



Conversation analysis and conversational technologies: finding the common ground between academia and industry

LSE Research Online URL for this paper: <http://eprints.lse.ac.uk/123945/>

Version: Accepted Version

Article:

Stokoe, Elizabeth ORCID: 0000-0002-7353-4121, Albert, Saul, Buschmeier, Hendrik and Stommel, Wyke (2024) Conversation analysis and conversational technologies: finding the common ground between academia and industry. *Discourse and Communication*, 18 (6). ISSN 1750-4821 (In Press)

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Conversation Analysis and Conversational Technologies: Finding the Common Ground between Academia and Industry

Elizabeth Stokoe

Department of
Psychological and
Behavioural Science,
The London School of
Economics and
Political Science,
Houghton Street,
London WC2A 2AE,
UK.

E.Stokoe@lse.ac.uk

Saul Albert

Discourse And
Rhetoric Group
(DARG),
Communication &
Media, School of
Social Sciences and
Humanities,
Loughborough
University, Epinal Way,
Loughborough, LE11
3TU, UK.

S.B.Albert@lboro.ac.uk

Hendrik Buschmeier

Digital Linguistics
Lab, Faculty of
Linguistics and
Literary Studies,
Bielefeld University,
Bielefeld, Germany.

[hbuschme@uni-
bielefeld.de](mailto:hbuschme@uni-bielefeld.de)

Wyke Stommel

Centre for Language
Studies &
Interdisciplinary Hub
for Digitalization and
Society, Radboud
University, Houtlaan 4,
6525 XZ Nijmegen,
The Netherlands.

wyke.stommel@ru.nl

Accepted version May 2024: Stokoe, E., Albert, S., Buschmeier, H., & Stommel, W. (2024) Conversation analysis and conversational technologies: Finding the common ground between academia and industry. *Discourse and Communication*, 18(6). DOI tbc

1.0 Introduction

The purpose of this Special Issue of *Discourse and Communication* is to bring academia and industry together in a conversation. More specifically, it aims to showcase cutting-edge ethnomethodological and conversation analytic research on conversational technologies to enable an interaction between academic and industry expert voices whose work involves analyzing and developing such technologies. Given the fast-moving landscape around, for example, large language models (LLMs), natural language processing (NLP), issues of trust, ethics, and bias – and given the proliferation of problems in science itself (e.g., the replication crisis, paper mills, AI-augmented or AI-generated research outputs) – it is especially important to foster close collaboration between academia and industry and ensure that developments are underpinned by evidence. We recognise that the pace of development is rapid and so another aim of this Special Issue is to mark a moment in time; one in which academic and industry experts reflect on the current state of their fields.

The Special Issue contains eleven academic papers with corresponding commentaries written by industry professionals who work as, for example, conversation designers, startup founders, or researchers, across organizations including Google, IBM, and Microsoft. We refer to the research throughout this opening essay, in which we set out some implications of taking conversation analysis seriously for the development not just of conversational technologies but of any product, practice, or process that (claims to) leverage ‘conversation.’ Such products may include communication guidance, tools to assess communication skills,

sales scripts, focus group schedules, and so on. While such products seem a world away from conversational technologies, we argue that stereotypical notions of what ‘conversation’ is may be mobilized and reproduced in both (e.g., Atkins, 2019; Stokoe et al., 2020). This is partly because, in conversation-leveraging products, ‘conversation’ often manifests as ‘conversational’ (Stokoe, 2021); that is, informal, casual, relaxed, friendly, colloquial, or idiomatic. When it comes to technology, ‘conversational’ may be realized through ‘personas’ or ‘tone of voice’ and using “phrases and sentences that would be comfortably spoken” (Hunter, 2018). However, from the perspective of conversation analysis, ‘conversational’ refers to the systematics and mechanics of social interaction. There is no interaction between or with humans that is not ‘conversational’ in the sense that, even in standardized forms of talk (e.g., Houtkoop-Steenstra, 1995), form-filling (e.g., Heritage, 2002), or legally-proscribed encounters (e.g., Richardson et al., 2019), the underpinning machinery drives the conversational imperative.

We set the scene for this Special Issue by starting with an illustration of the core concerns of conversation analysis via an example of human-human interaction: the analogue telephone call. We compare this to the newest kind of human-computer interaction in which a large language model (LLM) appears as a ‘participant.’ We refer to the Special Issue papers through this section. We then draw out the implications of our empirical demonstration in the context of omnirelevant questions about whether conversational technologies ‘understand’ their human interlocutors.

2.0 “Guess what?”

We start by comparing the opening sequences of two interactions. Extract 1 is a classic of conversation analysis: an analogue US-situated telephone call between Donny and Marsha (see Schegloff, 1995). Extract 1 starts with Marsha answering a ringing telephone summons. Extract 2 was generated by ‘prompt engineering’ an LLM-based chatbot to ‘role-play’ as Marsha while we, for the sake of demonstration and comparison, typed in Donny’s original turns using standard orthography. Lines 01 and 05 are the only ‘Marsha’ turns that turned out to be identical in both the original recording and LLM-generated version.

Extract 1: Stalled

01 Marsha: Hello?
02 Donny: H’lo Marsha,
03 Mar: ↑Ye:a[h.]
04 Donny: [It’s D]onny.
05 Marsha: ↑Hi: Do:nny.
06 Donny: Guess what. hh
07 Marsha: What.
08 Donny: .hh my ca:r is sta:lled.h

Extract 2: Stalled/ChatGPT

01 LLM: Hello.
02 This is Marsha speaking.
03 How can I assist you.
04 Donny: Hello Marsha, it’s Donny.

05 LLM: Hi Donny.
 06 How can I assist you. Is there something
 07 specific you need help with?
 08 Donny: Guess what. hh
 09 LLM: Oh. I'm not very good at guessing games.
 10 But I'll give it a try. Hmm.
 11 Did you win a contest or receive some good news.
 12 Donny: My car is stalled.

It may be instructive to consider these extracts in terms of the “generic orders of organization in talk-in-interaction” (Schegloff, 2007: xiii) on which many of the foundational advances in conversation analytic research over the past six decades have focused. Schegloff (2007: xiv) describes these as “the various organizations of practice that deal with the various generic organizational contingencies of talk-in-interaction without which it cannot proceed in an orderly way”, including turn-taking, action formation, sequence organization, trouble (including ‘repair’), word selection, and overall structural organization. Whitehead et al. (frth.) further include “the epistemics problem”: “how are relative rights to knowledge displayed and managed in sequentially organized activities?” We make some brief observations about Extracts 1 and 2 for each of these “generic orders.”

Our first observation focuses on *turn-taking*, that is, “who should talk next and when should they do so?” (Schegloff, 2007: xiv). One striking difference between the two extracts is that, in the original call (Extract 1), participants use the affordances of spoken (though not co-present) talk to initiate and complete turns (Sacks et al., 1974). The unfolding sequence comprises several ‘turn constructional units’ (TCU, e.g., “Hello”, “guess what”). Each TCU ends with possible ‘transition relevance places’ (TRP) indicated through grammatical and prosodic completeness of the TCU and thereby the action (e.g., greeting, pre-announcement). Furthermore, in Extract 1, and as Schegloff (1995: 193) observes, Donny’s opening turns across lines 02 and 04 are “rushed, charged, almost breathless.” Rather than wait for Marsha to recognize who is speaking, Donny pre-emptively identifies himself (line 04) and, at line 06, pre-empts the initiation of a ‘how-are-you’ sequence which may typically occur in this position. These turns, with precision placement in overlap (lines 03–04), help to constitute an ‘urgent’ situation. By contrast, in Extract 2, rather than generate one TCU/action per ‘turn’ (i.e., per post or message), LLM generates ‘multi-TCU’ ‘turns’ as paragraphs without TRPs and, of course, without the overlaps, speech perturbations, gaps, pauses, or phonetic resources that are fundamental to turn-taking and the other ‘generic orders’ of interaction. The LLM generates text, not talk.

Next, we come to *action formation*; that is, how “the resources of the language, the body, the environment of the interaction, and position in the interaction fashioned into conformations designed to be, and to be recognizable by recipients as, particular actions” (Schegloff, 2007: xiv). In the original call, Marsha answers the phone with a greeting token (only), while LLM immediately adds two further ‘TCUs’ in addition to a greeting. The first announces who is speaking, although this is not an unusual way to answer the telephone. However, at line 03, the third TCU (“How can I assist you”) is more fitted to institutional interaction (e.g., service provision) than the personal/domestic one it turns out to be in Extract 1. Furthermore, in Extract 1, in response to Donny’s self-identification, Marsha’s

second greeting also confirms recognition of Donny (“↑Hi: Do:nny.”, line 05). In Extract 2, LLM produces the same confirmation (line 05). However, without leveraging either the intonational affordances of spoken talk, or the resources of written interaction such as punctuation or emojis (e.g., Meredith, 2014), LLM does not use one action (i.e., ‘greeting’) as the vehicle for another (i.e., ‘recognition’). LLM extends this turn with two further TCUs: wh- and interrogative-formatted offers of assistance (lines 06-07). Donny has no opportunity to respond to the first (“How can ...”) since, without TRPs, LLM fails to enable Donny to convey urgency through sequence, pace, and turn design, and thus also fails to design its turns for a specific recipient.

The most striking difference in terms of action, however, is in the responses to Donny’s pre-announcement (“Guess what”). In Extract 1, Marsha provides a preferred response, a go-ahead “what” (line 07). In so doing, she displays an understanding that Donny did a pre-announcement and aligns to the unfolding sequence by not, for instance, attempting to initiate other actions. In Extract 2, while LLM ‘understands’ ‘Guess what?’ as *language*, it fails to recognize the *action* of pre-announcing and instead treats it as the start of a “guessing game” (line 09). While LLM’s guess – “did you receive some good news” – may be apposite in some other encounter, as a fitted response in Extract 2 it is neither projectable from Donny’s pre-announcement nor *epistemically* warranted: this is Donny’s call, and Donny’s news. Rather than managing, turn-by-turn, how “relative rights to knowledge [are] displayed and managed in sequentially organized activities” (Whitehead et al., frth), LLM’s turns have an unwaveringly steep “epistemic gradient” (Heritage, 2012), whereby each is designed ‘for anyone’ rather than ‘for the person I am currently interacting with.’

In terms of *sequence organization*, we can see in Extract 1 how “successive turns [are] formed up to be ‘coherent’ with the prior turn” as well as “the nature of that coherence” (Schegloff, 2007: xiv). Key to Extract 1 is the mutual orientation to the ‘contiguity’ of turns (Sacks, 1987) and the ‘progressivity’ of sequences (Stivers and Robinson, 2006). Donny’s pre-emptive placement of a ‘go-ahead’ where one might expect a ‘how-are-you’ informs “the construction and understanding of the talk” (Schegloff, 2007, p. xiv) and, although we do not present the entire interaction here (see Schegloff, 1995), it takes barely 30 seconds to complete. In contrast, the complete LLM version (not presented here) takes three minutes to verbalize, since the *overall structural organization* is massively expanded across lengthy multi-TCU turns and the LLM gets stuck in several service/offer phases.

Furthermore, and as we noted earlier regarding turn-taking, LLM’s paragraph-type turns are generated without TRPs. Of course, the interaction is written, and LLM training data is not typically derived from multimodal datasets of embodied “in the wild” social interaction or transcripts thereof (cf. Cooney and Reece, 2024). Yet, given that messaging apps already afford single- or multi-TCU turns, LLM’s turn organization is not simply attributable to its written modality (cf. Schlangen and Skantze, 2011 on incremental dialogue systems). As researchers have shown, online ‘chat’ interaction often displays similar orientations to contiguity and progressivity as found in spoken interaction (e.g., König, 2019; Meredith, 2014; Stommel and te Molder, 2015). In Extract 2, ‘Donny’ and ‘Marsha’ have reduced opportunities to take turns or build adjacency pairs, and thus to leverage the ‘next-turn-proof procedure.’ Consequently, this limits the scope for ‘intersubjectivity’, both in terms of shared understanding in an ‘informational’ or ‘semantic’ sense and of what kinds of

initial and responsive actions can be slotted together sequentially. LLM thus fails to identify, co-construct, and adapt to the urgency of the situation. This lack of affordance places the burden of maintaining LLM's turns as intelligible unilaterally on human participants (e.g., Rudaz and Licoppe, this issue; Tisserand and Baldauf-Quilliatre, this issue).

In addition to failing to achieve the pace of the original call, or use overlap to build pre-emptive actions, LLM does not leverage other interactional resources such as silence, pauses, and speech perturbations (except for a possible word-search type “hmm” at line 10) that Donny and Marsha use in Extract 1. Although there is no *trouble* in terms of repair initiation or resolution, ‘intersubjectivity’ is neither established nor maintained, since in Extract 2 LLM fails to treat Donny's pre-announcement as a pre to a potentially upcoming base sequence. And another kind of trouble is ‘*category trouble*’ (Whitehead et al., frth). By beginning with “This is Marsha speaking” (line 02), LLM produces an action of the right type in the right place. But rather than next creating a TRP for Donny to take a turn, LLM adopts the category incumbency of ‘service provider’ and subsequently fails the task of *recipient design* (see Housley and Dahl, this issue). Fundamental to the issue of *recipient design* is *word selection*; that is, how components of a turn get selected, and how selections “inform and shape the understanding achieved by the turn's recipients” (Schegloff, 2007: xiv). In Extract 2, LLM's selection of ‘assist’ contributes to its ‘service provider’ incumbency.

While it is clear from Extract 1 that both speakers leverage vocal cues to enable identification of their interlocutor – not just in some abstract sense but also in terms of category incumbency in ways that support recipient design (e.g., Flinkfeldt et al 2022) – it is equally clear that LLM cannot and does not. In prompt engineering Extract 2, we did not ‘help’ LLM by formulating what we take to be Donny and Marsha's relationship, since in actual interaction, such category incumbencies are achieved, maintained, and updated turn-by-turn. These achievements, which are crucial for social interaction, are non-existent in Extract 2 (also see Pütz and Esposito, this issue; Relieu, this issue). How category incumbencies are oriented to and made relevant for interaction are emerging questions for conversational technologies (De Rijk et al, this issue), alongside how encoding them may generate the same biases as those emerging across AI (e.g., Bender et al., 2021; Schlesinger et al. 2017; Yin et al., 2024).

The case of Donny and Marsha illustrates features of the LLM relevant to the broader issue of leveraging ‘conversation’ introduced earlier. For example, LLM is poor at simulating authentic interaction (see also Brandt et al., this issue). Indeed, it took several prompts for LLM to role play as a single participant rather than instantly generating both sides of the conversation, suggesting the fundamentally non-dialogical model of language it is using. Yet, there is nothing special about LLMs being unable to generate authentic interaction through role-play, not only because people do things differently when role-playing during communication training compared to when doing their jobs (also see Atkinson, 2019; Stokoe et al., 2020) but also because the kinds of practices that are trained for or assessed often bear no relation to what the ‘equivalent’ practices look like in their situated and endogenous environment. For example, one practice thought to ‘build rapport’ is to use one's interlocutor's name. Police officers are encouraged to use suspects' names in training environments; salespeople are encouraged to use the names of prospective clients, and conversational technologies elicit users' names to be ‘piped’ into the user journey later (see

Stokoe, 2018). Yet we know that the misuse, misplacement, or overuse of people’s names can create jarring moments in interaction that create disaffiliation rather than ‘rapport’ (ibid). Therefore, it is not just interactions in LLM-based ‘conversational’ products that may differ wildly to authentic human conversation.

In terms of industry developments, we argue that neither our imaginations nor role-play need be primary sources of material for products that leverage conversation. This is not just because real and role-played interactions differ in important ways but, more importantly, because (in)effective practices can be identified and described by analyzing actual interactions with enough precision to offer actionable insights. There are six decades of conversation analytic research to draw on. By scrutinizing human and LLM-generated extracts side-by-side using Schegloff’s ‘generic orders’, we have demonstrated how conversation analysis (CA) can identify, describe, and share specific component features of interactions in ways that can enable developers to build and modify conversational technologies with reference to precise instances, practices, and mechanisms of interaction. Where industry lacks the terminology to describe CA’s rich array of documented interactional practices, it will not be able to design them into purportedly ‘conversational’ products. The aim of this Special Issue has been to develop common ground between conversation analysis and emerging conversational technologies. In the following section, we reflect on what achieving common ground between CA and conversational technology would entail, starting with the psycholinguistic concept of ‘common ground’ itself.

3.0 Finding the common ground

By focusing on the “technology of conversation” (Sacks, 1984: 413), CA proposes a radically different starting point to other theories of communication that underpin ‘conversational’ technology. Yet CA’s conceptualization of ‘shared understanding’ is also surprisingly compatible with – and may be useful for – the development of generative AI.

Technologists’ conceptualizations of ‘successful’ communication have often centred on the process of establishing and maintaining shared understanding through a process of ‘grounding’ (Clark, 1996). From this perspective, communication is possible when interlocutors (including machines) have either an explicit or tacitly shared basis for interaction. ‘Common ground’, then, consists of propositions that follow from this shared basis. According to Clark (1996), in conversation, novel propositions, if understood by the interlocutor, are ‘grounded’ through spoken (and, in face-to-face conversation, multimodal) acts that provide evidence of ‘understanding’, thus enabling ongoing interaction. Various formal and computational models of grounding and related phenomena have been proposed for dialogue systems development (for an overview, see Buschmeier, 2018, sect. 2.2), and although these are rarely applied in system design (Kopp and Krämer, 2021), they are often cited as fundamental for conversational technology (Stokoe et al., 2021).

However, such theories miss the fundamental point that, in interactions between humans and with machines, it is “trust” rather than “truth” that matters (Garfinkel, 1963). As Garfinkel’s early work on chatbots like ELIZA argued: “what allowed communication with a machine to sometimes have the feel of human interaction was that the machine was exploiting human social competencies to get its work done” (Eisenmann et al., 2023: 2; also

see Edwards, 1997). From this and from users' perspectives, it does not much matter what is 'going on under the hood' of conversational technologies. LLMs, for example, can often surprise users by achieving a 'good enough' show of 'understanding' by displaying an "adequate interpretation of social action within interaction" (Moore and Arar, 2019: 23). As long as the interaction progresses, human participants can bring off 'understanding' as having been ongoingly achieved. As Garfinkel (2021: 25) puts it, on the question of representations and mental states, "there's nothing in your head but brains."


In many of the situations explored in this Special Issue, the technology of conversation operates to sustain the progressivity of the interaction. Where a disabled person and their homecare worker encounter trouble interacting with a smart home system (Hall and Albert, this issue), adaptive, collaborative work is occasioned to move things along. Similarly, participants continually adapt and redesign their requests when interacting with a smart speaker, taking system responses into account while pursuing practical-interactional goals (Reineke and Helmer, this issue). That the work of 'doing understanding' is built into interaction is illustrated by Klowait et al. (this issue) who found that an 'explainer AI', developed to help users understand AI technology, is 'successful' simply when the user engages in interaction with the explainer as an agent, irrespective of the system's intended explanations. These cases suggest that the *progressivity* of interaction, rather than the security of intersubjective understanding or informational 'common ground' – action rather than 'intent' – is sufficient to establish the trust conditions for interaction with conversational technologies.

There is already a substantial track record in CA research, conducted with organizational and industrial partners across many sectors, in which its empirical power is used to identify communicative challenges and opportunities for intervention (Antaki, 2011; Stokoe, 2018) – including in conversational technologies (e.g., Albert et al., 2023; Brandt et al., this issue). Furthermore, the evaluation of emerging conversational technologies could use CA to, say, target the action-level of interaction rather than using simplistic response methods (Liesenfeld and Dingemanse, this issue). While developers and engineers of conversational technologies do most of their work 'under the hood', there are myriad insights from CA to draw upon.

4.0 Conclusion

Just as Garfinkel might have responded to contemporary hyperbole about 'machine sentience' by saying that "there's nothing in your LLM but numbers", we should also take Sacks's corollary position: that we should be cautious about anthropomorphizing humans in our analysis of social interaction (Sacks, 1967, cited in Jefferson, 1989). The common ground of conversation analysis and conversational technologies is the focus on the object of *conversation*. Many papers in this Special Issue show how the generic orders of conversational structure can be analyzed as technology that we co-produce in and through social interaction. By focusing on the *technology of conversation* rather than the '*conversationality of technology*', we dispense with cognitivist presuppositions and the restrictive assumptions they impose on our conceptualizations of humans, machines, and conversation itself.

5.0 References

- Albert S, Pearl C and Stokoe E (2023) Should Conversational User Interfaces have a personality? Conversation Design Institute, July. Available online: <https://www.youtube.com/watch?v=OegIMvaYT6U>
- Antaki C (ed.) (2011) *Applied Conversation Analysis. Intervention and Change in Institutional Talk*. London, UK: Palgrave.
- Atkins S (2019) Assessing health professionals' communication through role-play: An interactional analysis of simulated versus actual general practice consultations. *Discourse Studies*, 21, 109–134.
- Bender EM, Gebru T, McMillan-Major A, and Shmitchell S (2021) On the dangers of stochastic parrots: Can language models be too big? . In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623. DOI: [10.1145/3442188.3445922](https://doi.org/10.1145/3442188.3445922)
- Brandt A, Hazel S, McKinnon R, Sideridou K, Tindale J, and Ventoura N (2024) Educating Dora: Teaching a conversational agent to talk. *Discourse and Communication* 18(6): DOI tbc
- Buschmeier H (2018) *Attentive Speaking. From Listener Feedback to Interactive Adaptation*. PhD-thesis. Bielefeld University, Bielefeld, Germany. DOI: [10.4119/unibi/2918295](https://doi.org/10.4119/unibi/2918295)
- Clark HH (1996) *Using Language*. Cambridge, UK: Cambridge University Press.
- Cooney G and Reece A (2024). NaturalTurn: A Method to Segment Transcripts into Naturalistic Conversational Turns. *arXiv preprint arXiv:2403.15615*. DOI: [10.48550/arXiv.2403.15615](https://doi.org/10.48550/arXiv.2403.15615)
- De Rijk L, Breukelman M, Dalmaijer E and Stommel W (2024) ‘This uh... young lady young man’: Gender attribution in the context of a gender ambiguous robot. *Discourse and Communication* 18(6): DOI tbc
- Eisenmann C, Mlynář J, Turowetz J, et al. (2023) “Machine Down”: Making sense of human–computer interaction – Garfinkel’s research on ELIZA and LYRIC from 1967 to 1969 and its contemporary relevance. *AI & SOCIETY*. DOI: [10.1007/s00146-023-01793-z](https://doi.org/10.1007/s00146-023-01793-z).
- Edwards D (1997) *Discourse and Cognition*. London: Sage.
- Flinkfeldt M, Parslow S, and Stokoe E (2022) How categorization impacts the design of requests: Asking for email addresses in call-centre interactions. *Language in Society* 51: 693–716.
- Garfinkel H (1963) A conception of and experiments with “trust” as a condition of concerted stable actions. In: O’Brien J (ed) *The Production of Reality: Essays and Readings on Social Interaction*. New York, NY, USA: Ronald Press, pp. 379–390.
- Garfinkel H (2021) Ethnomethodological misreading of Aron Gurwitsch on the phenomenal field. *Human Studies* 44: 19–42.
- Hall L and Albert S (2024) Distributed agency in smart homecare interactions: A conversation analytic case study. *Discourse and Communication* 18(6): DOI tbc
- Heritage J (2002) Ad hoc inquiries: Two preferences in the design of routine questions in an open context, In: Maynard DW and Houtkoop-Steenstra H, Schaeffer NC and van der Zouwen, J (eds) *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*. New York, NY, USA: Wiley, pp. 313–334.
- Heritage J (2012) Epistemics in action: Action formation and territories of knowledge. *Research on Language & Social interaction* 45(1): 1-29.
- Houtkoop-Steenstra H (1995) Meeting both ends: Standardization and recipient design in telephone survey interviews, In: ten Have P and Psathas G (eds) *Situated Order: Studies*

- in Social Organization and Embodied Activities*. Washington, DC, USA: University Press of America, pp. 91–106.
- Housley W and Dahl P (2024) Membership categorisation, sociological description and role prompt engineering with Chat GPT. *Discourse and Communication* 18(6): DOI tbc
- Hunter P (2018) 10 things to fix in your conversational app UX right now. *Medium*. <https://medium.com/@designoutloud/11cee93ab959>
- Jefferson G (1989) Letter to the editor Re: Anita Pomerantz' epilogue to the special issue on sequential organization of conversational activities, spring 1989. *Western Journal of Speech Communication* 53(4): 427–429.
- Klowait N, Erofeeva M, Lenke M, Horwath I, and Buschmeier H (2024). Can AI explain AI? Interactive co-construction of explanations among human and artificial agents. *Discourse and Communication* 18(6): DOI tbc
- König K (2019) Sequential patterns in SMS and WhatsApp dialogues: Practices for coordinating actions and managing topics. *Discourse & Communication* 13(6): 612-629.
- Kopp S and Krämer N (2021) Revisiting human-agent communication: The importance of joint co-construction and understanding mental states. *Frontiers in Psychology* 12. DOI:[10.3389/fpsyg.2021.580955](https://doi.org/10.3389/fpsyg.2021.580955)
- Liesenfeld A and Dingemanse M (2024) Interactive probes: Towards action-level evaluation for dialogue systems. *Discourse and Communication* 18(6): DOI tbc
- Meredith J (2014). *Chatting Online: Comparing Spoken and Online Written Interaction between Friends*. Unpublished PhD Dissertation, Loughborough University, UK.
- Moore RJ and Arar R (2019) *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. New York, NY, USA: Association for Computing Machinery. DOI:[10.1145/3304087](https://doi.org/10.1145/3304087)
- Pütz O and Esposito E (2024) Performance without understanding: How ChatGPT relies on humans to repair conversational trouble. *Discourse and Communication* 18(6): DOI tbc
- Reineke S and Helmer S (2024) User practices in dealing with trouble in interactions with virtual assistants in German: Repeating, altering, and insisting. *Discourse and Communication* 18(6): DOI tbc
- Relieu M (2024) How Lenny the bot convinces you that he is a person: Storytelling, affiliations and alignments in multi-unit turns. *Discourse and Communication* 18(6): DOI tbc
- Richardson E, Stokoe E, and Antaki C (2019) Establishing intellectually impaired victims' understanding about 'truth' and 'lies': Police interview guidance and practice in cases of sexual assault. *Applied Linguistics*, 40: 773–792.
- Rudaz D and Licoppe C (2024) "Playing the robot's advocate": Bystanders' descriptions of a robot's conduct in public settings. *Discourse and Communication* 18(6), DOI tbc
- Sacks H (1987) On the preferences for agreement and contiguity in sequences in conversation. In Button G and Lee JRE (eds) *Talk and Social Organization* London, UK: Routledge, pp. 54–69.
- Sacks H (1984) On doing "being ordinary". In: Atkinson JM (ed.) *Structures of Social Action*. Studies in Emotion and Social Interaction. Cambridge, UK: Cambridge University Press, pp. 413–429.
- Sacks H, Schegloff EA and Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696–735.
- Schegloff EA (1995) Discourse as an Interactional Achievement III: The Omnirelevance of Action. *Research on Language and Social Interaction* 28: 185–211.
- Schegloff EA (2007) *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge, UK: Cambridge University Press.

- Schlangen D and Skantze G (2011) A general, abstract model of incremental dialogue processing. *Dialogue and Discourse* 2: 83–111.
- Schlesinger A, Edwards WK and Grinter RE (2017) Intersectional HCI: Engaging identity through gender, race, and class. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, CO, USA, pp. 5412–5427.
- Stivers T and Robinson JD (2006) A preference for progressivity in interaction. *Language in Society* 35: 367–392.
- Stokoe E (2018) *Talk: The Science of Conversation*. Little, Brown.
- Stokoe E (2021). What does it mean to be conversational? *Deployed*.
<https://www.deployed.co/blog/what-does-it-mean-to-be-conversational>
- Stokoe E, Albert S, Parslow S and Pearl C (2021) Conversation analysis and conversation design: Where the moonshots are. *Medium* <https://elizabeth-stokoe.medium.com/c2a2836cb042>
- Stokoe E, Sikveland RO, Albert S, Hamann M and Housley W (2020) Can humans simulate talking like other humans? Comparing simulated clients to real customers in service inquiries. *Discourse Studies* 22: 87–109.
- Stommel W and te Molder H (2015) Counseling online and over the phone: when preclosing questions fail as a closing device. *Research on Language and Social Interaction* 48, 281–300.
- Tisserand L and Baldauf-Quilliatre H (2024) Rejecting a robot’s offer: An analysis of preference. *Discourse and Communication* 18(6): DOI tbc
- Whitehead K, Stokoe E and Raymond G (forthcoming) *Categories in Social Interaction*. Routledge.
- Yin L, Alba D and Nicoletti L (2024) OpenAI GPT sorts resume names with racial bias, test shows. *Bloomberg*. <https://www.bloomberg.com/graphics/2024-openai-gpt-hiring-racial-discrimination/>

Biographies

Elizabeth Stokoe is a professor in the Department of Psychological and Behavioural Science at The London School of Economics and Political Science, UK. Her research in conversation analysis and membership categorization currently focuses on the ways that ‘conversation’ is leveraged in both research and industry processes and products. **Saul Albert** is a senior lecturer in Social Science (Social Psychology) in the Department of Communication and Media at Loughborough University, UK, and a member of the Discourse And Rhetoric Group. His research at the intersection of cognitive science and conversation analysis explores how efforts to manage miscommunication underpin human social interaction. **Hendrik Buschmeier** is a junior professor for Digital Linguistics at the Faculty of Linguistics and Literary Studies at Bielefeld University, Germany. He is interested in empirical and computational modeling of dialogue phenomena, speech and multimodality, and conversational interaction between humans, and between humans and artificial agents. **Wyke Stommel** is a Full Professor of Language and Interaction in Digital Contexts at the Centre for Language Studies, Radboud University, The Netherlands. Her research focuses on the analysis of spoken and digital interaction, including chat, human-robot and video-mediated interaction across institutional and mundane settings, using and advancing conversation analytic methods.