Conversation considered harmful?

Stuart Reeves
stuart@tropic.org.uk
School of Computer Science, University of Nottingham

ABSTRACT

As a concept, 'conversation' is rife with troublemaking potential. It is not that we should necessarily abandon use of 'conversation' in conversational user interface (CUI) research, but rather treat it with a significant measure of care due to the varied conceptual problems it introduces—problems sketched in this paper. I suggest an alternative, possibly safer articulation and conceptual shift: *conversation-sensitive design.*

CCS CONCEPTS

• **Human-centered computing** → **Natural language interfaces**.

KEYWORDS

Conversational design, ethnomethodology, conversation analysis

1 INTRODUCTION

The *concept of conversation* is habitually deployed when we talk about all manner of natural language, agent based interfaces. The argument goes that a whole class of such systems, whether using text-based chat or voice, either are—or aspirationally need to be—designedly 'conversational'. The sense in which the notion of conversation is used tends to imply that successful conversational user interfaces (CUIs) will need the computational detection, recognition, and understanding of at least some features of human conversation. But, it is very unclear what 'conversational' might mean, in spite of attempts to delineate types, e.g., "social talk" or "non-transactional conversation" [3]. Confusing matters further, 'conversational design' can also stand in for conversational *style* interfaces, encompassing a much broader design space. Here, 'conversation' acts more as a generative metaphor for design (e.g., [4]) than an attempt to recognise and handle conversational phenomena as tends to be the case for CUIs. It is this latter characterisation that I run with for the rest of this paper. (Note I reserve comment exclusively for voice-driven CUIs, too.)

The problem: the very idea of conversational design and CUIs trade on various category mistakes (cf. [9, pp. 8-12]). When designers and researchers state they are designing conversational systems, immediate questions arise. Do they mean that they are creating bone fide 'conversational' partners— agents that offer a *simulacrum*

of human conversationalists' competencies (e.g., to replicate actual, everyday, conversation via computational models of conversational turn-taking or repair practices)? Or, are they using a loose 'conversational metaphor' for design instead? These are very different things. Further, what are the analytic costs (and those to rigour) of leveraging conversation as a concept? And what about the *users* of CUIs themselves? How do they actually treat them, and what does this tell us not only about conversational design but also about everyday conversations as interactional phenemona?

2 THE TROUBLE WITH CONVERSATION

To begin with I want to talk about the words we are using here. If CUI research aims to build conversational interactions with people, then CUIs must also enable their users to 'talk to' them. It doesn't feel unreasonable to say that someone is talking to a CUI, just as we might say someone is talking to another person or talking to a voice recorder. A reason for this is clearly because it is possible to use 'talking' without necessarily also meaning 'understanding': talking out loud or talking to one's cat being cases in point. Crucially, it is also not that we mean 'talking' in the same sense here in spite of using the same word: and we tend *not* to confuse one sense—e.g., talking to a friend—with another—e.g., talking to a CUI (although we might imagine some edge cases where confusion *is* possible. 'Conversation' in turn is more problematic than 'talking' because—adopting an ethnomethodological point—conversation comes with an implication of intersubjective understanding and the mutual production of meaning. Thus it seems unlikely that we might hear it said earnestly that someone is 'having a conversation' with a CUI. Or at least if this were the case, we might take this as a joke, said ironically *precisely because* of the absence of a process of conventional 'understanding' taking place. Again, as with talking, thoroughly different senses of conversation are meant in spite of using the same word.

This is the first hint that something is awry when we discuss CUIs and the concept of conversation too blithely. Uses of 'conversation' in CUI research tend to be *unconventional uses*. Formalisations like 'social conversation' or 'task-based conversation' tend to be researcher-imposed categorisations at odds with how co-conversationalists themselves categorise their own talk activities. It is for conversationalists to decide what kind of social object it is that they are engaging in. The problem is that when we say 'conversational design' or 'conversational user interfaces' we probably mean a certain kind of machine-oriented, structurally-formatted *version* of conversation that is different to the sequential organisation of everyday conversation and the joint achievement of mutual understanding as exemplified by conversation analytic studies [10]. This a different position to that taken in some CUI research examining the structure of social talk [3]. There are a number of reasons I make this assertion; particularly because I want to stay with the

challenges posed by ethnomethodological conversation analysis to CUI research.

Firstly, encounters between people and CUIs are not conversational in the conventional sense because they are not *treated* as part of a conversation by their users. This view is informed by recent work analysing the organsation of conversation around actual CUIs which demonstrates that conversationalists treat CUIs as categorically different to their fellow co-conversationalists [2, 6–8]. This research shows how interactions with CUIs come to be embedded in turns-at-talk through the work of conversationalists. The CUI *does not do this embedding work*[1]; competent conversationalists do. This happens because input to the CUI must be staged and formulated appropriately, while output from the CUI must be accounted for and accommodated in the ongoing flow of talk. While users of CUIs must employ talking and listening practices to interact with the CUI, it just so happens that CUIs offer opportunities for input and output as sequential turns. This is decidedly not the same as mutual understanding as accomplished routinely in conversation. In effect we are confusing turn-based input and output for conversation. While it might be fair to call it 'a kind of conversation', it is not the same thing.

The difference is sharp for reasons discussed next, although first I want to head off two typical rebuttals. The first is it is argued that the state of the art in dialogue systems is far better than what is available commercially, so our studies cited above are unfair. The second is that the problems of conversation with CUIs will be solved in the near future with better technology or more training data or new as yet undiscovered approaches. But this is missing the point: interactional fragmentations will still be present in the same way as Suchman's well-known studies of AI-driven photocopier expert help systems illustrated over 30 years ago. Suchman showed how interactions *fragmented* across various divisions that are still present today: in social circumstances available between humans, resources available to machines, machine resources available to humans, and associated design rationales [13].

Secondly, beyond how members of conversations treat CUIs, there is a broader intellectual challenge mounted by the ethnomethodological conversation analysis view of conversation. This is that conversation is an ongoingly produced context-shaping, context-renewing intersubjective activity [5]. As a practice of establishing mutual intelligibility, prior turns-at-talk are heard and made sense of (analysed by members of the situation, *not researchers*) in light of present ones, present turns offer analyses of priors, and present turns also embed projections of next actions. This is how meaning is produced: leveraging mundane conversational competencies and a vast range of unspoken, joint understandings. Wittgenstein's later language philosophy underpins this move: linguistic referentialism is abandoned in favour of a conventionalist view of meaning in language, discarding cognitive explanations [12] in favour of the irremediably situated production of meaning.

This challenge leads to a critical implication for CUIs of Sacks et al.'s argument about conversation's next turn "proof procedure" [11, pp. 728-9]. Consider a core part of conversational turn-taking that all CUIs must support: adjacency pairings, like question-answer.

There is *no* intrinsic feature of a first-pair-part of a question-answer pair that determines—without recourse to how it is subsequentially treated by other conversationalists—that an utterance is a definitively a question. The 'question-ness' of the first-pair-part only becomes a question in and as this treatment ("answers make the question" [6]). This suggests that "simple question/answer routines" [3]—the bedrock of current commercial CUIs—really are not so simple at all, and carry within them the germ of much broader challenges to be faced in CUI research.

It seems unlikely that these conundrums are solvable with statistical pattern matching (machine learning techniques) or symbolic approaches. There are strong extant arguments against computational models of conversation that have yet to be addressed by proponents of a 'hard' conversational design view (see [1]). Presently, CUIs—state of the art or otherwise—don't anything other than offer pre-configured ways to treat conversational phenomena. What is absent is meaning and understanding.

Now we can return to the title of this paper: conversational design may be harmful, leading the design of CUIs towards a dead end, building ever more elaborate conversational structures in computational miniature, set against a backdrop of conceptually confused and / or utopian readings of the capabilities of artificial intelligence techniques.

3 CONVERSATION CONSIDERED USEFUL

A retort might follow from the previous section's critique that this treatment of the concept of conversation is mistaken: that we know we don't *really*, literally mean 'conversation' when we talk about CUIs and conversational design. This may be true, but firstly it is unclear whether this is a prevalent view and secondly if so, that fact needs to be relentlessly underlined, and the dimensions of the metaphor outlined clearly. Ideas that are brought along with invocation of 'conversation', when poorly delimited and uncaveated, can set up hugely unrealistic expectations for research, design, and end users. Blurring the technical with vernacular (conventional) uses of conversation does no-one any favours.

But, I don't want to argue entirely against using the concept of conversation in the context of CUI design, nor the advancements achieved by speech technology research from following this view. Rather, I am arguing against a free, easy and uncritical use of a 'conversation' that is laden with traps. Maybe, treated with sensitivity, conversation could be useful for research on and design for CUIs. So instead I want to respecifying conversational design and CUIs as concepts that *themselves* need continuously working on and greater care. Conversational design could instead be about how CUIs fit *into* and *around* conversation as as opposed to embuing conversational capabilities to systems.

To conclude, instead of talking about 'designing conversation' or 'conversational design' I propose **conversation-sensitive design**. This better captures the relationship between the CUIs we build, the actual ways in which they become embedded in everyday conversation as it unfolds, and avoids the conceptual muddle being set up.

REFERENCES

[1] Graham Button, Jeff Coulter, John R. E. Lee, and Wes Sharrock. 1995. *Computers, Minds and Conduct.* Polity Press.

---

[1]AI systems never can do this work: there is always human work to make the AI work. Such human work is sublimated in accounts of AI systems with negative effect.

[2] Joel Fischer, Stuart Reeves, Martin Porcheron, and Rein Sikveland. 2019. Progressivity for Voice Interface Design. In *Conversational User Interfaces 2019*.

[3] Emer Gilmartin, Benjamin R. Cowan, Carl Vogel, and Nick Campbell. 2018. Explorations in multiparty casual social talk and its relevance for social human machine dialogue. *Journal on Multimodal User Interfaces* 12, 4 (01 Dec 2018), 297–308. https://doi.org/10.1007/s12193-018-0274-2

[4] Erika Hall. 2018. *Conversational Design*. A Book Apart.

[5] John Heritage. 1984. *Garfinkel and Ethnomethodology*. Polity Press, Cambridge.

[6] Martin Porcheron, Joel E. Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life. In *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems (CHI '18)*. ACM, ACM, New York, NY, USA. https://doi.org/10.1145/3173574.3174214

[7] Stuart Reeves and Martin Porcheron. 2018. Talking with Alexa. *The Psychologist* 31 (December 2018), 37–39.

[8] Stuart Reeves, Martin Porcheron, and Joel Fischer. 2018. 'This is Not What We Wanted': Designing for Conversation with Voice Interfaces. *Interactions* 26, 1 (Dec. 2018), 46–51. https://doi.org/10.1145/3296699

[9] Gilbert Ryle. 1949. *The Concept of Mind*. The University of Chicago Press.

[10] Harvey Sacks. 1992. *Harvey Sacks: Letures on Conversation*. Basil Publishing, Oxford.

[11] Harvey Sacks, Emanuel Schegloff, and Gail Jefferson. 1974. A Simplest Systematics for the Organisation of Turn-Taking for Conversation. *Language* 50 (1974), 696–735.

[12] Severin Schroeder. 2006. *Wittgenstein: The Way Out of the Fly-Bottle*. Polity Press.

[13] Lucy A. Suchman. 2006. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press, New York, NY, USA.