



<https://bulletin.vals-asla.ch> · e-ISSN: 2813-7973 · p-ISSN: 1023-2044

Tschannen, J. (2023). Multimodale beendigungsstrategien in audio-postings. *Bulletin suisse de linguistique appliquée*, 117, 207-225. <https://doi.org/10.5169/seals-1063001>

This article is publish under a *Creative Commons Attribution 4.0 International* (CC BY): <https://creativecommons.org/licenses/by/4.0>



© Jana Tschannen, 2023

Multimodale beendigungsstrategien in audio-postings

Jana TSCHANNEN

Universität Basel, Deutsches Seminar
Nadelberg 4, 4051 Basel, Suisse
jana.tschannen@unibas.ch
ORCID: 0009-0007-3175-9609

With voice messages, spoken language is integrated into chat communication that has so far tended to be written. In public discourse, the ends of voice messages are often caricatured; the closing of voice messages apparently represents a communicative problem.

In this paper, I use a conversation analytical approach to show that users do not simply deactivate the microphone for this, but instead use multimodal resources to process the closing and further organization of the text-form-based interaction (Beißwenger 2020). It will be shown that users sometimes use proven grammatical and prosodic resources as strategies, but sometimes also develop new, voice message-specific strategies to cope with the communication situation.

Keywords: digital Conversation Analysis, text-form-based interaction, voice message, messenger chat, temporality, multimodality, prosody, answering machine communication.

Stichwörter: Digitale Konversationsanalyse, Textformen-basierte Interaktion, Sprachnachricht, Messenger-Chat, Zeitlichkeit, Multimodalität, Prosodie, Anrufbeantworterkommunikation.

1. Einleitung

Über WhatsApp¹ werden nach eigenen Angaben² weltweit täglich sieben Milliarden Sprachnachrichten³ verschickt; daraus lässt sich schliessen, dass Audio-Postings für viele Menschen zum kommunikativen Alltag gehören. Gleichzeitig stören sich einige Menschen daran und diese Abneigung wird medial regelmässig aufgegriffen. In allen sozialen Medien lassen sich ironische, belustigte und kritische Kommentare finden, die sich u.a. auf die sprachliche Ausgestaltung von Audio-Postings beziehen. Es finden sich dazu vor allem selbstironische Kommentare aus Sicht der Produzierenden.

¹ WhatsApp-Audio-Postings bilden das Korpus des vorliegenden Papers. Die Funktion gibt es auf vielen Kommunikationsplattformen wie bspw. WeChat, Telegram, Facebook oder Instagram. Die Produktions- und Rezeptionsbedingungen unterscheiden sich teilweise, z.B. sind Audio-Postings bei Instagram, Facebook oder WeChat auf eine Minute limitiert.

² Vgl. z.B. auf dem WhatsApp-Blog: <https://blog.whatsapp.com/making-voice-messages-better?lang=de> [18.06.2023].

³ WhatsApp selbst und auch viele Nutzer*innen verwenden den Begriff *Sprachnachricht*. König & Hector (2019) stellen fest, dass dieser Begriff aus linguistischer Sicht etwas unpassend sei, da Sprache auch per Textnachricht versendet werde und zudem nicht nur Sprache aufgenommen und versendet würde, sondern beispielsweise auch Geräusche (ebd. 2017). Sie schlagen daher in Anlehnung an Beißwenger (2015) den Begriff *Audio-Posting* vor, den ich im Folgenden übernehmen werde.



Abb. 1a und b: Screenshots zweier Beispiele für selbstironische Kommentare.

Aus Memes wie *Was wir am Ende einer Sprachnachricht sagen: 1% "Tschüss.", 99% "und ..joaa..."* (Abb. 1a; vgl. auch König 2020: 253) oder *Sprachnachrichten beenden: Keiner: Sich am Ende verabschieden und "liebe Grüße" wünschen. Alle: Einfach nur "ja" stammeln, weil man nicht weiß, was man sonst noch sagen soll.* (Abb. 1b)⁵ lässt sich schliessen, dass am Ende von Audio-Postings eigentlich eine Beendigungsformulierung wie *Tschüss* oder *Bis dann* erwartet wird, aber die aufnehmenden Personen diese eigene Erwartung oft nicht erfüllen. Die Vielzahl dieser Memes, die sich für Eröffnungen von Audio-Postings in diesem Ausmass nicht finden lassen, deuten darauf hin, dass die Beendigung von Audio-Postings offenbar ein kommunikatives Problem darstellt, dessen sich Nutzer*innen bewusst sind.

Aus technischer Sicht lassen sich Audio-Postings direkt auf der Chat-Oberfläche über das Mikrofon-Symbol aufnehmen und versenden. Die Aufzeichnung läuft so lange⁶, bis das Symbol wieder losgelassen wird. Alternativ lässt sich das Mikrofon über eine Feststelltaste einrasten und durch das Tippen auf den Sende-Button versenden. Das Beenden und Versenden eines Audio-Postings funktioniert also technisch ähnlich wie das Versenden eines Text-Postings.

Wie aus linguistischer Forschung zu Anrufbeantworter-Kommunikation hervorgeht (vgl. dazu bspw. Alvarez-Caccamo & Knoblauch 1992; Liddicoat 199; Dingwall 1995; Lange 1999; Knirsch 2005), kommt es auch bei Nachrichten, die Anrufende hinterlassen (im Folgenden *AB-Nachricht*) am Ende zu

⁴ <https://www.pinterest.de/pin/31103053665451152/> [18.06.2023].

⁵ <https://www.picuki.com/media/2327765198289172847> [18.06.2023].

⁶ Seit Anfang 2022 lässt sich die Aufnahme zwischendurch pausieren; das war zum Zeitpunkt meiner Datenerhebung noch nicht möglich.

kommunikativen Problemen. Knirsch (2005: 188) sieht in der Beendigung von AB-Nachrichten das Problem, dass

dort im Falle des Telefongesprächs im Gegensatz zur telefonischen Gesprächseröffnung ein potenziell höherer Aushandlungsbedarf gegeben ist, der dann auf dem AB nicht erfüllt werden kann.

Die kommunikativen Probleme in AB-Nachrichten lassen sich auf die dialogisch ausgerichtete, gesprochene Sprache zurückführen, die asynchron verwendet wird. Es ist naheliegend, diese Ursache auch für Audio-Postings anzunehmen, auch wenn sie einige Unterschiede zu AB-Nachrichten aufweisen: erstere sind intendiert, lassen sich in einen visuellen Austausch von Postings integrieren und von dort aus direkt rezipieren, so dass eine wechselseitige Kommunikationssituation entsteht. AB-Nachrichten sind meist nicht die intendierte Kommunikationsform. Auch eine Wechselseitigkeit ist nur beschränkt gegeben: Anrufende werden mit einem vorab aufgenommenen und automatisch abgespielten First Pair Part konfrontiert, der sich i.d.R. an alle Anrufenden richten, also prinzipiell mehrfachadressiert ist (Knirsch 2005: 52). Nach dem Hinterlassen der AB-Nachricht muss die Kommunikationsform gewechselt werden, um die Kommunikation weiterzuführen.

Im Folgenden werde ich zeigen, dass ähnliche Beendigungsprobleme wie in AB-Nachrichten auch in Audio-Postings vorkommen können, die sich auf die Affordanzen der Kommunikationsform Messenger-Chat zurückführen lassen, die ich in Abschnitt 2 beschreibe. Ich werde dann in Anlehnung an Beißwenger (2020) dafür argumentieren, den Austausch von Text- und Audio-Postings als Textformen-basierte Interaktion zu betrachten (Abschn. 3). Davon ausgehend werde ich eine konversationsanalytische Herangehensweise reflektieren, mit der ich mein Korpus untersucht habe (Abschn. 4). Ich werde dann zeigen, welche multimodalen Strategien Nutzer*innen für die Beendigung ihrer Audio-Postings einsetzen und auf welche Ressourcen sie dabei zurückgreifen (Abschn. 5).

2. Affordanzen der Kommunikationsform Messenger-Chat

In Anlehnung an König & Hector (2017: 7; vgl. auch Herring 2015) lässt sich WhatsApp als Kommunikationsplattform bezeichnen, über die unterschiedliche Kommunikationsformen realisiert werden können. Nach Dürscheid (2005: 8) lassen sich sechs Merkmale anführen, durch die sich verschiedene Kommunikationsformen unterscheiden: der Zeichentyp, die Kommunikationsrichtung, die Anzahl der Kommunikationspartner*innen, die räumliche sowie die zeitliche Dimension und schliesslich das Kommunikationsmedium, mit dem hier das technische Endgerät gemeint ist. Anhand dieser Merkmale lassen sich bei WhatsApp verschiedene Kommunikationsformen ausmachen; dazu zählen z.B. Video- oder Audio-Telefonie, Statusmitteilungen und Chats. Alle Kommunikationsformen unterliegen spezifischen Affordanzen, also bestimmten

technischen Handlungsmöglichkeiten (vgl. boyd 2014: 10), die die Kommunikation zwar nicht determinieren, aber beeinflussen können (vgl. König & Duan i.V.).

Im Zusammenhang mit den Affordanzen gehe ich auf drei der oben genannten Merkmale der Kommunikationsform Messenger-Chat genauer ein, die sich in Bezug auf multimodale Beendigungsstrategien als relevant erweisen werden.

2.1 Zeichentyp

Die Kommunikationsform Messenger-Chat zeichnet sich dadurch aus, dass sie nicht auf einen Zeichentyp beschränkt ist. Im Vergleich zu traditionellen Chats⁷, für die Dürscheid (2005: 8) nur geschriebene Sprache als Zeichentyp aufführt, listet Imo (2015: 7) für mobile Messenger-Chats daneben auch die Zeichentypen Bild, Audio und Video auf. Mit Zeichentypen sind sowohl bei Dürscheid als auch bei Imo weniger verschiedene Zeichenmodalitäten gemeint (verbale oder nonverbale Zeichen), sondern eher unterschiedliche Sinnesmodalitäten (visuell, auditiv, audiovisuell) (vgl. für diese Unterscheidung Imo & Lanwer 2019: 247). Da der Modalitätsbegriff innerhalb der Linguistik teilweise Zeichenmodalitäten, teilweise Sinnesmodalitäten umfasst (vgl. z.B. Mondada 2014; Imo & Lanwer 2019: 246-248), lassen sich je nach Definition sowohl die Kommunikationsform Messenger-Chat als auch Audio-Postings als multimodal bezeichnen. Im Folgenden nutze ich den Multimodalitätsbegriff im Sinne von Zeichenmodalitäten, wie es innerhalb der Konversationsanalyse üblich ist (vgl. Mondada 2014: 138) und beziehe ihn nicht auf die Kommunikationsform, sondern ausschliesslich auf Audio-Postings. Audio-Postings bzw. darin eingesetzte Beendigungsstrategien lassen sich demnach als multimodal bezeichnen, da Nutzer*innen neben verbalen auch auf nicht-verbale und prosodische Gestaltungsmittel zurückgreifen können. Inwiefern nicht-verbale Gestaltungsmittel als Ressource eingesetzt werden, beschreiben König & Hector (2017: 28-31) an Beispielen wie die Geräuschkulisse einer Party, Tierlaute (vgl. Hector 2017: 25-26) oder auch Flatulenzgeräusche, die gezielt in Audio-Postings eingesetzt werden. Zu den prosodischen Gestaltungsmitteln⁸ gehören u.a. der Tonhöhenverlauf, Lautstärke, Sprechtempo und Rhythmus (Selting 1995: 1). Imo & Lanwer (2019: 248) bezeichnen beispielsweise Verfahren der Redewiedergabe als multimodal, da neben den lexikalisch-grammatischen Ressourcen auch ein Wechsel der Sprechstimme nutzbar gemacht und damit als prosodische Ressource eingesetzt werden kann. Wie sich zeigen wird, greifen Nutzer*innen sowohl auf verbale als auch auf

⁷ Damit ist die Chat-Kommunikation der 1980er und 1990er gemeint, die hauptsächlich über Internet Relay Chats oder Webchats erfolgte. Für diese traditionelle Chat-Kommunikation war die zeitliche Ko-Präsenz der Chattenden erforderlich, da die Nachrichten nur für den Verlauf der Sitzung gespeichert wurden (vgl. Beißwenger 2007: 22-23).

⁸ Ob prosodische Gestaltungsmittel zu den verbalen Zeichen zählen oder als paraverbale Zeichen eine eigene Kategorie bilden, ist umstritten (vgl. dazu ausführlich Selting 1995: 16-20).

prosodische Ressourcen zurück, um die Beendigung ihres Audio-Postings zu bearbeiten.

2.2 Zeitliche Dimension

Mit der zeitlichen Dimension ist die Zeitspanne gemeint, über die sich die Kommunikation erstreckt und in der Produktion und Rezeption von Beiträgen erfolgen können. Es ist möglich, dass Postings unmittelbar aufeinander folgen und Nutzer*innen gleichzeitig auf den Austausch ausgerichtet sind. In dem Zusammenhang wird die Kommunikationsform Chat häufig als *quasi-synchron* (vgl. Garcia & Baker Jacobs 1999; Dürscheid 2016) bezeichnet. Dieser Begriff verdeutlicht, dass Chats zwar nicht synchron sein können, da Produktion und Rezeption von Postings immer nacheinander und nie zeitgleich erfolgen können, aber auch nicht asynchron sind wie monologische Texte. Der quasi-synchrone Austausch wird häufig als (proto)typisch für Chats dargestellt (vgl. z.B. Dürscheid 2016: 443; Imo 2019: 52) und in vielen konversationsanalytischen Forschungsbeiträgen zur Chat-Kommunikation werden sequenzielle Phänomene häufig auf diesen Zeitlichkeitscharakter zurückgeführt (z.B. *chunking*, also das Aufteilen von Handlungen oder Teilhandlungen auf einzelne Postings oder *Phantomadjazenz*, die vorliegt, wenn zwei Postings visuell aufeinander folgen, aber nicht durch konditionelle Relevanz verbunden sind; vgl. Garcia & Baker Jacobs 1999; Imo 2015; König 2015; Wyss & Hug 2016). Rückblickend lassen sich die Forschungsbeiträge, in denen Chat-Kommunikation mit konversationsanalytischen Methoden untersucht werden, der digitalen Konversationsanalyse zuordnen. Die Entwicklung dieser Forschungsdisziplin basiert nach Meredith et al. (2021: 12) auf der Erkenntnis, dass Konzepte wie Mechanismen des Sprecherwechsels, das Konzept des *turns* sowie adjazente Organisationsstrukturen nicht eins zu eins auf Chat-Kommunikation übertragbar sind. Eine Analyse von Chat-Kommunikation darf nicht in Abhängigkeit zu synchroner Interaktion geschehen, bei der typische Eigenschaften der Chat-Kommunikation als Versuch gesehen werden, Face-to-Face-Bedingungen zu simulieren und auf das Fehlen der visuell-räumlichen Kopräsenz zurückgeführt werden (vgl. Arminen et al. 2016: 294, 302). Stattdessen muss Chat-Kommunikation unter Einbezug der spezifischen Affordanzen, denen die Kommunikation unterliegt, analysiert werden. Beendigungen in Audio-Postings dürfen also nicht als Simulation eines Face-to-Face-Gesprächs gesehen werden, sondern müssen unter Berücksichtigung der Affordanzen, allen voran den Zeitlichkeitsbedingungen, analysiert werden.

Die Möglichkeit des schnellen Austauschs und die damit verbundene dialogische Struktur von Postings führt dazu, dass Chat-Kommunikation von jeher eher als Interaktion denn als Text konzeptualisiert wird; in der deutschsprachigen Forschung schlägt sich das in Begriffen wie *geschriebene Mündlichkeit* (Kilian 2001), *getippte Dialoge* (Dürscheid & Brommer 2009) oder *interaktionsorientiertes Schreiben* (Beißwenger & Storrer 2012) nieder. Virtanen

et al. (2021: 2) zeigen in Bezug auf aktuellere Forschungsergebnisse (vgl. z.B. Baron 2013; Lyons & Tagg 2019; Meredith 2019; König 2021), dass der quasi-synchrone Austausch nicht der Standard ist, sondern auch asynchrone Kommunikation per Messenger-Chat stattfindet. Es lässt sich also festhalten, dass die Möglichkeit zum quasi-synchronen bis hin zum asynchronen Austausch von Postings eine wesentliche Affordanz von Messenger-Chats (auch im Vergleich zu traditionellen Chats) darstellt. Aus technischer Sicht stellt die zeitliche Persistenz von Postings in einem Verlaufsprotokoll die Voraussetzung für den asynchronen Austausch dar (Beißwenger 2020: 307). Die Gründe dafür, wann Nutzer*innen auf welche zeitliche Dimension zurückgreifen, lassen sich jedoch nicht technisch bestimmen, sondern müssen aus den sozialen und situativen Nutzungspraktiken abgeleitet werden (vgl. Androutsopoulos 2007: 92).

2.3 Kommunikationsrichtung

Die Kommunikationsform Messenger-Chat zeichnet sich durch ihren dialogischen Charakter aus. Damit ist nicht nur gemeint, dass Nutzer*innen abwechselungsweise Postings produzieren und rezipieren können, sondern vor allem, dass die Postings aufeinander Bezug nehmen und nur in diesem Bezug einen Sinn ergeben. Entscheidend für den dialogischen Charakter (vgl. zur Unterscheidung von Dialog und Interaktion Abschnitt 3) von Chats ist die Einbettung von Postings und die gemeinsame Hervorbringung von Bedeutung und Struktur. Postings sind global in die Kommunikationsgeschichte der Beteiligten eingebettet; gleichzeitig nehmen sie lokal eine bestimmte Position innerhalb eines konkreten Austauschs ein. In diesem Sinne sind Postings sequenziell geordnet und durch konditionelle Relevanzen miteinander verbunden: Jedes Posting baut auf vorangegangene Postings und/oder die Kommunikationsgeschichte auf und öffnet Erwartungen für bestimmte Folgepostings (vgl. Imo 2016: 66). Zudem sind Postings bzw. sequenzielle Abfolgen von Postings in grössere kommunikative Projekte, wie z.B. *ein Treffen vereinbaren* oder *vom (All)Tag erzählen*, eingebettet, die die Beteiligten gemeinsam bearbeiten. Mit der gemeinsamen Bearbeitung ist zum einen gemeint, dass sich die Beteiligten stetig gegenseitig zeigen, wie sie einander verstehen und verstanden werden wollen (*display*; vgl. Schegloff & Sacks 1973). Zum anderen arbeiten sie auch gemeinsam an der Organisation von Postings, indem sie Bezüge herstellen⁹ und den weiteren Dialogverlauf verhandeln.

3. Interaktion

Dass der Austausch von Postings tatsächlich nach dialogischen Prinzipien verläuft, steht seit Beginn der traditionellen Chat-Forschung und auch der

⁹ Die Bezüge können z.B. verbal oder durch die Replyfunktion hergestellt werden (vgl. Virtanen et al. 2021).

modernen Messenger-Forschung immer wieder im Fokus linguistischer Analysen (vgl. z. B. Schönfeld 2001; Beißwenger 2007; Berglund 2009; Dürscheid & Brommer 2009; Beißwenger & Storrer 2012; Baron 2013; König 2015; Arminen et al. 2016; König & Hector 2019; Androutsopoulos & Busch 2021). Umstritten ist, ob es sich bei einem Austausch von Postings um Interaktion handelt. Grund dafür ist, dass die Begriffe *Dialog* und *Interaktion* nicht einheitlich verwendet werden und unterschiedlich weit gefasst sein können.

Die Interaktionale Linguistik (Selting & Couper-Kuhlen 2000; zusammenfassend Imo & Lanwer 2019) vertritt einen eher weiten Interaktionsbegriff. Interaktion in diesem Sinne basiert auf Dialogizität und ist als Gegenbegriff zum Monolog zu verstehen. Die Begriffe *Interaktion* und *Dialog* können nach dieser Auffassung synonym verstanden werden. Imo & Lanwer (2019: 37) beschreiben die begriffliche Unterscheidung wie folgt:

[M]an [kann] mit dem Dialog sehr griffig das sprachliche >Produkt< einer Interaktion im Kontrast zum Monolog benennen [...] und [hat] mit Dialogizität vs. Monologizität ein anschauliches Gegensatzpaar zur Verfügung [...], das weniger >holprig< als Interaktionalität vs. Monologizität klingt. (Herv. im Org.)

Ein Monolog ist ein sprachliches Produkt, das eine Person allein herstellt, ohne dass eine Reaktion erwartet oder gar ermöglicht wird (Imo & Lanwer 2019: 11-12). Monolog und Interaktion bilden nach dieser Auffassung zwei Pole, dazwischen liegt ein Kontinuum: einzelne Interaktionen, aber auch einzelne Postings oder Sequenzen innerhalb einer Interaktion können mehr oder weniger Dialogizität aufweisen (ebd.: 12). Voraussetzung für eine Interaktion ist die gemeinsame Orientierung an einem kommunikativen Projekt; weder der zeitliche noch der mediale Aspekt (mündlich vs. schriftlich) sind dafür entscheidend (ebd.).

Eine weitaus engere Auffassung vertreten Hausendorf (2015) und Hausendorf et al. (2017) innerhalb der Interaktionslinguistik. Interaktion in diesem engeren Sinne ist an raumzeitliche¹⁰ Anwesenheit und damit verbunden an Wahrnehmungswahrnehmung geknüpft. Mit Wahrnehmungswahrnehmung ist die synchrone und wechselseitige Wahrnehmbarkeit gemeint. Es bedeutet, dass man wahrnimmt, vom Gegenüber wahrgenommen zu werden und umgekehrt (vgl. Hausendorf 2015: 46). Der Gegenbegriff zur Interaktion ist die Textkommunikation. Im Gegensatz zu den Kontrastbegriffen *Interaktion* und *Monolog* wird innerhalb der Interaktionslinguistik zwischen Interaktion und Textkommunikation kein Kontinuum angenommen. Der Austausch von Postings ist nach dieser Auffassung also keine Interaktion, da keine Wahrnehmungswahrnehmung gegeben ist, sondern es handelt sich um Textkommunikation (Hausendorf et al. 2017: 135). Der grundlegende Unterschied zur Interaktion ist

¹⁰ Die einzige Ausnahme ist das Telefonat, weil hier nur eine zeitliche, aber keine räumliche Anwesenheit gegeben ist (Hausendorf 2015: 46).

die fehlende Anwesenheit und damit verbunden die fehlenden sequenziellen Zugzwänge (ebd.: 136).

Angesichts der beiden Perspektiven auf den Austausch von Postings argumentiert Beißwenger (2020: 295) dafür, diesen Austausch als "ein Drittes" neben Interaktion und Text zu konzipieren. Für die Analyse von Audio-Postings übernehme ich Beißwengers Konzipierung, die ich im Folgenden kurz skizziere: Wenn der Begriff *Interaktion* gleichermassen auf synchrone wie auch auf Chat-Kommunikation bezogen werde, wie es im Rahmen der Interaktionalen Linguistik getan wird, so könnten damit die charakteristischen Zeitlichkeitsbedingungen von Chat-Kommunikation nicht erfasst werden. Auch mit dem Textbegriff, wie Hausendorf (2015) ihn im Rahmen der Interaktionslinguistik für Kommunikation unter Lesbarkeitsbedingungen definiert, könne die Zeitlichkeitsbedingung nicht angemessen beschrieben werden. Nach Beißwenger (2020: 296) ist die Spezifik von Chat-Kommunikation weder als "Interaktion mit Abstrichen" noch als "Textkommunikation mit gewissen Besonderheiten" zufriedenstellend erfasst.

Da die Chat-Kommunikation aber sowohl Merkmale der Interaktion als auch der Textkommunikation teilt, schlägt Beißwenger vor, von *Textformen-basierter Interaktion* auszugehen. Damit entwickelt er zum einen das Interaktionskonzept der Interaktionalen Linguistik weiter und macht es fruchtbar für die Beschreibung des sequenziell geordneten Austauschs von Postings. Zum anderen betont Beißwenger (2020: 297; Herv. im Org.) damit

die neuen, ganz und gar Interaktions-untypischen Möglichkeiten der Gestaltung und Organisation kommunikativen Austauschs, die dann erkennbar werden, wenn man die ausgetauschten Produkte und die Strukturen, in die sie eingebettet sind, unter dem Gesichtspunkt ihrer Textformenhaftigkeit (und nicht nur ihrer medialen Schriftlichkeit) betrachtet.

Nach Beißwenger (2020: 305) ist jedes Posting eine Textform, die "als flächiges Ganzes am Bildschirm präsentiert und dort visuell gegen Einheiten gleichen Typs abgegrenzt wird". Eine weitere Textform stellt das Verlaufsprotokoll dar,

das die Verarbeitungsgrundlage bildet und in dem Postings [...] in eine lineare, ggf. auch in Threads verzweigende Ordnung gebracht und in dieser Organisation für die wiederholte Rezeption vorgehalten werden. (ebd.)

Mit dem Konzept der Textformen-basierten Interaktion kann erfasst werden, dass sich Nutzer*innen zwar an einer synchronen Interaktion orientieren; Produktion und Rezeption von Postings sind aber jeweils zeitlich entkoppelt, sodass die Sequenzorganisation damit eine individuelle Leistung darstellt und nicht in gleichem Masse kollaborativ stattfindet wie in einer synchronen Interaktion (ebd.: 301).

4. Methode und Korpus

Die Konzipierung von Chat-Kommunikation als Textformen-basierte Interaktion ermöglicht es, den Fokus auf das gemeinsame, aber zeitversetzte Handeln der

Beteiligten zu legen. Die Frage nach den Ressourcen, auf die Interagierende bei der Bewältigung von Beendigungen in Audio-Postings unter den in Abschnitt 2 beschriebenen Affordanzen zurückgreifen, reiht sich in die Forschungstradition der digitalen Konversationsanalyse (*Digital Conversation Analysis*; vgl. Giles et al. 2015; Meredith et al. 2021) ein. Bereits in den späten 1990ern wurden Konzepte und Methoden der Konversationsanalyse auf die heute meist als traditionell bezeichnete Chat-Kommunikation übertragen. Die Frage, wie Nutzer*innen ihren Austausch organisieren und strukturieren, stellt von Anfang an einen Schwerpunkt der Analysen dar (vgl. z.B. Storrer 2001; Schönfeldt & Golato 2003; Beißwenger 2007).

Meine Ergebnisse basieren auf 42 Chat-Verläufen, die ich 2018 und 2020 im Rahmen meiner Dissertation (Tschannen o.J.) nach Aufrufen in sozialen Medien, in meinen Lehrveranstaltungen und im privaten Umfeld erhalten habe. Mir haben 23 Personen Teile ihrer Chats zur Verfügung gestellt, an denen sich insgesamt 72 verschiedene Nutzer*innen beteiligten. Mir liegen insgesamt 233 Audio- und 567 Text-Postings¹¹ vor, die sich auf die insgesamt 42 Chats verteilen. Davon sind 33 Chats dyadisch, bei den restlichen neun handelt es sich um Gruppenchats. 21 Chats stammen von Nutzer*innen aus der Schweiz und sind in einem schweizerdeutschen Dialekt realisiert. Alle Audio-Postings liegen mir als Audiodateien vor. Die Audio-Postings sind nach GAT2 transkribiert. Aus den transkribierten Audio-Postings und den mir als Screenshot vorliegenden Text-Postings sowie den Metainformationen habe ich insgesamt 42 anonymisierte Gesamttranskripte angefertigt.

5. Multimodale Beendigungsstrategien

Im Folgenden zeige ich einige Ressourcen, auf die Nutzer*innen zurückgreifen, um sie als multimodale Beendigungsstrategie in ihrem Audio-Posting einzusetzen. Unter Ressourcen verstehe ich in Anlehnung an Stöckl (2016: 9) die Bestandteile gesprochener Sprache, wie sie in Audio-Postings zum Einsatz kommt. Zu den möglichen Ressourcen gehören beispielsweise die Sprechgeschwindigkeit, Stimmhöhe oder Lautstärke. Zu multimodalen Strategien werden diese Ressourcen, wenn sie kombiniert für einen bestimmten Zweck (hier: Beendigungen) eingesetzt werden. Als Beendigung fasse ich alles auf, was in einem Audio-Posting zwischen der letzten thematischen Handlung und der Deaktivierung des Mikrofons passiert. Wenn die Deaktivierung des Mikrofons unmittelbar auf die letzte thematische Handlung folgt, wird die Beendigung ausschliesslich technisch gelöst. In diesem Fall fällt die Themenbeendigung mit dem Ende des Postings zusammen. Im Folgenden zeige ich vier ausgewählte Beispiele, die besonders anschaulich demonstrieren, wie die Beendigung im Anschluss an die letzte Themenbeendigung mit multimodalen Strategien

¹¹ Darunter fasse ich an dieser Stelle auch Emojis.

bearbeitet wird. Die hier dargestellten Strategien sind das Resultat einer rein qualitativen Analyse; quantitative Aussagen zur Verteilung der Strategien oder zu den Unterschieden zwischen den schweizerdeutschen und bundesdeutschen Daten stehen noch aus.

Das folgende Beispiel (1) stellt einen Ausschnitt aus einem Audio-Posting der Nutzerin Vera dar. Sie greift darin mehrere Themen (*lines of talk*, vgl. Virtanen et al. 2021) auf, die bereits im vorherigen Austausch zwischen ihr und ihrer Freundin aufkamen. In ihrer letzten line of talk geht es um das Computerspiel *Die SIMS*, das Vera nach dem Abendessen noch spielen möchte.

(1) Transkript "Vampire" (dyadisch)

Posting #6: 16:41 Uhr; Vera (Audio, 01:22; Ausschnitt 0030-Ende)

0030 =<<schnell> ich wollt heut ma n bisschen was a m > irgendwas
ANderes ausprobieren,=
0031 =irnwass anderes BAUen **oder so**-
0032 (2.0) **ahm (.) mal GUCKen;**
0033 ich werd auch glaub ich nicht ZU lange bauen,=
0034 =weil ich will die ja noch EINziehen lassen noch **und so;**
0035 (2.0) <<leise> **joa > ma GUCKen;**
0036 (-) vielleicht das HAUS auch auf meine familie (ja) dann_auch
direkt abSTIMM-
0037 (1.0) **joa;**
0038 h° (2.0)

Vera formuliert in ihrer letzten line of talk den Plan, im Computerspiel etwas zu bauen (0031). Das *oder so*, das hier als Diskursmarker fungiert, projiziert bereits das Themenende (zu *oder so* als Grenzmarker vgl. König & Stoltenburg 2013), das durch die Pause und den folgenden Diskursmarker *mal gucken* bestätigt wird (0032). Es folgt dann ein kurzer Einschub (0033-0034), in dem es noch einmal um das Spiel geht. Die fallende letzte Tonhöhe und die erneute Verwendung einer Etcetera-Formel (König & Stoltenburg 2013), diesmal *und so*, projizieren erneut das Themenende (0034). Nach einer Pause erfolgt eine Wiederholung des *mal gucken*, diesmal eingeleitet durch ein *joa* und deutlich leiser formuliert (0035). Darauf folgt wieder ein Einschub (0036), in dem es noch einmal um das Spiel geht. Prosodisch ist nach diesem Segment eine Fortführung erwartbar, weil sich die letzte Tonhöhe nicht verändert. Nach einer Pause folgt mit dem *joa* erneut ein Beendigungssignal (vgl. Hector 2017: 35; 42) (0037), das gleichzeitig auch das technische Ende des Postings einleitet. Es vergehen noch etwa zwei Sekunden¹², bis das Posting tatsächlich endet.

¹² Diese hier als Pause transkribierte Zeitspanne, in der nichts mehr zu hören ist, ist in meinen Daten kein Einzelfall. Es kommt vor allem bei längeren Nachrichten häufig dazu, dass die Aufnahme noch ein bis zwei Sekunden weiterläuft, bevor das technische Ende erfolgt. Da ich keine näheren Angaben zu Nutzungsgewohnheiten erhoben habe, kann ich nur mutmassen, dass diese Personen vermutlich das Mikrofon für die Aufnahme nicht gedrückt halten, sondern die Feststellsfunktion nutzen und diese wenigen Sekunden benötigen, um den Sende-Button auszulösen.

Aus diesem Beispiel geht exemplarisch hervor, dass kommunikative Aufgaben, die in synchroner Interaktion kollaborativ verhandelt werden wie hier die Themenbeendigung, in Audio-Postings allein bearbeitet werden müssen. Vera steht im Laufe des Postings mehrmals vor dem kommunikativen Problem, die aufgegriffenen lines of talk (hier sind nur die letzten zwei aufgeführt) wieder zu beenden. Sie greift dabei auf in synchroner Interaktion typische Diskursmarker wie *joa* als Ressource zurück (Auer 2017). Im Unterschied zur synchronen Interaktion, in der diese Diskursmarker Pre-Closing-Signale darstellen (vgl. Schegloff & Sacks 1973), die die Bereitschaft zur Themenbeendigung markieren, vom Gegenüber aber bestätigt werden müssen, damit die Beendigung vollzogen ist, fehlt der kollaborative Aspekt in Audio-Postings. Zwar kommt den hier verwendeten Diskursmarkern wie *mal gucken* oder *joa* sowie den prosodischen Hinweisen auch eine Pre-Closing-Funktion zu, denn das Themenende und auch das technische Ende werden dadurch projiziert, aber nicht ausgehandelt, sondern durch die Deaktivierung des Mikrofons einseitig vollzogen.

Hier lässt sich auch eine Parallele zu AB-Nachrichten erkennen, in denen es nach der Anliegensformulierung¹³ ebenfalls zu Pre-Closing-Signalen kommt, beispielsweise durch *okay* oder *ja* (Liddicoat 1994: 305; Alvarez-Caccamo & Knoblauch 1992: 485), die die Funktion haben, das Ende des letzten Themas zu markieren und das Ende der Nachricht einzuleiten. Knirsch (2005: 296) stellt in Bezug auf AB-Nachrichten klar, dass der Hinweis auf ein abgeschlossenes Thema eigentlich unnötig ist, da ohne Gegenüber keine Aushandlung möglich ist, also auch nicht die Bereitschaft zur Beendigung angezeigt werden müsste. Trotzdem kommt es in AB-Nachrichten regelmässig zu resümierenden Phänomenen, die die Themenbeendigung markieren. Knirsch (2005: 296-297) führt diese Beendigungssignale auf die Abwesenheit eines Gegenübers zurück:

Sie [die Anrufenden, die eine Nachricht hinterlassen] bestätigen sich damit selbst, dass es [das Thema] abgeschlossen ist, und leiten gleichsam die Beendigung ein. Sie überbrücken damit einen im Telefongespräch hochgradig erwartbaren TRP [Transition Relevance Place] und helfen sich mangels Partnerhilfe selbst, wieder aus dem Text¹⁴ 'herauszukommen'. (Ergänzungen J.T.)

Sowohl Knirsch (2005) als auch Lange (1999) stellen fest, dass sich Anrufende beim Hinterlassen einer AB-Nachricht an der interaktiven Struktur von synchroner Interaktion, genauer gesagt von Telefongesprächen, orientieren.

Das vorherige Beispiel (1) zeigt, wie multimodale Ressourcen in Audio-Postings als Beendigungsstrategie eingesetzt werden, die in synchroner Interaktion eine ähnliche Funktion haben. Auch im nächsten Beispiel (2) wird eine aus der

¹³ Zum sequenziellen Aufbau von AB-Nachrichten vgl. Lange (1999).

¹⁴ Knirsch (2005) bezeichnet AB-Nachrichten wegen ihrer Asynchronität und Monologizität als Texte.

synchronen Interaktion bekannte Strategie eingesetzt, die aber im Audio-Posting eine andere Funktion hat.

(2) Transkript "Schlüssel" (dyadisch)

Posting #5: 17:17 Uhr; Sina (Audio, 00:46; Ausschnitt 0017-Ende)

- 0017 (-) em:: und JA-
 0018 (1.5) e: SCHLUSSl,
 äh schlüssel
 0019 dr ÜSnä?
 den rausnehmen
 0020 oder WAS,
 0021 also AUen i briefchaschte tüen oder was,
 also einfach in den briefkasten tun oder was
 0022 (1.2) u::nd wo w was heit ihr am mOre für proGRAMM,
 und wo was habt ihr am morgen für programm
 0023 em oder chonsch du mor EU cho wandere oder NID oder ja;
 ähm oder kommst du morgen auch wandern oder nicht oder ja
 0024 <<leiser> ich ga MIT uf jede fall >;
 ich gehe mit auf jeden fall
 0025 <<höher> TSCHÜhüs > ,

Das Audio-Posting von Sina ist das fünfte Posting innerhalb der mir vorliegenden Interaktion und folgt zeitlich unmittelbar auf das vorherige Audio-Posting #4 von Sinas Schwester Ilenia (hier nicht aufgeführt). Es handelt sich hier um Sinas erstes Audio-Posting innerhalb der Interaktion. Es geht darin um zwei unterschiedliche Themen: Zum einen hat Ilenia ihren Schlüssel im Auto vergessen, das die beiden Schwestern gemeinsam nutzen, zum anderen erkundigt sich Sina, ob Ilenia am folgenden Tag bei der Wanderung mit den Eltern dabei sein wird. Sina markiert zunächst einige Verständnisschwierigkeiten in Bezug auf den Schlüssel (0019, 0021), bevor sie dann eine neue line of talk einführt (0022). Sie fragt ihre Schwester, ob diese an der Wanderung teilnehmen wird, gibt ihr die möglichen Antwortoptionen vor (0023), informiert sie über ihre eigene Teilnahme und beantwortet damit bereits eine mögliche Gegenfrage (0024). Die geringere Lautstärke dieser letzten Handlung markiert das Ende dieser line of talk, ebenso die fallende Tonhöhenbewegung in diesem Segment. Sina beendet dann das Posting verbal mit einer Abschiedspartikel (0025), die – wie es in synchroner Interaktion und auch in AB-Nachrichten typisch ist – prosodisch durch die Stimmhöhe und die Treppenintonation (Auer 2017: 46) deutlich von der restlichen Nachricht abgegrenzt ist. Genau wie in synchroner Interaktion oder AB-Nachrichten hat das *tschüss* auch hier eine beendende Funktion; im Unterschied zu den anderen Kommunikationsformen markiert die Abschiedspartikel im Audio-Posting aber nicht das Ende der Interaktion, sondern nur das Ende des Postings. Sina etabliert im Posting mehrere konditionelle Relevanzen, die von Ilenia im zeitlich unmittelbar folgenden Audio-Posting (hier nicht aufgeführt)

durch die passenden Second Pair Parts eingelöst werden. Die Interaktion und auch die lines of talk werden also nahtlos weitergeführt.

In meinen Daten kommen diese Posting-beendenden Abschiedspartikeln häufig in den Fällen vor, in denen Nutzer*innen zunächst Text-Postings ausgetauscht haben und dann zum ersten Mal innerhalb der Interaktion ein Audio-Posting produzieren. Alle weiteren Audio-Postings von Sina und Ilenia enden nicht mit einer Abschiedspartikel, was die Vermutung nahelegt, dass sich diese Beendigungen nicht nur auf die Zeitlichkeitsbedingungen zurückführen lassen, sondern auch auf den Wechsel des Postingtyps (*Mode-Switching*, vgl. König & Hector 2017: 35).

Das nächste Beispiel (3) zeigt, dass Nutzer*innen für die Beendigung ihrer Audio-Postings auch Strategien einsetzen, die nicht an synchroner Interaktion orientiert sind.

(3) Transkript "Fahnenfest" (dyadisch)

Posting #1: 21:38 Uhr; Marko (Audio, 00:14; vollständig transkribiert)

0001 (0.5) ä: JO,
 0002 ä SCHREIB mir ma noch ma bitte: öm,
 0003 (.) wann dieses: FAHNfest is ä: wo wir hingehn wolln;
 0004 ich hab jetzt vergessn grad <<schneller werdend> was fürn DATum
 das war>,
 0005 (1.0) <<höher, gepresst> Ö::m joa >,
 0006 (1.0) <<schnell, tiefer> das WARS au schon >;

In dem Audio-Posting erkundigt sich Marko bei seinem Freund Noel nach dem Datum eines Festes, das beide zusammen besuchen wollen. Marko fordert seinen Freund zunächst auf, ihm das Datum zu nennen (0002-0003) und begründet diese Aufforderung (0004). Der Anstieg der letzten Tonhöhe und das schneller werdende Sprechtempo in diesem Segment machen eine Weiterführung des Postings erwartbar. Im Anschluss folgt zunächst eine Pause und dann der gedehnte Häsitationsmarker *öhm* (0005), der Planungsaktivität anzeigt (vgl. Clark & Fox Tree 2002). Das anschliessende *joa* markiert zwar ein Ende der vorherigen Handlung und auch die Veränderung der Stimmqualität ist ein Hinweis auf den Abschluss dieser Handlung; durch den Anstieg der letzten Tonhöhe wird hier aber eine weitere Handlung, also wiederum eine Fortführung der Nachricht projiziert (0005). Nach einer weiteren Pause beendet Marko sein Anliegen explizit mit dem verbalen Hinweis darauf, dass keine weiteren Handlungen mehr folgen werden (0006). Die prosodischen Ressourcen, also vor allem die veränderte Stimmqualität und der Fall der letzten Tonhöhe, markieren zusätzlich die Beendigung.

Wie auch im vorherigen Beispiel (2) lassen sich die Beendigungsstrategien auf die asynchrone Produktionssituation zurückführen: Die Anliegensformulierung stellt einen First Pair Part dar, der einen Sprecherwechsel erforderlich macht,

damit die konditionellen Relevanzen durch einen passenden Second Pair Part eingelöst werden können. Der Sprecherwechsel kann aber nicht erfolgen, bevor Marko sein Posting versendet. An der Position, an der ein Sprecherwechsel in synchroner Interaktion erfolgen würde (nach 0004), entsteht hier zunächst eine Pause. Die Planungsaktivität, die sich durch das gedehnte *öhm* ausdrückt, lässt sich darauf zurückführen, dass Marko offenbar überlegt, ob es ein weiteres Anliegen zu klären gibt. Dass das nicht der Fall ist, markiert Marko dann verbal (0006). Diese multimodale Beendigungsstrategie hat hier nicht nur die Funktion, das Audio-Posting zu beenden und das technische Ende einzuleiten, sondern auch, den Sprecherwechsel zu organisieren, also den kommunikativen 'Ball' weiterzuspielen und zu markieren, dass es dieses eine Anliegen gibt, das nun beantwortet werden muss¹⁵.

Das vorherige Beispiel (3) illustriert, wie durch die Beendigung auch die Organisation der Interaktion bearbeitet wird. Marko zeigt, dass sein Posting beendet ist und nun ein Folgeposting des Gegenübers erforderlich ist. Im folgenden Beispiel (4) wird am Ende des Postings auch die Organisation, genauer gesagt das Rederecht, bearbeitet. Anders als im vorherigen Beispiel wird aber nicht ein intendierter Sprecherwechsel markiert, sondern ein weiteres eigenes Posting projiziert und damit markiert, dass das Rederecht noch nicht übergeben wird¹⁶.

(4) Transkript "Wann ist dein Geburtstag?" (dyadisch)

Posting #6: 16:42 Uhr; Luisa (Audio, 04:47; Ausschnitt 0107-0108; 0116-Ende)

- 0107 <<smile voice> aber bitte SÄG mir doch ob du no hesch >-
aber bitte sag mir doch ob du noch hast
- 0108 auso WENN du überhaupt geburtstag HESCH;=
also wann du überhaupt geburtstag hast
- [...]
- 0116 mou wär i scho FROH etz;
doch wär ich schon froh jetzt
- 0117 <<lachend> hm es mau z WÜSse >;
es mal zu wissen
- 0118 da chan ich mirs würkch ma UFSchribe;
dann kann ich mir es wirklich mal aufschreiben
- 0119 (-) ((schmatzt)) genau;
- 0120 (1.8) ja ich schicke das mau AB?
ja ich schicke das mal ab

¹⁵ Es handelt sich hier um das initiale Posting der mir vorliegenden Interaktion. Die konditionellen Relevanzen werden durch das Folgeposting von Noel tatsächlich eingelöst.

¹⁶ Das ist zwar technisch jederzeit möglich und anders als in synchroner Interaktion, in der es dann zu Überlappungen kommt, würden die Postings trotzdem linear im Verlaufsprotokoll angezeigt. Das führt aber u.U. zu Phantomadjazenzen (Garcia & Baker Jacobs 1999; Imo 2015), bei der die Postings nur zufällig adjazent, also visuell benachbart sind, aber keinen Bezug durch konditionelle Relevanzen aufweisen, sondern sich auf nichtadjazente Postings beziehen. Das erfordert Interpretationsleistungen; die Replyfunktion oder Strategien, wie hier beschrieben, tragen zur Bearbeitung der sequenziellen Ordnung bei.

Nachricht #7: 16:46 Uhr; Luisa (Audio, 03:21; Ausschnitt 0001-0002)

0001 (2.9) °h ehm-

0002 (1.0) mir si (.) geschter ou no chli ga spaZIre,
wir sind gestern auch noch etwas spazieren gegangen

Luisa beendet ihr fast fünfminütiges Audio-Posting #6, indem sie zunächst die letzte line of talk abschliesst. Sie fasst die Wichtigkeit ihres Anliegens (zu wissen, wann ihre Freundin Geburtstag hat) zusammen (0116-0118) und beendet diese line of talk mit dem Diskursmarker *genau* (vgl. Auer 2020), der auf eine kurze Pause und ein Schmatzen folgt (0119). Auf das *genau* könnte eine neue line of talk folgen oder das technische Ende des Postings. Stattdessen leitet Luisa das technische Ende des Audio-Postings explizit ein, indem sie in Aussicht stellt, das Posting nun abzuschicken (0120). Damit beendet sie zwar verbal das Posting; durch den starken Anstieg der letzten Tonhöhe markiert sie aber gleichzeitig, dass ihr Beitrag zur Interaktion noch nicht beendet ist. Obwohl sie also am Ende ihres Postings konditionelle Relevanzen etabliert (0107-0108) und damit eigentlich eine Redeübergabe an ihre Freundin bearbeitet, weist sie mit der Aussicht auf den Versand der Nachricht auf eine weitere eigene Nachricht hin. Diese Art der Beendigung lässt sich damit als eine Form von *turn-keeping* Signal (Sacks et al. 1974) interpretieren. Luisa stellt eine weitere Nachricht in Aussicht und bearbeitet damit die Organisation der Interaktion. Es folgt auch tatsächlich direkt im Anschluss ein weiteres Audio-Posting #7, in dem Luisa eine noch offene line of talk aufgreift. Sie teilt also die unterschiedlichen lines of talk auf verschiedene Audio-Postings auf und stellt am Ende des ersten Postings das folgende Posting #7 in Aussicht.

6. Fazit

Die Beispiele haben gezeigt, dass Interagierende durch die Affordanzen der Kommunikationsform Chat am Ende von Audio-Postings vor spezifischen kommunikativen Problemen stehen, für deren Lösung sie unterschiedliche multimodale Strategien einsetzen. Dabei greifen sie auf verbale und prosodische Ressourcen zurück, die teilweise auch in synchroner Interaktion eingesetzt werden, teilweise aber auch nicht vergleichbar mit anderen Kommunikationsformen sind. Während in Beispiel (1) und (2) auf Strategien zurückgegriffen wird, die auch in synchroner Interaktion mit ähnlicher Funktion zum Einsatz kommen, sieht man in Beispiel (3) und (4), dass Nutzer*innen neue, Audio-Posting-spezifische Strategien einsetzen, für die sie ebenfalls verbale und prosodische Ressourcen einsetzen, die aber in dieser Form in keiner anderen Kommunikationsform belegt sind.

Die kommunikativen Probleme lassen sich auf die Zeitlichkeitsbedingungen zurückführen, denen der Einsatz von gesprochener Sprache in Textformen-basierter Interaktion unterliegt: Zum einen führen die asynchronen Produktionsbedingungen dazu, dass eigentlich kollaborative Aufgaben wie beispielsweise Themenbeendigungen oder Sprecherwechsel allein bearbeitet werden müssen. Zum anderen sorgt die zeitlich-visuelle Persistenz von Postings in einem Verlaufsprotokoll für eine lineare Abfolge, die durch das Eintreffen der Postings auf dem Server bestimmt ist und die die Rezeptionsgrundlage bildet.

Die hier präsentierten Beispiele sind alle auf der Kommunikationsplattform WhatsApp entstanden. Offen bleibt, ob die zeitliche Limitierung von Audio-Postings wie beispielsweise bei Facebook oder WeChat einen Einfluss auf die Beendigungsstrategien hat. Es stellt sich die Frage, ob das technische Ende eines Audio-Postings überhaupt verbal bearbeitet werden muss, wenn es durch das Zeitlimit gesetzt ist.

LITERATUR

- Alvarez-Caccamo, C. & Knoblauch, H. (1992). 'I was calling you': Communicative patterns in leaving a message on an answering machine. *Text - Interdisciplinary Journal for the Study of Discourse*, 12(4), 473-505. doi.org/10.1515/text.1.1992.12.4.473.
- Androutsopoulos, J. (2007). Neue Medien – neue Schriftlichkeit? *Mitteilungen des Deutschen Germanistenverbandes*, 54(1), 72-97.
- Androutsopoulos, J. & Busch, F. (2021). Digital punctuation as an interactional resource: The message-final period among German adolescents. *Linguistics and Education*, 62, 1-9. <https://doi.org/10.1016/j.linged.2020.100871>.
- Arminen, I., Licoppe, C. & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction*, 49(4), 290-309. doi.org/10.1080/08351813.2016.1234614.
- Auer, P. (2017). Anfang und Ende fokussierter Interaktion: eine Einführung. *Interaction and Linguistic Structures*, 59, 1-56.
- Auer, P. (2020). Genau! Der auto-reflexive Dialog als Motor der Entwicklung von Diskursmarkern. In B. Weidner, K. König, L. Wegner & W. Imo (Hgg.), *Verfestigungen in der Interaktion – Konstruktionen, sequenzielle Muster, kommunikative Gattungen* (S. 263-294). Berlin/Boston: de Gruyter.
- Baron, N. S. (2013). Instant messaging. In S. Herring, D. Stein & T. Virtanen (Hgg.), *Pragmatics of Computer-Mediated Communication* (S. 135-162). Berlin/ Boston: de Gruyter.
- Beißwenger, M. (2007). *Sprachhandlungskoordination in der Chat-Kommunikation*. Berlin/ Boston: de Gruyter.
- Beißwenger, M. (2015). Praktiken der internetbasierten Kommunikation. In A. Deppermann, H. Feilke & A. Linke (Hgg.), *Sprachliche und kommunikative Praktike* (S. 279-309). Berlin/ Boston: de Gruyter.
- Beißwenger, M. (2020). Internetbasierte Kommunikation als Textformen-basierte Interaktion: ein neuer Vorschlag zu einem alten Problem. In K. Marx, H. Lobin & A. Schmidt (Hgg.), *Deutsch in Sozialen Medien* (S. 291-318). Berlin/ Boston: de Gruyter.

- Beißwenger, M. & Storrer, A. (2012). Interaktionsorientiertes Schreiben und interaktive Lesespiele in der Chat-Kommunikation. *Zeitschrift für Literaturwissenschaft und Linguistik*, 42(4), 92-124. <https://doi.org/10.1007/BF03391261>.
- Berglund, T. Ö. (2009). Disrupted turn adjacency and coherence maintenance in Instant Messaging conversations. *Language@Internet*, 6(2), 1-25.
- Boyd, D. (2014). *It's complicated: The social lives of networked teens*. Yale: Yale University Press.
- Clark, H. H. & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84, 73-111.
- Dingwall, S. (1995). Telephone answering machine messages: Dialogues or monologues? *Cahiers de Linguistique Française*, 16, 113-143.
- Dürscheid, C. (2005). Medien, Kommunikationsformen, kommunikative Gattungen. *Linguistik Online*, 22(1), 3-16. <https://doi.org/10.13092/lo.22.752>.
- Dürscheid, C. (2016). Neue Dialoge – alte Konzepte? *Zeitschrift für germanistische Linguistik*, 44(3), 437-468. <https://doi.org/10.1515/zgl-2016-0023>.
- Dürscheid, C. & Brommer, S. (2009). Getippte Dialoge in neuen Medien. Sprachkritische Aspekte und linguistische Analysen. *Linguistik Online*, 37(1), 3-20. <https://doi.org/10.13092/lo.37.511>.
- Garcia, A. C. & Baker Jacobs, J. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction*, 32(4), 337-367. https://doi.org/10.1207/S15327973rls3204_2.
- Giles, D., Stommel, W., Paulus, T., Lester, J. & Reed, D. (2015). Microanalysis of online data: The methodological development of "digital CA". *Discourse, Context & Media*, 7, 45-51. <https://doi.org/10.1016/j.dcm.2014.12.002>.
- Hausendorf, H. (2015). Interaktionslinguistik. In L. Eichinger (Hg.), *Sprachwissenschaft im Fokus* (S. 43-70). Berlin/Boston: de Gruyter.
- Hausendorf, H., Kesselheim, W., Kato, H. & Breitholz, M. (2017). *Textkommunikation: ein textlinguistischer Neuansatz zur Theorie und Empirie der Kommunikation mit und durch Schrift*. Berlin/ Boston: de Gruyter.
- Hector, T. M. (2017). Nutzungskontexte und Dialogizität von WhatsApp-Sprachnachrichten. *Studentische Arbeitspapierreihe Sprache in Interaktion (StudSpin)*, 27, 1-61.
- Herring, S. C. (2015). New frontiers in interactive multimodal communication. In A. Georgakopoulou & T. Spilioti (Hgg.), *The Routledge handbook of language and digital communication* (S. 398-402). London: Routledge.
- Imo, W. (2015). Vom Happen zum Häppchen ... Die Präferenz für inkrementelle Äußerungsproduktion in internetbasierten Messengerdiensten. *Networx*, 69, 1-35. <https://doi.org/10.15488/2960>.
- Imo, W. (2016). Dialogizität – eine Einführung. *Zeitschrift für germanistische Linguistik*, 44(3), 337-356. <https://doi.org/10.1515/zgl-2016-0019>.
- Imo, W. (2019). Das Medium ist die Message: Interaktion und ihre situativen und technischen Rahmenbedingungen. In K. Marx & A. Schmidt (Hgg.), *Interaktion und Medien: Interaktionsanalytische Zugänge zu medienvermittelter Kommunikation* (S. 35-58). Heidelberg: Winter.
- Imo, W. & Lanwer, J. (2019). *Interaktionale Linguistik: eine Einführung*. Berlin/ Heidelberg: J.B. Metzler.
- Kilian, J. (2001). T@stentöne. Geschriebene Umgangssprache in computervermittelter Kommunikation. Historisch-kritische Ergänzungen zu einem neuen Feld der linguistischen Forschung. In M. Beißwenger (Hg.), *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld* (S. 55-78). Stuttgart: Ibidem.
- Knirsch, R. (2005). *"Sprechen Sie nach dem Piep": Kommunikation über Anrufbeantworter. Eine gesprächsanalytische Untersuchung*. Tübingen: Niemeyer.

- König, K. (2015). Dialogkonstitution und Sequenzmuster in der SMS- und WhatsApp-Kommunikation. *Travaux neuchâtelois de linguistique*, 63, 87-107. <https://doi.org/10.26034/tranel.2015.2971>.
- König, K. (2020). Sprachnachrichten als Zugang zu authentischer Mündlichkeit im Kontext Deutsch als Fremdsprache. In S. Günthner, J. Schopf & B. Weidner (Hgg.), *Gesprochene Sprache in der kommunikativen Praxis. Analysen authentischer Alltagssprache und ihr Einsatz im DaF-Unterricht* (S. 247-277). Tübingen: Stauffenburg.
- König, K. (2021). Mehrsprachigkeit in transmodaler Kommunikation – Praktiken der Sprachalternation in deutsch-arabischen WhatsApp-Chats. *Linguistik Online*, 112(7), 39-64. <https://doi.org/10.13092/lo.112.8259>.
- König, K. & Duan, L. (i.V.). Mündlichkeit im Posting-Format: Praktiken der dialogischen Kommunikation über Sprachnachrichten per WhatsApp und WeChat. In K. König & O. Lämke (Hgg.), *Zugänge zur Mündlichkeit. Sprechen als Gegenstand und Methode im Kontext DaF* (S. 175-208). Münster: Waxmann.
- König, K. & Hector, T. M. (2017). Zur Theatralität von WhatsApp-Sprachnachrichten. Nutzungskontexte von Audio-Postings in der mobilen Messenger-Kommunikation. *Networx*, 79, 1-41. <https://doi.org/10.15488/2970>.
- König, K. & Hector, T. M. (2019). Neue Medien – Neue Mündlichkeit? Zur Dialogizität von WhatsApp-Sprachnachrichten. In K. Marx & A. Schmidt (Hgg.), *Interaktion und Medien* (S. 59-84). Heidelberg: Winter.
- König, K. & Stoltenburg, B. (2013). "oder so", "und so", "und so was", "und so weiter" etc. Eine Interaktionale Perspektive auf Etcetera-Formeln. *Grammatik in der Interaktion*, 48(7), 1-35.
- Lange, I. (1999). Die sequenzielle Struktur von Anrufbeantworter-Kommunikation. *Interaction and Linguistic Structures*, 14, 1-213.
- Liddicoat, A. (1994). Discourse routines in answering machine communication in Australia. *Discourse Processes*, 17(2), 283-309. <https://doi.org/10.1080/01638539409544870>.
- Lyons, A. & Tagg, C. (2019). The discursive construction of mobile chronotopes in mobile-phone messaging. *Language in Society*, 48(5), 657-683.
- Meredith, J. (2019). Conversation analysis and online interaction. *Research on Language and Social Interaction*, 52(3), 241-256.
- Meredith, J., Giles, D. & Stommel, W. (Hgg.) (2021). *Analysing digital interaction* (= Palgrave studies in discursive psychology). Cham, Switzerland: Palgrave Macmillan.
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137-156. <http://dx.doi.org/10.1016/j.pragma.2014.04.004>
- Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735. <https://doi.org/10.2307/412243>.
- Schegloff, E. A. & Sacks, H. (1973). Opening up Closings. *Semiotica*, 8(4), 289-327. <https://doi.org/10.1515/semi.1973.8.4.289>.
- Schönfeldt, J. (2001). Die Gesprächsorganisation in der Chat-Kommunikation. In M. Beißwenger (Hg.), *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in computervermittelter Kommunikation* (S. 25-54). Stuttgart: Ibidem.
- Schönfeldt, J. & Golato, A. (2003). Repair in chats: a conversation analytic approach. *Research on Language and Social Interaction*, 36(3), 241-284. https://doi.org/10.1207/S15327973RLSI3603_02.
- Selting, M. (1995). *Prosodie im Gespräch: Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.
- Selting, M. & Couper-Kuhlen, E. (2000). Argumente für die Entwicklung einer "interaktionalen Linguistik". *Gesprächsforschung*, 1, 76-95.

- Stöckl, H. (2016). Multimodalität – Semiotische und textlinguistische Grundlagen. In N.-M. Klug & H. Stöckl (Hgg.), *Handbuch Sprache im multimodalen Kontext* (S. 3-35). Berlin/Boston: de Gruyter.
- Storrer, A. (2001). Sprachliche Besonderheiten getippter Gespräche: Sprecherwechsel und sprachliches Zeigen in der Chat-Kommunikation. In M. Beißwenger (Hg.), *Chat-Kommunikation. Sprache, Interaktion, Sozialität und Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld* (S. 3-24). Stuttgart: Ibidem.
- Tschannen, J. (o.J.). *Kommunikative Aufgaben in Sprachnachrichten. Eine konversationsanalytische Untersuchung von Sprachnachrichtenanfängen und -enden*. Unveröffentlichte Dissertation, Universität Bern.
- Virtanen, M. T., Vepsäläinen, H. & Koivisto, A. (2021). Managing several simultaneous lines of talk in Finnish multi-party mobile messaging. *Discourse, Context & Media*, 39, 1-14. <https://doi.org/10.1016/j.dcm.2020.100460>.
- Wyss, E. & Hug, B. (2016). WhatsApp-Chats. Neue Formen der Turn-Koordination bei räumlich-visueller Begrenzung. In C. Spiegel & D. Gysin (Hgg.), *Jugendsprache in Schule, Medien und Alltag* (S. 259-274). Frankfurt a. M.: Lang.