



Social Interaction. Video-Based Studies of Human Sociality.
2025 Vol. 8, Issue 2
ISBN: 2446-3620
DOI: 10.7146/si.v8i2.140128

Social Interaction

Video-Based Studies of Human Sociality

Multilingual Participation in Accepting or Declining an Offer Made by a Monolingually Programmed Robot

Anna Filipi¹, Rosalyn Langedijk² & Kerstin Fischer²

¹Monash University

²University of Southern Denmark

Abstract

This paper reports an investigation of the interactions between groups of German and Danish speakers and a service robot which was programmed to produce English at an international university campus. We analysed three sets of interactions that involved an offer of water by the robot, and we used Conversation Analysis to track the human participants' responses to the robot in examining how their language choice featured in their participation. We found that the overall organisation of the interactions was monolingual: participants used German and Danish with each other to express wonderment, frustration and confusion, and to comment on the robot's actions, and English to respond to the robot's offer and to ridicule it when acceptance of the offer was missed. Language choice and variations in volume when speaking each language, as added dimensions in recipient design, thus established monolingual participation frameworks. We argue that these findings reveal a different orientation to the robot as co-participant and question the extent to which robots are oriented to as social members in settings that mirror real-life contexts. Findings also raise design issues in the future development of robots.

Keywords: human-robot interaction, multilingualism, Conversation Analysis, monolingual participation frameworks, orientation to robots as co-participants, attributions of interactional competence

1. Introduction

Robots are increasingly being placed in human social spaces. They are being developed with human-like behaviours to achieve effortless sharing of social spaces with people (e.g., Del Duchetto et al., 2019; Mizumaru et al., 2019; Mutlu & Forlizzi, 2008). This notwithstanding, from an ethnomethodological perspective, even though the number of studies on the interactional organisation of human and robot interaction (HRI) is rising (e.g., Fischer, 2023; Licoppe & Rollet, 2020; Majlesi et al., 2023; Pelikan & Broth, 2020; Pelikan et al., 2016; see also Mlynář et al., 2024 for a recent review), there remain gaps in our knowledge about how people respond to robots in naturalistic contexts. In particular, the investigation of how groups of multilingual human speakers who share the same languages interact with each other and with a robot, is one area of HRI that, as far as we have been able to ascertain, has not yet received extensive research attention. Investigations of the organisation of multilingual interactions can shed light on aspects of HRI that might otherwise be missed. They can elucidate what human speakers' multilingual practices reveal about how they orient to the robot as a co-participant in real-life or in quasi real-life settings.

This paper is concerned with the study of the multilingual participation of human speakers as they interact with a robot. The concept of participation framework, used as one component of the analytical framing for the study, enables the analysis of how participants locally manage their participation in interaction as they recipiently design their turns for their addressee. In talk between multilingual speakers, language choice adds a dimension to recipient design (Filipi, 2015) and is a resource that enables speakers to manage their participation (Mondada, 2004) to accomplish a range of actions including exclusion, alignment, or in-group talk. A second component of the analytical framing is Gafaranga's (2005, 2018) notion that language choice is an activity "in its own right" (2005, p. 291-292). Language choice provides "a level of talk organisation" (2018, p. 43) beyond the single turn that can determine whether the participation framework is: 1) monolingual where speakers speak in the same language, 2) in "parallel mode" (Gafaranga & Torras, 2001) where speakers consistently speak a different language from each other for an entire conversation or episode, or 3) multilingual where some turns are in a mix of two or more languages (or to use Gafaranga and Torras' (2001) term, in "mixed mode"). The combined framing will be used in the analysis of the participation of groups of multilingual English and Danish or German speakers in their interactions with a robot programmed to use English to determine how the human participants orient to the robot as a co-participant and how language choice makes their orientations visible.

Two research questions will guide the analysis:

1. How do multilingual Danish and English or German and English speakers interact with each other and with the robot in an offer sequence and manage problems that arise through their language choice?

2. What does the human participants' language choice reveal about their orientation to the robot as a co-participant?

2. Background

2.1 Human-robot interaction

In research institutions and industry across the globe, robots are being designed to assist in, or even take over, functions currently fulfilled by humans. Consequently, to ensure that people can interact with robots in intuitive ways (Bartneck et al., 2020; Fischer, 2020), robots are being equipped with 'social signal processing' that enables them to recognise and process social signals like gaze, speech or body orientation. That is, they are equipped to recognise human social behaviours and to respond in human-like ways to these signals (for instance, they recognise and produce speech in response to the recognised speech). At the same time, while the processes by means of which this happens are still under discussion, humans often interact with robots in social ways (Clark & Fischer, 2023; Groom et al., 2011); that is, they are responded to as if they are sentient beings and social actors. Indeed, there is heated discussion about the causes of apparent misattributions of characteristics of human beings to artefacts like robots (e.g., Lombard & Xu, 2021; the commentaries in Clark & Fischer, 2023).

From the perspective of Conversation Analysis (CA), many human social practices have been found to be instantiated in similar ways when they are being enacted by robots. For instance, Pitsch et al. (2009) have shown that when a robot uses a restart (Goodwin, 1980), museum visitors look at and listen to the robot more often and for longer. Similarly, in line with CA research by Pillet-Shore (2012) that shows that greeting sequences are longer and more extended in interactions with family members and close friends, Fischer et al. (2020) found that if a robot uses longer greetings, people treat the robot with more familiarity. Furthermore, when initiating service encounters, people have been found to apply similar practices as identified in CA work (Mortensen & Hazel, 2014) on initiating service encounters (Fischer, 2023). For example, they coordinate different levels of action including greeting exchanges, gaze engagement and moving into a spatial position to indicate readiness to interact.

However, while there are many similarities in the interactions between people and between people and robots, there are also many differences (cf. Clark & Fischer, 2023; Fischer, 2016; Rudaz et al., 2023). Especially in naturalistic, real-life situations, in contrast to laboratory studies, in which people are not necessarily prepared to encounter a robot, it is not clear how a robot will be received. Previous work has demonstrated both interpersonal variation (e.g., Clark & Fischer, 2023; Fischer, 2011) and changes over time in the way people display anthropomorphising behaviour (Fischer, 2021; Rudaz et al., 2023). Furthermore, most research on HRI still takes place in the laboratory where the

participants are volunteers, often students, who have explicitly consented to interacting with a robot (e.g., Lee et al., 2022). Thus, most work on human-robot interaction has been conducted in laboratory studies rather than naturalistic settings where people have often been found to respond to robots' behaviour as if it was produced by another human (e.g., Jung & Hinds, 2014). This may not necessarily be the case when they encounter a robot in real life, especially when participants are not specifically recruited to interact with a robot.

In naturally occurring and uncontrolled contexts outside the laboratory, findings have started to emerge about people's attention to social order in HRI. In Majlesi et al.'s (2023) study, for example, violations to turn-taking by robots and the way violations were managed by human participants, manifested through overlaps. Majlesi et al. conclude that such findings not only have implications that can inform engineers' robot designs but that they provide evidence of people's orientations to the normative order of turn-taking. As a further example, Rudaz et al. (2023) showed how a robot emerged as a possible co-participant by means of a sequential analysis of the interactions that occurred while a robot was being started up. The authors report that the timing of the robot's preprogrammed utterances influenced whether participants understood the robot to be responsive and interactive. Similarly, Bu et al. (2025) described a broad range of behaviours towards trashcan robots deployed in a public space in New York City and found that not everyone responded to the robot as a co-participant. Thus, while studies on how robots are oriented to as co-participants are starting to emerge, there is still a need for research that involves participants with different linguistic backgrounds and in different constellations.

2.2 Studies of bilingual interaction in CA

The study of bilingual interaction is now a well-established strand of research in CA since Auer's (1998) pioneering work on the local organisation of the bilingual turn. This is the case particularly in the context of language learning both within and outside the classroom. (For recent reviews, see the collections in Filipi & Markee (2018) on language learning and in Wong & Waring (2022) on multilingual storytelling in the home and in the classroom.) Core to this research is the premise that talk is social. Speakers draw on a reciprocity of perspectives based on speaker expectations that are displayed in their turn-taking and derived from prior experience, and it is in this sense that they become a set of norms. These norms are used to make sense of speakers' actions as they engage in talk (Bonacina-Pugh, 2020). From this perspective, the practice of language alternation is a social action that speakers deploy as a resource (2002) to achieve a range of social functions, including identity work, and to claim membership (Gafaranga, 1999; Mondada, 2004; Vöge, 2011).

A major preoccupation in multilingual CA research has been the issue of determining if there is a base code or primary language that characterises the

overall interaction (Auer, 1998; Gafaranga, 1999). Auer (1998) maintained that language alternation occurs at the local level of the turn. Gafaranga (1999) extended Auer's notion by suggesting that language alternation extends beyond the single turn to an episode or entire conversation, which he referred to as the overall order (Gafaranga, 2018). By adopting a language or medium that can be monolingual or multilingual, speakers' language alternation works as a 'scheme' (Gafaranga & Torras, 2001) or 'grid of interpretation' (Heritage, 1984) that determines a speaker's language choice. The choice can be the same language as a co-speaker's, in a mixed mode when part of a turn or some turns are in a mix of two or more languages, also referred to as translanguaging, or in parallel mode where speakers consistently interact in a different language from each other for an entire conversation (Gafaranga & Torras, 2001).

Where some speakers in a group share a common language, speakers commonly speak with each other in that language. Such a practice establishes a participation framework (Goodwin, 2007) based on language choice that temporally unfolds moment-by-moment. Chen and Bonacina (2021, p. 114) define participation framework in multilingual talk as involving speakers in "engag(ing) in the enterprises of maintaining shared attention and achieving mutual understanding with one another. Therefore, people in a multilingual speech community can draw upon multilingual resources in constructing such enterprises". Speakers on occasion need to deal with delicate matters, resolve interactional problems such as lack of understanding without openly admitting to not understanding, as happens in the foreign language classroom for example (Filipi, 2018), or address a matter that is not relevant to a wider audience that does not share the speaker's language, or that does not have the same access to the event being described or the language proficiency to understand the nuances (Skårup, 2004). In university foreign language, multilingual and mainstream educational settings (e.g., Chen & Bonacina-Pugh, 2021; Cromdal, 2005; Kunitz, 2018; Reichert & Liebscher, 2018; Filipi, in press; Filipi & Chuang, 2023), multilingual students have been shown to use their languages in highly organised ways to conduct different aspects of assigned learning tasks. These include to check understanding about requirements before engaging in group work (Filipi & Chuang, 2023) and to plan tasks, rehearse or to talk about the task product (Kunitz, 2018; Reichert & Liebscher, 2018).

The above studies show how language alternation provides a highly organised way of dealing with a range of interactional matters both in the classroom and more broadly. Research has also revealed that the practice of alternating language can be organised in prosodically different ways. For example, speakers may consistently whisper in one of their languages and speak in normal volume in the other (e.g., Amir, 2013; Filipi, in press; Filipi & Chuang, 2023). Through such practices, speakers may display an orientation to policy (such as resisting an English only policy in a foreign language classroom (Amir, 2013)) or to a breach in social norms (Filipi, in press; Filipi & Chuang, 2023). Notably, while establishing participation frameworks in a language shared by some speakers

can exclude speakers who do not speak that language, a speaker can also invite the participation of an excluded speaker as an act of “bilingual brokering” (Skårup, 2004, p. 41) to (re)engage them in talk when they alternate to a shared language. They may also return to the language used in the wider context once the business conducted in the other language has been accomplished (Filipi & Chuang, 2023).

This review has drawn attention to HRI, to the practices of multilingual speakers as they draw on their linguistic resources to accomplish specific social actions, and to the organisation of language choice in establishing different participation frameworks. The review has also highlighted how multilingualism can be a resource for managing participation, and how language choice can function as an observable action that sheds light on how participation is construed in each moment. In our investigation, the analysis of human participants’ multilingual practices in their interactions with a robot provides an additional analytical focus to uncover 1) how human participants respond to and orient to the robot as a co-participant when accepting or declining an offer, 2) how they manage problems that arise, and 3) how their language choice is deployed and organised in such management.

3. Methods

3.1 The study

The data for the study was recorded in the cafeteria of an international university which is a public open space in which international students and staff gather for lunch. While the official language at this university is English, many students and members of staff speak either German or Danish as their L1 due to the location of the university in the Danish-German border region. The data was gathered over three days using a goPro camera mounted on the head of a large service robot, which drove around with a tray with glasses of water that it offered to people, as illustrated below.

Figure 1. *The robot in three different locations in the canteen; left: the robot drives around; middle: the robot is interacting with a group of four at a table; right: the robot interacts with people standing in the hallway*



The robot's actions of driving around, stopping by groups of people for a certain amount of time and then leaving again, allowed clearly delimited phases to emerge in which it addressed and responded to them. In these exchanges, the robot positioned itself in people's field of view and initiated the interaction with a greeting. It generally completed the interaction by producing a farewell message. Authors 2 and 3 recorded a total of 197 such interactions, but for this study, we only focus on lunchtime sessions (88 in total) and specifically, on the interactions in which the human participants used either German or Danish (16 in total, selected by all three authors), as summarised in Table 1. Three of the 16 (as a representative sample of interactions that were either unproblematic or problematic) were then selected, which generated a total 14 minutes and 28 seconds of data with the average length of each interaction being 54.3 seconds.

Table 1. *The dataset analysed*

Total # of lunchtime interactions	# of interactions in which German or Danish was used	Total length of the recordings analysed
88	16	14:28

The interactions comprise episodes in which the human participants were interacting in pairs or small groups and were either moving along corridors (Extract 1) or sitting and sharing a coffee or lunch (Extracts 2 and 3) and conversing in German or Danish as the robot approached them with an offer of water. In two extracts, the robot's offer was accepted (but problematically in Extract 3), while in the other (Extract 2) it was not.

3.1.1 The robot

In the study, we used the SMOOTH-robot (Krüger et al., 2021), which is a large service robot developed to take over several tasks in elderly care facilities, such as transporting laundry or serving drinks. The robot carries its load on the back, which means that it must turn around to show people its load; in this case the cups with water. The robot's head is equipped with a microphone, speakers, cameras and two touchscreens, one in the front and one in the back. The front touchscreen displays a pair of simulated eyes (see pictures below).

Figure 2. *The robot with displayed eyes, the tray with glasses of water on its back, and a goPro camera between the loudspeakers on its head*



The robot is equipped with autonomous navigation and dialog capabilities; however, to ensure participants' safety and to be able to adapt the dialog to the right circumstances, the robot was remotely operated by two operators, called 'wizards' (Riek 2012), one for the navigation and one for the dialog.

3.1.2 The wizards and the dialog

The two robot operators, i.e. wizards, were located on a higher floor of the atrium with clear sight of the robot, but they were not immediately visible to the participants (see pictures below). The navigation wizard drove the robot around with a joystick while the dialog wizard listened through the robot's microphone and selected pre-recorded utterances from a script.

Figure 3. *The robot operators located on the balcony above the canteen*



Authors 2 and 3 designed a set of comparable utterances within seven categories so that overhearers did not hear the same utterances over and over again; namely: greetings (e.g. *hi there, sorry to bother you but...*), offers (e.g. *can I offer you some water, would you like some water*), persuasive utterances (e.g. *most women actually do take something to drink; research shows that it is important to drink enough water during the day*), sayings/jokes/day-specific utterances (e.g. *you can lead a horse to water but you can't make it drink—are you sure you don't want something after all; what did the ice cube say to the water—I was water before it was cool*), directives to take a drink (e.g. *take your drink please*), toasts (e.g. *cheers*) and closings (e.g. *it was nice meeting you, enjoy your drink*). The utterances were based on the pre-defined script, but the selection of the specific greeting, offer or other utterance was determined by the dialog-wizard based on the situation. As the official language used at the international campus is English, during the lunchtime recordings analysed here, the robot's utterances were in English.

The operators' actions were restricted in ways that rendered the resulting interactions very similar to the performance limitations of current state-of-the-art robot technology; i.e. the navigation wizard steered the robot's movements only from his bird's eye view from the gallery making it very hard for him to see whether, for instance, someone had already taken a glass of water or who, in a group of people, was interacting with the robot. Similarly, the dialog wizard was only provided with audio feedback and had only very limited visual access to the scene. These limitations caused problems similar to those arising in autonomous robots. Furthermore, for the participants, the robot seemed autonomous; not a single participant indicated that they believed the robot was being operated by a person. Thus, from an emic perspective, people in this study interacted with the robot as if it were autonomous.

3.2 Participants

Our participants in this study were either students and staff from the university or visitors. People were informed about the presence of a robot and the video recordings with posters at the entrances to the area. At the end of the interactions, i.e. when the robot had moved away again, participants were asked to sign informed consent forms to grant permission for use of the video recordings for research purposes. However, not all participants consented to the publishing of the video footage.

3.3 Transcription and analytic methods

The analysis of the data is informed by the microanalytic methods of CA. As outlined in section 2 above, a basic and central principle in CA is the notion that human interaction is an orderly accomplishment that emerges through speakers' actions in talk as they take turns, closely monitor them, and produce appropriately fitted next turns that progress the talk or that repair trouble that temporarily stymies a turn's progress (Sidnell, 2012; Stivers & Robinson, 2006). These actions constitute a "next turn proof procedure" (Hutchby & Wooffitt, 1998, p. 17) and provide evidence of speakers' orientation to what Sacks (1984, p. 22) referred to as "order at all points". To facilitate analysis of these features, the process of transcription in CA is an interpretative process and is highly detailed to enable the capture of the multimodal design of turns, including how speakers position themselves to engage as speakers.

Verbatim transcriptions were conducted as a first step by the second author and a student assistant. The first author then conducted further detailed transcription to include prosody, nonverbal features and pauses. Courier New font 10 has been used for the verbal transcription and the notations are from Jefferson (1984); nonverbal features are glossed. The curly brackets (from Filipi, 2007) are used to capture the onset and offset of nonverbal features. The turns produced in German or Danish (which appear in bold) were idiomatically translated by the second author and a student assistant. The English translations appear in italics directly below their respective lines.

4. Analysis and Discussion

4.1 Summary of findings

The overall organisation (Gafaranga & Torras, 2001) of the episodes in each extract was monolingual. The speakers spoke to each other in Danish or German, and in English to the robot. Of the three episodes selected for analysis, only one (Extract 1) is characterised as achieving a positive outcome. It is assessed as positive on the grounds that it is a trouble-free, nonverbal acceptance of the robot's offer that culminates in a smooth closing in English

through an exchange of cheers. Also noteworthy, the sequence ended with a strong assessment in Danish (addressed to the human co-participants) that expressed wonderment at the robot's action. Trouble emerged in Extracts 2 and 3 because of the robot's actions subsequent to the human participants' response to the offer: a rejection that was not initially accepted in Extract 2 leading to frustration and confusion expressed in German, and an initial acceptance missed by the robot in Extract 3 with subsequent pursuits by the robot leading to ridicule in English. There was also a variation in volume in recipient design: lower in Danish and German in comparison to English (owing in part, but not entirely, to the proximity of the speaker who engaged with the robot in English).

Aligning with former research in multilingualism (e.g., Chen & Bonacina-Pugh, 2021; Filipi, in press; Filipi & Chuang, 2023), the findings show that the action of alternating to a language not shared by the robot, established a participation framework that excluded the robot, revealing a different orientation to the robot as participant. However, alternating to English could also be used to exclude the robot. In Extract 3, the human participants recipiently designed their turns in English, an action that invited participation of the robot through a shared language, but the robot's participation was used as an occasion for ridicule. The shared language thus shifted the moral order of the interaction and manifested a different orientation to the robot as participant even when the language was shared. Lending further weight to this claim, such bald actions of ridicule are not normally encountered in human-to-human interaction and would be deemed to be socially unacceptable behaviour.

4.2 Analysis and discussion

Analysis begins with Extract 1 which provides a comparative basis for describing the trouble that surfaced in the interactions in Extracts 2 and 3.

Extract 1

Extract 1 involves a group of three Danish speakers. Danish is used by two of the three human participants (HP1 and HP2) to confirm that there is water and to produce a positive assessment at the end of the sequence as they react to the robot's (R) actions, while the third participant, HP3, remains verbally uninvolved. English is used only in a single word turn between HP2 (the self-selected speaker) and R in the toasting exchange.

R approaches HP1, HP2, and HP3. All three are looking at R. HP1 is on the left, HP2 is in the middle and HP3 is on the right.

1 R: hi there,
 2 (0.8) ((all three look at R, HP2 steps closer to R,
 then HP3 smiles and looks at HP1 and HP2 then back at R))
 3 i wonder if you would like something to drink;
 (2.0) ((HP1 moves to check what is behind R))
 4 HP1: °mh::: (0.3) {**der er vand med**°}
there is water
 {((returns to the group but his back is to
 R and the camera))}
 5 HP2: {°**ja**°}
yes
 {((smiling))}
 6 HP1: °huh huh°
 7 (1.0) ((R turns away))
 8 HP2: huh huh [huh]
 9 R: [take] your drink please.
 10 6.0 ((HP2 takes the drink))
 11 R: cheers.
 12 HP2: {cheers.}
 {((she lifts her glass and smiles at R))}
 13 {**huh huh °kan den høre os**°}
can it hear us
 {((smiling, looks slightly up to the left in HP1's
 direction with her body still oriented to R)}
 15 HP1: ((also smiles and then starts to record with his phone))
 16 (4.2) ((HP2 drinks the water and R starts to move away))
 17 °**huh huh fantastisk. huh huh**°
fantastic
 ((indecipherable talk continues in Danish))

R initiates the interaction by producing a greeting in English. A gap ensues during which the attention of all three participants is mobilised through their gaze directed at R; HP2 and HP3 are smiling and looking at R. In line 3, R formulates the offer of water. None of the three human participants verbally engages immediately. Instead, HP1 moves behind R to establish that there is a source of water. He then returns to the group and reports in Danish, and in low volume, that there is in fact water. HP2 then produces a whispered minimal response through the token *ja* (yes), addressed to HP1, that is uttered with a slightly rising intonation while she visually continues to focus on R. This visual action projects her self-selection as speaker to interact with R. Next, HP1 laughs briefly, an action that invites and receives return laughter from HP2.

In terms of sequence structure this is a side-sequence that excludes R through its recipient design features—Danish and low volume. The turn design features establish a human alignment in which the talk is about R (its mechanical ability). Research on human group interaction with social robots shows how robots can exert an influence on the interactions (Oliveira et al., 2021), neatly displayed here through language choice and prosody. Together with the laughter, these actions set up the conditions for an assessment which is subsequently produced in line 17. HP2 then nonverbally self-selects to make herself available as the recipient

of the offer in response to R's request to take the water. This is a nonverbal acceptance of the offer. HP2 continues to gaze at R showing continued reciprocity, an action that also defers to R to complete the exchange. In fact, R then launches a third and socially appropriate concluding turn—the toast through *cheers*, which is returned by HP2. The sequence then continues in Danish through HP2's laughter, her second confirmation check uttered softly in line 13 (*{°kan den høre os_i° (can it hear us_i)}*) which is left unanswered, and an assessment (Pomerantz, 1984; Pomerantz & Heritage, 2013), (*fantastisk (fantastic)*, line 17) which expresses her wonderment. This second confirmation check attributes the possibility of hearing to R.

The overall organisation of the episode is monolingual (Gafaranga & Torras, 2001). It is in Danish in the interactions between HP1 and HP2 who talk about R and react to its ability to provide water, and it is in English between HP2 and R through the exchange of cheers. The two participation frameworks (Goodwin, 2007) (between the humans, and between the robot and human) are marked not only through language choice but also prosodically: they are conducted in normal volume with R in English, but in lower volume in Danish so that prosody becomes an additional resource in recipient design, adding yet another dimension to the ways that recipient design can be achieved. The changes in the participation framework are thus both manifested and realised multimodally in alignment with Mondada (2012)—through the different prosodic features that are aligned with the two languages, the embodied actions of the human speakers, and the shifts in language.

Extract 2

Two German-speaking students (HP1 and HP2) are seated and eating. HP1 is also drinking from a bottle as R approaches. Both HP1 and HP2 laugh briefly. HP1, who is closest to R, then turns to face R to self-select as the main speaker to interact with R.

```

1 R:   hi there.
2      (0.4)
3 HP1: {hello.}
        {(nods, smiles, puts his bottle down and looks down.))}
4      (1.0)
5 R:   {how about some water.
        {(HP1 looks at R)}
6 HP1: {no thanks.
        {(shakes his head, waves his hand at the same time)}}
7      (2.5) {(looks down and takes his sandwich, looks briefly
        at HP2 then back at R and {shakes his head)}}
8 R:                                     {most men actually do take
9      something to {drink.
10 HP1:                                {(lifts his bottle briefly to indicate

```

11 that he has one, (1.0) looks at R then ahead at
 someone off camera))

12 HP1: °huh was passiert hier°?
 what is happening here?

13 {°huh was \$soll ich jetzt {machen°?}
 what should i now do?

 {((starts to look down and towards HP2 on his right
 briefly, laughs and then looks down and eats)))}

14 HP2: {(°° °°)}
 {((shrugs, laughs briefly, looks away)))}
 {((HP1 waves his hand in a gesture of dismissal))}

15 (2.4) ((HP1 puts his sandwich down))

16 HP1: [({)}]
 {((waves his hand towards HP2)))}

17 R: [what did] the ice cube say to the water,
 18 (2.0) ((HP1 nods and eats, faces R))

19 HP1: °good°.

20 (0.4) ((HP1 looks behind him))

21 R: i was water before it was cool.
 ((in lines 22-25 HP1 and HP2 speak softly but inaudibly in
 German, and both laugh))

26 HP1: [{((()))]
 {((looking ahead, and speaking in German))}

27 R: [what did] the ice cube say to the water,
 ((HP1 and HP2 continue to converse softly and indecipherably in
 German, lines 28-30))

31 R: {i was water before it was cool.}
 {((moving away)))}

The extract opens with an exchange of greetings (lines 1–3) in English followed by R's offer of water and HP1's rejection (line 6). His rejection is constructed both verbally (*no thanks*) and nonverbally through the shaking of the head and waving of the hand in a dismissive gesture. The nonverbal action attributes to R the ability to see and understand the gesture. It is also produced directly without hesitation. HP1's subsequent actions of looking away, returning to his sandwich, and producing a headshake (line 7), display a completion, and therefore disengagement from further talk. However, R keeps the interaction open in lines 8–9 through a saying—*most men actually do take something to drink*—which projects an account or compliance as a possible response. In response, HP1 simply repeats his rejection which is expressed nonverbally through his action of holding up his bottle (lines 10–11). This action again displays an expectation that R can see and therefore understand that his action is a rejection (i.e. that he already has water).

At this point HP1 begins to display confusion and frustration. He does so by turning away from R and addressing a speaker off-camera in German. He initiates his turn by producing an equivocal laugh or 'quasi-laugh' (Lavin & Maynard, 2001) at the start of his turn and then produces a question about what

is going on and what he should do in lines 12 and 13 (*°huh was passiert hier? {huh was \$soll ich jetzt {machen°? (what is happening here? {what should I now do?}*)), conveyed in a smiley voice and in low volume. The position of the halting laughter at the beginning of the turn, and the prosodic quality of the utterance suggest frustration and confusion about how he should respond. Majlesi et al.'s (2023) recent study on human orientations to robots' violations in turn-taking similarly report speaker displays of frustration that are also marked by laughter. To be noted is that HP1 does not wait for a response to his questions, which are structured as recruitments of assistance (Kendrick & Drew, 2016). Instead, he turns to HP2, and through his laughter, elicits return laughter and produces a dismissive wave of his hand as he resumes eating his sandwich. These actions work as affiliative actions to possibly invite an agreement from HP2 that he is justified in feeling confused. They also exclude R's participation (an exclusion conveyed multimodally and through language choice) and project that HP1 is concluding his interaction with R.

In lines 15 and 16, HP1 stops eating but continues to address HP2 in German. R interrupts by initiating a joke in line 17 which indicates that for R, the interaction is ongoing. HP1 re-engages, by nodding and looking at R (line 18). He then produces a response to the joke: *good* in line 19, and then looks behind him. His subsequent embodied stance of turning away from R, and his whispered talk to HP2 in German, project his disengagement. Laughter is again produced but it is not possible to interpret what prompts the laughter and whether it is affiliative or in response to R's joke. We note, too, that HP1's assessment in line 19 is ignored by R who repeats the joke but starts to move away, and there is no formal closure as might be expected at this point.

R's repeated failure to accept the rejection of the offer of water has become the source of HP1's display of frustration and confusion (expressed through his appeal for help, his quasi-laugh, and prosodically within the turn (line 13)) about how to manage R's continued pursuit of an acceptance. These design features align with Glenn's (2013) analysis of participants in interviews who were shown to use laughter to manage delicate matters. Here the delicate matter was how often and in what other ways the rejection could be made salient to progress the talk to a conclusion.

The analysis has provided a display of HP1's orientation to the norms of turn-taking and the design rules governing an offer and acceptance or rejection. HP1's actions therefore attribute interactional competence to R. By producing his nonverbal actions together with his open verbal rejection of the offer, there are grounds to suggest that HP1 expects R to register or to 'see' that he already has a bottle of water. Therefore, R's pursuit is a breach of turn-taking norms. In other words, R's action of not accepting the rejection has violated the norms surrounding the need to display speaker understanding of the prior turn that comes through the production of a fitted, conditionally relevant response to it.

Turning to language choice, as in Extract 1, the overall organisation of the episode was monolingual. As a consequence of the displayed projections of interactional competence attributed to R not being met, HP1's language choice shaped two distinct participation frameworks manifested through the recipient design of his turns: in English with R, and in German to interact with HP2 by commenting and reacting. HP1 also interacted briefly in German with someone off-camera to seek clarification and understanding in German that treated R's pursuit as a cause of confusion. As in Extract 1, the talk in German was conducted in lower volume in comparison to the talk in English. Embodiment also marked the shifts in language as HP1 physically turned away from R both through his physical orientation and his gaze. In sum, the participation frameworks provided a way to locally manage the interactional repositioning of R as a result of the gap in R's interactional competence, and as in extract 1, they were manifested and realised multimodally.

Extract 3

In the next extract, analysis draws attention to the absence of a response to the acceptance of an offer of water that causes trouble as the speakers react to R's actions. As in Majlesi et al.'s (2023) study, and Extract 2, there is an orientation by the human participants to the violation of turn-taking norms. This time however, the speakers treat the lack of response from R as an occasion for open and direct ridicule by issuing an unrealistic request and providing a negative formulation of R, both actions that invite complicit laughter directed at R. This is a breach of socially acceptable behaviour that is fundamental to maintaining and achieving sociality (Silverman, 1997).

A group of five (HP1, HP2, HP3, HP4 and HP5) is seated having lunch and speaking to each other in Danish. HP1, HP2 and HP3 each address R.

```

1 HP1  {°det er nemlig rigtig [(          )°]}
        it is actually correct
        {(turns towards R, but is speaking to his companions))}
2 R:                                     [sorry to bother you]
3      (0.3) but can i offer you some water,
4 HP1: {yes.}
        {(moves closer to R))}
5      (0.3)
6      tanks. (0.2) {huh huh huh}
        {(turns to the others, who also laugh))}
7 HP1: {(turns back to R, still laughing))
8      (0.6) ((R turns away))
9 R:   take your drink [please.
10      [(all laughing))]
11      (0.4) ((HP1 takes the drink))

```

12 HP2: {can i have some whiskey.
 {((arms folded behind his head, looking at R and smiling,
 and moves closer to it)}, HP4 and HP5 are talking in
 Danish to each other))

13 (2.2)

14 R: {research shows that it is important to drink enough
 {((HP2 turns back to the others))

15 water during the day.

16 (1.0)

17 HP2: {can i have a rum and cola.
 {((looks at R))

18 (6.2) ((HP1 and HP2 are facing R, HP4 and HP5 are talking
 inaudibly to each other in Danish while HP3 is not visible,
 R turns back again))

19 R: you can lead a horse to water {but you can't make it}
 {((HP2 looks at HP3 who
 has moved closer)))}

20 drink.

21 HP3: ()

22 R: {are you sure you don't want something after all.}

23 {((HP2 is looking at R)))}

24 HP3: are you {retarded,}

25 {((HP2 and HP3 laugh, HP3 moves back, HP2 looks
 at R)))}

26 (0.8)

27 ((HP4 and HP5 still talking to each other inaudibly in
 Danish, HP4 stands up))

28 HP4: **hvad laver den**)
what is it doing

29 HP2: **så blev den fornærmet.**
now it is offended

30 HP1: **{nu tager den tilløb til.}**
now it is preparing to take off
 {((looks down at R who is starting to move off, and then
 at HP2)))}

31 HP4: **hvad er det?**
what is this

32 R: have a lovely day.

33 HP1: {huh f*** off.}
 {((looking at HP2))

((they speak to each other in Danish, HP4 leaves))

The episode starts off with an interruption by R's overlapped offer (lines 1 and 2). However, it is prefaced with an apology for the interruption. HP1 turns, moves closer to R and self-selects as first speaker because of his proximity to R. He starts by verbally accepting the offer (*yes*) and then follows up with a deliberate mispronunciation of thanks (*tanks*). It is accompanied by his laughter and gaze at the group, actions launched to generate a shared laughter (line 6) and that set a derisive tone for the interaction with R. As an action, this sets up an in-group stance (Oliveira et al., 2021) for how R will be treated. He subsequently turns

back to face R who issues the directive to take the drink in line 9 (received with laughter from the group in line 10). HP1 takes the proffered glass of water (line 11). Two members of the group (HP4 and HP5) then disengage from the interaction with R and HP4 only re-enters towards the end.

Unlike the above extracts, where only one speaker engages, two speakers (HP2 and HP3) also address R. HP2 issues two unrealistic requests for an alcoholic drink (*whiskey* in line 12 and *rum and cola* in line 17) that continue to build the derisive tone and in-group stance. They are launched to tease, and neither receives a contingent response from R. Instead, in lines 14–15 R continues the initial request by issuing a statement of fact (*research shows that it is important ...*) and a saying in lines 19–20 (*you can lead a horse to water but you can't make it drink.*). R then re-issues the request in line 22 (*are you sure you don't want something after all*), produced as a follow-up confirming question. R's turns here each treat HP1's action as a rejection of the offer of water so that R has clearly missed HP1's physical up-take of the glass of water (line 11) preceded by his verbal acceptance (*yes*) in line 4. R's actions also draw attention to the missing closure that would be appropriate and to the absent response from R to HP1, which is further augmented by HP3's negative formulation of R (*are you retarded?*) in line 24 which is an escalation of the teasing that was launched by HP2 in lines 12 and 17 when he requested alcohol. These actions suggest that the human participants are orienting to the missing responses from R as a violation. Furthermore, the shared laughter is produced in the context of humour at the expense of R. It is launched to mock and carries a negative assessment (Glenn, 2003) of R's actions. As in Extract 2, the ridicule could be construed as a reaction to the gaps in R's interactional competence. Thus, while R is included in the participation framework through their shared language, English, it is as a 'victim' or target of ridicule.

At this point HP4 stands up to leave, and HP1 and HP4 observe and comment on R's mechanical actions in Danish as R begins to move away. In line 29, HP2 also attributes the human quality of being offended to R, made relevant by HP3's negative assessment, with which HP1 appears to agree through his negatively constructed *f*** off* in response to R's leave-taking in line 32. It is not clear whether this is addressed to R however, as HP1's gaze is directed at HP2 not at R and it is followed immediately by inaudible talk in Danish.

The outcome in Extract 3 is the absence of a socially appropriate conclusion to the initial offer made by R. This has led to trouble evidenced through the speakers' actions produced in response to a lack of uptake by R (responding to HP1 and to HP2) which is dealt with through jest and ridicule by the human participants. There are three monolingual participation frameworks within the group as well which are all managed through language choice. The first is between HP1, HP2 and HP3 who engage together with R in English but in jest so that English, their shared language with R, becomes a tool for mocking R; the second is between HP4 and HP5 who are part of the group only initially and

speak in Danish exclusively and in lower volume, an action that excludes R but one that also defines their lack of participation with R; and the third is between HP1, HP2 and HP4 who comment on R's actions in Danish at the end in lower volume, which also excludes R. As in Extract 2, the human speakers' actions here display an orientation to a missing set of responses (to acknowledge the initial acceptance of R's offer and to respond to HP2's counter-offers) that an offer makes relevant, even when made in jest.

5. General Discussion and Conclusions

The purpose of this study was twofold; one was to examine how multilingual Danish and English or German and English speakers interacted with each other and with the robot in an offer sequence and managed problems that arose through their language choice; and the other was to examine what their language choice might reveal about their orientation to the robot as a co-participant.

Analysis revealed that the interactions with the robot were limited to engaging in short, paired turns, delimited to a large degree by the sequence organisation of the offer sequence: an opening turn, offer/acceptance or offer/rejection, and a follow-up third turn and/or post expansion sequence shaped by whether the offer was accepted or not. The response to the offer was always conducted in English, recipiently designed for the robot who was programmed to produce English. When the offer was accepted, as in Extract 1, the result was a smooth outcome, displayed by the production of a positive assessment in Danish, and therefore addressed to the human participants, to express wonderment at the robot's ability to provide water. When the offer was rejected or the initial acceptance missed, both actions that led to trouble, the robot conducted a pursuit in English (through a saying, a joke or a repetition of the offer). Trouble surfaced as a display of confusion and frustration in German addressed to a co-participant or member of the research team (Extract 2), or derision in English in which more than one human participant took part (Extract 3). As in Extract 1, alternation to Danish also occurred in Extract 3 to comment on the robot's movements, which amounted to talk about the robot.

In determining the scheme (Gafaranga & Torras, 2001) or bilingual grid of interpretation, the human speakers' language alternation practices occurred in monolingual mode in a language shared by the speaker(s) and resulted in monolingual participation frameworks and a monolingual overall organisation: they were entirely in Danish or German between the human participants and entirely in English with the robot. This finding aligns with previous research in university contexts where multilingual students used their languages in highly organised, monolingual ways to conduct different aspects of assigned learning tasks in each of their languages (e.g., Chen & Bonacina-Pugh, 2021; Kunitz, 2018; Reichert & Liebscher, 2018).

In turning to the second research focus, by using language choice and participation frameworks as lenses, the study has uncovered three issues pertinent to human participants' orientations to the robot as being different from their orientation to humans. First the human participants attributed human abilities to the robot: the ability to hear (Extract 1), to see (Extract 2), and to feel emotion (e.g. being offended) (Extract 3) as well as interactional competence (Extracts 2 and 3). Physical and emotional attributions have been previously reported (e.g., Fischer, 2016, 2023; Rudaz et al., 2023). However, by locating the study in language choice, the study extends these findings in two important ways. The first relates to the monolingual organisation of the human participants' interactions with the robot and with their peers where German or Danish was used to create in-group alignment (Mondada, 2004) that established a negotiated stance (Oliveira et al., 2021) for how the robot was treated. In conducting these actions in the language (German and Danish) or in routines that the robot was not programmed in (counter requests, negative formulations) that exposed gaps in interactional competence, the robot was oriented to as something to be talked about, to wonder at, to express frustration about and to use as an object of ridicule. These interactions created not only different participation frameworks but also differences in the participation that violated norms of behaviour.

Second, the attention to the interactional properties in the design of the turns in the languages uncovered a variation in volume associated with the languages; turns in German and Danish were produced in lower volume in comparison with English. While such a feature could have been accounted for by the physical positioning of the speakers selected to speak to the robot (i.e. their proximity to the camera and recording device), this did not explain the actions of all speakers. In multilingual interaction, differences in volume have been identified in studies by Amir (2013) in an English as a Foreign Language learning context and by Filipi (in press) and Filipi and Chuang (2023) in a Higher Education mainstream (i.e. not language learning) classroom context. Filipi and Chuang (2023) reported that in alternating to Mandarin to deal with an interactional matter relevant to their lack of understanding of task requirements, international students used a lower volume when speaking Mandarin. However, the students displayed an awareness that excluding speakers from their interactions in a language not shared by the other speakers at their table could be construed as being impolite and explicitly accounted for their action.

Accounting for behaviour (that might be construed as being impolite and therefore that violates the norms of social behaviour that are pivotal to human sociality) did not occur in the interactions in this study. In the displays of bald ridicule in Extract 3, the human speakers oriented to the robot in a different way from human-to-human interaction. This finding contrasts with overwhelming reports from laboratory studies that robots are oriented to in ways similar to human participants (e.g., Kahn et al., 2015; Rudaz et al., 2023). The examination of the organisation of language choice in a field trial where participants do not

interact with a robot in controlled one-on-one encounters but are free to choose whether to interact with it and how so, has brought to the surface a different set of human orientations to the robot (cf. also Bu et al., 2025). The selection of context together with the research approach adopted, points to the value and need for further studies on human speaker orientations to robots in natural contexts to capture differences in the ways in which they participate with them and what might give rise to the interactional differences that surface.

Third, speaker frustration to the robot's response to their rejection of the offer of water or the escalation into ridicule when the initial offer and acceptance were missed, extends the very recent work on breaches to turn-taking in HRI as reported by Majlesi et al. (2023) who similarly noted speaker frustration. This finding suggests that counter requests that might lie outside a routine managed by the robot or the ways in which rejection is accomplished or responded to, are not easily predicted in formulaic ways in the creation of scripts. When designing robots capable of interacting with humans, the management of rejection (including rejection produced nonverbally) and contingencies that arise, call for the need for greater attention to be paid to sequences of talk and their possible trajectories as they occur in naturally occurring interaction. For example, greater attention to how repair is accomplished both generally and in multilingual interaction to resolve matters of misunderstanding or mishearing is one area that needs to be a focus in the creation of scripts. In particular, understanding the participation status regarding how the robot is treated as co-participant, rather than as an object to be talked about, is crucial to determine whether or not, and how, the robot can interrupt an activity in the course of the interaction (Cao et al., 2025). However, the challenge of creating robots capable of interacting in natural and spontaneous ways remains. This raises the question about whether robots that can manage turns in the uncontrolled and spontaneous unfolding locus of interaction can ever be created.

Acknowledgments

The field study reported on was carried out in the framework of the Smooth Project, funded by the Danish Innovation Foundation, which we gratefully acknowledge. Furthermore, data elicitation was supported by several colleagues: Selina Eisenberger, Lotte Damsgaard Nissen, Matous Jelinek, Oskar Palinko and Eduardo Ruiz Ramirez. We also thank the anonymous reviewers for their constructive and insightful comments on earlier versions of this manuscript.

References

- Amir, A. (2013). Self-policing in English as a foreign language classroom. *Novitas-ROYAL (Research on Youth and Language)*, 7(2), 841–05.
- Auer, P. (1998). Introduction: Bilingual conversation revisited. In P. Auer (Ed.), *Code-switching in conversation* (pp. 1–28). Routledge. <https://doi.org/10.4324/9780203017883>
- Bartneck, C., Belpaeme, T., Eyssel, F., Kanda, T., Keijsers, M., & Šabanović, S. (2020). *Human-robot interaction: An introduction*. Cambridge University Press. <https://doi.org/10.1017/9781108676649>
- Bonacina-Pugh, F. (2020). Legitimizing multilingual practices in the classroom: The role of the ‘practiced language policy’. *International Journal of Bilingual Education and Bilingualism*, 23(4), 434–448. <https://doi.org/10.1080/13670050.2017.1372359>
- Bu, F., Fischer, K., & Ju, W. (2025). Making sense of robots in public spaces: A study of trash barrel robots. *Transactions in Human-Robot Interactions. ACM Transactions on Human-Robot Interaction*, 14(4), 1–21. <https://doi.org/10.1145/3731252>
- Cao, S., Moon, J., Mahmood, A., Antony, V. N., Xiao, Z., Liu, A., & Huang, C. M. (2025). Interruption handling for conversational robots. *arXiv preprint arXiv:2501.01568*.
- Chen, Q., & Bonacina-Pugh, F. (2021). Spotlights on ‘practiced’ language policy in the internationalised university. In D. Dippold & M. Heron (Eds.), *Meaningful teaching interaction at the internationalised university* (pp. 110–122). Routledge. <https://doi.org/10.4324/9780429329692-9>
- Clark, H. H., & Fischer, K. (2023). Social robots as depictions of social agents. *Behavioral and Brain Sciences*, 46, e21. <https://doi.org/10.1017/S0140525X22000668>
- Cromdal, J. (2005). Bilingual order in collaborative word processing: On creating an English text in Swedish. *Journal of Pragmatics*, 37, 329–53. <https://doi.org/10.1016/j.pragma.2004.10.006>
- Del Duchetto, F., Baxter, P., & Hanheide, M. (2019). Lindsey the tour guide robot-usage patterns in a museum long-term deployment. *28th IEEE international conference on robot and human interactive communication (RO-MAN)* (pp. 1–8). IEEE. <https://doi.org/10.1109/RO-MAN46459.2019.8956329>

- Dobrosovestnova, A., Babel, F., & Pelikan, H. (2025). Beyond the user: Mapping subject positions for robots in public spaces. In *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 163–173).
<https://doi.org/10.1109/HRI61500.2025.10974177>
- Filipi, A. (2007). A toddler's treatment of mm and mm hm in talk with a parent. *Australian Review of Applied Linguistics*, 30(3), 1–17.
- Filipi, A. (2015). The development of recipient design in bilingual child-parent interaction. *Research on Language and Social Interaction*, 48(1), 100–119. <https://doi.org/10.1080/08351813.2015.993858>
- Filipi, A. (2018). Making teacher talk comprehensible through language alternation practices. In A. Filipi & N. Markee (Eds.), *Conversation analysis and language alternation: Capturing transitions in the classroom* (pp. 183–202). John Benjamins Publishing Company.
<https://doi.org/10.1075/pbns.295.10fli>.
- Filipi, A. (in press). Chinese international students' language choice as a display of assessment anxiety in a higher education context in the Anglosphere. *Higher Education*. <https://doi.org/10.1007/s10734-025-01519-8>
- Filipi, A., & Markee, N. (Eds.). (2018). *Conversation analysis and language alternation: Capturing transitions in the classroom*. John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.295>
- Filipi, A., & Chuang, M.-S. K. (2023). Chinese whispers: International Chinese students' language practices in an anglophone Higher Education context. *Classroom Discourse*, 14(3), 238–257.
<https://doi.org/10.1080/19463014.2022.2072353>
- Fischer, K. (2011). Interpersonal variation in understanding robots as social actors. In *Proceedings of HRI'11*, March 6–9th, 2011. Lausanne, Switzerland (pp. 53–60). <https://doi.org/10.1145/1957656.1957672>
- Fischer, K. (2016). *Designing speech for a recipient: The roles of partner modeling, alignment and feedback in so-called 'simplified registers'*. John Benjamins Publishing Company.
- Fischer, K. (2020). Why collaborative robots must be social (and even emotional) actors. *Techné: Research in Technology and Philosophy*, 23(3), 270–289. <https://doi.org/10.5840/techne20191120104>
- Fischer, K. (2021). Tracking anthropomorphizing behavior in human-robot Interaction. *ACM Transactions on Human-Robot Interaction*, 11, 1, Article 4 (March 2022), 28 pages. <https://doi.org/10.1145/3442677>

- Fischer, K. (2023). Defining interaction as coordination benefits both HRI research and robot development: Entering service interactions. In *32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (pp. 213–219). IEEE.
<https://doi.org/10.1109/RO-MAN57019.2023.10309642>
- Fischer, K., Jung, M., Jensen, L. C., & aus der Wieschen, M. V. (2019). Emotional expression by robots: When and why. In *Proceedings of the International Conference on Human-Robot Interaction*, Daegu, Korea (pp. 29–38). <https://doi.org/10.1109/HRI.2019.8673078>
- Gafaranga, J. (1999). Language choice as a significant aspect of talk organisation. The orderliness of language alternation. *Text*, 19, 201–225.
<https://doi.org/10.1515/text.1.1999.19.2.201>
- Gafaranga, J. (2005). Demythologising language alternation studies: Conversational structure vs. social structure in bilingual interaction. *Journal of Pragmatics*, 37, 281–300.
<https://doi.org/10.1016/j.pragma.2004.10.002>
- Gafaranga, J. (2018). Overall order versus local order in bilingual conversation: A conversation analytic perspective on language alternation. In A. Filipi & N. Markee (Eds.), *Conversation analysis and language alternation: Capturing transitions in the classroom* (pp. 35–60). John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.295.03gaf>
- Gafaranga, J., & Torras, M. C. (2001). Language versus medium in the study of bilingual conversation. *International Journal of Bilingualism*, 5(2), 195–219.
<https://doi.org/10.1177/13670069010050020401>
- Glenn, P. (2003). *Laughter in interaction*. Cambridge University Press.
- Glenn, P. (2013). Interviewees volunteered laughter in employment interviews: A case of “nervous” laughter? In P. Glenn & E. Holt (Eds.), *Studies of laughter in interaction* (pp. 255–276). Bloomsbury.
<https://doi.org/10.5040/9781472542069.ch-013>
- Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological inquiry*, 50(3–4), 272–302.
<https://doi.org/10.1111/j.1475-682X.1980.tb00023.x>
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse & Society*, 18(1), 53–73.
<https://doi.org/10.1177/0957926507069457>
- Groom, V., Srinivasan, V., Bethel, C. L., Murphy, R., Dole, L., & Nass, C. (2011). Responses to robot social roles and social role framing. 2011

International Conference on Collaboration Technologies and Systems (CTS) (pp. 194–203). IEEE. <https://doi.org/10.1109/CTS.2011.5928687>

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Polity Press.

Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Blackwell Publishers Inc.

Jefferson, G. (1984). Transcription notation. In J. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. ix–xvi). Cambridge University Press. <https://doi.org/10.1017/CBO9780511665868.002>

Jung, M., & Hinds, P. (2018). Robots in the wild: A time for more robust theories of human-robot interaction. *ACM Transactions on Human-Robot Interaction (THRI)*, 7(1), 1–5. <https://doi.org/10.1145/3208975>

Kahn Jr, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Gary, H. E., & Ruckert, J. H. (2015). Will people keep the secret of a humanoid robot? Psychological intimacy in HRI. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction* (pp. 173–180). <https://doi.org/10.1145/2696454.2696486>

Kendrick, K. H., & Drew, P. (2016). Recruitment: Offers, requests, and the organization of assistance in interaction. *Research on Language and Social Interaction*, 49(1), 1–19. <https://doi.org/10.1080/08351813.2016.1126436>

Krüger, N., Fischer, K., Manoonpong, P., Palinko, O., Bodenhagen, L., Baumann, T., & Dalgaard, L. (2021). The smooth-robot: A modular, interactive service robot. *Frontiers in Robotics and AI*, 294. <https://doi.org/10.3389/frobt.2021.645639>

Kunitz, S. (2018). L1/L2 alternation practices in students' task planning. In A. Filipi & N. Markee (Eds.), *Conversation analysis and language alternation: Capturing transitions in the classroom* (pp. 107–28). John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.295.06kun>

Lavin, D., & Maynard, D. W. (2001). Standardization vs. rapport: Respondent laughter and interviewer reaction during telephone survey. *American Sociological Review*, 66, 453–479. <https://doi.org/10.2307/3088888>

Lee, H. R., Cheon, E., Lim, C., & Fischer, K. (2022). Configuring humans: What roles humans play in HRI research. *17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 478–492). IEEE. <https://doi.org/10.1109/HRI53351.2022.9889496>

- Li, W. (2002). "What do you want me to say?" On the conversation analysis approach to bilingual interaction. *Language in Society*, 31, 159–180. <https://doi.org/10.1017/S0047404501020140>
- Licoppe C., & Rollet N. (2020). "Je dois y aller". Analyses de séquences de clôtures entre humains et robot. *Réseaux*, 220–221(2–3), 151–193. <https://doi.org/10.3917/res.220.0151>
- Lombard, M., & Xu, K. (2021). Social responses to media technologies in the 21st century: The media are social actors paradigm. *Human-Machine Communication*, 2, 29–55. <https://doi.org/10.30658/hmc.2.2>
- Majlesi, A. R., Cumbal, R., Engwall, O., Gillet, S., Kunitz, S., Lymer, G., Norrby, C., & Tuncer, S. (2023). Managing turn-taking in human-robot interactions: The case of projections and overlaps, and the anticipation of turn design by human participants. *Social Interaction. Video-Based Studies of Human Sociality*, 6(1). <https://doi.org/10.7146/si.v6i1.137380>
- Mizumaru, K., Satake, S., Kanda, T., & Ono, T. (2019). Stop doing it! Approaching strategy for a robot to admonish pedestrians. *14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 449–457). IEEE. <https://doi.org/10.1109/HRI.2019.8673017>
- Mlynář, J., de Rijk, L., & Liesenfeld, A. (2024). AI in situated action: A scoping review of ethnomethodological and conversation analytic studies. *AI & Soc.* <https://doi.org/10.1007/s00146-024-01919-x>
- Mondada, L. (2004). Ways of 'doing being a plurilingual' in international work meetings. In R. Gardner & J. Wagner (Eds.), *Second language conversations* (pp. 18–39). Continuum.
- Mondada, L. (2012). The dynamics of embodied participation and language choice in multilingual meetings. *Language in Society*, 41(2), 213–235. <https://doi.org/10.1017/S004740451200005X>
- Mortensen, K., & Hazel, S. (2014). Moving into interaction—Social practices for initiating encounters at a help desk. *Journal of Pragmatics*, 62, 46–67. <https://doi.org/10.1016/j.pragma.2013.11.009>
- Mutlu, B., & Forlizzi, J. (2008). Robots in organizations: The role of workflow, social, and environmental factors in human-robot interaction. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (pp. 287–294). <https://doi.org/10.1145/1349822.1349860>

- Oliveira, R., Arriga, P., & Paiva, A. (2021). Human-robot interaction in groups: Methodological and research practices. *Multimodal Technologies and Interaction*, 5(10), 59. <https://doi.org/10.3390/mti5100059>
- Pelikan, H. R. M., & Broth, M. (2016). Why that nao? How humans adapt to a conventional humanoid robot in taking turns-at-talk. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 4921–4932). <https://doi.org/10.1145/2858036.2858478>
- Pelikan, H. R. M., Broth, M., & Keevallik, L. (2020). ‘Are you sad, Cozmo?’: How humans make sense of a home robot’s emotion displays. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction* (pp. 461–470). <https://doi.org/10.1145/3319502.3374814>
- Pelikan, H., Broth, M., & Keevallik, L. (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>
- Pitsch, K., Kuzuoka, H., Suzuki, Y., Sussenbach, L., Luff, P., & Heath, C. (2009). “The first five seconds”: Contingent stepwise entry into an interaction as a means to secure sustained engagement in HRI. *RO-MAN 2009—The 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 985–991). IEEE. <https://doi.org/10.1109/ROMAN.2009.5326167>
- Pomerantz, A. M. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 57–101). Cambridge University Press.
- Pomerantz, A. M., & Heritage, J. (2013). Preference. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 210–228). Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch11>
- Psathas, G. (1995). *Conversation analysis: The study of talk-in-interaction*. Sage.
- Reichert, T., & Liebscher, G. (2018). Transitions with “okay”: Managing language alternation in role-play preparations. In A. Filipi & N. Markee (Eds.), *Conversation analysis and language alternation: Capturing transitions in the classroom* (pp. 129–148). John Benjamins Publishing Company. <https://benjamins.com/catalog/pbns.295.07rei>
- Riek, L. D. (2012). Wizard of oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1), 119–136. <https://doi.org/10.5898/JHRI.1.1.Riek>

- Rudaz, D., Tatarián, K., Stower, R., & Licoppe, C. (2023). From inanimate object to agent: Impact of pre-beginnings on the emergence of greetings with a robot. *ACM Transactions on Human-Robot Interaction*, 12(3), 1–31. <https://doi.org/10.1145/3575806>
- Sacks, H. (1984). Notes on methodology. In J. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 21–27). Cambridge University Press.
- Sacks, H. (1992). *Lectures on conversation* (G. Jefferson, Ed., Vols. 1 & 2). Basil Blackwell.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511791208>
- Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8(4), 289–327. <https://doi.org/10.1515/semi.1973.8.4.289>
- Sidnell, J. (2012). Basic conversation analytic methods. In J. Sidnell & T. Stivers (Eds.) *The handbook of conversation analysis* (pp. 77–99). Wiley Blackwell. <https://doi.org/10.1002/9781118325001.ch5>
- Silverman, D. (1997). *Discourses of counselling: HIV counselling as social interaction*. Sage.
- Skårup, T. (2004). Brokering and membership in a multilingual community of practice. In R. Gardner & J. Wagner (Eds.), *Second language conversations* (pp. 40–57). Continuum.
- Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in talk. *Language in Society*, 35(3), 367–92. <https://doi.org/10.1017/S0047404506060179>
- Vöge, M. (2011). Employing multilingualism for doing identity work and generating laughter in business meetings: A case study. In G. Pallotti & J. Wagner (Eds.), *L2 learning as social practice: Conversation-analytic perspectives* (pp. 237–264). University of Hawai'i, National Foreign Language Resource Center.
- Wong, J., & Waring, H. Z. (Eds.). (2021). *Storytelling in multilingual interaction: A conversation analysis perspective*. Routledge. <https://doi.org/10.1080/02188791.2023.2257082>