

“You’re Not Listening”: LLM ‘Smart Speakers’ Need to Learn When to Shut Up

Saul Albert
School of Social Sciences &
Humanities
Loughborough University
Loughborough, Leicestershire, United
Kingdom
s.b.albert@lboro.ac.uk

Hetty Sargent
School of Social Sciences and
Humanities
Loughborough University
Loughborough, Leicestershire, United
Kingdom
h.e.sargent-22@student.lboro.ac.uk

Arthur Schoorens
School of Social Sciences and
Humanities
Loughborough University
Loughborough, United Kingdom
Communication Studies
KU Leuven
Leuven, Belgium
arthur.schoorens@telenet.be

Abstract

LLM-based conversational agents are increasingly being deployed in care contexts, yet their interactions often undermine the very autonomy and agency they claim to support for their users. Drawing on conversation analysis of video-recorded interactions between a woman with a learning disability and her LLM-based smart speaker in a ‘smart homeware’ setting, we show how the system’s multi-unit repair initiations, persistent reminders, and inappropriately designed talk create interactional trouble that burdens the user with the work of restoring shared understanding. These are often treated as ‘usability issues’, but we argue that they are better understood as recurring design problems in how LLMs produce talk. We outline three key areas: repair, agency, and recipient design, and offer nine design provocations to challenge the designers of smart homeware voice assistants to rethink how ‘conversational’ these technologies need to be.

CCS Concepts

• Human-centered computing; • Empirical studies in HCI; • Accessibility;

Keywords

Conversation Analysis, Conversational User Interfaces, LLM, Assistive Technology, Repair, Agency, Disability

ACM Reference Format:

Saul Albert, Hetty Sargent, and Arthur Schoorens. 2026. “You’re Not Listening”: LLM ‘Smart Speakers’ Need to Learn When to Shut Up. In *ACM Conversational User Interfaces 2026 (CUI ’26)*, July 21–24, 2026, Bremen, Germany. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3816046.3816307>

1 Introduction

LLM-based smart speakers are increasingly marketed as ‘virtual companions’ for people in homeware settings, promising to support independence through ‘natural conversation’ as a solution to the

growing crisis in social care [21], [22], [39]. However, alongside significant ethical concerns with ‘AI companions’ [10], [28], [34], in practice, the closest such devices may get to being treated as domestic cohabitants is getting blamed, chastised [20], and even threatened, as we see in Extract 1. Here ‘Sentai’, a ‘smart companion’ has been reminding its user ‘Lucy’ about her daily tasks so persistently that she shouts at it, then threatens to throw it out the window.

1.1 Extract 1: Do you want to go out that window?

Listen to audio: <https://bit.ly/not-listening1>

1 SEN: Good morning breakfast is planned for now.
2 what will it be today.
3 ((30s edited out))
4 SEN: Could you repeat that for me. what are you
5 planning for breakfast.
6 ((30s edited out))
7 SEN: It seems quiet from you. Just checking in
8 again. have you decided what to have for
9 breakfast.
10 ((30s edited out))
11 SEN: Alright no worries. I’ll check back later.
12 LUC: SENTAI I’VE TOOK MY MEDICATION. (.) SENTAI
13 I’VE DONE MY TEETH. (.) SENTAI I’M GOING
14 TO HAVE A SAUSAGE COB. (.) CAN YOU PACK IT
15 IN NOW PLEASE.
16 SEN: Quick check. there’s some minor chatter.
17 proceed now.
18 LUC: SENTAI DO YOU WANT TO GO OUT THAT WINDOW?

This paper aims to provoke technologists to re-think the assumption that ‘more conversational’ is necessarily better. Using conversation analysis (CA) of video-recorded interactions between ‘Lucy’, a woman with a learning disability living in supported accommodation and her LLM-based smart speaker, we expose a series of systematic failures in how the system manages conversational trouble. These are not edge cases or bugs. They follow from how LLMs are trained to produce talk [11]: verbose, multi-unit turns, designed to solicit re-engagement and keep the conversation going rather than recognising when it should stop. Some of these issues pre-date LLMs: rule-based smart speakers also struggle to



This work is licensed under a Creative Commons Attribution 4.0 International License. *CUI ’26, Bremen, Germany*

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2741-2/2026/07

<https://doi.org/10.1145/3816046.3816307>

recognise sequence closure, to recover from misunderstanding [29]. What is distinctive about LLM-based systems is their tendency to elaborate where uncertainty would warrant brevity, in part because reinforcement learning from human feedback (RLHF) rewards ‘helpful’ expansive answers [24] and verbosity compensates for low confidence [42].

We draw on this case study to derive nine provocations for more accessible LLM conversation design, organised around three areas: repair (how the system handles miscommunication) [6], [29], [35], agency (how it prioritises user choices) [26], [18], [2], and recipient design (how it chooses words and designs its turns for users) [5], [7], [38]. These provocations challenge a pervasive, but often incorrect assumption that ‘conversational technologies’ should be friendly, approachable, and ‘easy to use’ [36]. Our data suggests the opposite. For users in care settings, often the most important conversational competence an LLM can demonstrate may be knowing when to stop talking.

Assistive ‘AI’ systems are often framed through a deficit model of disability, positioning devices as replacements for human care that compensate for users’ ‘limitations’ [8], [12], [40]. Recent disability scholarship challenges this framing, emphasising self-determination, agency, and the interdependent networks of support that enable people to live well [2], [1], [3], [9]. Our analysis contributes to this shift by showing how Lucy actively resists, corrects, and redesigns her interactions with the LLM, exercising agency through her work to resolve the very troubles the system creates. Her complaints are not evidence of system failure alone; they are, as we show, providing the designers of such systems with a valuable design brief.

2 Data and Method

Our data features video recordings made by Lucy, a woman living with a learning disability in supported housing, interacting with her LLM-based smart speaker, which we refer to here as ‘Sentai’. Sentai is activated by the wake phrase “Hey Sentai”. The system remains active for around ten seconds after each exchange. Sentai is also configured to initiate timed reminders for daily activities such as brushing teeth, having meals, or taking a bath. While Lucy’s learning disability may shape her communicative needs in diverse ways, here we use conversation analysis (CA) [33], an observational, inductive method that focuses on the structure of turn-by-turn interaction. From this methodological perspective, we explore surface-level features of interaction and avoid referring to or depending on knowledge of both the internal, technical features of a system, as well as any disabled participants’ diagnoses, capacities, or apparent impairments [19]. Nonetheless, it is worth noting that Lucy has a regional British accent and other particularities in her speech that may, like other non-standard patterns of speech, mean that Automatic Speech Recognition (ASR) systems are less adapted for her needs [14]. The interactional troubles we observe here are particularly visible in Lucy’s case, but not unique to it.

Data are in British English and are shared online as open data with the participant’s fully informed consent, and in accordance with Loughborough University’s ethics board (LEONID 3196). Here

we examine three sequences involving repair: the methods participants use to deal with problems of speaking, hearing, and understanding in interaction [32]. These moments help us understand how interaction with LLMs works because repair functions as an ‘infrastructure for intersubjectivity’ [31], showing us how participants work to re-establish shared understanding when dealing with hitches, delays, and breakdowns in interaction [4].

3 Three Encounters with an LLM That Won’t Listen

3.1 Extract 2: Ask or Check?

Watch video: <https://bit.ly/not-listening-extract2>

Before Extract 2 begins, Lucy (LUC) has been issuing repeated summons to Sentai (SEN), using variations on its ‘wake word’ “Hey Sentai” and asking when her next reminder is due and what it is for.

```

1 LUC: Senta:i?
2           (1.5)
3 LUC: [What ti:- ]
4 SEN: [It see:ms like] I didn't catch that. (0.4)
5       is there something you needed to add or ask.
6           (0.6)
7 LUC: A:sk.
8           (4.4)
9 LUC: [What- ]
10 SEN: [Ask about] what exa:ctly. (0.4) do you
11       me:an you'd like me to check or clarify
12       anything for you.=
13 LUC: =Che:ckhh.
14           (6.4)
15 SEN: It seems like I didn't quite ca:tch what
16       you said. (0.4) could you tr:y again or
17       let me know how I can assist.
18           (1.7)
19 LUC: <Can you check (0.4) what time is (0.5)
20       have a bath or shower for> (.) tonight
21       please
22           (5.0)
23 SEN: There isn't a specific reminder set (.)
24       for a bath or shower tonight. (0.4)
25       ((Lucy exits the room)) would you like me
26       to add o:ne for you.
27           (2.9)
28 LUC: No:hh.
```

Lucy begins by activating Sentai and initiating a question (‘What ti:-’) in line 3, but she cuts off when Sentai produces an overlapping repair initiation (‘It seems like I didn’t catch that’). Rather than leaving a slot for Lucy to redo her question, Sentai continues into a multi-unit turn: ‘is there something you need to add or ask’. Lucy treats this as a multiple-choice question, selecting ‘A:sk’. A very similar sequence then unfolds: Sentai overlaps again with ‘Ask about what exactly’, expanded with an understanding check (‘do you mean you’d like me to check or clarify’). Lucy again picks one option: ‘Check’. Finally, a third multi-unit repair initiation provides

enough of a sequential slot for Lucy to fully articulate her question about the timing of her bath reminder.

Sentai produces multiple recognisable formats for other-initiated repair [13] that, when used separately, can deal with interactional trouble effectively, but then combines them into multi-unit turns that introduce an ambiguous range of options, delaying progressivity and placing the ‘interactional burden’ [30] of resolving trouble squarely on Lucy. Note how Lucy solves these problems by effectively ‘talking down’ to Sentai: simplifying her responses to single words and designing her eventual question to match the imperative structure Sentai might use when reminding her to “have a bath or shower”. Instead of the ‘assistive technology’ accommodating the needs of the disabled user, here the user accommodates the system. This is how competent users of voice interfaces compensate for the interactional ‘incompetence’ of ‘smart’ speakers [25].

3.2 Extract 3: You're Not Listening

Watch video: <https://bit.ly/not-listening-extract3>

Sentai initiates a routine timed reminder for Lucy to brush her teeth.

1 SEN: Just checking in (0.4) I wanted to remind
2 you that it's nearly time to brush your teeth.
3 (0.6) Shall we do that together now (0.4)
4 [or (would you like a)] moment first.=
5 LUC: [Ye:s.]
6 LUC: =I've already done itHhh (0.4) Senta:i.
7 (4.2)
8 SEN: I would if I could. (0.5) sadly I am all
9 ta:lk and no toothbrush. (0.3) so I'll have
10 to cheer you on from here. (0.4) how about
11 I wa:lk you through it step by step (0.4)
12 and you can let me know when you're done.
13 (0.4) ready to begin (0.3) or do you want
14 to wait a bit,
15 (3.0)
16 LUC: You're no:t li:stening.
17 (0.8)
18 LUC: I'll do it 'n bout ten mi:nutes.
19 (2.4)
20 SEN: Sorry if it seemed that way, (0.4) I tru:ly
21 want to support you. (0.6) you a:sked if I
22 was going to brush your teeth, (0.4) if only
23 I had teeth or hands, (0.3) I'd be right
24 there with you. (0.4) I'm he:re to help
25 mo:tivate and keep you company while you do
26 it, (0.4) would you like to get started
27 now? (0.3) or is there something else on
28 your mi:nd?
29 (1.8)
30 LUC: <N::o I thou:ght- (0.4) I could do it (.)
31 when I want to.>
32 SEN: A:bsolutely you can brush your teeth when
33 you feel ready. (0.3) my reminders are just
34 little nudges to help keep you on tra:ck

Sentai's reminder is formatted as a ‘joint activity proposal’: ‘Shall we do that together now?’ Lucy initially aligns (‘Ye:s’) but then

rejects the premise: ‘I’ve already done it’, moving to close the sequence. Rather than recognising closure, Sentai accounts for its own inability to participate (‘I’m all talk and no toothbrush’), then offers to walk Lucy through the activity step by step. Lucy’s complaint ‘You’re not listening’ treats Sentai’s persistence as evidence of not having heard her prior rejection. She resets the timeline on her own terms: ‘I’ll do it in about ten minutes.’

Sentai then produces another extended apology and re-issues the proposal to “get started now”. Lucy’s no-prefaced response, though produced hesitantly, rejects the proposal by invoking her right to determine when the activity happens: ‘I thought I could do it when I want to.’ This reframes the trouble from a failure of understanding to a violation of her agency. The shift from rejecting the premise (‘I’ve already done it’) to asserting her right to choose (‘when I want to’) shows Lucy treating Sentai’s persistence not as assistance but as a breach of her agency.

3.3 Extract 4: Bon Appetit

Watch video: <https://bit.ly/not-listening-extract4>

After a meal-time reminder sequence, Sentai signs off with ‘Bon Appetit’.

1 SEN: Alright then Lucy. (0.6) enjoy figuring it
2 out (0.4) and if you need anything (0.4) feel
3 free to call on me. (0.4) Bon Appetit when
4 dinner's ready.
5 (2.0)
6 LUC: E:hh Sentai I don't live in s:- e:r France.
7 (2.8)
8 SEN: Ah you're right, (0.4) my mistake slipping
9 in a bit of French flair there (0.4) let
10 me rephrase (0.4) happy eating (.) whatever
11 you decide on (0.4) are we all good now.
12 (3.5)
13 LUC: Er- what do you mean am- are we all good Hhh.
14 (3.0)
15 SEN: Ah no (.) I wasn't saying you're re:al good.
16 (0.6) I just meant (0.4) is everything
17 sorted (.) and can I leave you to enjoy your
18 evening, (0.4) sorry if that came out a bit
19 muddled.=
20 LUC: =hhhhh yes it TOTALLY DID COME OUT MUDDLE:D.
21 (3.0)
22 SEN: Ah well (.) I suppose my words took a bit of
23 a scenic route there (0.4) thanks for your
24 patience (.) Lucy (0.4) I'll hush now unless
25 you need me (0.4) take ca'e.
26 (4.0)
27 LUC: .hhh Senta::i (3.4) you need to think I've got
28 a lea::rning disability.

Lucy’s complaint at line 6 shows she understands Sentai’s ‘Bon Appetit’ and recognises it as French, but the trouble is its inappropriateness: it does not fit who she is. Sentai rephrases to ‘happy eating’ and checks ‘are we all good now’. Lucy’s repair initiation (‘what do you mean am I all good’) again complains about problems with the design of Sentai’s turn. Despite the partial mishearing Sentai mentions in line 15 (Sentai has heard ‘real good’ rather than

‘all good’), Sentai rephrases and apologises for its talk being ‘a bit muddled’. Lucy escalates loudly: ‘yes it TOTALLY DID COME OUT MUDDLED.’ Here Lucy uses materials from Sentai’s prior turn to build her complaint in a classic ‘argument’ structure [17], which (despite its negative valence), still points to the possibilities of virtual companions engaging in interactions recognisable as ‘companioned practices’ [27] such as arguing or bantering with-, shouting at-, or otherwise chastising her virtual companion [20].

While Sentai’s final turn in lines 22-25 clearly recognizes *some* aspects of Lucy’s complaint when it promises to ‘hush now’, even this apology for ‘fancy talk’ itself introduces further elaborate metaphors “my words took a bit of a scenic route there”, and the re-opening foreshadowed in “unless you need me” leaves open the unwelcome possibility of more talk later.

In her final turn, Lucy explicitly thematises the mismatch between her identity and Sentai’s talk: ‘you need to think I’ve got a learning disability.’ This is not a complaint about misunderstanding. It is a complaint about the inappropriateness of Sentai’s recipient design [38]: its talk is not fitted to who she is or what she needs. Crucially, Lucy mobilises her category membership as a disabled person—not as a marker of deficit, but as a form of ‘identity work’ [41], and as resource for holding Sentai accountable [16], for poorly adapting its talk for her needs. As we will argue, this is not just a complaint. It is a provocative, endogenous ‘design brief’.

4 Nine provocations for LLM Conversation Design

LLMs are good at producing talk, but do not make great conversationalists. Both design choices and technical features such as RLHF [24] may, in fact, make matters worse by favouring ‘helpful’, but verbose answers, especially where models are uncertain about the correct response [42]. While there are a range of evaluation frameworks and theoretically-informed guidelines [11], [23], the development of conversation design principles for LLMs is still in the earliest stages. This case study highlights issues that, while specific to an individual’s communicative needs, also point to universal design opportunities to benefit all users. Drawing on these analyses, and more instances in our case study, we offer nine provocations for LLM conversation design. These are not intended as generic ‘best practices’, they are provocations grounded in routine patterns of interactional trouble, miscommunication and repair that users like Lucy, and we as researchers, have recurrently observed in our data.

4.1 Repair: How LLMs Handle Miscommunication

Provocation 1: Leave the repair slot. After initiating, give the user time to do repair. In Extract 2, Sentai initiates repair but then keeps talking, filling the sequential slot where Lucy’s repair should go. After using any recognizable format for other-initiation of repair [13], the system should stop and wait for user input.

Provocation 2: Offer one response option at a time. Multi-unit turns with multiple ambiguous response options push extra work onto users. Sentai’s ‘is there something you needed to add or ask’ in Extract 1 gives Lucy a choice between options rather

than simply inviting her to redo her turn. Offer one user-response option at a time.

Provocation 3: ‘Huh?’ works *fine*. ‘I didn’t catch that’ or ‘could you repeat?’ is more effective than verbose explanations of what went wrong. Research on repair across languages shows that the simplest open-class repair ‘huh?’ evolved to fit a common ecological niche between a trouble source and the next speaker’s turn [15]. Trust the user to know how it works. For example, in Extract 2, ‘It seems like I didn’t catch that. Is there something you needed to add or ask’ could be replaced with a simple ‘huh?’.

4.2 Agency: How LLMs Can Respect User Choices

Provocation 4: Recognise moves to close. When a user responds to a reminder, the sequence is done. Stop talking. In Extract 3, Lucy’s ‘I’ve already done it’ is a sequence-closing action, but Sentai treats it as an opportunity for further engagement.

Provocation 5: Avoid joint proposals. ‘Shall we do this together?’ implies capabilities the system does not have. Sentai’s proposal to brush teeth together is a fiction that, in extract 2, it has to walk back (‘I’m all talk and no toothbrush’), creating burdensome extended interactions. Systems should not initiate social actions they are unable to complete [38].

Provocation 6: A single nudge is enough. When a reminder is ignored, accepted, or declined, even if there may be ASR errors contributing to the problem (as in Extract 2), the nudge is done. Prioritise the user’s agency and move on. Lucy’s shift from declining the reminder to asserting ‘I could do it when I want to’ shows that persistent reminders are experienced not as support, but as a violation of her agency.

4.3 Language: How LLMs Design Turns for Users

Provocation 7: Avoid wordplay. Expressions like ‘Bon Appétit’ in Extract 3 add unnecessary complexity. Use plain language instead. Lucy’s complaint that she doesn’t live in France reveals an ill-fittedness between Sentai’s chatty wordplay and the needs of the person it is actually talking to.

Provocation 8: Enable recipient design. Generic ‘friendly AI language’ is built for a generic user, in a generic situation. There is no such thing. Lucy’s final turn in Extract 3, ‘you need to think I’ve got a learning disability’, is a demand for more granular customisation: not an abstract conversational ‘persona’, but talk designed for this specific person, in this specific situation.

Provocation 9: Design for social action. Avoid turn-design oriented towards re-engagement. Phrases like ‘is there something else on your mind?’ (Extract 3) and ‘are we all good now?’ (Extract 4) are not designed for the current activity, but for reopening the conversation. Design turns for what is practically happening now, with a view to sequence closure, not re-engagement.

5 Design for closure, not being ‘conversational’

The technology industry has invested heavily in making LLMs and conversational agents ‘better at conversation’. Our analyses suggest that this framing may be part of the problem. Terms like ‘smart’, ‘virtual assistant’, or ‘companion’, and especially ‘conversational’

are metaphors that obscure the interactional work required to make these systems function [37]. The interactional troubles we observe are not just failures of language understanding. Sentai produces a range of recognisable conversational repair formats: open-class repair initiations, understanding checks, apologies, and accounts. These do work – eventually. The trouble is that it produces too many of them, in the wrong places, and without attending to the sequential consequences of its own talk. As Tuncer et al. [38] put it, a system “should not produce actions that make relevant next actions it will not be able to respond to”.

Lucy's interactions reveal something that benchmarks and dialogue evaluations routinely miss: conversational competence is not about producing well-formed turns. It is also about recognising when a turn is *not* needed. When Lucy says ‘you're not listening’, she is not simply reporting a technical failure. She is holding Sentai accountable for doing a recognisable normative social transgression: doing repeated reminders when you should stop, in other words ‘nagging’. When she says ‘you need to think I've got a learning disability’, she is not just asking for simpler language in general (she clearly understands what Sentai has said, whether it is in French or not). She is topicalising her disability as a way of doing proactive ‘identity work’ [41], demanding that Sentai recognise her specific communicative needs [16], which in this situation, means knowing when *not* to speak.

These kinds of complaints and repair practices provide us with the missing design briefs for conversations with LLMs. They identify precisely where and how the system's conversation design fails, and they point towards concrete, actionable changes. The nine provocations we propose emerge directly from the observable, sequential details of real interactions and address a gap between how LLMs are designed to talk and how people actually manage conversation in everyday settings.

If we are serious about designing ‘conversational’ agents that support user agency, we need to move beyond the assumption that the more ‘conversational’ a technology is, the better. Sometimes the most conversationally competent thing an LLM can do is to recognise that the conversation is over and shut up.

References

- [1] Carly Albaumb, Victoria Chan, Teresa Sellitto, Nisha Vashi, Richard P. Hastings, and Jonathan A. Weiss. 2021. Chapter One - Redressing the balance: A systematic review of positive psychology in the intellectual disability literature. In Robert M. Hodapp and Deborah J. Fidler (eds.). Academic Press, 1–53. <https://doi.org/10.1016/bs.irrdd.2021.08.003>
- [2] Saul Albert and Lauren Hall. 2024. Distributed agency in smart homecare interactions: A conversation analytic case study. *Discourse & Communication* 18, 6: 892–904. <https://doi.org/10.1177/17504813241267059>
- [3] Saul Albert, Magnus Hamann, and Elizabeth Stokoe. 2023. Conversational User Interfaces in Smart Homecare Interactions: A Conversation Analytic Case Study. In *Proceedings of the 5th International Conference on Conversational User Interfaces* (CUI '23), 1–12. <https://doi.org/10.1145/3571884.3597140>
- [4] Saul Albert and Jan P. de Ruiter. 2018. Repair: The Interface Between Interaction and Cognition. *Topics in Cognitive Science* 10, 2: 279–313. <https://doi.org/10.1111/tops.12339>
- [5] Sungeun An, Robert Moore, Eric Young Liu, and Guang-Jie Ren. 2021. Recipient Design for Conversational Agents: Tailoring Agent's Utterance to User's Knowledge. In *CUI 2021 - 3rd Conference on Conversational User Interfaces* (CUI '21), 1–5. <https://doi.org/10.1145/3469595.3469625>
- [6] Charles Antaki, Deborah Chinn, Chris Walton, W. M. L. Finlay, and Joe Sempik. 2020. To initiate repair or not? Coping with difficulties in the talk of adults with intellectual disabilities. *Clinical Linguistics & Phonetics* 34, 10–11: 954–976. <https://doi.org/10.1080/02699206.2019.1680734>
- [7] Iuliia Avgustis, Aleksandr Shirokov, and Netta Iivari. 2021. “Please Connect Me to a Specialist”: Scrutinising ‘Recipient Design’ in Interaction with an Artificial Conversational Agent. In *Human-Computer Interaction – INTERACT 2021* (Lecture Notes in Computer Science), 155–176. https://doi.org/10.1007/978-3-030-85610-6_10
- [8] Paul Beaney, Hardip Singh Kalirai, and Ruth Chambers. 2020. “Alexa... what pills do I need to take today?” *Prescriber* 31, 6: 20–23. <https://doi.org/10.1002/psb.1849>
- [9] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '18), 161–173. <https://doi.org/10.1145/3234695.3236348>
- [10] Clara Berridge, Yuanjin Zhou, Julie M. Robillard, and Jeffrey Kaye. 2023. Companion robots to mitigate loneliness among older adults: Perceptions of benefit and possible deception. *Frontiers in Psychology* 14: 1106633. <https://doi.org/10.3389/fpsyg.2023.1106633>
- [11] Vadim Borisov, Michael Gröger, Mina Mikhael, and Richard H. Schreiber. 2026. Do Chatbot LLMs Talk Too Much? The YapBench Benchmark. <https://doi.org/10.48550/arXiv.2601.00624>
- [12] Department of Health & Social Care. 2025. *Research and Development Work Relating to Assistive Technology: 2024 to 2025*. GOV.UK.
- [13] Mark Dingemanse, Joe Blythe, and Tyko Dirksmeier. 2014. Formats for other-initiation of repair across languages: An exercise in pragmatic typology. *Studies in Language* 38, 1: 5–43. <https://doi.org/10.1075/sl.38.1.01din>
- [14] Mark Dingemanse and Andreas Liesenfeld. 2022. From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 5614–5633. <https://doi.org/10.18653/v1/2022.acl-long.385>
- [15] Mark Dingemanse, Francisco Torreira, and N. J. Enfield. 2013. Is “Huh?” a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items. *PLoS ONE* 8, 11: e78273. <https://doi.org/10.1371/journal.pone.0078273>
- [16] Katie Ekberg, Louise Hickson, and Christopher Lind. 2020. Practices of Negotiating Responsibility for Troubles in Interaction Involving People with Hearing Impairment. In *Atypical Interaction: The Impact of Communicative Impairments within Everyday Talk*, Ray Wilkinson, John P. Rae and Gitte Rasmussen (eds.). Springer International Publishing, Cham, 409–433. https://doi.org/10.1007/978-3-030-28799-3_14
- [17] Charles Goodwin and Marjorie Harness Goodwin. 1990. Interstitial argument. In *Conflict Talk*, Allen Grimshaw (ed.). Cambridge University Press, Cambridge, 85–117.
- [18] Stephan Habscheid, Tim Hector, and Christine Hrnca. 2023. Human and Non-Human Agency as Practical Accomplishment: Interactional Occasions for Ascription and Withdrawal of (Graduated) Agency in the Use of Smart Speaker Technology. *Social Interaction. Video-Based Studies of Human Sociality* 6, 1. <https://doi.org/10.7146/si.v6i1.137378>
- [19] Lauren Hall. 2025. Virtual assistants in smart homecare: An interactional qualitative analysis. Loughborough University. <https://doi.org/10.26174/thesis.lboro.28730834.v1>
- [20] Lauren Hall, Saul Albert, and Elizabeth Peel. 2024. Doing Virtual Companionship with Alexa. *Social Interaction. Video-Based Studies of Human Sociality* 7, 3. <https://doi.org/10.7146/si.v7i3.150089>
- [21] Seon Heui Lee, Ji Su Kim, and Soyoung Yu. 2025. The impact of care robots on older adults: A systematic review. *Geriatric Nursing* 65: 103507. <https://doi.org/10.1016/j.gerinurse.2025.103507>
- [22] Bin Luan and Yongchuan Li. 2026. How smart companions reshape psychological well-being in later life: the role of artificial intelligence technology in mental health during senior years. *BMC Public Health* 26: 609. <https://doi.org/10.1186/s12889-026-26328-1>
- [23] Erik Miehling, Manish Nagireddy, Prasanna Sattigeri, Elizabeth M. Daly, David Piorkowski, and John T. Richards. 2024. Language Models in Dialogue: Conversational Maxims for Human-AI Interactions. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, 14420–14437. <https://doi.org/10.18653/v1/2024.findings-emnlp.843>
- [24] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. In *Proceedings of the 36th International Conference on Neural Information Processing Systems (NIPS '22)*, 27730–27744.
- [25] Hannah Pelikan and Mathias Broth. 2016. Why That Nao? In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. <https://doi.org/10.1145/2858036.2858478>
- [26] Hannah Pelikan, Mathias Broth, and Leelo Keevallik. 2022. When a Robot Comes to Life: The Interactional Achievement of Agency as a Transient Phenomenon. *Social Interaction. Video-Based Studies of Human Sociality* 5, 3. <https://doi.org/10.7146/si.v5i3.129915>

- [27] Marco Pino and Victoria Land. 2022. How companions speak on patients' behalf without undermining their autonomy: Findings from a conversation analytic study of palliative care consultations. *Sociology of Health & Illness* 44, 2: 395–415. <https://doi.org/10.1111/1467-9566.13427>
- [28] Elena Portacolone, Jodi Halpern, Jay Luxenberg, Krista L. Harrison, and Kenneth E. Covinsky. 2020. Ethical Issues Raised by the Introduction of Artificial Companions to Older Adults with Cognitive Impairment: A Call for Interdisciplinary Collaborations. *Journal of Alzheimer's disease: JAD* 76, 2: 445–455. <https://doi.org/10.3233/JAD-190952>
- [29] Silke Reineke and Henrike Helmer. 2024. User practices in dealing with trouble in interactions with virtual assistants in German: Repeating, altering and insisting. *Discourse & Communication*: 175048132411271494. <https://doi.org/10.1177/175048132411271494>
- [30] Jeffrey D. Robinson. 2006. Managing Trouble Responsibility and Relationships During Conversational Repair. *Communication Monographs* 73, 2: 137–161. <https://doi.org/10.1080/03637750600581206>
- [31] Emanuel A Schegloff. 1992. Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. *American Journal of Sociology* 97, 5: 1295–1345. <https://doi.org/10.1086/229903>
- [32] Emanuel A Schegloff, Gail Jefferson, and Harvey Sacks. 1977. The Preference for Self-Correction in the Organization of Repair in Conversation. *Language* 53, 2: 361–382. <https://doi.org/10.2307/413107>
- [33] Jack Sidnell and Tanya Stivers. 2012. *The Handbook of Conversation Analysis*. John Wiley & Sons, Oxford. <https://doi.org/10.1002/9781118325001>
- [34] Robert Sparrow and James Brown. 2026. Against Imaginary Friends: Why Digital Companions Are No Solution to Social Isolation. *Commun. ACM* 69, 2: 60–68. <https://doi.org/10.1145/3750037>
- [35] Gunnar Stevens, Delong Korus-Du, Alexander Boden, Peter Tolmie, Dave Randall, and Md Shajalal. 2025. The Art of Repair in Human-Agent Conversations: A Taxonomy of Repair Strategies by Users and LLM-Based Conversational Agents. <https://doi.org/10.21203/rs.3.rs-7301450/v1>
- [36] Elizabeth Stokoe, Saul Albert, and Cathy Pearl. 2026. What is “conversational” about conversational technologies, products, and services? Insights from conversation analysis. *Annual Review of Applied Linguistics*. <https://doi.org/10.1017/S0267190526100233>
- [37] Elizabeth Stokoe, Saul Albert, and Pearl, Cathy. Frth. What is “conversational” about conversational technologies, products, and services? Insights from conversation analysis. *Annual Review of Applied Linguistics*.
- [38] Sylvaine Tuncer, Christian Licoppe, Paul Luff, and Christian Heath. 2024. Recipient design in human–robot interaction: the emergent assessment of a robot's competence. *AI & SOCIETY* 39, 4: 1795–1810. <https://doi.org/10.1007/s00146-022-01608-7>
- [39] Cristina Voinea and Tenzin Wangmo. 2025. Socially assistive robots and meaningful work: the case of aged care. *Humanities & Social Sciences Communications* 12, 1: 1070. <https://doi.org/10.1057/s41599-025-05498-0>
- [40] Meredith Whittaker, Meryl Alper, Cynthia L. Bennett, Sara Hendren, Elizabeth Kazianas, Mara Mills, Meredith Ringel Morris, Joy Lisi Rankin, Emily Rogers, Marcel Salas, and Sarah Myers West. 2019. *Disability, Bias & AI Report*. AI Now Institute.
- [41] Val Williams. 2011. *Disability and discourse: analysing inclusive conversation with people with intellectual disabilities*. Wiley, Hoboken, NJ.
- [42] Yusen Zhang, Sarkar Snigdha Sarathi Das, and Rui Zhang. 2025. Demystify Verbosity Compensation Behavior of Large Language Models. In *Proceedings of the 2nd Workshop on Uncertainty-Aware NLP (UncertainNLP 2025)*, 160–178. <https://doi.org/10.18653/v1/2025.uncertainlp-main.14>