

Please, cite as: Sampietro, A., & König, K. (2024). The medium is accountable: Metacommunication and media ideologies about voice messages in WhatsApp chats. *Discourse & Communication*, 18(1), 51-71. <https://doi.org/10.1177/17504813231187109>

## **The Medium is Accountable: Metacommunication and Media Ideologies about Voice Messages in WhatsApp Chats**

Agnese Sampietro, Universitat Jaume I, Spain

Katharina König, Westfälische Wilhelms-Universität Münster, Germany

### *Abstract*

Voice messages (VMs), which allow users to send recorded messages to other contacts, are a popular feature of instant messaging applications. Despite their popularity, linguistic research on VMs is still in its infancy. This study analyses metacommunication around VMs in mobile messaging conversations among WhatsApp users in Germany and Spain. It focuses on participants' metacommunicative accounts (such as references, explanations, or reasons) of their preference for audio posting over other forms of communication (e.g., texting). Drawing on recent advances in digital conversation analysis, we examine how accounts placed in different sequential positions in messenger chats (preceding a VM, at the beginning or end of a VM, or after it has been sent) address diverse aspects of voice messaging (from either the sender's or recipient's perspective). We demonstrate that accounting accomplishes different social actions, such as framing a VM as something outstanding or worth apologizing for. We argue that a sequential analysis of accounting and metacommunication offers rich insights into users' media ideologies concerning the appropriateness and timing of text and voice messaging. Overall, these findings contribute to a better understanding of the growing importance of voice in mobile communication.

**Keywords:** voice messages, texting, WhatsApp, accountability, metacommunication, media ideologies, digital conversation analysis, computer-mediated discourse, transmodal interaction, mode-switching

---

Corresponding author: Agnese Sampietro, Department of European Languages and Cultures, University Jaume I, Av. Sos Baynat s/n, Castellón de la Plana 12071, Spain. [sampietr@uji.es](mailto:sampietr@uji.es)

# 1 Introduction

‘There is practically no demand for a recording attachment to telephones’, claimed Emile Berliner, the inventor of the gramophone, in 1900 (Morton, 2000: 110). The history of communication technology proves him to be wrong. Answering machines were a huge success and became an essential accessory to the landline telephone (Morton, 2000). Moreover, the current rise in popularity of voice messages (henceforth VMs),<sup>1</sup> that is, aural postings, which are recorded, sent, and playable in mobile messenger chats, suggests that despite claims that ‘[o]ver the past decade, smartphones became largely written language, rather than oral devices’ (Baron, 2020: 408), voice is becoming important again in digital communication.

Although voice-recording technologies were set out in the pre-digital era, recent advances in digital recording and voice-over IP protocols have reintroduced aural communication in mobile media. People can now speak on their phones, not only to call but also to record and send voice memos. The success of voice messaging seems to reverse the trend from the early 2000s, when the mobile phone became a device for writing rather than speaking, as texting surpassed phone calls (Baron, 2013; Fortunati, 2001).

Many instant messaging platforms (e.g., Facebook Messenger, KakaoTalk, Line, WeChat, WhatsApp) support the transmission audio-based messages. Voice messaging was first introduced by the Chinese app WeChat in 2011 and integrated into WhatsApp in 2013, followed by Apple Messages in 2016 (Bereznak, 2018). VMs are very popular among Chinese users; in 2015 alone, WeChat users sent an average of 280 million minutes of VMs per day (Wang, 2016: 39). As for WhatsApp, users worldwide send an average of 7.000 million voice messages every day (WhatsApp, 2022). Voice messaging is widespread in countries such as Argentina, Brazil, Germany, India, and Spain, to name a few (BVDW, 2019; CETIC, 2022). Sending VMs is necessary in some regions, such as Cambodia, due to the underdevelopment of keyboards that support local scripts (Elliott and Phorn, 2021). Strikingly, voice messaging is not as well liked in the US, where texting is still the preferred

---

<sup>1</sup> In our corpora, users refer to voice messages with various expressions (e.g., *audio* or *nota de voz* in Spanish or *Sprachnotiz*, *-nachricht* or *-memo* in German). In English, ‘voice message’, ‘voice note’, and ‘voice memo’ are the most common terms. We will use them interchangeably.

mode (Baron, 2020), and WhatsApp is less popular than other messaging platforms (Statista, 2022).

Despite the growing importance of aural communication, literature on computer-mediated discourse has mainly focused on ‘text appearing on screens’ (Baron, 2013: 135), i.e., writing on computers or texting. This study addresses the need for more empirical research on spoken digital interaction by exploring why users choose voice messaging over other communication modes. In particular, we analysed users’ metacommunication (i.e., communication about communication) around VMs in authentic German and Spanish WhatsApp interactions. In particular, the empirical analysis focuses on users’ metacommunicative accounts (such as references, explanations, or justifications for using VMs) of their mode choice in transmodal WhatsApp chats and media ideologies, i.e., beliefs and values people hold about a single medium (see Gershon, 2010) concerning voice messaging reflected in users’ metacommunicative conduct.

We begin by briefly locating mobile voice messaging in the history of recording technologies before outlining previous research on accounting and media ideologies about digital communication (Section 2). Section 3 describes the corpora used in this study and the analytical procedures. Section 4 presents an analysis of accounts for choosing voice messages as a contribution mode examining functional differences in relation to accounts’ positions in transmodal chats. We then discuss the findings with a view to the media ideologies in the accounts and finally point out avenues for future research on transmodal messaging (Section 5).

## **2 Literature review**

### **2.1 The past and the present of voice-recording technologies**

Mobile VMs are among the latest developments in voice-recording technologies, which began to develop shortly after the invention of the telephone. Although the earliest precursor was the ‘phonograph’, patented by Thomas Edison in 1877, the first viable device capable of recording telephone conversations was the ‘telegraphon’, patented in 1898 by Valdemar Poulsen (Mooney, 2009). However, the focus of work in the field of recording technology at that time was on music recording (Morton, 2000). The first successful telephone recorders

were manufactured in the 1930s, while answering machines, which allowed recording and replaying messages, were sold from 1949 and became common in Western homes by the 1970s (Fisher, 1992: 369), as the development of microelectronics enabled the production of cheaper devices. By the 2000s, answering services were ubiquitous even on personal cell phones, owing to the development of more efficient digital recording technologies (Ling 2004).

Answering machines introduced asynchrony in telephone communication. Recording a message proved convenient for the sender of the call, who could reach the callee even if he or she was not immediately available, and for the recipient, who did not miss calls or could avoid answering the phone on purpose. In Ling's (2004: 9) words, answering machines increased the recipient's agency and control over incoming calls.

One of the major technical differences between mobile VMs and previous recording technologies is that voice notes are sent over an Internet connection, which reduces costs and allows for the recording of messages of nearly unlimited size (16MB for recordings, equivalent to an 80-minute VM); the main limitation is the app and phone storage (Kleinholz, 2019). In other words, with answering machines, the device determines the length of the message, whereas with VMs, the person recording the message controls its duration. Furthermore, voice messaging is only one of the many features instant messaging systems offer, as users can also send text messages, pictures, images, links, and other files as well as make phone or video calls through the same interface.

Despite these differences, VMs share certain technical characteristics with previous technologies that affect their verbal design. First, VMs, answering machines, and telephones enable long-distance spoken communication. With VMs and answering machines, the exchange is usually asynchronous, and the addressee listens to the message later, waiting for the recording to be completed. The lexical and prosodic structures of messages left on answering machines are similar to synchronous telephone and face-to-face interactions (Gold, 1991), with the exception that some questions may be omitted or left unanswered (Liddicoat, 1994). Álvarez-Cáccamo and Knoblauch (1992) observed that the communicative structures of answering machine talk result from the three characteristics that are also applicable to VMs: the mediacy of time and space, one-sidedness and orality.

Messages left on answering machines lack an interactive turn-taking structure and are designed as ‘pre-packaged messages’ (Hutchby and Tanna, 2008), which often include an opening section with greetings, the body of the message, and closings with expressions of thanks, acknowledgements, or references to future meetings or calls. This structure may change depending on the nature of the message (requesting or giving information, arranging a meeting, etc.) or in institutional contexts where self-identification changes and farewells may be omitted (Hobbs, 2003). Greetings and farewells can also be omitted from the follow-up messages. Similar to the messages left on answering machines, many VMs have a clear structure with openings, a message body that may contain extended storytelling sequences (König, 2019), questions and requests, multiple topics, or lines of talk (Emery, 2018; König, 2021). Overall, messages on answering machines and VMs share the characteristics of telephone calls and written forms of communication, such as letters (Dingwall, 1992; Gold, 1991; König and Hector, 2017, 2019).

Another similarity between VMs and answering machine discourse is how users frame their activities. Early instruction manuals for answering machines were clearly business oriented: the owners should identify themselves and the company, give instructions on how to leave a message, and so on. In contrast, home users developed their customs, which often included an explanation or excuse for using the device (Morton, 2000: 131). The opening sections of answering machine greetings are often followed by a ‘warrant’, i.e., a justification for using the device (instead of directly answering the phone), expressed usually as a ‘statement of unavailability’ (Liddicoat, 1994: 290). Callers’ contributions can also include justifications for calling or leaving a recorded message, particularly in domestic answering machine messages. Liddicoat (1994) interprets these justifications as showing that the recordings are ‘dispreferred options’ compared with a preferred synchronous phone conversation. According to the author, ‘[...] to gain access to an interlocutor in his or her absence is in some way intrusive and is, therefore, accountable’ (Liddicoat, 1994: 301).

## **2.2 Accounting and media ideologies**

Accountability is a key concept in ethnomethodology (Garfinkel, 1967) and conversation analysis (Heritage, 1988). This captures the idea that social actors generally attribute meaning to each other’s actions and can thus be held accountable for their actions and hold others

accountable for theirs. Accountability usually relates to social conduct (referred to as accounts *of* action by Buttny and Morris, 2001: 286), and it is often negotiated in rather subtle ('off-record) ways, such as asking known-answer questions to make behaviours or opinions accountable (Raymond and Stivers, 2016) or tellings that 'fish' for the speaker's observation or invite further explanations about it (Pomerantz, 1980). However, accountability can also be the subject of metacommunicative discourse (accounts *for* action; Buttny and Morris, 2001: 286), for example, in the form of explanations or justifications (Heritage, 1988). Such accounts can lead to social action being framed as dispreferred or marked in a particular context (Scott and Lyman, 1968).

In media linguistics, the study of metacommunication and accounts helps identify participants' subjective conceptions of their mediated practices, what they treat as (un)acceptable conduct, and the value they attribute to different media, that is, their media ideologies (Gershon, 2010). For example, in WhatsApp chats, users have been reported to offer explanations and apologies for recording a VM, indicating that, in some contexts, this is the dispreferred communication mode in this application (König and Hector, 2017: 22). Some accounts for selecting an aural contribution are situational (e.g., when driving) or production-related (e.g., speaking is quicker than typing). However, accounts are uncommon, suggesting that VMs are now accepted as an appropriate mode of communication (König and Hector, 2019: 67). Yet, users' accounts give significant clues about their media ideologies,<sup>2</sup> i.e., 'the assumption that people hold about how a medium accomplishes a communicative task' (Gershon, 2010: 392). Rather than relying solely on what is technically possible, media understanding is shaped by users' perceptions, values, and beliefs regarding the proper use of technology (Boczkowski et al., 2018). For instance, in examining college students' breakup stories, Gershon (2010: 391) observed that interviewees reflected on which medium indexed more intimacy or was perceived as more formal, and how they switched from one medium to another upon ending their relationship. As for WhatsApp, in one of the case

---

<sup>2</sup> Although most people associate the term 'ideology' with politics, linguistic anthropologists have used it to refer to cultural belief systems related to language (language ideologies), signs (semiotic ideologies), and media (media ideologies) (Gershon, 2010; Thurlow, 2017). Scholars in this discipline use the term to acknowledge the new subtleties that the word 'ideology' has acquired in the late 20<sup>th</sup> century. Ideology today refers to a coherent set of ideas, beliefs, and values. Therefore, in this study we rely on this current definition, which 'relieves the concept of negative connotations' (Knight, 2006: 626).

studies analysed by Busch and Sindoni (2022: 251), a chat between two German teenagers in which one user explicitly asks the other to record a VM, the switch to voice is treated as more appropriate for some activities, such as explanations. Accounts thus prove to be an insightful key to users' conception of this new way of mediated aural communication.

This study follows this line of research by analysing accounts and other forms of metacommunication on the use of VMs in different datasets of authentic WhatsApp chats. Contrary to previous studies focusing on just one modality (Busch and Sindoni, 2022, on the text surrounding VMs; König and Hector, 2017, only on VMs) or relying on 'indirect' data only (interviews by Gershon, 2010, focus groups by Fernández-Ardévol et al., 2021, media discourse by Thurlow, 2017), this study examines *in situ* discourse both in text and voice messages. The analysis focuses on how users interpret and negotiate the modal opportunities afforded by mobile messaging platforms in metacommunicative comments, and the switch from one mode to another. Accounts are given when there is a departure from expected social conduct (e.g., Antaki, 1996). We are, thus, not only interested in what it is that users account for (e.g., the length of a voice message) or the metalanguage that they use (e.g., referring to their talk as *(voll)quatschen*, meaning 'gabbing') but also in the actions they accomplish with accounting for their mode choice. Drawing on recent approaches to digital conversation analysis (Giles et al., 2015; Meredith, 2019), we study the sequential unfolding of transmodal WhatsApp chats, explore when and how users' accounts frame their use of voice messages, and how senders and receivers take up different perspectives on audio postings. Accounting practices thus offer rich insights into the *in-situ* instantiation and negotiation of media ideologies about transmodal messaging chats.

### **3 Methodology**

This study draws on two private collections of WhatsApp voice messages retrieved in Germany and Spain. The German private corpus, compiled between 2017 and 2022, comprises 159 WhatsApp chats with 1,064 voice messages (totalling 535 minutes of spoken data). Most of it consists of dyadic and multiparty chats between students in their early twenties and their families. As for the Spanish data, the private corpus consists of messages (oral and written) received by the first author from acquaintances in 2021 and 2022, and the

log of another group chat, not including the researcher. Messages (written or spoken) sent by the author were excluded from the analysis. Participants were nine women and 16 men with an average age of 40 years. The dataset includes eight chats (six dyads and two groups) for over 26,000 postings (including text messages, pictures, and other attachments), of which 251 are VMs (total: 150 minutes of spoken data). All interactions were spontaneous and not elicited by the researchers since we compiled the data after the interactions took place. All users provided informed consent to use their chats in linguistic research. Usernames and other personal details were anonymised. Given the focus of this study, we use verbatim transcriptions of VMs.

Additionally, we analysed two large, open-access corpora of WhatsApp messages: the *Mobile Communication Database* (MoCoDa), a web-based corpus of German text messaging (Beißwenger et al., 2019), and the Spanish corpus *Microsintaxis del Español Actual 2.0* (MEsA) published under a Creative Commons Licence (Fuentes Rodríguez, 2021). The MoCoDa database provides access to more than 37,000 messages in 818 dialogues with more than 3,000 chatters. MEsA 2.0 includes the log of 10 group and dyadic WhatsApp chats, totalling around 28,000 single postings. Corpora include either no VMs or only a few (MoCoDa does not usually provide detailed transcriptions but rather short summaries). Nevertheless, these datasets are valuable sources of metacommunications concerning VMs, as we shall expound later.

Our methodological approach to accounting in VMs is informed by computer-mediated discourse analysis (Herring, 2019) and digital conversation analysis (see Giles et al., 2015; Meredith, 2019). Following Buttny and Morris (2001), we analyse explicit ('on record') accounts *for* voice messaging, expressed as justification or excuse, and accounts *of* the use of VMs ('off-record'), i.e., metacommunication around recording and listening to VMs. The first discourse-analytical coding of the corpus is then triangulated with the study of these accounts in the sequential unfolding of the interaction, in line with a digital CA approach. In fact, Buttny and Morris (2001: 288) explain that 'accounts are not recognisable by the form of the utterance alone, but rather by their sequential position in context'. Therefore, in addition to the actual content of the messages, we focused on how users offered and responded to different forms of metacommunication in the interaction. In other words, we



are interested in the sender and recipient timing or placement of accounts in transmodal messaging and their interactional functions in the given sequential slots (before a VM, inside a VM, or after sending a VM).

The analysis revealed the following forms of accounts and metacommunication in both text and voice messages:<sup>3</sup>

1. **Explicit accounts:** This category includes explanations, justifications, or even apologies for using VMs.
2. **Mentions of recording practices:** Comments around some characteristics of the recording of a VM, such as mentions of its (excessive) length, ‘side effects’ of the aural mode (e.g., commenting on background noise) or references to the act of recording.
3. **References to listening practices:** Listening can be made explicit when recipients explain when they have listened or will listen to a VM or when the sender recommends an appropriate point in time to play the message.

Following this procedure, we identified 382 instances of accounting and other forms of metacommunication within the corpora.

The following section analyses the sequential integration of accounts and metacommunication on VMs into the ongoing discourse, focusing on how they are typically introduced and negotiated in transmodal WhatsApp chats.

## 4 Analysis: Negotiating the choice of voice messages

The analysis of selected extracts presented in this section illustrates how different account forms (on- or off-record, written or spoken) integrate into the chats’ trajectory, which aspects of a VM are accountable, and how users negotiate the accounts in their subsequent messages. Following the sequential development of transmodal WhatsApp chats, we will first examine metacommunication before voice memos are posted to the chat (Section 4.1). We then move

---

<sup>3</sup> The distinction between the three categories is an analytic one. In the data, sometimes they combine. For instance, recording and listening are often connected: comments on background noise show awareness of the aural mode during the recording and later listening; likewise, comments on the length of a VM relate to demands for playing the message.

on to senders' accounts and apologies in VMs (4.2) or shortly after a VM (4.3). Finally, Section 4.4 turns to recipients' metacommunicative uptake of voice memos.

#### 4.1 Announcing voice messages

A metacommunicative practice that senders use before recording a voice memo is to announce a mode switch in a text message. Such announcements are rare in the datasets (11 in the MEsA 2.0 corpus, 3 in the MoCoDa, 10 in the Spanish private collection, and 6 in the German private corpus). Still, they offer rich insights into users' conceptualisation of the use of VMs.

In the MEsA 2.0 corpus, 9 of the 11 announcements state that a VM will be sent further along the line.<sup>4</sup> They include temporal adverbs and expressions such as *después* or *luego* (later), *ahora* (in a moment) or *en cuanto pueda* (as soon as I can). These messages typically indicate a particular upcoming communicative move, such as a detailed narration (König, 2019). In Example (1), the announcement of the recording is part of the opening of the dialogue.

#### Example (1): MEsA 2.0 corpus,<sup>5</sup> dyad 1, excerpt 14, text messages + voice message

Message	Timestamp	User	Messages <sup>6</sup>
01	20/11/2016 13:38	Ana	He estado en una exposición buenísima en Bratislava <i>I was in a very good exhibition in Bratislava</i>
02	20/11/2016 13:38		<b>Después te envío un audio contándote porque ha sido suuuper interesante</b> <i>I'll send you a voice message later on to tell you because it was veeery interesting</i>
03	20/11/2016 13:38		<M1 posts three pictures>
04	20/11/2016 13:41	Juan	pinta interesante <i>looks interesting</i>
05	20/11/2016 16:50	Ana	<voice message>

<sup>4</sup> In some instances, senders also consider the addressees' situation in their announcements when they notify recipients about an upcoming VM but leave it up to the addressee to listen to it later. This way, they reflect on the affordances of VMs, which allow later reception.

<sup>5</sup> The MEsA corpus includes only the text log of several WhatsApp chats retrieved in Spain. All media files have been excluded, so only the timestamp and the type of attachment is available. Details about the content of the media files or the length of VMs are therefore not available for this corpus.

<sup>6</sup> All text messages are given in their original form, i.e., spelling errors were not corrected.

The first three messages frame Ana’s experiences in an exhibition in Bratislava as highly tellable: she announces a more extensive telling for a future moment (#02) and sends pictures of the exhibition (#03). Juan aligns with her evaluation (#04) and thus gives Ana a ‘ticket’ to say more about it. Such metacommunicative announcements frame VMs as a time-consuming activity to carry out when the conditions allow for it, particularly if extended explanations or narrations are concerned (Busch and Sindoni, 2022; König, 2019).

Announcements do not always postpone VMs to a later point; they also project a mode switch that follows immediately, as illustrated in Example (2) from the German MoCoDa. Manuel gives an extensive account (and thus implicitly apologises) for not calling Thomas on his birthday.

**Example (2):** <https://db.mocoda2.de/view/fdbKG>

Message	Timestamp	User	Content
01	08/05/2018 09:00	Manuel	Ooooooh nein <i>Ooooooh no</i>
02	08/05/2018 09:00		ich habe dich quasi vergessen <i>I virtually forgot about you</i>
03	08/05/2018 09:00		also nicht wirklich <i>that is not really</i>
04	08/05/2018 09:00		habe an dich gestern gedacht <i>I thought about you yesterday</i>
05	08/05/2018 09:00		aber nach dem Fußball habe ich mir an der Wade nen Muskelfaserriss zugezogen <i>but I incurred a rupture of muscle fibre in my calf after the football game</i>
06	08/05/2018 09:01		und irgendwie war alles kacke und ich erst spät zu Hause <i>and somehow everything was shit and I came home late</i>
07	08/05/2018 09:01		und dann habe ich vergessen anzurufen <i>and then I forgot to call you</i>
<b>08</b>	<b>08/05/2018 09:01</b>		<b>deswegen nun</b> <b><i>that is why now</i></b>
09	08/05/2018 09:02		<Sprachnachricht, selbstgesungenes Geburtstagslied> < <i>voice message, sings a birthday song</i> >
10	08/05/2018 09:02		:)
11	08/05/2018 10:00	Thomas	👋 Hallo Manuel, Deine Frau hat doch für Dich mitgratuliert, redet ihr nicht miteinander 😊 Dir gute Besserung!

			📞 <i>Hello Manuel, your wife wished me happy birthday in your name, too, don't the two of you talk to each other 😊 get well soon!</i>
12	08/05/2018 10:02		Vielen lieben Dank für Deinen Gesang 😊 <i>Many thanks for your singing 😊</i>

After posting seven individual messages with minimal delay, message #08 foreshadows that something ‘compensatory’ is coming up. In contrast to the announcement in Example (1), Manuel’s projection (#08) does not explicitly state what mode is about to follow (text message, VM, picture, etc.). In this context, the mode switch is not framed as imposed by communicative activities, such as the detailed narration announced in Example (1). Here, the VM presents an audible performance (König and Hector, 2017) and thus functions as a sign of affection and personal commitment.

The previous excerpts show that users establish a particular contextual frame for the upcoming VM with their metacommunicative announcements. On the one hand, they project or pre-figure (more or less explicitly) what they intend to do with their message. On the other hand, they present VMs as something out of the ordinary (in our examples, a detailed narration, making up for not having called). At the same time, they consider issues of timing, convenience, and commitment that are relevant for both the sender and the recipient.

## 4.2 Framing mode choice within VMs: Describing, explaining, and apologising

Accounts within a VM, at a point where the mode choice has already been made and is accessible to the recipient, are a more common form of metacommunication (our corpora yield 160 such audio postings). The following analysis focuses on the two most prominent slots in which such VM-internal accounting occurs, namely in the opening and closing sections of the VM, and examines how the addressees respond to them. Metacommunicative comments range from mere descriptions of what users are doing to explanations and explicit apologies for communicating via VM. We will also point out the recurring lexical patterns found in metalingual practices.

### 4.2.1 Initial comments

The opening part of a voice message is the slot where most VM internal accounts occur in our data: we found 23 direct explanations for the user’s choice of voice messaging. Given that we coded 1,315 VMs, this is not a common practice. These results suggest that users consider VMs as a standard and legitimate practice for WhatsApp interactions (König and

Hector, 2017). This observation is consistent with studies in CA, which emphasise that people routinely do not account for their actions since interactants usually do what is expected (Buttny and Morris, 2001: 289). Thus, accounting is a marked strategy.

When users *explicitly* account for their mode choice, they mention different reasons for voice messaging. A recurrent argument in our data is the perceived convenience of talking compared to texting. VMs are described as simpler and quicker to produce or more suitable when in a hurry or multitasking, as Excerpts (3) to (5) illustrate.

### Examples (3) to (5) (private collections)

Example	Corpus	User	Voice message (transcription)
(3)	German, UDEGA- 02JEJU01	Julia	Okay, jetzt schick ich dir einfach mal ne Sprachnachricht zurück, weil ich das auch nicht alles schreiben kann [...] <i>Okay, I'll just send you a voice message in reply now, as I also can't write everything</i>
(4)	German, INARR- 02AMLM01	Anna	Hallo, ich mach jetzt mal ne Sprachnachricht, ich bin nämlich zu faul zu tippen [...] <i>Hello, I'm just recording a voice message as I am too lazy to type</i>
(5)	Spanish, WA2-31	Verónica	Mejor te lo digo hablao porque voy ya no que no me da tiempo. Ya voy a contrarreloj. [...] <i>I'll tell you better by talking 'cause I don't have time. I'm rushing [...]</i>

One striking difference between accounts in oral conversations and messaging interactions lies in the addressees' uptake; in face-to-face interactions, recipients usually respond to an account. If it is honoured, the accounting episode ends; otherwise, more accounting is needed (Buttny and Morris, 2001: 294-295). In contrast, accounts for mode choice do not usually receive a response in our data. In their subsequent messages, recipients comment on other issues mentioned in the memo, which are thus foregrounded as focal, i.e., accounts for voice messaging are not treated as in need of 'ratification'.

The same is true for metacommunicative comments, in which users merely describe what they are doing (i.e., recording a VM) at the opening of their VM, as shown in Examples (6) and (7). These comments, too, usually did not receive a response in the subsequent discourse.

### Examples (6) and (7) (private collections)

Example	Corpus	User	Voice message (transcription)
(6)	German, DUCMS- 02FRHA01	Franzi	Hi Süße, so, ich wollte dir das jetzt mal eben kurz erzählen [...] <i>Hi sweetheart; well, I just wanted to tell you this briefly</i>
(7)	Spanish, WA2-144	Patricia	Alicia, ¿qué tal? Como siempre te envío un audio que. Madre mía. Se ha pasado el tiempo volando. <i>Alicia, how are you? As always, I send you a memo so... My goodness! Time has gone by so fast.</i>

What requires further analysis here is that such accounts and comments are formulated as implicit framing devices, as users take a stance on what they are doing in their VM. They frequently introduce VMs as spontaneous, anchored in the situation the speakers are in while recording. This spontaneity is illustrated by the heavy use of temporal adverbs, such as *jetzt* in German and the equivalent *ahora* in Spanish (now) or *gerade* (at the moment) and *ahora mismo* (right now). In the German data, speakers repeatedly use the modal particles *einfach* and *mal* (just), which also frame voice messages as something done in a rather offhand, casual, or incidental manner. The same is true for the Spanish discourse markers *bueno* or *pues*, which are also common in our recordings. With the explanations above and such stance-rich descriptions, users consider their perspectives on mode choice.

#### 4.2.2 Closing comments

The other prototypical slot in which users place VM-internal metacommunicative comments is the closing of a VM. Closing a VM without direct interaction with the addressee poses a ‘communicative challenge’ for speakers (see König and Hector, 2017). Users often repeat or summarise what they have said before to (re-)organise their contribution (König and Duan, in preparation). Metacommunicative comments form yet another closing practice. In this slot, speakers usually comment on the length of their recording, as in Excerpts (8) and (9), or apologise explicitly for what they perceive as a lengthy message (see Example 2).

#### Examples (8) and (9) (private collections)

Example	Corpus	User	Voice message (transcription)
(8)	German, LAUDA- 02ANNA02	Bine	[...] <b>Also, ich will jetzt auch nicht zu lange dich hier vollquatschen.</b> Ganz viel Spaß, dann sehen wir uns ja

			auch nächste Woche noch mal. Viel Spaß mit den anderen. Bis bald [end of message] <i>So, I don't want to gab to you too long here. Have fun; we'll meet again next week. Have fun with the others. See you soon</i>
(9)	Spanish, WA2-18	Sara	Vale, pues un besito <b>y madre mía, tres minutos</b> . Me vas a matar. Chao. Felices fiestas. Y espero, espero que estéis muy bien también. Chao chao. [end of message] <i>Okay, so kisses and, my goodness, three minutes. You're gonna kill me. Bye. Happy holidays. And I hope, I hope you are very well too. Bye bye.</i>

The evaluative metalanguage for commenting on the length of audio postings also reflects an apologetic stance: users refer to their activity by expressions such as *(voll)quatschen* (gabbing), *chapa* (bore; see Example 10 below), *rollo* (drag), or the verb *enrollarse* (getting long-winded), which frame their talk as potentially annoying or pointless for the addressee. Apologies or comments about the length of the message show consideration to the listener. In contrast to comments at the beginning of a VM, here, users refer to the overall design of the audio posting and consider the recipient's perspective. Again, we do not find a systematic uptake of these accounts in the data, indicating that senders, rather than recipients, treat VMs as accountable.

### 4.3 Metacommunicative supplements

Most accounts and metacommunicative comments in our corpora were part of the VMs themselves. However, in very few instances (12 in our data), users send a (text or voice) message dealing with their mode choice directly following their VM. We refer to these chunks as metacommunicative supplements, framed as 'postscripts' to a user's previous aural contribution. Example (10) comes from a dyadic chat in which users talk about a political meeting that Ricardo had attended the previous day.

#### Example (10): MEsA 2.0 corpus, dyad 3, excerpt 9, text messages

Message	Timestamp	User	Content
01	23/03/2017 14:34	Ricardo	<image>
02	23/03/2017 14:37	María	¡En primera fila! 🤔👁 <i>In the front row!</i> 🤔👁

03	23/03/2017 14:38	Ricardo	<voice message >
<b>04</b>	<b>23/03/2017 14:39</b>		<b>Perdon por la chapa</b> <b><i>Sorry for pestering you</i></b>
05	23/03/2017 14:39	María	Nada <i>No worries</i>
06	23/03/2017 14:39		Si me gusta <i>I like it</i>
07	23/03/2017 14:39		
08	23/03/2017 14:39	Ricardo	<image>
09	23/03/2017 14:40	María	Yo esta tarde tengo clases particulares y después voy al teatro <i>I have private lessons this afternoon and then I go to the theatre.</i>
10	23/03/2017 14:43	Ricardo	<voice message >
11	23/03/2017 14:44	María	Muy buen resumen <i>Very good summary</i>
12-16			<i>[They chat about a politician attending the meeting]</i>
17	23/03/2017 14:45	Ricardo	<voice message >
<b>18</b>	<b>23/03/2017 14:45</b>		<b>Perdon por los audios pero estoy preparando la comida</b> <b><i>Sorry for the voice messages, but I'm making lunch</i></b>
19	23/03/2017 14:46	María	Levantarse del sillón para todo <i>Getting off the chair for everything</i>
20	23/03/2017 14:46		Eso seguro <i>That's for sure</i>
<b>21</b>	<b>23/03/2017 14:46</b>		<b>Tu comida dieta</b> <b><i>Your diet lunch</i></b>

Within a short interval, Ricardo repeatedly sends VMs (#03, #10, #17) and supplements two of them with metacommunicative comments in the form of an apology (#04 and #18).<sup>7</sup> María's replies are written in this dialogue; therefore, this sequence makes the mismatch between modalities noticeable. By accounting for his mode choice, Ricardo implies that texting is their default communication mode. In contrast to the VM-internal accounts and apologies, the addressee replies to metacommunicative supplements in this case. In messages #05 to #07, María offers an 'absolution' (Robinson, 2004), declaring that she likes VMs, thus

<sup>7</sup> Ricardo's second account (#18) is also cast in the form of an excuse. He avoids negative attributions by transferring the cause of his mode choice to situational factors (Robinson, 2004: 311), in this case, being busy preparing lunch.



ratifying the mode switch. The standalone emoji in #07 can either strengthen her appreciation or mark an (unsuccessful) closing of the interaction, often marked with a kissing emoji in Spanish WhatsApp messages (Sampietro et al., 2022). In #21, she accepts that Ricardo is sending VMs because he is busy preparing lunch, and implicitly acknowledges the apology by referring to its content. While accepting an apology is one way to mitigate a potential offense, it is still recognised it as such. Therefore, both users consider long VMs in this sequence to be accountable.

In our data, metacommunicative supplements to VMs, such as those presented in the excerpt above, work better in inviting uptake than those that form part of a more elaborate VM. Not only do they consider the recipient's needs, but their presence in a separate post also provides an immediate response slot, as recipients can reply or refer to them in the adjacent message. However, they were relatively sparse in our data, suggesting that the default way to comment on one's mode choice is to announce an upcoming VM or in the VM's opening and closing sections.

#### **4.4 Recipients' comments: Deferring and apologising for a delayed reply**

So far, the analysis has focused on how senders of a VM comment on their communicative behaviour. The analysis now turns to metacommunication in response to VMs. As indicated above, recipients rarely reply to other users' accounts; however, they can put forward their meta-perspective on senders' VMs, which usually concern their listening practices. In our data, we observed such metacommunicative comments in voice and text messages. In VMs, users recurrently refer to the moment they (finally) find the time to listen to an earlier memo; when texting, they repeatedly state that they cannot immediately listen to (and thus respond to) a VM. In both cases, listeners account for the lack of immediate uptake. These responses suggest that, although messaging generally affords asynchronous interaction, users still expect a timely response from their chat partners.

Sometimes, addressees defer their response in a text message because they cannot listen to a VM immediately or announce that they will listen to it later. In Example (11), the addressee of the VM (Mateo) explains that he cannot access Sonia's VM for the time being.

**Example (11): MEsA 2.0 corpus, dyad 4, excerpt 24, text messages**

Message No	Timestamps	User	Content
01	04/12/2017 23:00	Sonia	<voice message >
02	04/12/2017 23:00	Sonia	<voice message >
03	04/12/2017 23:01	Sonia	<picture >
04	04/12/2017 23:09	Sonia	Tengo ganas de verte <i>I look forward to seeing you</i>
05	04/12/2017 23:09	Sonia	Voy a Granada mañana pero me vuelvo al día siguiente porque trabajo el 7 <i>I'm going to Granada tomorrow but I'm going back the next day 'cause I'm working on the 7<sup>th</sup></i>
06-07			<i>[Sonia and Mateo comment on their next trips and a possible future encounter in Granada]</i>
08	04/12/2017 23:24	Mateo	Creo que voy a tener un par de semanas <i>I think I'll have a couple of weeks</i>
<b>09</b>	<b>04/12/2017 23:24</b>	<b>Mateo</b>	<b>Escucho los audios luego, que estoy fuera</b> <b><i>I'll listen to the memos later, as I'm out.</i></b>
10	05/12/2017 01:50		Si, me han felicitado al final <i>Yes, I finally got my birthday greetings</i>

Mateo refers to his current activity and explains why he cannot listen to Sara's VMs (#09) immediately, but replies later (#10). The message documents that the sender expects a timely response to her VMs, similar to text messages, which are answered promptly. In his response (#10), Mateo does not comment again on listening to the VM, but only addresses its content.

Recipients also referred to having just listened to the previous memo or explaining why they could not answer sooner. Example (12) shows a dialogue between two friends who study in different places in Germany, in which Sonja comments on her late reply.<sup>8</sup>

**Example (12): DUCMC\_02LISO-01 (German private collection)**

Message No	Timestamps	User	Content
01	06/02/2018 14:25	Lisa	<voice message, 1:54> ((talks about a test she passed, possible next meeting in person, reasons why Sonja is moving out))

<sup>8</sup> See König and Duan (in preparation) for an extensive discussion (and a detailed conversation-analytic transcription) of this dialogue.

02	02/04/2018 17:52		<voice message, 0:23> ((birthday wishes, tells Sonja that she is in town for a few days and that they might meet in person))
03	02/04/2018 23:29	Sonja	<p>&lt;voice message, 1:48&gt; [0:00-0:15] Hallo Lisa, ähm erstmal vielen, vielen, vielen lieben Dank. ((Räuspern)) Ähm ich hab mir jetzt endlich mal deine alte Sprachmemo angehört. Das war auch wirklich Zeit. Und ja, das is jetzt alles schon bisschen veraltet, aber trotzdem nochmal herzlichen Glückwunsch zu deinem bestandenen Testat.</p> <p><i>Hi Lisa, first of all thank you very very very much. ((clears her throat)) Uhm I finally managed to listen to your old voice message. It really was about time. And yes, all of this is a already a bit outdated, but anyway congratulations that you passed your test.</i></p> <p>[1:31-1:44] Sorry, dass das so lang gedauert hat, ne? Ich hatte in den Semesterferien irgendwie, weiß ich nich, ich hatte irgendwie so viel zu tun. Das war irgendwie viel mehr als während des Semesters. Aber naja, egal. Ich antworte jetzt auf jeden Fall wieder schneller. Und joa, ich hoffe, du hast noch schöne Restferien.</p> <p><i>Apologies that it took so long. I had so much during the semester break somehow, I don't know, I had so much to do somehow. It was somehow much more than during the semester. But well, anyway. I will definitely answer more quickly now. And yeah, I hope you can enjoy the rest of your break.</i></p>
04	04/04/2018	Lisa	<p>&lt;voice message, 2:00&gt; [0:00-0:14] Ja gar kein Problem. Mir passiert das auch echt oft, und vor allem, wenn man eh nich so mega viel Kontakt hat, dass man dann irgendwie sagt, okay, man nimmt sich dann Zeit zum Antworten und macht das irgendwann anders. Und dann schiebt man so ganz viel vor sich her. Und deswegen. Äh mir passiert das auch schonmal öfters ((lachen)). Ähm aber freut mich, dass du schön gefeiert hast.</p>

			<p><i>Yeah no problem at all. This happens to me really often, too, and especially if you don't have such a close contact that you somehow say, okay you set a time for responding and you do it sometime else. And then you push along quite a lot. And that is why. Uh this also happens to me often ((laughs)). Uhm but I am glad that you had a nice party.</i></p>
--	--	--	---

Nearly two months passed between messages #01 and #02. In #03, Sonja first replies to Lisa's birthday wishes in message #02 before referring to the long time lag between the first two messages. First, she describes how she found the time to listen to the VM Lisa sent in February. Sonja also stressed the necessity of doing so before dealing with the issues mentioned in Lisa's older VM. Sonja closes her audio posting with an explicit apology for her delayed reply and an account of why she could not find the time to respond earlier. She promises to reply at a faster pace. Sonja's metacommunicative comments also express consideration for Lisa's perspective; Sonja 'brackets' her VMs with these apologetic comments.

In contrast to apologies that deal with the length of VMs, comments on the time lag between messages invite a response in the subsequent message. In #04, Lisa acknowledges and relativises Sonja's apology, stressing that she understands from her own experience how such delays can occur. Therefore, again, we see how recipients' metacommunication leads to explicit negotiation of the timing of responsive moves in messaging.

## **5 Discussion and conclusions**

This paper has studied VMs, a popular feature of many instant messaging applications that have proliferated in many countries worldwide. In particular, this study focused on voice messaging in Germany and Spain, two countries where sending VMs is an established practice among users of instant messaging applications. Although voice messaging has introduced aural communication to hitherto primarily text-based mobile messaging, recorded voice communication is by no means new, having been around since the late 19<sup>th</sup> century. Analogies between VMs on WhatsApp and earlier forms of recorded communication lie in the perception and use of the devices, as well as the discursive design of the messages themselves. We found that users feel accountable for using VMs, similar to messages

recorded on answering machines (see Liddicoat, 1994). Nevertheless, unlike phones and answering machines users, WhatsApp customers can choose the posting mode (text, VM, photos, and videos). Despite this free choice of contribution mode, the users provided explanations for the use of VMs. Likewise, WhatsApp users can decide the length of the message; however, they usually apologise if a recording is deemed too long. Recipients of VMs in our data provide explanations for not being able to listen to a VM immediately, which violates the expectation of a timely response.

Our analysis focused on these forms of metacommunication regarding voice messaging in transmodal WhatsApp chats to capture the joint negotiation of meta-perspectives on VMs. In addition to analysing the content and linguistic form of *in situ* metadiscourse, our approach conceptualises metacommunication as a sequentially embedded practice. Such an approach highlights that different sequential slots can be linked to specific actions that accounting performs in messenger chats. Although not a common phenomenon, metacommunicative accounts offer deep insights into how WhatsApp users perceive and negotiate the shift from written to aural modes, and how their media ideologies are shaping digital conduct.

The analysis shows that switching to the aural mode on WhatsApp can be treated as accountable in different slots and that accounting is associated with different participation roles: it is usually volunteered by senders but not made relevant by VM recipients. Announcing a VM frames the upcoming audio posting as something remarkable or unusual (e.g., a lengthy narration) for the sender or recipient. With VM-internal accounts at the beginning of audio postings, users often frame the upcoming verbal activity as spontaneous or explain their mode choice, mainly from the sender's perspective. For instance, users point out that a VM is quicker than texting and allows them to multitask (e.g., recording a VM while walking). In contrast, with VM-internal accounts in the closing section of a VM, users often express an apologetic stance (e.g., commenting on length), considering the recipient's perspective. Unlike face-to-face interactions, such accounts do not usually invite any uptake by recipients. This differs from apologetic metacommunicative supplements, which tend to elicit responses in which the recipient accepts the sender's apology to choose voice messages. In the response slot, recipients' accounts mainly deal with the expectation of a timely reply (they postpone their response or apologise for a late reply). Thus, accounts can be used in

several ways. They can project, frame, comment, or respond to the choice of aural modality in transmodal messaging.

At the same time, users' metacommunication reveals their beliefs about the nature of this mediated mode of interaction, that is, their media ideologies. First, as already observed in previous research, VMs are considered to be better suited for certain actions, such as long explanations and narrations or for 'vocal performances' (see Busch and Sindoni, 2022; König and Hector, 2017). Second, on a more general level, we see that by accounting for using a VM, users frame their mode choice as deliberate, not as a practice determined solely by the application's technical affordances, individual preferences, or external factors. Note, however, that users account for VMs rather than text messages, and thus frame audio postings as a marked alternative to texting, the default or non-accountable mode on WhatsApp.<sup>9</sup>

Another critical ideological perspective of the media concerns timing. Meta (no date) describes WhatsApp voice messaging as a way to 'instantly communicate with contacts and groups', i.e., a means 'to deliver important and time-sensitive information'. However, despite designers' intentions, our research shows that voice messaging, albeit sometimes framed as spontaneous, is treated and understood as a asynchronous communication mode. Senders announce that they will be recording a VM at another time or 'allow' recipients to listen to their VMs later. Receivers announce when they are able to listen or apologise for not responding earlier. This implies that users treat VMs as a contribution mode requiring specific circumstances for reception and response (e.g., time and attention), as opposed to the sense of immediacy usually associated with texting.

In summary, our analysis shows how users conceptually differentiate between text and VMs. Either contribution accomplishes different actions (e.g., detailed narrations, non-interruptive deferrals), sets in motion various response relevancies (concerning the time lag between messages), and requires different efforts (easy to produce vs. requiring specific conditions to listen to them).

In future studies, research on media ideologies should combine the results of the sequential analysis of *the in situ* metacommunication on VMs with interview and survey data about this particular contribution mode. While VMs have received considerable negative attention in

---

<sup>9</sup> This is different from the exchange of photos, and, to a lesser extent, videos, which are commonly not announced or accounted for in WhatsApp chats (see König, forthcoming).

media discourse (e.g., in memes, see König, forthcoming), strikingly, downright rejections of VMs do not appear in our data. Such discrepancies between the discursive meta-reflection of VMs and users' metacommunication in actual messaging chats could further complement the scarce literature on the role of spoken language in digital interaction (Baron, 2013). Research on voice messaging would also benefit from comparative studies, as user data point to cultural disparities in the adoption of voice messaging on WhatsApp and other platforms. Further research is needed to determine whether these preferences are consistent with cultural perceptions of mobile phone use in public (i.e., the appropriateness of speaking on the phone in public spaces) or privacy concerns. Larger and more culturally diverse corpora of transmodal chats are required to gain access to the broader network of media ideologies involved in voice recording and to understand the success of voice messaging worldwide.

### **Acknowledgements**

This research was partially supported by funding from Generalitat Valenciana (Travel Grant CIBEST/2021/174).

## **6 References**

- Álvarez-Cáccamo C and Knoblauch H (1992) 'I was calling you': Communicative patterns in leaving a message on an answering machine. *Text* 12(4): 473–505.
- Antaki C (1996) Explanation slots as resources in interaction. *British Journal of Social Psychology* 35: 415–432.
- Baron N (2013) Do mobile technologies reshape speaking, writing, or reading? *Mobile Media & Communication* 1(1): 134–140. DOI: 10.1177/2050157912459739.
- Baron N (2020) Talking, reading, and writing on smartphones. In: Ling R, Fortunati L, Goggin G, Lim SS and Li Y (eds) *Mobile Communication and Society*. New York: Oxford University Press, pp. 408–421.
- Beißwenger M, Imo W, Fladrich M and Ziegler E (2019) <https://www.mocoda2.de>: A database and web-based editing environment for collecting and refining a corpus of mobile messaging interactions. *European Journal of Applied Linguistics* 7(2): 333–344. DOI: 10.1515/eujal-2019-0004.

- Bereznak, A (2018) Is the era of voice texting upon us? *The Ringer*, 5 November. Available at: <https://www.theringer.com/tech/2018/11/5/18056776/voice-texting-whatsapp-apple-2018> (accessed 17 May 2023).
- Boczkowski P, Matassi M, and Mitchelstein E (2018) How young users deal with multiple platforms: the role of meaning-making in social media repertoires, *Journal of Computer-Mediated Communication* 23(5): 245–259, DOI:10.1093/jcmc/zmy012.
- Bundesverband Digitale Wirtschaft (BVDW) (2019) Digital Trends Umfrage zur Nutzung von Sprachnachrichten. Available at: [https://www.bvdw.org/fileadmin/bvdw/upload/studien/BVDW\\_Nutzung\\_Sprachnachrichten.pdf](https://www.bvdw.org/fileadmin/bvdw/upload/studien/BVDW_Nutzung_Sprachnachrichten.pdf) (accessed 10 July 2022).
- Busch F and Sindoni M (2022) Metapragmatics of mode-switching: Young people’s awareness of multimodal meaning-making in digital interaction. In: Groff C, Hollington A, Hurst-Harosh E, Nassenstein N, Nortier J, Pasch H and Yannuar N (eds) *Global Perspectives on Youth Language Practices*. Berlin, Boston: De Gruyter Mouton, pp. 245–264. DOI: 10.1515/9781501514685-016
- Buttny R and Morris GH (2001) Accounting. In: Robinson WP and Giles H (eds) *The New Handbook of Language and Social Psychology*. Chichester: John Wiley & Sons, pp. 285–301.
- Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (CETIC) (2022). Penetration rate of sending voice messages or video calls among internet users in Brazil from 2017 to 2022. Available at: <https://www.statista.com/statistics/1083507/brazil-sending-voice-messages-video-calls-penetration-rate/> (accessed 30 May 2023).
- Dingwall S (1992) Leaving telephone answering machine messages: Who’s afraid of speaking to machines? *Text* 12(1): 81–101.
- Elliott V and Phorn B (2021) Fifty percent of Facebook Messenger’s total voice traffic comes from Cambodia. Here’s why. Available at: <https://restofworld.org/2021/facebook-didnt-know-why-half-of-messengers-voice-traffic-comes-from-cambodia-heres-why/> (accessed 1 July 2022).
- Emery R (2018) Whatsapp voice messaging as an emergent digital practice: A multi-method analysis. *Cahiers de l'ILSL* 55: 135–157.



- Fisher, C (1992) *America Calling: A Social History of the Telephone to 1940*. Berkeley and Los Angeles: University of California Press.
- Fernández Ardévol M, Bellotti F, Mulargia S, Rosales A and Comunello F (2022) ‘I do it my way’: Idioms of practice and digital media ideologies of adolescents and older adults. *New Media & Society* 24(1): 31–49. DOI: 10.1177/146144482095929.
- Fortunati L and Baron N (2017) Evolving patterns of mobile call openings and closings. In Serrano Tellería A (ed) *Between Public and Private in Mobile Communications*. New York: Routledge, pp.37–59.
- Fortunati L (2001) The mobile phone between orality and writing. In *E-usages – 3rd International Conference on Uses and Services in Telecommunications*. Paris: France Telecom, pp. 312–321.
- Fuentes Rodríguez C (2021) (dir) Corpus MEsA 2.0. Available at: <http://www.grupoapl.es/materiales-corpus/corpus-mesa> (accessed 1 July 2022).
- Garfinkel H (1967) *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
- Gershon I (2010) Breaking up is hard to do: Media switching and media ideologies. *Journal of Linguistic Anthropology* 20(2): 389–405. DOI: 10.1111/j.1548-1395.2010.01076.x.
- Giles D et al. (2015) Microanalysis of online data: The methodological development of ‘digital CA’. *Discourse, Context & Media* 7: 45–51. DOI: 10.1016/j.dcm.2014.12.002.
- Gold R (1991) Answering machine talk. *Discourse Processes* 14(2): 243–260. DOI: 10.1080/01638539109544784
- Heritage J (1988) Explanations as accounts: A conversation analytic perspective. In: Antaki C (ed) *Analyzing Everyday Explanation: A Casebook of Methods*. London: SAGE, pp. 127–144.
- Herring S (2019) The Co-evolution of computer-mediated communication and computer-mediated discourse analysis. In: Bou-Franch P and Garcés-Conejos P (eds) *Analyzing Digital Discourse: New Insights and Directions* London: Palgrave Macmillan, pp. 25–67.
- Hobbs P (2003) The medium is the message: politeness strategies in men’s and women’s voice mail messages. *Journal of Pragmatics* 35: 243–262.

- Hutchby I and Tanna V (2008) Aspects of sequential organisation in text message exchange, *Discourse & Communication* 2(2): 143–164.
- Knight K (2006) Transformations of the concept of ideology in the twentieth century. *American Political Science Review* 100(4): 619–626.
- König K (forthcoming) Transmodal messenger interactions: The sequentiality of text and audio postings in WhatsApp chats.
- König K (2019) Narratives 2.0 – A multi-dimensional approach to semi-public storytelling in WhatsApp voice messages. *Journal für Medienlinguistik* 2(2): 30–59. DOI: 10.21248/jfml.2019.10
- König K (2021) Sprachnachrichten als Zugang zu authentischer Mündlichkeit im Kontext Deutsch als Fremdsprache. In: Günthner S, Schopf J and Weidner B (eds) *Gesprochene Sprache in der kommunikativen Praxis. Analysen authentischer Alltagssprache und ihr Einsatz im DaF-Unterricht*. Tübingen: Stauffenburg, pp. 247–277.
- König K and Hector TM (2017) Zur Theatralität von WhatsApp-Sprachnachrichten. Nutzungskontexte von Audio-Postings in der mobilen Messenger-Kommunikation. *Networx* 79. Available at: <https://www.mediensprache.net/networx/networx-79.pdf> (accessed 8 July 2022).
- König K and Hector TM (2019) Neue Medien – neue Mündlichkeit? Zur Dialogizität von WhatsApp-Sprachnachrichten. In Marx K and Schmidt A (eds) *Interaktion und Medien*. Heidelberg: Winter, pp. 59–84.
- König, K and Duan, L (in preparation) Mündlichkeit im Posting-Format. Praktiken der dialogischen Kommunikation über Sprachnachrichten per WhatsApp und WeChat.
- Liddicoat AJ (1994) Discourse routines in answering machine communication in Australia. *Discourse Processes* 17(2): 283–309. DOI:10.1080/01638539409544870
- Ling R (2004) *The Mobile Connection: The Cell Phone's Impact on Society*. San Francisco: Morgan Kaufmann.
- Meredith J (2019) Conversation analysis and online interaction. *Research on Language and Social Interaction* 52(3): 241–256. DOI: 10.1080/08351813.2019.1631040

- Meta (2022) New voice message features on WhatsApp. Available at: <https://about.fb.com/news/2022/03/new-voice-message-features-on-whatsapp/> (accessed 7 July 2022).
- Meta (no date) How to send voice messages. Available at: [https://faq.whatsapp.com/779704622749936/?helpref=uf\\_share](https://faq.whatsapp.com/779704622749936/?helpref=uf_share) (accessed 28 July 2022).
- Mooney M (2009). The history of magnetic recording. Reprinted from *Tape Recording*. Available at: <https://www.historyofrecording.com/support-files/thehistoryofmagneticrecording.pdf> (accessed 30 May 2023).
- Morton D (2000) *'Off the record': The technology and culture of sound recording in America*. Piscataway NJ: Rutgers University Press.
- Pomerantz A (1980) Telling my side: 'Limited Access' as a 'Fishing' Device. *Sociological Inquiry* 50(3-4): 186–198. DOI: 10.1111/j.1475-682x.1980.tb00020.x
- Raymond CW and Stivers T (2016) The omnirelevance of accountability. Off-record accounts solicitations. In: Jeffrey D Robinson (ed) *Accountability in Social Interaction*. Oxford: Oxford University Press, pp. 321–353.
- Robinson J (2004) The sequential organisation of 'explicit' apologies in naturally occurring English. *Research on Language and Social Interaction* 37(3): 291–330, DOI: 10.1207/s15327973rlsi3703\_2
- Sampietro A, Felder S and Siebenhaar B (2022) Do you kiss when you text? Cross-cultural differences in the use of the kissing emojis in three WhatsApp corpora. *Intercultural Pragmatics* 19(2): 183–208. DOI: 10.1515/ip-2022-2002
- Scott, M. B., & Lyman, S. M. (1968). Accounts. *American Sociological Review* 33(1): 46–62. DOI:10.2307/2092239
- Statista (2022). Messenger usage by brand in the US in 2022. Available at <https://www.statista.com/forecasts/997099/messenger-usage-by-brand-in-the-us> (accessed 8 July 2022).
- Thurlow C (2017) 'Forget about the words'? Tracking the language, media and semiotic ideologies of digital discourse: The case of sexting. *Discourse, Context & Media* 20: 10–19. DOI: 10.1016/j.dcm.2017.06.008
- Wang X (2016) *Social media in industrial China*. London: UCL University Press.