Improbable Conversations: Interactional Dynamics in TikTok Duets

Susan C. Herring & Ashley R. Dainas
Indiana University Bloomington

Abstract

Reciprocal conversational interaction is seemingly a logical impossibility on the asynchronous short-form video platform TikTok. Yet users overcome the constraints and affordances of the platform to create the appearance of synchronous interaction by co-opting the *duet* feature, which allows users to respond to a previously recorded video, generating a new video in which the original video (OV) and the response video (RV) are displayed side by side. We examine how TikTok duetters create the illusion of conversing together in real time by applying conversation analysis methods to a judgment sample of TikTok duets, with a focus on strategies of turn-taking and overlapping. OVs and RVs collaborate to various extents in co-constructing duet conversations; OVs by explicitly or implicitly inviting duets, and RVs by orienting to OVs through the positioning and timing of their responses. Through these interactional dynamics, duetters performatively evoke participant roles, personae, and situational contexts, while simultaneously pursuing broader communicative goals of entertaining and growing an audience of followers.

Keywords: addressivity; affordances; (a)synchronicity; conversation; duets; interaction management; overlap; performativity; turn-taking; video

1. Introduction

There is a basic human urge to be sociable and exchange turns of talk, even when the content is banal. This urge is sometimes referred to as the conversational imperative (Baron, 1990), and it can be observed when strangers make small talk while waiting for a bus or engage in conversation during a long flight. Humans also feel impelled to converse with other humans over technological systems. Notably, they try to converse even when the technology in question does not support conversation well or, in some cases, was not designed for human-to-human communication at all, as was the case for the Internet itself (or rather, its predecessor, the ARPANET) (Lukasik, 2010). Within the scope of this article collection, we ask: How do people overcome physical, technical, and logistical challenges and route around obstacles posed by interfaces and systems that are unwelcoming/unfriendly to conversation? What (new) kinds of conversation emerge through this process?

We address these questions through an analysis of TikTok, a mobile app that features user-created, short-form, asynchronous videos. Of special interest is TikTok's *duet* feature, a platform affordance that allows two or more videos to play at the same time in a split screen display. Because of this feature (along with the *stitch* feature, which allows a user to incorporate and respond to a snippet of another user's video in their video, and the possibility for viewers to leave comments on videos), TikTok cannot be considered altogether unfriendly to conversation. However, TikTok is an asynchronous platform, and duetters do not engage reciprocally or in real time with the video they duet; rather they record themselves over a pre-existing video. This makes conversational interaction seemingly a logical impossibility. Yet some TikTok duets are virtually

indistinguishable from synchronous back-and-forth conversational exchanges. How is this effect achieved, and what do duetters gain from engaging in this behavior?

This study explores the phenomenon of conversational TikTok duets through the lens of conversation analysis (Sacks et al., 1974), focusing particularly on turn-taking and the sequential organization of exchanges. Our systematic qualitative analysis is based on a judgment sample of 100 TikTok duets that exhibit certain a priori conversational characteristics. The findings reveal that users obey the conversational imperative by creatively exploiting TikTok's duet feature to simulate conversational interaction. Response videos (RVs) orient to spoken conversational norms in duetting, inserting their turns at appropriate places in the speech of the original videos (OVs) or creatively constructing overlap as cooperative or uncooperative. Through these interactional dynamics, participant roles, personae, and situational contexts are performatively evoked.

2. Background

2.1 TikTok Duets

TikTok is the global version, released in 2018, of the Chinese app Douyin, which was launched in 2016. Many of TikTok's features, including duets, were inherited from Musical.ly, an app acquired by Douyin in 2017, which let users create and share 15-second lip-sync videos.[1] Since then, TikTok has evolved to include many other kinds of content, including dancing, cooking, do-it-yourself, and micro-lectures.

The traditional layout of TikTok duets is represented schematically in Figure 1, with the response video played alongside the previously recorded video, although since 2020 other layouts are available. Duets extend the activity of lip syncing or singing along with recorded music, for example, by allowing two users to perform in separate videos together. The paradoxical nature of the duet affordance (both together and not-together) can make it seem challenging and fun, contributing to its popularity – almost half of TikTok users duet occasionally.[2] When duets are done skillfully, the output is a clever illusion, the appearance that the creator of the OV and the creator of the video that responds to it, the RV, recorded the duet video collaboratively in real time, for example, that they are dancing or singing together. Users may feel rewarded by a sense of accomplishment for achieving this illusion. At the same time, duets serve as entertainment performed for the benefit of TikTok viewers, who appreciate clever and entertaining content.
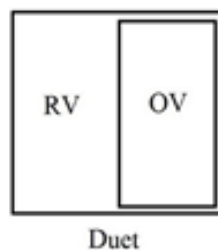

Figure 1: Schematic representation of embedding of OV in an RV Duet

---

[1] https://www.big3.sg/blog/a-brief-history-of-tiktok-and-its-rise-to-popularity, accessed 1/24/24
[2] https://www.houseofmarketers.com/47-tiktok-user-stats-and-tiktok-statistics-2022/, accessed 1/24/24

From its origins in lip syncing and dancing together, duet use has expanded to include myriad other ways an RV can react to or engage with an OV. Young content creators have built careers around duetting and reacting to videos in signature ways,[3] and celebrities duet to react to fan videos and to increase their fan base, for example by posting "duet challenges" (Abidin, 2020). Some TikTok users view duets as "a second-rate form of creative content" that "mooches" off original content (Herman, 2023, p. 17). Still, the structure of duets undeniably lends itself to practical uses. Musicians utilized duets to record collaboratively during the Covid-19 pandemic lockdowns (Kaye, 2022). Duets are also well-suited to acting and foreign language learning; OVs design their video to be duetted by leaving pauses for RVs to respond and by scripting or partially scripting what the RV should say (Herman, 2023; Kåhlman, 2023). In such duets, the OV is clearly an active participant, as illustrated in Kåhlman's (2023) multimodal conversation analysis of how the gestures and vocalizations in a language teacher's OVs prompt young children to respond in kind in their RVs.

A more sophisticated use of the TikTok duet feature, albeit one that has received little scholarly attention, is to duet an OV in such a way that it creates the illusion that the OV and RV are conversing with each other in real time (Figure 2).
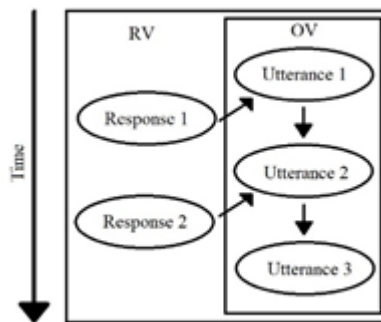

Figure 2: Schematic representation of turn-taking in a conversational duet

Constructing such a conversational duet requires interactional skill on the part of the TikToker recording the RV. Crucially, it involves managing the mechanics of turn-taking (Sacks et al., 1974), that is, identifying when it is appropriate to take a turn at talk when someone else is speaking, as well as precisely timing the insertion of one's contributions so as to create a smooth exchange of turns (cf. Sacks et al., 1974). While most adults manage turn-taking in speech with little conscious effort, turn-taking in TikTok duets is complicated by the fact that the creator of the OV might not have been aware that their video would become part of a conversation; they might, for example, speak quickly, leaving little space for an RV to insert responses. Moreover, since the OV was recorded earlier, its creator cannot respond dynamically to the RV or adjust the timing of their utterances; OV's turns are fixed. This fact problematizes the assumption of conversation analysts that conversation is a collaborative ongoing achievement in which the role of speaker alternates dynamically between participants (Schegloff, 1982). Especially given these challenges, the illusion of natural-seeming conversation in duets can be considered an "interactional achievement" (Schegloff, 1982).

---

[3] https://newsroom.tiktok.com/en-us/trending-on-tiktok-duet-me, accessed 1/24/24

2.2 The Conversational Imperative

The term 'conversational imperative' has been invoked sporadically across different disciplines but has not been theorized by any of them. In rhetoric and composition, it has been used to argue that there is a necessary relationship between conversation and collaborative learning (cf. Duffy, 2014). In linguistic pragmatics, conversational appropriateness (King & Sereno, 1984) and politeness (Sarfo-Adu & Osei, 2021) have been claimed to be conversational imperatives, in the sense of necessary and expected aspects of cooperative conversations. Our usage follows Baron (1990), who defines the term broadly as "the desire to speak … when in the presence of other people … even when we have nothing much to say" (p. 31). Baron's comments were made in reference to the primacy of conversation in child language acquisition. Indeed, Jordania (2009) argues that the imperative to converse has an evolutionary basis, in that for primitive humans, silence was a signal of danger, and that an unease with silence remains with us today.

One type of evidence for the conversational imperative is the prevalence of "small talk," or what Malinowski (1923) termed phatic communion. According to Malinowski, phatic communion primarily serves a bonding function, whereby "ties of union are created by a mere exchange of words" (1923, p. 478). Small talk is common across the globe, although in some cultures, such as Finland and Japan, people tend to be more comfortable with silence (Petkova, 2015), and some neurodivergent populations prefer interest-based sociality to talking for the sake of talking (Rosqvist, 2019). Small talk is also included in fictional dialogue in novels and films to simulate unscripted spoken conversation (Urbanová, 2007). Urbanová (2007) observes that in addition to phatic communion on conventional or stereotypical topics, which mainly function to "establish[...] a social contact" independent of context, many literary uses are individualistic or idiosyncratic, embedded in a peculiar context of situation, to express humor, irony, or mockery. Fictional dialogue is similar to TikTok conversational duets in that both mimic spontaneous spoken conversation. Like fictional dialogue, the duets illustrate both conventional and idiosyncratic conversations, and many are humorous.

Another type of evidence for the conversational imperative is when people converse even when they must overcome obstacles in order to do so. This can be seen on websites and platforms intended for activities such as file sharing, broadcasting, and gaming, whose design discourages conversational exchanges. Yet users of these platforms, driven by the urge to converse in some form, manipulate the available technological affordances and constraints of the platforms in order to do so (e.g., Herring et al., 2009; Zelenkauskaite & Herring, 2008). In the present article, we refer to this as the computer-mediated conversational imperative, and we illustrate it by analyzing the ways in which users of TikTok co-opt the duet feature to create the illusion of conversing in real time in videos that do not technically allow for turn exchanges.

2.3 Conversation in CMC

In this study, we draw on the broad definition of conversation proposed by Herring (2010) for text-based computer-mediated communication (CMC): "any exchange of messages between two or more participants, where the messages that follow bear at least minimal relevance to those that preceded or are otherwise intended as responses" (n.p). Thus, conversations involve back-and-forth exchanges, and responses to initiating turns are typically topically and pragmatically relevant

(cf. Grice, 1975). We also expect that participants will address one another, explicitly or implicitly, to indicate for whom individual turns are intended. This is especially important in environments with multiple potential addressees and in textual CMC in general, where the visual and gestural cues that signal turn allocation in face-to-face communication are absent (cf. Kendon, 1967).

Herring (2010) further proposes that the definition of 'conversation' depends on the available communication technologies, and that technological change is driving an expansion of its meaning. While both synchronous and asynchronous text-based CMC are now generally accepted as conversational, new forms are emerging on live streaming, gaming, and broadcasting platforms (for discussion of other emergent forms of computer-mediated conversation, see Herring, 2013). For example, an analysis of text messages (SMS) sent to an Italian interactive television (iTV) program found that some texters used the TV interface like a chat forum (Zelenkauskaite & Herring, 2008). However, iTV poses numerous obstacles to conversation, including the need for users to monitor the TV constantly for briefly displayed text messages and to determine for whom they are intended. While more traditional conversational modes (e.g., regular text messages or email) would have been easier to use, some iTV SMS users appeared to derive satisfaction from the challenge of communicating in improbable ways. There are parallels in this regard between iTV SMS and conversational TikTok duets.

In order to create conversational coherence in "noisy" environments such as iTV SMS and multi-participant chat rooms, CMC users have innovated strategies to explicitly relate turns and participants to one another. These include addressivity, e.g., invoking another participant by name (Herring & Chae, 2021; Werry, 1996), and quoting, or embedding one message into another to create the impression of turn adjacency (Severinson Eklundh, 2010). In duets, each participant has at least two potential addressees: the other participant (OV or RV) and the audience of TikTok viewers. (Hereafter, OV and RV are used to refer both to the video creator and the video.) Addressivity in duets can thus be an indicator of conversationality: Rather than addressing the audience, in natural spoken conversation RV and OV should address each other – indirectly, through cohesive devices, or directly, by name, second-person pronouns such as 'you,' and/or imperative verb forms.

Video-mediated communication (VMC) is another productive source of new conversational forms, including "talking heads" (Licoppe & Morel, 2012), mode-switching (Sindoni, 2021), and cross-modal exchanges (Rosenbaun et al., 2016). In synchronous video chat via Skype, Zoom, or Apple FaceTime, people take turns speaking and need to minimize overlaps and interruptions, like in spoken conversation. The interfaces pose particular challenges, though; for example, in Zoom interactions it can be difficult to gain the speaking floor through gaze and gesture (cf. Kendon, 1967); instead, users signal possible turn transition places (TRPs) by means of vocal indicators, such as pausing, lowered speech volume, and laughter (Hollingsworth, 2022).

In contrast, although asynchronous VMC can involve an exchange of videos where each video constitutes a turn, "conversations" of this sort are uncommon. Pilhaja (2011) analyzed a "video thread" that occurred between a Christian and an Atheist YouTuber; however, the thread only involved the exchange of three videos. Adami (2009) analyzed a lengthy and popular YouTube thread in which YouTubers shared their video responses to the question: "Where Do You Tube?"; however, interaction among the responding videos was minimal. The relative paucity of

asynchronous video conversations may be due to the (typically) longer length and the greater amount of time and effort traditionally required to create a video, and the fact that YouTubers do not always indicate which prior video(s) they are responding to.

TikTok has inherited some characteristics of older video sharing platforms such as YouTube. Both are asynchronous, broadcast in a one-to-many format, allow comments below the videos, and tend to feature one person who delivers the primary content. At the same time, TikTok's many multimodal affordances – text captioning (i.e., converting audio into text displayed on a video), hashtags (i.e., use of "#" before a word or phrase to categorize content), text to speech (i.e., a default synthesized voice that reads text inputted by the user), audio/video filters (i.e., special effects such as beautification or altering the speaker's voice), background music, and graphics – set it apart from older video sharing platforms. Moreover, the ease of short-form video creation on TikTok's smartphone app facilitates the exchange of whole videos, and features such as stitches and duets encourage TikTokers to interact within a single video. How turn-taking and overlap are managed in these new forms of conversation is a question that has yet to be explored. Indeed, TikTok as a new form of VMC raises a unique set of issues related to conversation management that call for investigation.

2.4. Conversation analysis

This research draws on conversation analysis, a qualitative research methodology that originated in sociology (e.g., Sacks et al., 1974; Schegloff, 1982) but that is also used extensively in linguistics, especially discourse analysis, as well as in applied fields (White, 2019) and for microanalysis of online data (Giles et al., 2015). A major focus of the methodology is the organization of turn-taking. A key insight about natural spoken conversation is that turn exchange typically takes place at a Transition Relevance Place (TRP), a point of possible completion of an utterance where speaker change is a possible next action. Speakers tend to identify TRPs and insert their turns in those places. Moreover, the normative tendency in turn exchange is that there will be be "no gap and no overlap" – meaning the conversation flows seamlessly from speaker to speaker without interlocutors leaving long pauses between turns or interrupting each other (Sacks et al., 1974).

This tendency notwithstanding, overlap of turns in speech occurs rather often. Studies report that 30–50% of all turn exchanges in multi-party meetings contain some overlap (e.g., Holler et al., 2016). Overlap can be inadvertent (e.g., when a second speaker anticipates a TRP) or intentional, to interrupt the first speaker or to show engagement. Competitive overlaps that interrupt another speaker mid-utterance are typically perceived as rude (Goldberg, 1990). However, minimal response overlaps such as "uh-huh" and "yeah" are considered cooperative and supportive of the person speaking (Schegloff, 2000). Cooperative overlap can also be evidence of what Tannen (1984) calls a high involvement style. However, the interpretation of this style varies by culture and region. Tannen (1984) analyzed a dinner table conversation between Californians and New York Jews and observed that the latter overlapped each other's speech extensively, using a high involvement style. This behavior was perceived as impolite by the Californians but as evidence of engagement and interest by the Jewish New Yorkers.

Most CMC platforms, including video-sharing sites such as YouTube, are characterized by strict sequential message exchange and typically do not allow overlap or interruption, unlike spoken conversation (Schneider et al., 2002). In contrast, although TikTok duets are technically asynchronous, RV can easily overlap OV; in fact, overlap is difficult to avoid. As we will see, duet conversations include performances of both competitive and cooperative overlap.

3. Methods

3.1 Research Questions

Our first research question asks: Within the limitations and possibilities afforded by the TikTok interface, what user behaviors and discursive practices promote the illusion that the OV and the RV are conversing together in real time? Following Urbanová's (2007) observations on fictional dialogue, we also ask what kinds of situational contexts are invoked through the interactional dynamics and content of duets (e.g., conventional/stereotypical, as opposed to idiosyncratic)? How, if at all, do these practices support the computer-mediated conversational imperative?

We expect that in order to produce a natural-seeming, cooperative conversational TikTok duet, an RV needs to identify potential TRPs in the OV and time the insertion of their responses precisely in order to avoid gaps and overlaps, and in particular, the appearance of interrupting. If overlap is unavoidable, they should time it so as to make the overlap appear intentional (e.g., as purposeful impoliteness or a high involvement style [Tannen, 1984]). The RV should address OV directly, avoiding indirect, 3rd-person forms of reference. Finally, we expect that duetters will tend to invoke humorous or entertaining situations, given that the constructed conversations are performed in part for the benefit of TikTok audiences.

3.2 Data

To address the above questions, we manually collected a judgment sample of 100 English-language videos that were publicly available on the TikTok app in fall 2021 and spring 2022. Variants of the word "duet," e.g., #duets, #duetted were inputted into TikTok's search function. In order to be included in the sample, a video had to 1) utilize TikTok's duet function; 2) include at least two consecutive "exchanges" between RV and OV in any modality (except singing); and 3) at least one of the participants had to directly address the other at least once. This sample is not generalizable, since TikTok's algorithm adapts to user preferences in real time, feeding researchers more of the kinds of videos they already favorited or saved. For this reason, we do not attempt to quantify the phenomena described in this study, as the numbers would not be meaningful beyond our sample. Rather, our goal was to assemble and qualitatively explore a sample of duets that were potentially conversation-like, in that they included directed verbal exchanges between RV and OV, which, following Herring (2010), we considered to be a minimal requirement for a conversation.

3.3 Analytical Approach

The approach to conversation analysis we adopt in this study is qualitative systematic analysis. Systematic conversation analysis "involves transcribing a sequence and then analyzing it systematically using previous principal findings in CA to "unpack" the sequence and identify

patterns in the data (White, 2019, p. 480). We adapted this approach to take into account the unique affordances of TikTok duets (cf. Giles et al., 2015) by incorporating an iterative data-driven step prior to transcription to allow bespoke patterns to emerge from our data.

We first viewed the videos in the sample repeatedly and jointly categorized them according to characteristics that emerged from the corpus related to turn-taking (e.g., timing and overlap), and addressivity (e.g., personal pronoun use). We also examined the contributions of OV to the duets (e.g., monologue, one-sided dialogue), as well as the degree to which OV anticipates being part of a conversation with an RV. From this, we identified recurrent patterns and transcribed duets that best illustrated those patterns using a modified and simplified version of Jefferson's (2004) transcription system for spoken conversation (see Appendix). Finally, we analyzed each transcribed duet systemically, working turn-by-turn, with a close focus on their interactional dynamics. These dynamics are discussed in detail in the Findings section below.

4. Findings

4.1 Contribution of the OV

A basic precondition for a duet is that the OV must have set permissions to allow duetting. A video creator can disable the duet feature, although in order to be part of the duets sampled in this study, all OVs had to have this setting enabled. Beyond this, TikTokers can choose to invite duets explicitly in their spoken dialogue (by saying, e.g., "duet this" or "[Name,] please respond to this video") or implicitly by leaving gaps in their speech or by speaking slowly. They can also script how RV should respond using overlaid text.

Scriptedness is a matter of degree. The OV can fully script both sides of a conversation while performing one half of it. A potential duetter would then only need to perform their assigned lines, as described by Herman (2023) for acting and foreign language learning duets. We excluded such examples from our data, as they did not seem like natural conversations. An OV can also partially script a conversation, as when the OV is a one-sided conversation (e.g., as if talking on the phone to someone we cannot hear) but does not explicitly tell a potential duetter what to say. Example 1 below features a partially-scripted OV. An OV can also be unscripted, as in Example 2.

Additionally, the OV shapes the nature of the duet conversation as regards overlap. In both scripted and partially scripted OVs, the OV intentionally leaves gaps for an RV to fill. An OV can also unintentionally leave gaps in their natural flow of speech, as for example when sharing a personal anecdote or illustrating how to do an activity (Example 2). An OV may also leave no gaps in their speech if they speak quickly (Example 5), making it difficult for RVs to insert contributions without appearing to interrupt or talk over OV. Conversely, OVs with slower speech tend to have more gaps, which RVs sometimes exploit to insert their utterances. In the examples below, overlapping speech is indicated in square brackets.

4.2 Strategies employed by the RV

Equally if not more important are the strategies employed by the RV, since the RV has the last opportunity to shape the duet. In this section, we discuss three patterns as regards RV's

construction of interactional dynamics: orderly turn-taking, cooperative overlap, and uncooperative overlap. We use the term uncooperative overlap to refer broadly both to overlaps that are competitive and to non-aligned overlaps that occur for other reasons, such as randomness and talking over someone in order to be heard. The duet examples presented below were selected because they constitute some of the clearest illustrations of the patterns observed in our sample.

4.2.1 Orderly turn-taking

Some RVs manage to insert their turns such that there is mostly no gap and no overlap, in keeping with the expectation of orderly turn alternation (Sacks et al., 1974). In Example 1,[4] this is possible because OV, a mother, leaves obvious and intentional gaps to be filled by her son,[5] whom she explicitly invites to duet her video.

1)

1. OV: Only duet this if you're my son. **$**
2. RV: ((waves))
3. OV: HI ZEE::! **$**
4. RV: Hi Mom **$**
5. OV: Oh::, I miss you so much. How's LA treating you. **$**
6. RV: I miss you too, and it's .. amazing **$**
7. OV: Goo::d! Well I hope you're staying hydrated and healthy! **$**
8. RV: I am ((nods)) **$**
9. OV: Doing a lot of working and less playing, I hope. **$**
10. RV: Working every day **$**
11. OV: Good **$**

The mother's questions expect certain kinds of answers, such that the son is limited in his response options if he is being cooperative, that is, responding in ways that follow the pattern established by the OV, as opposed to simply collaborating in constructing a duet conversation. Despite (or because of) this one-sided control, the effect is quite natural – except for the first line, the duet in Example 1 resembles conventional small talk between a Mom and her son who is away at college.

The duet in Example 2, in which a TikTok personality with the user ID 'angryreactions' reacts to a young woman showing how she attaches her prosthetic legs, illustrates orderly turn-taking with an unscripted OV. The RV reacts to this informative video primarily through minimal responses (e.g., Schegloff, 2000), short utterances that show listenership and engagement, which he fits in the gaps between OV's utterances.

---

[4] TikTok users have the ability to limit the visibility, use, and spread of their videos. All examples shared or referred to are excerpts from longer TikTok videos that were publicly available at the time of our data collection. Since then, some videos have been made private or deleted by their creators. To respect all the creators' privacy, we have chosen not to provide links to the remaining videos. The authors may be contacted for access to private copies of the videos.

[5] Tiktokers' gender, age, and assumed identities are mentioned when they are known to the authors (e.g., because the Tiktokers are well-known and/or because others orient to those identities) and relevant to the analysis, as well as to provide missing visual context.

2)

1. OV: Hi Tiktok! **$**
2. RV: Hey
3. OV: I just wanted to give you a little glimpse on how I put on my two prosthetic legs! **$**
4. RV: Ok
5. OV: First step: getting into a chair, **$** definitely the hardest part. **$**
6. RV: Be careful, be careful!
7. OV: Next I grab my inserts, these connect my leg to my prosthetic,
8. and they have a screw [at the bottom, which=]
9. RV: [Oh! I never knew they had a screw]
10. OV: =makes sure my prosthetic leg does not come off when I'm walking, **$**
11. RV: Makes sense
12. OV: /So/ clip it in **$**
    (etc.)

By contributing minimal responses, the RV mostly avoids overlapping OV, except when he makes a longer comment in line 9 ("Oh! I never knew they had a screw"). Even here, RV might have jumped in (or created the impression of jumping in) thinking that OV had reached a TRP after the grammatically complete utterance "and they have a screw" in line 8, resulting in the overlap in square brackets. Although this overlap violates the "rule" that turn alternation will typically be orderly, it still seems natural.

4.2.2 Cooperative overlap

Since overlap is difficult to avoid, it is sometimes leveraged for interactional effects. Example 3 illustrates cooperative overlap (Tannen, 1984). It is the resolution of a longer duet in which a self-described gamer girl (RV) pretends to pick a fight with another gamer girl whose "girly" persona speaks in a high-pitched voice (OV). This information is inferrable from the duet as well as signaled by text overlaid on the videos of OV (POV: Another gamer girl tries to fight) and RV (POV: Gamer girl tries to fight 🥺). The POV ('point of view') label on TikTok indicates performance of a role or scenario.

3)

1. OV: You know as girls why don't we just .. become friends? Why don't we game together. **$**
2. RV: OK I'm [down] **$**
3. OV: [I understand] the voice can be annoying but you know=
4. RV: I can deal with it **$**
5. OV: =we, we need the wallets. **$**
6. RV: ((Laughs)) You're funny **$**
7. OV: You understand? **$**
8. RV: Yeah I'm so sor[ry for acting hard] and all that **$**

9. OV: [Awesome.   No!=]
10. OV: =no no I just wanted I literally just came in here because I wanted to be your friend $
11. RV: Really?
12. OV: Yeah, [I-]
13. RV: [Oh my god] $
14. OV: no, let's [play] $
15. RV: [Let's play]. [/??/]
16. OV: [I'm] totally down to play! $
17. RV: Okay

The OV is this example is partially scripted. Of special note, it includes interactive utterances that seem to assume overlapping speech, as when OV says "Awesome. No! no no" in lines 9-10 and "Yeah I- no, let's play" in lines 12 and 14. RV obliges by overlapping OV, as well as by timing some utterances such that OV appears to anticipate RV's TRPs out of eagerness, such as when RV begins "Yeah I'm so sorry -" in line 8 and OV appears to jump in with "Awesome." The effect is of a highly engaged, bespoke yet natural interaction (cf. Tannen's [1984] high involvement style).

4.2.3 Uncooperative overlap

Very few truly uncooperative RVs were found in the sample that we analyzed. An exception is a duet with a young woman (OV) who is performing half of a semi-scripted conversation between a "weird" girl and a football player who are paired to work together on a class project. One young man (RV) dressed as a football player simply repeated "I play football" throughout the duet, apparently randomly. Nonetheless, he still mostly inserted his contributions near or immediately after OV's TRPs.

RVs also sometimes strategically time their turns so that they or the OV appears uncooperative. The latter is the case in Example 4. The OV is a 21-second video of an approximately 11-month-old baby pretending to talk on a smartphone and saying a new (swear) word it had learned. Because the baby is just babbling, what it is saying can be interpreted in myriad ways, and the baby can be cast in various roles. In this particular duet, a young man (RV) plays the role of a desperate person who has borrowed money, casting the baby (OV) in the role of a heartless money lender.

4)

1. RV: C'mon, I just need a little bit more time,
2. RV: [alright?] $
3. OV: [((laug[hs))] $
4. RV: [Stop,] I- I will get you your money, [ok?] $
5. OV: [((sighs))] $
6. RV: I WILL! I promise, I have some [money coming in and-]
7. OV: [((babbles))] $
8. RV: {angrily} How could a person even say that, YOU ARE AWFUL! $
9. OV: ((indistinct)) $

10. RV: You just- I'm so[rry] **$**
11. OV: [((lau[ghs))]
12. RV: [Stop] laughing, stop la[ughing!] **$**
13. OV: [((sighs))]
14. OV: {softly} fuck

In this example, the baby's otherwise innocuous laughs and sighs in lines 3, 5, and 11 are reinterpreted as turns that competitively interrupt RV, creating the effect of rudeness (cf. Goldberg, 1990) and/or asymmetrical power relations. This idiosyncratic exchange (Urbanová, 2007) creates a humorous effect due to the incongruity of an infant in an (evil) adult role.

The difference between uncooperative overlap and unintentional overlap is not always clear-cut. In one duet, celebrity chef Gordon Ramsay (RV), an active TikTok duetter, repeatedly talks over a Young Asian-American man (OV) cooking one of Ramsay's recipes, so much so that the OV is sometimes hard to hear. The OV speaks quickly, leaving little room for an RV to respond. However, Ramsay has cultivated a rude persona, and so appearing to talk over OV could be intentional, at least in part. It also reflects the unequal power dynamic between Ramsay and the creators of the videos he duets – Ramsay may consider that he is doing them a favor by lending his celebrity to their videos (cf. Abidin, 2020) and may not take the time or make the effort to insert his utterances more carefully.

4.3 Beyond canonical conversation

Some duets that met our sampling criteria violate one or more of the norms of ordinary spoken conversation. We describe two trends in this section: mixed addressivity and mixed modality.

4.3.1 Mixed addressivity

To maintain the illusion of a real-time dyadic conversation, OV and RV should normally address each other using 2nd person forms. Referring to one's addressee in the 3rd person in a dyadic conversation can seem odd. However, OV and/or RV sometimes refer to each other in the 3rd person or address the TikTok audience in the 2nd person, breaking the conversational frame. In Example 5, James Charles (RV), a beauty influencer and active duetter, responds to a young woman (OV) putting on make-up inspired by Charles. The woman addresses her audience directly, asking them to tag Charles to attract his attention, before addressing him directly. Charles also switches from referring to the woman in the 3rd person to addressing her directly. In the example, *italics* indicate direct address between RV or OV, **bolding** indicates references to either RV or OV in the 3rd person, and ***italicized bolding*** indicates direct address to the audience.

5)

1. RV: Oh my gahd **her** eyebrows already look so:: good **$**
2. (some speech not transcribed)
3. OV: = I was doing my makeup almost every day for months **$**
4. OV: [And since **he** inspired me to do makeup=]
5. RV: [/???/ I love that for *you* ((laughs))]

6.   OV: =I've always wanted **him** to see at least one of my looks **$**
7.   OV: Would really appreciate it [if ***you guys*** tagged **James Charles** in the comments=]
8.   RV: [Oh my gahd looks so good wo::w, look at the blend?]
9.   OV: =and did everything ***you*** can for **him** to see this video.
10. OV: This is the final [look]
11. RV: [((Screams))] OH MY GAWD! /???/
12. RV: [Oh my gawd it looks so good]
13. OV: [And James if *you're*] watching this, *sister*, I love *you*
            and thank *you* [for inspiring me to do makeup]
14. RV: {high pitched} [I LOVE *YOU* TOO!]

The referential shifts in this duet highlight the performative nature of conversational duets by reminding us of the presence of an audience, for whom these videos are being produced as entertainment. Indeed, rather than orienting their posture or gaze towards each other, OV and RV in most of the examples presented here are looking directly into the camera. Further, like the Gordon Ramsay duet described above, James Charles' extensive overlapping of the OV suggests that as an internet celebrity invited to duet by the OV, Charles does not feel obligated to time his comments to avoid overlap; his participation is enough. Such overlaps seem *careless*, rather than unintentional or uncooperative.

4.3.2 Mixed modalities

Some conversational duets involve the mixing of modalities. Canonical conversation involves reciprocal speech; even in CMC conversations in chat rooms and web forums, all parties communicate in the same modality, text. (But cf. Sindoni [2021] on mode-switching between video and text in video chat and Rosenbaun et al. [2016] on 'cross-modal' Google Hangout conversations involving video, audio, and text chat.) In some duets, however, OV or RV is spoken, while the other half communicates simultaneously in printed out signs, text overlays, gesture, facial expressions, and/or sign language. An advantage of this strategy from the point of view of interaction management is that it enables extensive overlap without interfering with the intelligibility of the other half of the duet.

In Example 6, a young woman (OV) narrates an amusing sexual encounter in response to a prompt stitched into her video; we refer to the prompt as the initiating video (IV). The man the woman had sex with (RV) duets the OV and corroborates her story. Both the IV and the OV speak, and the IV also has a text overlay that repeats the prompt verbatim. However, RV communicates only via text, emojis, gesture, and by mouthing words. In this example, the contents in each row align temporally.

6)

#	Speech	Text	Nonverbal
1	IV: What's an embarrassing thing that you've said during sex? $	IV: What's an embarrassing thing you have said during sex?	
2	OV: So I'm making out with this guy from Tinder, never met him before the night,	RV: The guy from tinder in question 👇	RV: ((Repeatedly points to own face))
3	OV: Um, we're on my bed, I'm straddling him		
4	OV: and I pa- I start laughing, and I pause		
5	OV: and he's like what, and I go "look what I can do."		RV: ((nods))
6	OV: And I go, {Grinch voice} Four o'clock, Wallow in self pity; Four thirty, Stare into the abyss; Five o'clock, Solve world hunger, Tell no one; Five thirty, Jazz-ercise (...)	RV: this was adorable i love the grinch	RV: ((mouths Grinch monologue in unison with OV))
7	OV: {Normal voice} And what was interesting about that is like he was- he was cool with that,	RV: I was totally cool with it	RV: ((mouths)) I was totally cool
8	OV: and he like laughed, and we started making out again (...)		RV: ((mouths)) I thought it was hilarious

α The Grinch is a character in the animated film *How the Grinch Stole Christmas* (2000).

Strictly speaking, the RV in Example 6 is not conversing with the OV; he does not address her directly. Rather, OV is responding to IV, and RV is an overhearer (cf. Goffman, 1981) who co-narrates OV's story as if in an aside (to the audience and/or to IV). Although one can imagine similar face-to-face situations involving three participants, the use of multimodal resources in the TikTok duet, as well as the temporal separation of OV and RV, lend this discourse dynamic a novel flavor. More generally, Example 6 suggests the potential for stitches and duets to be embedded recursively in duets to invoke increasingly complex participation frameworks (Dynel, 2014) and to foster the use of mixed modalities to differentiate layers of embedded interaction. This proposition could be explored in future research by analyzing multimodal aspects of recursively embedded duets that have a conversational component.

5. Discussion

In this study, we asked how TikTok duetters create the illusion of conversing together in real time, given the constraints and affordances of the TikTok platform. We found that despite the technical impossibility of reciprocal interaction inside duets, OVs and RVs collaborate to various degrees in co-constructing duet conversations – OVs by inviting duets explicitly or implicitly and pre-structuring the interaction through scripting, pacing, or one-sided conversations, and RVs by orienting to OVs through the positioning and timing of their responses. Cases of an uncooperative RV resisting the OV's lead were rare in our sample, and even in those cases, RVs were typically only uncooperative with some elements of the OV.

RVs in our sample oriented to the 'no gap, no overlap' norm in spoken conversation, generally inserting their contributions as close as possible to transition relevance places (TRPs) in OV's speech. When overlap was unavoidable, for example because OV did not leave gaps, many RVs managed the overlaps strategically to create the appearance of (competitive) interruption or (cooperative) engagement (Schegloff, 2000; Tannen, 1984), thereby indexing social dynamics. Even celebrities' careless overlaps convey social meanings, for example, of status-based privilege.

We also asked what kinds of situational contexts are invoked through the duets. The duets in our sample illustrate both conventional/stereotypical (e.g., Example 1) and idiosyncratic or peculiar (e.g., Example 4) conversational situations, although other situations are more heterogeneous in nature. Situational contexts are established through the personae (e.g., gamer girl) and registers (e.g., conflict talk) that OV and RV perform, as well as the topics (e.g., embarrassing sex) they introduce. Interaction management strategies also index roles, relationships, and character attributes. For example, some OVs in our sample sought to attract the attention of a particular individual, usually a celebrity, since being duetted by a famous content creator can enhance one's visibility (cf. Herman, 2023). These OVs invoked the celebrity by name and presented themselves performing an activity for which the celebrity was known, casting themselves in the role of acolytes (or fans) and the celebrities in the role of experts (e.g., Example 5). The celebrities in our sample obliged by evaluating OV's activities, but they also interrupted or otherwise spoke over the OV carelessly, due to their privileged status – since they are famous, they can afford to devote less attention to turn-positioning. At the same time, their overlapping expressed the personal styles for which they are known: rude in the case of Gordon Ramsey, and over-the-top affectionate in the case of James Charles.

One-sided dialogue OVs dictate where and how a cooperative RV should respond; some also index relationships through interactional dynamics. For example, the highly engaged overlaps projected by the gamer girl OV in Example 3 suggest friendship, or at least friendliness. The one-sided phone conversation in the baby OV elicited many duetting responses, in most of which RVs constructed positive interactions featuring small talk on infant-related subjects such as breast milk, napping, crying, and pooping. In contrast, in Example 4, the RV recast the baby idiosyncratically as a money lender, creating a humorous effect. This role was reflected in RV's timing of overlap, such as when he made it appear that the baby (as money lender) was interrupting, evidence of its heartless character. In other duets with this video (not included in our sample), the baby was cast in other unlikely adult roles – including as a new lieutenant, a mob boss, and a tenant who owes back rent.

These observations underscore the performative nature of conversational duets. Most of the duets we analyzed function as displays of wit or humor – or in the case of the celebrities, as identity performances – more than as sincere attempts to converse with the OV. This performative tendency is especially pronounced in duets that deviate from the norms of spoken conversation – for example, because the two interlocutors are communicating in different modes or switching between referential frames. When TikTokers talk *about* each other in verbal asides to the audience, rather than *to* each other, it disrupts the illusion of natural conversation, while at the same time invoking other spoken registers, such as talk shows in which the host addresses the audience as well as the guest on stage (Dynel, 2014). Indeed, conversational duets, talk shows, radio shows, podcasts, and film dialogue all suggest that the conversational imperative extends to more than just the urge to talk; it also manifests as a desire to listen to other people talking.

6. Conclusions

This study contributes to our understanding of the intersection between digital design and user agency by exploring the ways in which users route around obstacles in TikTok's design to create conversation. Further, it contributes methodologically by responding to Giles et al.'s (2015) call for the development of CA approaches for the microanalysis of online conversation that take the novel affordances and contexts of digital channels into account.

We began by asking how people route around obstacles posed by interfaces and systems that appear prima facie unwelcoming to conversation, and what new kinds of conversation might emerge through this process. We argued that TikTok duets, although they afford a degree of user interaction, do not prima facie support conversation. This is because OVs are integral wholes, pre-recorded at a spatio-temporal remove from (usually unknown and unspecified) RVs, effectively ruling out the possibility of reciprocal turn exchanges. Yet despite this, in the various ways we have described, OVs and RVs can be said to collaboratively construct conversational exchanges, which are displayed as duets. This raises the question: Are the TikTok duets of the sort we analyzed a new type of conversation, or merely simulations performed to entertain TikTok audiences?

Our position is that conversational duets are both performances and computer-mediated conversations. They are performances, because like fictionalized dialogues, they are simulations of conversation as a concept, as well as invoking different kinds of conversational situations to achieve different communicative effects (Urbanová, 2007). They are also conversations of a sort, in that they exhibit conversational mechanics such as turn-taking and overlap (Sacks et al., 1974). They also satisfy Herring's (2010) definition of computer-mediated conversation, which includes asynchronous exchanges such as emails, as well as forum messages where a single message initiates multiple turns and an interlocutor responds later to each turn in a single message (cf. Condon & Cech, 2001).

In addition to satisfying the formal requirements of conversation, the duets fulfill conversational functions. These include information exchange, such as angryreactions learning about prosthetics and sharing that information with his followers in Example 2. Duets also create a shared social space where relational work can be done. In Example 5, the OV sends a message to her influencer inspiration, James Charles, with the help of her followers and is duetted back. Moreover, some OVs and RVs duet each other multiple times, such as in the case of the Gamer Girls in Example

3, who featured in several duets in our sample. Over time, these kinds of exchanges can build a sense of community among the video makers, as well as with their audiences.

According to Herman (2023), TikTok users are driven to create videos that provide visibility on the platform in order to grow an audience of followers. Tiktokers believe that the algorithm favors duets; thus, the users Herman interviewed create duets, even though they consider them "a second-rate form of creative content" that "mooches" off other users' content (p. 17). In contrast, the conversational duets we analyzed exhibit considerable creativity. Moreover, since the use of duets to mimic spoken conversation is not the originally intended use of the duet feature – the feature was created by Musical.ly for lip syncing and dancing alongside pre-existing videos – it is not clear if conversational duets are a usage that the algorithm favors. Rather, duetters seem to create them, at least in part, for the challenge. It is not a simple effect to achieve, and there is an intrinsic satisfaction in "conversing" in such an improbable venue (cf. Zelenkauskaite & Herring, 2008).

TikTok's algorithm does not permit representative sampling; thus, we do not know how popular the strategies we observed are in conversational duets, nor how popular conversational duets are compared to other duet types. Our judgment sampling technique privileged conversational duets, yet many other kinds of duets appear on TikTok, including those where an RV reacts to an OV nonverbally or monosyllabically. Many of these indeed seem more derivative than creative. Moreover, conversation analysis as a method cannot reveal TikTokers' underlying motives for producing conversational duets. Further insight could be gained by recording the process of duet creation from RV's perspective, as well as by interviewing or surveying conversational duet producers. Finally, although TikTok is an embodied medium, we did not systematically analyze how different creators use embodied resources in the duets, for example, to perform listenship. This remains a rich topic for future multimodal research.

Meanwhile, as regards our research questions, TikTok users can and do overcome the challenges posed by the platform's interface to produce conversational duets that seem natural from the point of view of turn-taking and other conversational mechanics. These findings suggest that users want to converse in, as well as through, short-format videos, in support of the computer-mediated conversational imperative. That is, the conversational imperative operates even in technological environments not well-suited for conversation, underscoring that conversation is a fundamental human drive (Jordania, 2009). At the same time, our analysis has implications for our understanding of conversation itself, a key concept in discourse studies. The duet conversations are quite different from conversations that take place over video conferencing systems. Their performative, temporally asymmetrical, and multimodal nature suggests a new form of video-mediated conversation, evidence that technological change continues to extend the definition of conversation itself (cf. Herring, 2010).

REFERENCES

Abidin, C. (2020). Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours. *Cultural Science Journal, 12*(1), 77-103.

Adami, E. (2009). Do You Tube? When communication turns into video e-nteraction. In D. Torretta, M. Dossena, & A. Sportelli (Eds.), *Migration of forms, forms of migration. Proceedings of the 23rd AIA Conference*, pp. 371-386.

Baron, N. S. (1990). Pigeon-birds and rhyming words: The role of parents in language learning. *Language in education: Theory and practice* (pp. 35-36). Prentice Hall Regents.

Condon, S. L., & Cech, C. G. (2001). Profiling turns in interaction: Discourse structure and function. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*. IEEE.

Duffy, W. (2014). Collaboration (in) theory: Reworking the social turn's conversational imperative. *College English, 76(*5), 416-435.

Dynel, M. (2014). Participation framework underlying YouTube interaction. *Journal of Pragmatics, 73*, 37-52.

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.

Goffman, E. (1981). *Forms of talk*. University of Pennsylvania Press.

Goldberg, J. A. (1990). Interrupting the discourse on interruptions: An analysis in terms of relationally neutral, power-and rapport-oriented acts. *Journal of Pragmatics, 14(*6), 883-903.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 41–58). Academic Press.

Herman, L. (2023). For who page? TikTok creators' algorithmic dependencies. *International Association of Societies of Design Research Congress*, Milan, 9-13 October. https://dl.designresearchsociety.org/iasdr/iasdr2023/fullpapers/224/

Herring, S. C. (2010). Computer-mediated conversation: Introduction and overview. *Language@Internet, 7*, article 2.

Herring, S. C. (2013). Discourse in Web 2.0: Familiar, reconfigured, and emergent. In D. Tannen & A. M. Tester (Eds.), *Georgetown University Round Table on Languages and Linguistics 2011: Discourse 2.0: Language and new media* (pp. 1-25). Georgetown University Press.

Herring, S. C., & Chae, S. (2021). Prompt-rich CMC on YouTube: To what or to whom do comments respond? In *Proceedings of the Fifty-fourth Hawai'i International Conference on System Sciences*, pp. 2906-2915.

Herring, S. C., Kutz, D. O., Paolillo, J. C., & Zelenkauskaite, A. (2009). Fast talking, fast shooting: Text chat in an online first-person game. In *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences*, pp. 1-10. IEEE Press.

Holler, S. C., Kendrick, J., Casillas, K. H., & Levinson, M. (2016). *Turn-taking in human communicative interaction*. Frontiers Media SA. https: //doi.org/10.3389/978-2- 88919-825-2

Hollingsworth, K. (2022). *Transition-relevance places in video-mediated conversations*. Honors thesis, Baylor University. https://baylor-ir.tdl.org/handle/2104/11837

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). John Benjamins.

Jordania, J. (2009). Times to fight and times to relax: Singing and humming at the beginnings of human evolutionary history. *Kadmosmi, 1*, 272–277.

Kaye, B. (2022). Please duet this: Collaborative music making in lockdown on TikTok. *Networking Knowledge: Journal of the MeCCSA Postgraduate Network, 15*(1), 59-77.

Kåhlman, R. (2023). *Informal language learning through social media platform TikTok.* Bachelor's thesis, Faculty of Humanities, University of Oulu, Finland.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26*, 22-63.

King, S. W., & Sereno, K. K. (1984). Conversational appropriateness as a conversational imperative. *Quarterly Journal of Speech, 70*(3), 264-273.

Licoppe, C., & Morel, J. (2012). Video-in-interaction:"Talking heads" and the multimodal organization of mobile and Skype video calls. *Research on Language & Social Interaction, 45*(4), 399-429.

Lukasik, S. (2010). Why the Arpanet was built. *IEEE Annals of the History of Computing, 33*(3), 4-21.

Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. Ogden & I. A. Richards (Eds.), *The meaning of meaning* (pp. 451-510). Harcourt, Brace & Company.

Petkova, D. P. (2015). Beyond silence: A crosscultural comparison between Finnish "quietude" and Japanese "tranquility." *Eastern Academic Journal, 4*, 1-14.

Pihlaja, S. (2011). Cops, popes, and garbage collectors: Metaphor and antagonism in an atheist/Christian YouTube video thread. *Language@Internet, 8,* article 1.

Rosenbaun, L., Rafaeli, S., & Kurzon, D. (2016). Participation frameworks in multiparty video chats: Cross-modal exchanges in public Google Hangouts. *Journal of Pragmatics, 94*, 29-46.

Rosqvist, H. B. (2019). Doing things together: Exploring meanings of different forms of sociality among autistic people in an autistic work space. *Alter, 13*(3), 168-178.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest semantics for the organization of turn-taking in conversation. *Language, 50*, 696-735.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 71-93). Georgetown University Press.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society, 29*, 1–63.

Schneider, S. J., Kerwin, J., Frechtling, J., & Vivari, B. A. (2002). Characteristics of the discussion in online and face-to-face focus groups. *Social Science Computer Review, 20*(1), 31-42.

Severinson Eklundh, K. (2010). To quote or not to quote: Setting the context for computer-mediated dialogues. *Language@Internet, 7*, article 5.

Sindoni, M. G. (2021). Mode-switching in video-mediated interaction: Integrating linguistic phenomena into multimodal transcription tasks. *Linguistics and Education, 62,* 100738.

Sarfo-Adu, K., & Osei, C. E. (2021). Politeness in media talk shows: The case of media panel discussions in Ghana. *Journal of Media and Communication Studies, 13*(3), 81-90.

Tannen, D. (1984). *Conversational style: Analyzing talk among friends*. Ablex.

Urbanová, L. (2007). Phatic communion and small talk in fictional dialogues. In C. S. Butler, R. H. Downing, & J. Lavid (Eds.), *Functional perspectives on grammar and discourse* (pp. 349-358). John Benjamins.

Werry, C. (1996). Linguistic and interactional features of Internet Relay Chat. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 47-64). John Benjamins.

White, S. J. (2019). Conversation analysis: An introduction to methodology, data collection, and analysis. In P. Liamputtong (Ed.), *Handbook of research methods in health social sciences* (pp. 471–490). Springer Singapore.

Zelenkauskaite, A., & Herring, S. C. (2008). Television-mediated conversation: Coherence in Italian iTV SMS chat. In *Proceedings of the Forty-First Hawai'i International Conference on System Sciences*. IEEE Press.

Appendix I: Transcription notations used in examples

Notation	Meaning
$	Transition Relevance Place (TRP)
ye[ah] 　[okay]	Overlapping speech
=	Latching or no pause between utterances
…	Pause
word-	False start; word cut off
ALL CAPS	Louder than surrounding speech
(( ))	Non-speech expression, e.g., sighs, nods, laughs
{ }	Special voice quality, e.g., high-pitched
( )	Transcriber comments
/?/	Transcriber uncertainty. Number of ? indicates number of syllables