *Die Mediatisierung sozialer Welten: Synergien empirischer Forschung*. Springer VS.
- Krotz, F., Despotović, C., & Kruse, M.-M. (Hrsg.). (2017). *Mediatisierung als Metaprozess: Transformationen, Formen der Entwicklung und die Generierung von Neuem*. Springer VS.
- Kupetz, M. (2014). Empathy Displays as Interactional Achievements: Multimodal and Sequential Aspects. *Journal of Pragmatics*, 61, 4–34. <https://doi.org/10.1016/j.pragma.2013.11.006>
- Labov, W. (1971). Das Studium der Sprache im sozialen Kontext. In W. Klein & D. Wunderlich (Hrsg.), *Aspekte der Soziolinguistik* (S. 111–194). Athenäum.
- Larsen, J. (2005). Families Seen Sightseeing: Performativity of Tourist Photography. *Space and Culture*, 8(4), 416–434. <https://doi.org/10.1177/1206331205279354>
- Lasen, A. (2005). *Understanding Mobile Phone Users and Usage*. Vodafone Group R and D.
- Lasen, A. (2006). How to Be in Two Places at the Same Time? Mobile Phone Use in Public Places. In J. R. Höflich & M. Hartmann (Hrsg.), *Mobile Communication in Everyday Life: Ethnographic Views, Observations and Reflections* (S. 227–251). Frank & Timme GmbH.
- Latour, B. (1986). Visualisation and Cognition: Drawing Things Together. *Knowledge and Society: Studies in the Sociology of Culture and Present*, 6, 1–40.
- Latour, B. (2009). Die Logistik der immutable mobiles. In J. Döring & T. Thielmann (Hrsg.), *Mediengeographie: Theorie—Analyse—Diskussion* (S. 111–144). transcript Verlag.
- Latour, B., & Woolgar, S. (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton UP.
- Laurier, E. (2001). Why People Say Where They Are During Mobile Phone Calls. *Environment and Planning D: Society and Space*, 19(4), 485–504. <https://doi.org/10.1068/d228t>
- Laurier, E. (2004). Doing Office Work on the Motorway. *Theory, Culture & Society*, 261–277. <https://doi.org/10.1177/0263276404046070>
- Laurier, E., & McGregor, M. (2016). Mediated Pedestrian Mobility: Walking and the Map App. *Mobilities*, 11(1), 117–134. <https://doi.org/10.1080/17450101.2015.1099900>
- Laurier, E., & Philo, C. (2006). Natural Problems of Naturalistic Video Data. In H. Knoblauch & B. Schnettler (Hrsg.), *Video Analysis: Methodology and Methods* (S. 183–192). Peter Lang.
- Lerner, G. H. (1989). Notes on Overlap Management in Conversation: The Case of Delayed Completion. *Western Journal of Speech Communication*, 53(2), 167–177. <https://doi.org/10.1080/10570318909374298>
- Lerner, G. H. (1992). Assisted Storytelling: Deploying Shared Knowledge as a Practical Matter.

- Qualitative Sociology*, 15(3), 247–271. <https://doi.org/10.1007/BF00990328>
- Licoppe, C. (2017). Showing Objects in Skype Video-Mediated Conversations: From Showing Gestures to Showing Sequences. *Journal of Pragmatics*, 110, 63–82. <https://doi.org/10.1016/j.pragma.2017.01.007>
- Licoppe, C., & Figeac, J. (2015). Direct Video Observation of the Uses of Smartphone on the Move: Reconceptualizing Mobile Multi-Activity. In A. de Souza e Silva & M. Sheller (Hrsg.), *Mobile Communication in Hybrid Spaces* (S. 48–64).
- Licoppe, C., & Figeac, J. (2018). Gaze Patterns and the Temporal Organization of Multiple Activities in Mobile Smartphone Uses. *Human–Computer Interaction*, 33(5–6), 311–334. <https://doi.org/10.1080/07370024.2017.1326008>
- Licoppe, C., & Tuncer, S. (2019). The Initiation of Showing Sequences in Video-Mediated Communication. *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion*, 20, 545–571.
- Ling, R. (1997). „One Can Talk about Common Manners!“: The Use of Mobile Telephones in Inappropriate Situations. *Themes in Mobile Telephony Final Report of the COST 248 Home and Work group*.
- Ling, R. (2002, Juli). *The Social Juxtaposition of Mobile Telephone Conversations and Public Spaces*. The Social Consequences of Mobile Telephones, Chunchon, Korea.
- Ling, R. (2004). *The Mobile Connection: The Cell Phone’s Impact on Society*. Morgan Kaufmann.
- Ling, R. (2005). Das Mobiltelefon und die Störung des öffentlichen Raums. In J. Höfllich & J. Gebhardt (Hrsg.), *Mobile Kommunikation: Perspektiven und Forschungsfelder* (S. 115–133). Peter Lang.
- Ling, R. (2008). *New Tech, New Ties: How Mobile Communication is Reshaping Social Cohesion*. MIT Press.
- Ling, R., & Yttri, B. (2002). Hyper-Coordination via Mobile Phones in Norway. In J. E. Katz & M. Aakhus (Hrsg.), *Perpetual Contact: Mobile Communication, Private Talk, Public Performance* (S. 139–168). Cambridge UP.
- Luckmann, T., & Bergmann, J. R. (1995). Reconstructive Genres of Everyday Communication. In U. M. Quasthoff (Hrsg.), *Aspects of Oral Communication* (S. 289–304). De Gruyter.
- Luff, P., & Heath, C. (1998). Mobility in Collaboration. *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*, 305–314. <https://doi.org/10.1145/289444.289505>
- Luhmann, N. (1997). *Die Gesellschaft der Gesellschaft* (1. Aufl Auflage). Suhrkamp.
- Lundby, K. (2009). *Mediatization: Concept, Changes, Consequences*. Peter Lang.
- Lynch, M. (1985). *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. Routledge & Kegan Paul.
- Lynch, M. (2002). From Naturally Occurring Data to Naturally Organized Ordinary Activities: Comment on Speer. *Discourse Studies*, 4(4), 531–537. <https://doi.org/10.1177/14614456020040040801>
- Mandelbaum, J. (1991). Conversational Non-Cooperation: An Exploration of Disattended Complaints. *Research on Language and Social Interaction*, 25(1–4), 97–138. <https://doi.org/10.1080/08351819109389359>
- Mandelbaum, J. (1993). Assigning Responsibility in Conversational Storytelling: The Interactional Construction of Reality. *Text*, 13(2), 247–266.

- Mandelbaum, J. (2013). Storytelling in Conversation. In J. Sidnell & T. Stivers (Hrsg.), *The Handbook of Conversation Analysis* (S. 492–510). Wiley-Blackwell.
- Mantere, E. (2020). What Smartphones, Ethnomethodology, and Bystander Inaccessibility Can Teach Us About Better Design? In F. Loizides, M. Winckler, U. Chatterjee, J. Abdelnour-Nocera, & A. Parmaxi (Hrsg.), *Human Computer Interaction and Emerging Technologies: Workshop Proceedings from the INTERACT 2019 Workshops*. (S. 91–99). Cardiff UP. <https://cardiffuniversitypress.org/site/chapters/e/10.18573/book3.k/>
- Mantere, E., & Raudaskoski, S. (2017). The Sticky Media Device. In A. R. Lahikainen, T. Mälkiä, & K. Repo (Hrsg.), *Media, Family Interaction and the Digitalization of Childhood* (S. 135–154). Edward Elgar Publishing.
- Mantere, E., Raudaskoski, S., & Valkonen, S. (2018). Parental Smartphone Use and „Bystander Ignorance“ on Child Development. *Les Cultures Médiatiques de l'enfance et de La Petite Enfance*, 98–113.
- Mayes, P. (1990). Quotation in Spoken English. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language"*, 14(2), 325–363. <https://doi.org/10.1075/sl.14.2.04may>
- Maynard, D. W. (2013). Defensive Mechanisms: I-mean Prefaced Utterances in Complaint and Other Conversational Sequences. In M. Hayashi, G. Raymond, & J. Sidnell (Hrsg.), *Conversational Repair and Human Understanding* (S. 198–233). Cambridge UP.
- McGregor, M. (2020). *Social Order of the Co-Located Mobile Phone: Practices of Collaborative Mobile Phone Use* [University of Stockholm]. <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-178020>
- McLuhan, M. (1968). *Die magischen Kanäle: Understanding Media*. Econ.
- Meredith, J. (2019). Conversation Analysis and Online Interaction. *Research on Language and Social Interaction*, 52(3), 241–256. <https://doi.org/10.1080/08351813.2019.1631040>
- Meyer, C., & Meier zu Verl, C. (2017). Epistemische Regime der neuen Medien: Eine kultursoziologische Perspektive auf digitale Bildkommunikation. *Navigationen - Zeitschrift Für Medien- und Kulturwissenschaften*, 17(1), 77–94.
- Mondada, L. (2006). Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis. In H.-G. Soeffner, H. Knoblauch, B. Schnettler, & J. Raab (Hrsg.), *Video analysis: Methodology and Methods* (S. 51–67). Peter Lang.
- Mondada, L. (2007). Commentary: Transcript Variations and the Indexicality of Transcribing Practices: *Discourse Studies*, 9(6), 809–821. <https://doi.org/10.1177/1461445607082581>
- Mondada, L. (2008). L' Analyse de „Collections“ de Phénomènes Multimodaux en Linguistique Interactionnelle: À Propos de l'Organisation Systématique des Ressources Gestuelles en Début de Tour. *Cahiers de praxématique*, 50, 21–66.
- Mondada, L. (2009). The Methodical Organization of Talking and Eating: Assessments in Dinner Conversations. *Food Quality and Preference*, 20(8), 558–571. <https://doi.org/10.1016/j.foodqual.2009.03.006>
- Mondada, L. (2012a). Talking and Driving: Multiactivity in the Car. *Semiotica*, 2012(191), 223–256. <https://doi.org/10.1515/sem-2012-0062>
- Mondada, L. (2012b). The Conversation Analytic Approach to Data Collection. In J. Sidnell & T. Stivers (Hrsg.), *The Handbook of Conversation Analysis* (S. 32–56). John Wiley & Sons. <https://doi.org/10.1002/9781118325001.ch3>

- Mondada, L. (2013a). Conversation Analysis: Talk and Bodily Resources for the Organization of Social Interaction. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Hrsg.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction* (Bd. 1, S. 218–226). De Gruyter. <https://doi.org/10.1515/9783110261318>
- Mondada, L. (2013b). Embodied and Spatial Resources for Turn-Taking in Institutional Multi-Party Interactions: Participatory Democracy Debates. *Journal of Pragmatics*, 46(1), 39–68. <https://doi.org/10.1016/j.pragma.2012.03.010>
- Mondada, L. (2014a). The Temporal Orders of Multiactivity: Operating and Demonstrating in the Surgical Theater. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Hrsg.), *Multiactivity in Social Interaction: Beyond Multitasking* (S. 33–75). John Benjamins.
- Mondada, L. (2014b). The Local Constitution of Multimodal Resources for Social Interaction. *Journal of Pragmatics*, 65, 137–156. <https://doi.org/10.1016/j.pragma.2014.04.004>
- Mondada, L. (2016a). Challenges of Multimodality: Language and the Body in Social Interaction. *Journal of Sociolinguistics*, 20(3), 336–366. [https://doi.org/10.1111/josl.1\\_12177](https://doi.org/10.1111/josl.1_12177)
- Mondada, L. (2016b). *Conventions for Multimodal Transcription*. [https://franzoestik.philhist.unibas.ch/fileadmin/user\\_upload/franzoestik/mondada\\_multimodal\\_conventions.pdf](https://franzoestik.philhist.unibas.ch/fileadmin/user_upload/franzoestik/mondada_multimodal_conventions.pdf)
- Mondada, L. (2018). Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Morek, M. (2018). Multimodalität als Ressource der Platzierung von Erzählungen in Peer-Interaktionen von Kindern. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 19 (Themenheft Erzählen multimodal), 304–345.
- Mortensen, K. K. (2017). Flirting in Online Dating: Giving Empirical Grounds to Flirtatious Implicitness. *Discourse Studies*, 19(5), 581–597. <https://doi.org/10.1177/1461445617715179>
- Müller-Doohm, S. (1993). Visuelles Verstehen: Konzepte kultursoziologischer Bildhermeneutik. In T. Jung & S. Müller-Doohm (Hrsg.), *„Wirklichkeit“ im Deutungsprozess: Verstehen und Methoden in den Kultur- und Sozialwissenschaften*. (S. 438–457). Suhrkamp.
- Müller-Doohm, S. (1997). Bildinterpretation als struktural-hermeneutische Symbolanalyse. In R. Hitzler & A. Honer (Hrsg.), *Sozialwissenschaftliche Hermeneutik: Eine Einführung* (S. 81–108). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-663-11431-4\\_4](https://doi.org/10.1007/978-3-663-11431-4_4)
- Müller-Doohm, S., & Neumann-Braun, K. (1995). Kulturinszenierungen: Einleitende Betrachtungen über die Medien kultureller Sinnvermittlung. In S. Müller-Doohm & K. Neumann-Braun (Hrsg.), *Kulturinszenierungen* (S. 9–23). Suhrkamp.
- Murtagh, G. M. (2002). Seeing the „Rules“: Preliminary Observations of Action, Interaction and Mobile Phone Use. In B. Brown, N. Green, & R. Harper (Hrsg.), *Wireless World* (S. 81–91). Springer VS. [https://doi.org/10.1007/978-1-4471-0665-4\\_6](https://doi.org/10.1007/978-1-4471-0665-4_6)
- Nardi, B., & O’Day, V. (1999). *Information Ecologies: Using Technology with Heart*. MIT Press.
- Neumann-Braun, K., & Müller-Doohm, S. (Hrsg.). (2000). *Medien- und Kommunikationssoziologie: Eine Einführung in zentrale Begriffe und Theorien*. Juventa.
- Niemelä, M. (2011). Resonance in Storytelling: Verbal, Prosodic and Embodied Practices of

- Stance Taking. *Oulu B* 95, 112.
- Norman, D. A. (1988). *The Psychology of Everyday Things*. Basic Books.
- Norman, D. A. (1989). *Dinge des Alltags: Gutes Design und Psychologie für Gebrauchsgegenstände*. Campus Verlag.
- Norrick, N. R. (2006). On the Conversational Performance of Narrative Jokes: Toward an Account of Timing. *Humor - International Journal of Humor Research*, 14(3), 255–274. <https://doi.org/10.1515/humr.2001.003>
- Nyiri, K. (Hrsg.). (2005). *A Sense of Place: The Global and the Local in Mobile Communication*. Passagen.
- Ochs, E. (1979). Transcription as Theory. In E. Ochs & B. B. Schieffelin (Hrsg.), *Developmental Pragmatics* (S. 43–72). Academic Press.
- O'Hara, K. P., Massimi, M., Harper, R., Rubens, S., & Morris, J. (2014). Everyday Dwelling with WhatsApp. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1131–1143. <https://doi.org/10.1145/2531602.2531679>
- Oksman, V. (2005). MMS and Its „Early Adopters“ in Finland. In K. Nyiri (Hrsg.), *A Sense of Place: The Global and the Local in Mobile Communication* (S. 349–362). Passagen.
- Oloff, F. (2019a). Das Smartphone als soziales Objekt: Eine multimodale Analyse von initialen Zeigesequenzen in Alltagsgesprächen. In A. Schmidt & C. Marx (Hrsg.), *Interaktion und Medien: Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation* (S. 191–218). Universitätsverlag Winter.
- Oloff, F. (2019b). Some Systematic Aspects of Self-Initiated Mobile Device Use in Face-to-Face Encounters [Diskussionspapier]. *Journal für Medienlinguistik*, 9. <http://dp.jfml.org/2019/opr-oloff-some-systematic-aspects-of-self-initiated-mobile-device-use/>
- Orlikowski, W. J. (2000). Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. *Organization Science*, 11(4), 404–428. <https://doi.org/10.1287/orsc.11.4.404.14600>
- Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits Make Smartphone Use More Pervasive. *Personal and Ubiquitous Computing*, 16(1), 105–114. <https://doi.org/10.1007/s00779-011-0412-2>
- Palen, L., Salzman, M., & Youngs, E. (2001). Discovery and Integration of Mobile Communications in Everyday Life. *Personal and Ubiquitous Computing*, 5(2), 109–122. <https://doi.org/10.1007/s007790170014>
- Palen, L., Salzman, M., & Youngs, E. (2000). Going Wireless: Behavior & Practice of New Mobile Phone Users. *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 201–210. <https://doi.org/10.1145/358916.358991>
- Plant, S. (2001). *On the Mobile: The Effects of Mobile Telephones on Social and Individual Life* [Research Report]. Motorola. <http://classes.dma.ucla.edu/Winter03/104/docs/splant.pdf>
- Pollner, M. (1974). Mundane Reasoning. *Philosophy of the Social Sciences*, 4(1), 35–54. <https://doi.org/10.1177/004839317400400103>
- Pomerantz, A. M. (1978). Compliment Responses: Notes on the Co-Operation of Multiple Constraints. In J. Schenkein (Hrsg.), *Studies in the Organization of Conversational Interaction* (S. 79–112). Academic Press.

- Pomerantz, A. M. (1984). Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes. In J. M. Atkinson & J. Heritage (Hrsg.), *Structures of Social Action* (S. 57–101). Cambridge UP.
- Pomerantz, A. M. (1986). Extreme Case Formulations: A Way of Legitimizing Claims. *Human Studies*, 9(2–3), 219–229. <https://doi.org/10.1007/BF00148128>
- Pomerantz, A. M., & Fehr, B. J. (1997). Conversation Analysis: An Approach to the Study of Social Action as Sense Making Practices. In T. Van Dijk (Hrsg.), *Discourse Studies: A Multidisciplinary Introduction* (2. Aufl., S. 165–190). SAGE.
- Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018). Voice Interfaces in Everyday Life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3173574.3174214>
- Porcheron, M., Fischer, J. E., & Sharples, S. (2016a). Using Mobile Phones in Pub Talk. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1649–1661. <https://doi.org/10.1145/2818048.2820014>
- Porcheron, M., Fischer, J. E., & Sharples, S. (2016b). „Do Animals Have Accents?“. Talking with Agents in Multi-Party Conversation. *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*, 207–219. <https://doi.org/10.1145/2998181.2998298>
- Potter, J. (2004). Discourse Analysis. In M. Hardy & A. Bryman (Hrsg.), *Handbook of Data Analysis* (S. 607–624). SAGE.
- Przybylski, A. K., & Weinstein, N. (2013). Can You Connect With Me Now? How the Presence of Mobile Communication Technology Influences Face-to-Face Conversation Quality. *Journal of Social and Personal Relationships*, 30(3), 237–246. <https://doi.org/10.1177/0265407512453827>
- Psathas, G. (1995). *Conversation Analysis: The Study of Talk-in-Interaction*. SAGE.
- Quasthoff, U. M. (1980). *Erzählen in Gesprächen: Linguistische Untersuchungen zu Strukturen und Funktionen am Beispiel einer Kommunikationsform des Alltags*. Narr.
- Quasthoff, U. M. (2001). Erzählen als interaktive Gesprächsstruktur. In K. Brinker, G. Antos, W. Heinemann, & S. Sager (Hrsg.), *Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung: Bd. 2. Halbband* (S. 1293–1309). De Gruyter.
- Raab, J. (2014). *Erving Goffman* (2. Aufl.). Herbert von Halem Verlag.
- Raclaw, J., DiDomenico, S. M., & Robles, J. S. (2016). Providing Epistemic Support for Assessments through Mobile-Supported Sharing Activities. *Research on Language and Social Interaction*, 49(4), 362–379. <https://doi.org/10.1080/08351813.2016.1199089>
- Raudaskoski, S., Mantere, E., & Valkonen, S. (2019). Smartphone and the Changing Practices of Face-to-Face interaction. *arXiv e-prints, 1910*, arXiv:1910.13175.
- Rawls, A. W. (1987). The Interaction Order Sui Generis: Goffman's Contribution to Social Theory. *Sociological Theory*, 5(2), 136–149. <https://doi.org/10.2307/201935>
- Rawls, A. W. (2002). Editor's Introduction. In H. Garfinkel, *Ethnomethodology's Program: Working Out Durkheim's Aphorism* (S. 1–64). Rowman & Littlefield.
- Relieu, M. (2009). Mobile Phone „Work“: Disengaging and Engaging Mobile Phone Activities with Concurrent Activities. In R. Ling & S. W. Campbell (Hrsg.), *The Reconstruction of Space and Time: Mobile Communication Practices* (S. 215–230). Transaction Publishers.
- Rivière, C. A. (2006). Mobile Camera Phones: A New Form of „Being Together“ in Daily

- Interpersonal Communication. In R. Ling & P. E. Pedersen (Hrsg.), *Mobile Communications: Re-Negotiation of the Social Sphere* (S. 167–186). Springer VS.
- Rivière, C. A., & Licoppe, C. (2005). From Voice to Text: Continuity and Change in the Use of Mobile Phones in France and Japan. In R. Harper, L. Palen, & A. S. Taylor (Hrsg.), *The Inside Text: Social, Cultural and Design Perspectives on SMS* (S. 103–126). Springer VS. [https://doi.org/10.1007/1-4020-3060-6\\_6](https://doi.org/10.1007/1-4020-3060-6_6)
- Roberts, J. A., & David, M. E. (2016). My Life has Become a Major Distraction from my Cell Phone: Partner Phubbing and Relationship Satisfaction Among Romantic Partners. *Computers in Human Behavior*, 54, 134–141. <https://doi.org/10.1016/j.chb.2015.07.058>
- Robinson, J. D. (2007). The Role of Numbers and Statistics within Conversation Analysis. *Communication Methods and Measures*, 1(1), 65–75. <https://doi.org/10.1080/19312450709336663>
- Robles, J. S., DiDomenico, S. M., & Raclaw, J. (2018). Doing Being an Ordinary Technology and Social Media User. *Language & Communication*, 60, 150–167. <https://doi.org/10.1016/j.langcom.2018.03.002>
- Ryave, A. L. (1978). On the Achievement of a Series of Stories. In J. Schenkein (Hrsg.), *Studies in the Organization of Conversational Interaction* (S. 113–132). Academic Press.
- Sacks, H. L. (1963). Sociological Description. *Berkeley Journal of Sociology*, 8, 1–16.
- Sacks, H. L. (1971). Das Erzählen von Geschichten innerhalb von Unterhaltungen. In R. Kjolseth & F. Sack (Hrsg.), *Zur Soziologie der Sprache: Ausgewählte Beiträge vom 7. Weltkongreß der Soziologie* (S. 307–314). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-663-05383-5\\_24](https://doi.org/10.1007/978-3-663-05383-5_24)
- Sacks, H. L. (1972). On the Analyzability of Stories by Children. In J. J. Gumperz & D. H. Hymes (Hrsg.), *Directions in Sociolinguistics: The Ethnography of Communication* (S. 329–345). Holt, Rinehart and Winston.
- Sacks, H. L. (1974). An Analysis of the Course of a Joke's Telling in Conversation. In J. Sherzer & R. Bauman (Hrsg.), *Explorations in the Ethnography of Speaking* (S. 337–353). Cambridge UP.
- Sacks, H. L. (1978). Some Technical Considerations of a Dirty Joke. In J. Schenkein (Hrsg.), *Studies in the Organization of Conversational Interaction* (S. 249–269). Academic Press.
- Sacks, H. L. (1984). Notes on Methodology. In J. Heritage & J. M. Atkinson (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (S. 21–27). Cambridge UP.
- Sacks, H. L. (1985). On Doing „Being Ordinary“. In J. M. Atkinson (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (S. 413–429). Cambridge UP. <http://dx.doi.org/10.1017/CBO9780511665868.024>
- Sacks, H. L. (1987). On the Preference for Agreement and Contiguity in Sequences in Conversation. In G. Button & J. R. E. Lee (Hrsg.), *Talk and Social Organization* (S. 54–69). Multilingual Matters.
- Sacks, H. L. (1995). *Lectures on Conversation* (G. Jefferson, Hrsg.). Wiley-Blackwell.
- Sacks, H. L., & Schegloff, E. A. (2002). Home Position. *Gesture*, 2(2), 133–146.
- Sacks, H. L., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696–735.
- Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist*,

- 70(6), 1075–1095. <https://doi.org/10.1525/aa.1968.70.6.02a00030>
- Schegloff, E. A. (1979). Identification and Recognition in Telephone Openings. In G. Psathas (Hrsg.), *Everyday Language: Studies in Ethnomethodology* (S. 23–78). Irvington.
- Schegloff, E. A. (1980). Preliminaries to Preliminaries: „Can I Ask You a Question?“ *Sociological Inquiry*, 50(2–3), 104–152.
- Schegloff, E. A. (1982). Discourse as an Interactional Achievement: Some Uses of ‘uh huh’ and Other Things that Come Between Sentences. In D. Tannen (Hrsg.), *Analyzing Discourse: Text and Talk* (S. 71–93). Georgetown UP.
- Schegloff, E. A. (1984). On Some Questions and Ambiguities in Conversation. In J. M. Atkinson (Hrsg.), *Structures of Social Action: Studies in Conversation Analysis* (S. 28–52). Cambridge UP. <https://doi.org/10.1017/CBO9780511665868.006>
- Schegloff, E. A. (1987a). Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis. *Social Psychology Quarterly*, 50(2), 101–114. <https://doi.org/10.2307/2786745>
- Schegloff, E. A. (1987b). Recycled Turn Beginnings: A Precise Repair Mechanism in Conversation’s Turn-Taking Organization. In G. Button & J. R. E. Lee (Hrsg.), *Talk and Social Organisation* (S. 70–85). Multilingual Matters.
- Schegloff, E. A. (1987c). Some Sources of Misunderstanding in Talk-in-Interaction. *Linguistics*, 25(1), 201–218. <https://doi.org/10.1515/ling.1987.25.1.201>
- Schegloff, E. A. (1988). Goffman and the Analysis of Conversation. In T. Wooton & P. Drew (Hrsg.), *Erving Goffman: Exploring the Interaction Order* (S. 89–135). Polity Press.
- Schegloff, E. A. (1991). Reflections on Talk and Social Structure. In D. Boden & D. H. Zimmerman (Hrsg.), *Talk and Social Structure* (S. 44–70). Polity Press.
- Schegloff, E. A. (1992). Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. *American Journal of Sociology*, 97(5), 1295–1345.
- Schegloff, E. A. (1995). Introduction. In H. L. Sacks (Hrsg.), *Lectures on Conversation* (S. ix–lxiii). Blackwell Publishers.
- Schegloff, E. A. (1996a). Confirming Allusions: Toward an Empirical Account of Action. *American Journal of Sociology*, 102(1), 161–216.
- Schegloff, E. A. (1996b). Turn Organization: One Intersection of Grammar and Interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Hrsg.), *Interaction and Grammar* (S. 52–133). Cambridge UP.
- Schegloff, E. A. (1996c). Issues of Relevance for Discourse Analysis: Contingency in Action, Interaction and Co-Participant Context. In E. H. Hovy & D. R. Scott (Hrsg.), *Computational and Conversational Discourse* (S. 3–35). Springer VS. [https://doi.org/10.1007/978-3-662-03293-0\\_1](https://doi.org/10.1007/978-3-662-03293-0_1)
- Schegloff, E. A. (1997). Whose Text? Whose Context? *Discourse & Society*, 8(2), 165–187. <https://doi.org/10.1177/0957926597008002002>
- Schegloff, E. A. (2002). Beginnings in the Telephone. In J. E. Katz & M. Aakhus (Hrsg.), *Perpetual Contact: Mobile Communication, Private Talk, Public Performance* (S. 284–300). Cambridge UP. <http://dx.doi.org/10.1017/CBO9780511489471.022>
- Schegloff, E. A. (2005). On Complainability. *Social Problems*, 52(4), 449–476. <https://doi.org/10.1525/sp.2005.52.4.449>
- Schegloff, E. A. (2007a). *Sequence Organization in Interaction: Volume 1. A Primer in Conversation Analysis*. Cambridge UP.

- Schegloff, E. A. (2007b). A tutorial on membership categorization. *Journal of Pragmatics*, 39(3), 462–482. <https://doi.org/10.1016/j.pragma.2006.07.007>
- Schegloff, E. A., Jefferson, G., & Sacks, H. L. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, 53(2), 361–382. <https://doi.org/10.2307/413107>
- Schegloff, E. A., & Sacks, H. L. (1973). Opening up Closings. *Semiotica*, 8(4), 289–327.
- Schegloff, E. A., & Sacks, H. L. (1979). Two Preferences in the Organization of Reference to Persons in Conversation and Their Interaction. In G. Psathas (Hrsg.), *Everyday Language: Studies in Ethnomethodology* (S. 15–21). Irvington Publishers.
- Schenkein, J. (1978). Sketch of an Analytic Mentality for the Study of Conversational Interaction. In J. Schenkein (Hrsg.), *Studies in the Organization of Conversational Interaction* (S. 1–6). Academic Press.
- Schlote, I., & Linke, C. (2010). Interaction and Individual Patterns of Mobile Communication in Public Places. In J. R. Höflich, G. F. Kircher, C. Linke, & I. Schlote (Hrsg.), *Mobile Media and the Change of Everyday Life* (S. 97–128). Peter Lang.
- Schlote, I., & Linke, C. (2014). *Bewegter Alltag: Eine kontextsensitive Betrachtung von Mobilität und Medienkommunikation*. Springer VS.
- Schmidt, G., & Androutsopoulos, J. (2004). „löbbe döch“: Beziehungskommunikation mit SMS. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 5, 51–76.
- Schmitt, R., & Deppermann, A. (2009). „damit sie mich verstehen“: Genese, Verfahren und recipient design einer narrativen Performance. In S. Jautz, S. Habscheid, & M. Buss (Hrsg.), *Theatralität des sprachlichen Handelns: Eine Metaphorik zwischen Linguistik und Kulturwissenschaften* (S. 79–107). Fink. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/277>
- Schmitt, R., & Knöbl, R. (2013). Recipient Design aus Multimodaler Sicht. *Deutsche Sprache. Zeitschrift für Theorie, Praxis, Dokumentation*, 41(3), 242–276.
- Scholtz, H. (2020). *Mediensoziologie: Eine systematische Einführung*. Springer VS.
- Schu, J. (2001). Formen der Elizitation und das Problem der Natürlichkeit von Gesprächen. In K. Brinker, G. Antos, W. Heinemann, & S. Sager (Hrsg.), *Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung: Bd. 2. Halbband* (S. 1013–1021). De Gruyter.
- Schubert, D. (2009). *Lästern: Eine kommunikative Gattung des Alltags* (1. Aufl.). Peter Lang.
- Schulz, W. (2004). Reconstructing Mediatization as an Analytical Concept. *European Journal of Communication*, 19(1), 87–101.
- Schütz, A. (1971). *Gesammelte Aufsätze Band 1: Das Problem der sozialen Wirklichkeit*. Nijhoff.
- Schütz, A., & Luckmann, T. (2017). *Strukturen der Lebenswelt* (2., überarb. Auflage). UVK.
- Schwitalla, J. (1991). Das Illustrieren: Eine narrative Textsorte mit zwei Varianten. In J. Dittmann, H. Kästner, & J. Schwitalla (Hrsg.), *Erscheinungsformen der deutschen Sprache: Literatursprache, Alltagssprache, Gruppensprache, Fachsprache. Festschrift zum 60. Geburtstag von Hugo Steger* (S. 189–204). Erich Schmidt.
- Selting, M. (2010). Affectivity in conversational storytelling: An analysis of displays of anger or indignation in complaint stories. *Pragmatics*, 20(2), 229–277. <https://doi.org/10.1075/prag.20.2.06sel>
- Selting, M. (2012). Complaint Stories and Subsequent Complaint Stories with Affect Displays.

- Journal of Pragmatics*, 44(4), 387–415. <https://doi.org/10.1016/j.pragma.2012.01.005>
- Selting, M. (2017). The Display and Management of Affectivity in Climaxes of Amusing Stories. *Journal of Pragmatics*, 111, 1–32. <https://doi.org/10.1016/j.pragma.2017.01.008>
- Selting, M., Auer, P., Barden, B., Bergmann, J. R., Couper-Kuhlen, E., Günthner, S., Meier, C., Quasthoff, U. M., Schlobinski, P., & Uhlmann, S. (1998). Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte*, 91–122.
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J. R., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U. M., Schütte, W., ... Uhlmann, S. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 10, 353–402.
- Sharrock, W., & Coulter, J. (1998). On What We Can See. *Theory & Psychology*, 8(2), 147–164. <https://doi.org/10.1177/0959354398082001>
- Shotter, J. (1983). „Duality of Structure“ and „Intentionality“ in an Ecological Psychology. *Journal for the Theory of Social Behaviour*, 13(1), 19–44. <https://doi.org/10.1111/j.1468-5914.1983.tb00460.x>
- Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language and Social Interaction*, 39(4), 377–409. [https://doi.org/10.1207/s15327973rlsi3904\\_2](https://doi.org/10.1207/s15327973rlsi3904_2)
- Sidnell, J. (2012). Basic Conversation Analytic Methods. In J. Sidnell & T. Stivers (Hrsg.), *The Handbook of Conversation Analysis* (S. 77–99). Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch5>
- Sidnell, J., & Stivers, T. (Hrsg.) (2012). *The Handbook of Conversation Analysis*. Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch5>
- Silverman, D. J. (1998). *Harvey Sacks: Social Science & Conversation Analysis*. Oxford Univ Pr.
- Smith, D. E. (1978). ‚K is Mentally Ill‘: The Anatomy of a Factual Account. *Sociology*, 12(1), 23–53. <https://doi.org/10.1177/003803857801200103>
- Speer, S. A., & Hutchby, I. (2003). From Ethics to Analytics: Aspects of Participants’ Orientations to the Presence and Relevance of Recording Devices. *Sociology*, 37(2), 315–337. <https://doi.org/10.1177/0038038503037002006>
- Spreckels, J. (2008). Identity Negotiation in Small Stories Among German Adolescent Girls. *Narrative Inquiry*, 18(2), 393–413. <https://doi.org/10.1075/ni.18.2.11spr>
- Stec, K., Huiskes, M., & Redeker, G. (2015). Multimodal Analysis of Quotation in Oral Narratives. *Open Linguistics*, 1(1). <https://doi.org/10.1515/opli-2015-0018>
- Stec, K., Huiskes, M., & Redeker, G. (2016). Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics*, 104, 1–17. <https://doi.org/10.1016/j.pragma.2016.07.008>
- Stivers, T. (2008). Stance, Alignment, and Affiliation During Storytelling: When Nodding Is a Token of Affiliation. *Research on Language and Social Interaction*, 41(1), 31–57. <https://doi.org/10.1080/08351810701691123>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., Ruitter, J. P. de, Yoon, K.-E., & Levinson, S. C. (2009). Universals and Cultural Variation in Turn-Taking in Conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>

- Stivers, T., & Sidnell, J. (2005). Introduction: Multimodal Interaction. *Semiotica*, 2005(156), 1–20.
- Streeck, J. (1993). Gesture as Communication I: Its Coordination with Gaze and Speech. *Communication Monographs*, 60(4), 275–299. <https://doi.org/10.1080/03637759309376314>
- Stukenbrock, A. (2009). Herausforderungen der multimodalen Transkription: Methodische und theoretische Überlegungen aus der wissenschaftlichen Praxis. In K. Birkner & A. Stukenbrock (Hrsg.), *Die Arbeit mit Transkripten in Fortbildung, Lehre und Forschung* (S. 144–169.). Verlag für Gesprächsforschung.
- Stukenbrock, A. (2014). Pointing to an ‚Empty‘ Space: Deixis am Phantasma in Face-to-Face Interaction. *Journal of Pragmatics*, 74, 70–93. <https://doi.org/10.1016/j.pragma.2014.08.001>
- Suchman, L. (1985). *Plans and Situated Actions: The Problem of Human-Machine Communication*. Xerox.
- Suderland, D. (2019). „oh isch FIND\_s nich;“: Eine konversationsanalytische Untersuchung sprachlicher Bezugnahmen auf smartphone-gestützte Suchanfragen in Alltagsgesprächen. *Journal für Medienlinguistik*, 2(2), 90–122. <https://doi.org/10.21248/jfml.2019.17>
- Tannen, D. (1989). *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge UP.
- Taylor, A. S. (2005). Phone Talk. In *Mobile Communications: Re-Negotiation of the Social Sphere* (S. 149–166). Springer VS. [https://doi.org/10.1007/1-84628-248-9\\_10](https://doi.org/10.1007/1-84628-248-9_10)
- Taylor, A. S., & Harper, R. (2003). The Gift of the Gab?: A Design Oriented Sociology of Young People’s Use of Mobiles. *Computer Supported Cooperative Work (CSCW)*, 12(3), 267–296. <https://doi.org/10.1023/A:1025091532662>
- ten Have, P. (2007). *Doing Conversation Analysis*. SAGE.
- Thiedeke, U. (2012). *Soziologie der Kommunikationsmedien: Medien - Formen - Erwartungen*. VS Verlag für Sozialwissenschaften.
- Thompson, J. B. (1995). *The Media and Modernity: A Social Theory of the Media*. Stanford UP.
- Tifentale, A., & Manovich, L. (2015). Selfiecity: Exploring Photography and Self-Fashioning in Social Media. In D. M. Berry & M. Dieter (Hrsg.), *Postdigital Aesthetics: Art, Computation and Design* (S. 109–122). Palgrave Macmillan UK. [https://doi.org/10.1057/9781137437204\\_9](https://doi.org/10.1057/9781137437204_9)
- Torres Cajo, S. (2017). „das is SO lächerlich; ohne SCHEISS jetz ma“: Zur affektiven Äußerungsmodalisierung durch ohne Scheiß-Konstruktionen im gesprochenen Deutsch. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 18, 223–240.
- Traverso, V. (2009). The Dilemmas of Third-Party Complaints in Conversation Between Friends. *Journal of Pragmatics*, 41(12), 2385–2399. <https://doi.org/10.1016/j.pragma.2008.09.047>
- Tuncer, S. (2016). The Effects of Video Recording on Office Workers’ Conduct, and the Validity of Video Data for the Study of Naturally-Occurring Interactions. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(3), 1–23. <https://doi.org/10.17169/fqs-17.3.2604>
- Urry, J. (2000). *Sociology Beyond Societies: Mobilities for the Twenty-first Century*. Routledge.
- Vanden Abeele, M. M. P., Hendrickson, A. T., Pollmann, M. M. H., & Ling, R. (2019). Phubbing Behavior in Conversations and its Relation to Perceived Conversation Intimacy

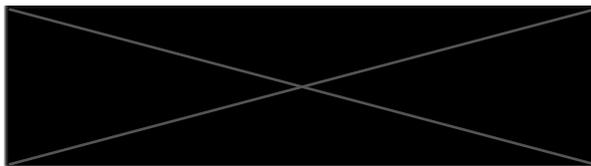
- and Distraction: An Exploratory Observation Study. *Computers in Human Behavior*, 100, 35–47. <https://doi.org/10.1016/j.chb.2019.06.004>
- Vorderer, P., Hefner, D., Reinecke, L., Klimmt, C., Hefner, D., Reinecke, L., & Klimmt, C. (2017). *Permanently Online, Permanently Connected: Living and Communicating in a POPC World*. Routledge. <https://doi.org/10.4324/9781315276472>
- Vyas, D., Chisalita, C. M., & van der Veer, G. C. (2006). Affordance in Interaction. *Proceedings of the 13th European Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-technical Systems*, 92–99. <https://doi.org/10.1145/1274892.1274907>
- Waber, B. (1993). Die Bildanalyse in den Sozialwissenschaften: Das Beispiel der typisierenden Strukturanalyse topographischer Karten. *Soziographie*, 6, 77–96.
- Wang, X., Xie, X., Wang, Y., Wang, P., & Lei, L. (2017). Partner Phubbing and Depression among Married Chinese Adults: The Roles of Relationship Satisfaction and Relationship Length. *Personality and Individual Differences*, 110, 12–17. <https://doi.org/10.1016/j.paid.2017.01.014>
- Weber, H. (2008). *Das Versprechen mobiler Freiheit: Zur Kultur- und Technikgeschichte von Kofferradio, Walkman und Handy* (1. Aufl.). transcript Verlag.
- Weilenmann, A. (2003). „I Can’t Talk Now, I’m in a Fitting Room“: Formulating Availability and Location in Mobile-Phone Conversations. *Environment and Planning A*, 35(9), 1589–1605. <https://doi.org/10.1068/a34234>
- Weilenmann, A., & Hillman, T. (2019). Selfies in the Wild: Studying Selfie Photography as a Local Practice. *Mobile Media & Communication*, 1–20. <https://doi.org/10.1177/2050157918822131>
- Weilenmann, A., Hillman, T., & Jungselius, B. (2013). Instagram at the Museum. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1843–1852. <https://doi.org/10.1145/2470654.2466243>
- Weilenmann, A., & Larsson, C. (2000, 25.10). Collaborative Use of Mobile Telephones. *NordiCHI2000 Proceedings*. <http://www.comlab.hut.fi/opetus/040/weilenmann.pdf>
- Weilenmann, A., & Larsson, C. (2002). Local Use and Sharing of Mobile Phones. In B. Brown, N. Green, & R. Harper (Hrsg.), *Wireless World: Social and Interactional Aspects of the Mobile Age* (S. 92–107). Springer VS. [https://doi.org/10.1007/978-1-4471-0665-4\\_7](https://doi.org/10.1007/978-1-4471-0665-4_7)
- Wiggins, S., & Potter, J. (2003). Attitudes and Evaluative Practices: Category vs. Item and Subjective vs. Objective Constructions in Everyday Food Assessments. *The British Journal of Social Psychology*, 42(4), 513–531. <https://doi.org/10.1348/014466603322595257>
- Wimmer, J., & Hartmann, M. (Hrsg.). (2014). *Medienkommunikation in Bewegung: Mobilisierung – Mobile Medien – Kommunikative Mobilität*. Springer VS.
- Wooffitt, R. (1992). *Telling Tales of the Unexpected: The Organization of Factual Discourse*. Rowman & Littlefield.
- Wyss, E., & Hug, B. (2016). WhatsApp-Chats: Neue Formen der Turnkoordination bei räumlich-visueller Begrenzung. In C. Spiegel & D. Gysin (Hrsg.), *Jugendsprache in Schule, Medien und Alltag* (S. 259–274). Peter Lang.
- Ziemann, A. (2014). *Soziologie der Medien* (2. überarb. Aufl.). transcript Verlag.
- Zillien, N. (2008). Die (Wieder-)Entdeckung der Medien: Das Affordanzkonzept in der Mediensoziologie. *Sociologia Internationalis*, 46(2), 161–181. <https://doi.org/10.3790/sint.46.2.161>

- Zima, E. (2017). Multimodale Mittel der Rederechtaushandlung im gemeinsamen Erzählen in der Face-to-Face Interaktion. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 18, 241–273.
- Zimmerman, D. H. (1998). Identity, Context and Interaction. In C. Antaki & S. M. Widdicombe (Hrsg.), *Identities in Talk* (S. 87–106). SAGE.

## **Eidesstattliche Versicherung**

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen.

Außerdem erkläre ich hiermit, dass die elektronische und die schriftliche Fassung dieser Arbeit übereinstimmen.



David Suderland

Mainz, den 12.01.2021