



OPEN ACCESS

EDITED BY

Cordula Schwarze,
University of Marburg, Germany

REVIEWED BY

Tim Hector,
University of Siegen, Germany
Laurenz Kornfeld,
Heidelberg University, Germany

*CORRESPONDENCE

Carolin Dix
✉ carolin.dix@uibk.ac.at

RECEIVED 08 March 2025

ACCEPTED 15 May 2025

PUBLISHED 23 July 2025

CITATION

Dix C (2025) 'Now it tells you to press
continue': multimodal strategies for
navigating between machine-human and
human-human instructions in Thermomix®
experience meetings.
Front. Commun. 10:1589894.
doi: 10.3389/fcomm.2025.1589894

COPYRIGHT

© 2025 Dix. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

'Now it tells you to press continue': multimodal strategies for navigating between machine-human and human-human instructions in Thermomix® experience meetings

Carolin Dix*

University of Innsbruck, Faculty of Language, Literature and Culture, Department of German
Philology, Innsbruck, Austria

This paper focuses on the interactive multimodal management and orchestration of human-to-human instructions and machine-to-human instructions when learning how to use the smart cooking aid Thermomix® during cooking experience meetings. These peer-to-peer events, hosted in private homes, combine elements of cooking workshops, product presentations, and sales events. Unlike traditional cooking classes, cooking experience meetings involve a human (the representative) and a machine which both give instructions to the participants. The study focuses on how Thermomix® representatives as experts spatially and interactively position themselves as intermediary between the machine and the users and at the same time as authoritative and legitimate instructors. The examples show how they coordinate their instructions with the digital guidance provided by the machine, elaborate these instructions or add new ones. Using video data from two cooking experience meetings, the analysis shows that while the Thermomix® provides step-by-step directives, the representative plays a crucial role in explaining, evaluating, and adapting the instructions to ensure participants understand the cooking process. The study contributes to research on human-machine interaction by illustrating how instructional authority is distributed between human experts and smart technology in interactive learning environments.

KEYWORDS

instruction, cooking, Thermomix®, multimodality, human-machine-interaction, space

1 Introduction

Giving and following instructions is a complex social action that has been described in various social settings, including car driving lessons (De Stefani and Gazin, 2014; Björklund-Flärd, 2024; Schubert, 2024), theatre rehearsals (Schmidt and Deppermann, 2023), orchestra rehearsals (Messner, 2023), therapy sessions (Ortner, 2023), and horse-riding lessons (Szczepek Reed, 2023, 2024), as well as in the use of technical devices, such as smartphones and smart speakers (Albert and Hall, 2024). This paper contributes to this body of research by analyzing instructional interaction within the context of so called *Erlebniskochen* ("Cooking Experience," hereafter: CE-meetings), during which participants learn how to use the smart kitchen appliance Thermomix® (hereafter: TM) and prepare various dishes. During these peer-to-peer events, an experienced TM user (the representative or expert, hereafter: TM-R), invited by a host into a private home, shares knowledge about the food processor and provides instructions

to a group of interested prospective users (the participants). These meetings are a key component of the company's sales strategy.¹ So, the events have the goal to (a) introduce the food processor, (b) give participants the opportunity to experience firsthand the capabilities of the TM by preparing a multi-course meal (e.g., bread rolls, spreads, boiled vegetables), and (c) offer the option to buy the food processor at the end of the event. Thus, CE-meetings combine different genres of social interaction: cooking workshop, product presentation and sales event.

Learning how to cook involves mastering the handling of objects (e.g., ingredients such as vegetables, meat, cheese) and tools (e.g., knives, pots, mixer) in the right way and in the right order (Mondada, 2014; Lilja and Piirainen-Marsh, 2022). Additionally, learners need to acquire cooking-specific vocabulary in order to interpret and follow recipes effectively and successfully prepare meals. The process of cooking often involves following instructions, which may be presented in multiple forms: written recipes with or without pictures (Sanchez-Stockhammer, 2021), online (tutorial) videos, and, of course, instructions given by an experienced individual (e.g., a professional chef, see Mondada, 2014). These instructions guide the cooking process by specifying its sequential order (telling what actions and tasks to perform in what order) and explaining how to carry out different preparation techniques. Since the CE-meetings are not classical cooking classes, basic elements are different. First, this affects knowledge about preparation techniques such as cutting and stirring as they are not performed by the cook anymore, but by the food processor. Consequently, specific instruction topics such as how to handle kitchen utensils, how to chop ingredients, and how to recognize different textures (see Mondada, 2014, p. 201), are no longer relevant. Second, not only does the TM-R provide instructions on how to prepare certain dishes using the machine, but the TM itself becomes an instructional agent giving directives to the users. This process of instructing people is known as “guided cooking.” That is, the machine directs the user step by step through the cooking process (Graf, 2023, p. 54). Thus, both the TM-R and the food processor serve as instruction giving agents. However, they use different interactional resources: while the TM-R gives verbal and visual-embodied instructions, the Thermomix® provides written instructions and images. As a result, the learners must attend to these two different instructional agents simultaneously. Therefore, the cooking events appear to be instructional interactions not only about cooking but also on how to follow instructions and interact with (smart) machines.

As the analysis will show, this setting challenges the interactive role of the TM-R in multiple ways: she (all representatives in the data are women) has to display her expertise in using the TM while teaching participants how to operate the machine independently. So, the TM-R needs to stay in her position as instructor while displaying, teaching and supporting the participant's autonomy. While each participant of the CE-meetings engages with the machine individually, the TM-R observes their actions, corrects and redirects them, explains the directives provided by the machine, and formulates new instructions. Therefore, she must position herself (both interactively

and spatially) as an intermediary between the machine and the learners.

The primary aim of this paper is to explore the multimodal strategies the TM-R uses to legitimize her role as instructor, considering the step-by-step instructions provided by the cooking machine. To address this, the paper asks the following questions:

1. How do TM-Rs spatially and interactively display and account for their role as instructor? What interactive resources do they use?
2. How do the TM-Rs coordinate the multimodal instructions of the Thermomix® (text, pictures, iconic and deictic signs) and their own multimodal instructions (verbal and embodied)?

The analysis is based on 6 h of video data from two cooking experience meetings recorded in Germany and follows the theoretical and methodological approaches of Multimodal Interaction Analysis (Mondada, 2016) and Interactional Linguistics (Couper-Kuhlen and Selting, 2018).

In the following section, this paper provides a brief overview of research on instructions in general and on dealing with technical devices and tools in particular. Section 3 then presents the data and introduces the methodological and theoretical frame the analysis is based on. The analysis and the results are presented in section 4. Section 5 finally summarizes the results.

2 Humans, machines, and instructions in interaction

The analysis of instructions and instructed actions has a long-standing research tradition within the fields of EMCA and Multimodal Interaction Analysis (Lynch and Lindwall, 2024). Research in this area has developed along four major strands of interest. The first one focuses on how humans teach and instruct other humans in co-present face-to-face settings either on how to use and move their own body (see, for example, Szczepek Reed, 2023, 2024 on horse-riding lessons; Schmidt and Deppermann, 2021, 2023 on theatre rehearsals; Smart and Szczepek Reed, 2025; Krug, 2025 on dance lessons) or on how to use tools, objects and technologies (see De Stefani and Gazin, 2014; Helmer and Reineke, 2021 on car driving lessons; Messner, 2023 on orchestra rehearsals; Mondada, 2014; Lilja and Piirainen-Marsh, 2022 on cooking and Lindwall and Ekström, 2012 on crocheting).

The second research strand examines the verbal, vocal, and visual-embodied forms used when humans instruct non-human animals, such as horses (Szczepek Reed, 2025).

Thirdly, research addresses how humans interact with and instruct (smart) machines, like Alexa and Siri (Habscheid et al., 2025; Albert and Hall, 2024; Pelikan et al., 2022). Among other aspects, researchers focus on the interactive construction and ascription of agency. That is, how humans discuss, attribute or reject agency toward smart machines. Pelikan et al. (2022) identified five different “forms of robotic agency” (Pelikan et al., 2022, p. 20). A prominent one they call “hybrid agency” which they define as a form “in which a human and a robot form a collaborative unit that consists of the human launching relevant actions and the robot performing them through its own body” (Pelikan et al., 2022, p. 21). Following this, research focuses on how

¹ <https://www.vorwerk.com/gb/en/c/home/product-demonstration/Thermomix> (latest access: 28.08.2024).

humans design and adjust instructions when interacting with smart machines such as smart speakers or voice assistant technologies within mundane and institutional settings (see [Albert and Hall, 2024](#) for home care interaction and [Habscheid et al., 2025](#) for private contexts).

The fourth strand of research explores interactive contexts in which instructions are provided to humans by technical devices, smart machines, and robots. This perspective is still relatively rare and mainly focuses on mobile contexts involving GPS systems, which have “the ability to display maps and [...] provide turn-by-turn instructions through visual and audio guidance” ([Brown and Laurier, 2012](#), p. 1621). This shift of roles, where the instructor is a machine rather than a human, leads to variations and changes in how instructions are given and followed ([Lynch and Lindwall, 2024](#), p. 7). In its simplest and unproblematic structure, “the GPS gives an instruction for a road maneuver, which the driver then follows” ([Brown and Laurier, 2012](#), p. 1623). Nevertheless, the driver has to manage the actions necessary to drive safely while simultaneously monitoring the map of the GPS, aligning it with the actual road in the field of vision, and listening to the verbal instruction ([Brown and Laurier, 2012](#), p. 1623). Moreover, drivers have to decide whether to follow, ignore, or modify the instruction and perform the instructed action accordingly ([Brown and Laurier, 2012](#), p. 1621). This paper follows this line of interest by investigating how humans deal with the instructions provided by the smart kitchen machine Thermomix® displayed on its digital interface.

Research in teaching and learning contexts, both in private and institutional settings, describe instructions as “instructions-in-education” ([Lindwall et al., 2015](#), p. 145), that is, as complex, multimodally achieved, and coordinated social actions designed to teach “people (how) to do things, (how) to do things with things and (how) to do things with their body” ([Stukenbrock, 2014](#), p. 80). They identify two main participants: the teacher, who is positioned as expert with the epistemic and deontic right to give instructions (the instructor), and the unexperienced learner (the instructee), who is willing to learn and follows the instructions ([Lindwall et al., 2015](#), p. 145; [Lilja and Piirainen-Marsh, 2022](#), p. 2). The purpose of instructions in educational settings is to facilitate the transition of the instructee “from a state where they do not know to a state where they do know” ([Lindwall et al., 2015](#), p. 145) and enable them to carry out the learned actions independently and without further supervision. Prototypically, instruction sequences consist of three parts: the instruction given by the instructor, the instructed action performed by the instructee, and the evaluation realized by the instructor ([Lindwall and Mondada, 2025](#), p. 12). Instructions then often appear in clusters, so called instructional chains ([De Stefani and Gazin, 2014](#), p. 71), where several instruction sequences are performed one after the other, splitting one larger task into smaller sections ([Lindwall and Mondada, 2025](#)). The instructor as well as the instructee can then use multiple verbal, vocal and bodily-visual resources to accomplish the instruction sequences ([Stukenbrock, 2014](#); [Evans and Lindwall, 2020](#); [Lilja and Piirainen-Marsh, 2022](#); [Schubert, 2024](#); [Lindwall and Mondada, 2025](#)).

Beside instructions-in-education, instructions can also be initiated as instructions-as-directives, which structure and coordinate social interaction but “are not specifically tied to the business of teaching” ([Lindwall et al., 2015](#), p. 145). Even though they “make a complying second action conditionally relevant” ([Lilja and Piirainen-Marsh, 2022](#), p. 2) and “aim to bring about a future action” ([Lilja and Piirainen-Marsh, 2022](#), p. 3), they are not primarily oriented

to change the epistemic status of the instructee ([Lindwall et al., 2015](#), p. 145).

Both forms, instructional and directive, can be combined and intertwined, and even instructions with an educational purpose can become directive in nature during an event (see [Lindwall and Ekström, 2012](#), p. 28). Within the setting of CE-Meetings, educational and directive instructions not only occur together but are often initiated by different instructing agents. While the TM provides instructions-as-directives, the TM-R initiates both instructions-as-directives and instructions-in-education, being oriented to the participants as learners in using the TM.

3 Data and method

The analysis is based on video recordings of two cooking-experience meetings, each of them lasting at least 3 h. The participants consist of one invited expert (the TM-R), the host that invited the expert and the other participants into the private home, and three invited people interested in the TM. In the two meetings analysed in the following sections, all participating people are women. Except one, all are German native speakers. Consequently, the language of instruction is German.

The first meeting took place in 2018 and was recorded with a 360° camera and one steady cam. The four participants use the representative's Thermomix®, which is the model TM5. The second meeting took place in 2021, this time using the TM6 model. This group used two machines, as not only the TM-R brings her machine, but the host already owned a TM and provided it for the CE-meeting (see [Table 1](#)). Both meetings had a similar structure, in line with the company's specifications:²

1. The TM-R welcomes the participants and introduces herself, explains the goals for the meeting and the meeting procedure.
2. The participants introduce themselves (by name) by explaining their individual cooking experiences (e.g., for how many people they regularly prepare food) and telling what is important for them when it comes to cooking (e.g., healthy food, quick cooking, delicious dishes, etc.). According to their self-descriptions, four participants are highly experienced, regularly preparing meals for more than two people. Three participants are moderately experienced, typically cooking for one or two people, and one participant is a beginner, cooking only for herself and not on a regular basis.
3. The TM-R introduces and explains the TM: which parts the food processor consists of, how the machine works, what cooking functions it has (e.g., caramelizing, steaming, emulsifying, kneading, cooking, grinding, mixing, stirring, heating, blending,

² As one of the representatives explained in a private conversation, representatives receive a “table display” when they start organising CE-meetings. This provides a rough timetable and contains all the topics that the customer should learn about. It is desirable that the representatives use this, but everyone should add their own touch.

TABLE 1 Overview data material.

Year	Participants	Recording	TM model	Duration
2018	1 host 1 TM-	360° camera 1 steady cam	TM5	3 h
2021	representative 3 active participants cooking	3 steady cams	TM6	3 h

weighing and chopping³) und which kind of dishes one can prepare with the help of the TM. Furthermore, the TM-R shows how to start the TM, how to scroll through the display, how to choose a recipe and how to start the process of “guided cooking.”

4. The participants start to cook and prepare different kinds of dishes by using the TM.
5. The participants eat the prepared meal together, talk about their experience while cooking and discuss possibilities and limitations of the TM. The TM-R mentions different purchasing options.

The analysis is data driven and theoretically and methodologically based on the principles of Multimodal Interaction Analysis and Interactional Linguistics. The transcripts follow the GAT-2 conventions for the verbal and vocal conduct (Selting et al., 2009) and use the inventory of the International SignWriting Alphabet (ISWA, Parkhurst and Parkhurst 2008) for visual and bodily resources as proposed by Dix (2023). The signs used in the transcripts are explained in the [Supplementary material \(Data Sheet 1\)](#).

4 Analysis

The analysis focuses on the transition of (instructional) space between step three (introducing the TM) and step four (cooking process), as well as on the instructive sequences within step four. During this part of the CE-meeting, the participants follow both the instructions of the TM and the hints, instructions and advice of the TM-R. Although all participants have prior cooking experience and are therefore not complete novices when it comes to cooking in general, only one participant in the 2021 data (the host) is familiar with the Thermomix® and already owns one. In this context, the TM-R assumes the role of the experienced practitioner and is positioned as the expert.

4.1 Cooking ensemble and instruction space

At the beginning of the CE-meetings (see part 3 of this paper for more detail), both groups show a similar spatial arrangement, building

a prototypical platform configuration (Kendon, 1990; Goffman, 1983) which is characterized by “someone [...] performs as a spectacle for others to watch whether on an elevated platform or encircled by the group of watchers” (Scollon and Scollon, 2003, p. 62). Here, the TM-R stands in front of the participants, positioned next to the TM, while the participants are oriented toward the TM-R, either standing in a slight circular arrangement (Figure 1, CE-meeting 2018) or sitting in a line (Figure 2, CE-meeting 2021), watching her performance on introducing herself, providing an overview of the machine, and outlining the procedure of the CE-meeting. Consequently, the TM-R is the focus person.

Nevertheless, the position of the TM next to the TM-R already refers to the fact that the TM will play a key role in the further course of the meeting and become the major focus point of the participants. In part 3 of the CE-meetings (introduction of the TM) then, the focus slightly shifts toward the TM, embodied through a re-arrangement of the F-formation (Kendon, 1990): Both the participants and the TM-R are now oriented toward the TM, forming a semi-circular arrangement around the machine, with the TM-R positioned closest to the food processor, showing and demonstrating, e.g., how to scroll through the display (see Figure 3).

The participants closely observe the TM-R, monitoring her movements and handling of the machine. Although the procedure of starting the TM, scrolling through the menu, accessing the recipe library, and selecting a recipe are fundamental steps when using the machine, both TM-Rs perform these actions themselves as demonstrations, accompanied by verbal explanations. Following this demonstration, the TM-R invites the participants to reposition themselves by stepping forward toward the machine, encouraging them to interact with the TM independently. This pivotal moment in the CE-meeting is illustrated by the following example from the 2021 data. It shows how the TM-R initiates the shift of interactional tasks (for the participants from watching her operate to dealing with the machine themselves and for her from operating the machine to monitoring the actions of the participants) and coordinates the transition of interactional space (for the representative: from being the focus person to becoming the guiding instructor; for the participants: building an instruction ensemble with the machine).

The transcript captures the visual movements of the arms (A), the hands and fingers (F), the head (H), the legs (L), the whole body (M) and the bodily-spatial orientation with the TM marking the “front” orientation (O). The participants are the Thermomix® representative (TM-R) and two participants of the CE-meeting (P2 and P3). The participants use one food processor each simultaneously so that the TM-R has to deal with two cooking ensembles (Example 1).

The example illustrates the transition of interactional spaces and the projection of future action, especially the participant’s following the instructions of the machine. After the TM-R has finished her introduction of the machine (not in the transcript), she verbally terminates the sequence with the particles “genau” (*okay*, l.001) and “also” (*so*, l.002; see Oloff, 2017 for a more detailed description of “genau”). Furthermore, she shifts her bodily orientation from facing

the machines (, l.001) to facing the participants by turning right (, l.002) after the “also” (*so*). She then addresses the participants verbally by formulating the directive “IHR fangt jetzt gleich beide mal mit an” (*You start right now both of you*, l.003–004) requesting them to now actively engage with and experience the cooking machine. This

³ <https://www.vorwerk.com/at/de/c/home/produkte/thermomix/tm6> (latest access: 19.08.2024).



FIGURE 1
TM_2018_participants (video still by the author).



FIGURE 2
TM_2021_participants (video still by the author).

verbal instruction is accompanied by visual cues: She changes her body orientation toward the participants (⊗, 1.003) and points to P2 and P3 using manual deictic gestures (↑, ↓, ↗, 1.003).

Afterwards, the TM-R steps back (↖, 1.004) and moves aside, allowing the participants to step forward, approach the machines and position themselves in front of a TM each. However, the participants maintain their initial position as listeners and observers during the request for participation. P3 remains seated with her legs crossed left

over right (×, 1.001), her fingers folded (✂, 1.001) and placed on her left knee (↵, 1.001). Only after the TM-R has finished her verbal request she acknowledges the request verbally (hm_mh, 1.005) and physically by uncrossing her legs moving her right leg to the right (→, 1.005) so that both legs are parallel. She opens her hands and moves them backward (↓, 1.005), placing them on her thighs (↵, 1.1000). With these actions she prepares to stand up, but she only



FIGURE 3
TM_2018_demonstration (video still by the author).

performs the movement ($\uparrow$, 1.006) and puts her hands on her lower back to rearrange her trousers and shirt ($\downarrow$, 1.006), after the TM-R elaborates her directive by specifying that she does not care which two of the three participants start using the machine (1.006). She continues with another directive that each of the chosen two participants should use one TM on their own (*jeder einen thermomix, each of you one thermomix*, 1.007), accompanied by a deictic manual gesture pointing toward the machines and establishing the space the participants should occupy ($\rightarrow$, 1.007). P3 then moves forward ($\uparrow$, 1.008) toward one of the machines. This is when also P2 stands up ($\uparrow$, 1.008) and moves forward ($\uparrow$, 1.008). She acknowledges the request verbally with a change of state token (*aha, oh*, 1.009). While P3 lets her arms hang when stepping forward, P2 displays her expectation of having to use her hands in the upcoming part of the CE-meeting and her readiness to be instructed: she bents her arms ($\downarrow$, 1.009) and holds her hands in an open palm up position in front of her body ($\uparrow$, 1.009). She holds this position while the TM-R formulates another instruction to press the green button on the TM's display to start the cooking process (*und ihr drückt auf kochvorgang starten, and you press start cooking*, 1.010, see [Supplementary material image 1](#)).

P3 immediately follows the instruction by pressing the green button on the TM display with her left index finger ($\downarrow$, 1.010). P2, on the other hand, continues with further preparation activities (such as putting on her glasses, not in the transcript) before she also touches the button on the machine's screen. Therefore, P2 and P3 collaboratively construct the role of the TM-R as leading and timing the structure of the event. Simultaneously they position themselves as inexperienced TM-users by waiting until the TM-R has formulated

the directive of starting the process. By touching the green continue-button on the screen, the participants start the process of guided cooking and therefore also initiate and establish the machine-to-human instruction chain. This line of sequences then becomes a constitutive part of the human-to-human instruction chain. Within the CE-meetings these two chains are intertwined, they are mutually dependent on each other and cause negotiation and sometimes troubles as the following examples will demonstrate in more detail.

Although the TM-R has previously explained and demonstrated how to navigate the machine's display (not in the transcript), the participants are more oriented toward the TM-R than to the TM at this point of the meeting. They only interact with the machine and touch the display to start the cooking process after the TM-R has initiated and instructed it. Thus, they attribute more instructional weight to the verbal instruction of the TM-R than to the machine's written directive. In both CE-meetings, this slightly shifts during the progression of the meeting, particularly as the participants become more experienced in using the machine.

As a result of the transition activities, each participant adopts a face-to-machine-formation, positioning themselves directly in front of the machine. This newly established cooking ensemble remains stable throughout the participant's engagement with the TM. Here, the participants display their interpretation of the TM as a single-person-machine and the mode of guided cooking as a process where instructions are directed to the cook alone. At the same time, the CE-meetings are a social encounter with co-present other participants and a mutually accepted expert. These conditions create the need for additional verbal and bodily actions, as following instructions becomes a collaborative activity. The TM-R displays her understanding of these conditions by positioning herself spatially in relation to this newly established cooking ensemble. When monitoring only one participant (as in the 2018 data), she takes over an over-shoulder-formation ([Figure 4](#)) and sometimes a side-by-side formation. When dealing with two cooking ensembles (as in the 2021 data), the TM-R steps behind and between the two participants. In both arrangements,

the TM-R watches over the shoulders of the participants how they (try to) follow the instructions of the machine and sometimes actively intervenes in the cooking process.

With this spatial arrangement, the TM-R accomplishes several tasks: (a) She opens the floor for the cooking ensemble and enables that the participants can read and follow the instructions of the TM; (b) She is able to read the instructions of the TM written on the display herself; (c) She holds herself available for support, explanations and further instructions offstage, claiming the right to interrupt and modify the instructed actions, as the following example will show.

4.2 Managing instructions and the cooking process

Although the TM-R withdraws from her position as focus person and steps behind or at the side of the cooking ensemble, she remains

a substantial instruction giving entity. As described for other contexts, the TM-R as human instructor takes on the task of “monitoring whether the novice has adequately completed one step and is ready to move to the next” (Lindwall and Mondada, 2025, p. 13). However, the TM-R not only monitors whether the participants have successfully completed her instructions but also whether they have completed the instruction the machine has given to them, performed the instructed action adequately, and are ready to move on to the next step within the process of guided cooking.

The next extract shows the participants working on preparing bread rolls. P2 and P3 each stand in front of a TM. They first select the recipe from the recipe library. The whole cooking process is segmented into 21 smaller steps, which are displayed successively by the TM. Each step is displayed on the machine’s screen. After having started the cooking process, the participants are instructed to preheat the oven (step 1), line the backing tray with paper (step 2), put cheese in the pot (step 3), place the lid on the pot (step 4), activate the machine to grate

001 TMR: geNAU;
 right

TMR	F	
	H	
	O	
P3	F	
	L	
S		#1

002 °h ALso (-)
 so

TMR	F	
	O	

003 Ihr fangt jetzt GLEICH (---);
 you start right now

TMR	F			
	O			
S		#2	#3 #4	

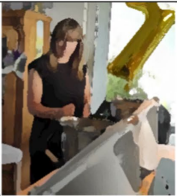
EXAMPLE 1 (Continued)

Frontiers in Communication

07

frontiersin.org

004 beIde ma mit AN,
 both of you

TMR	H	
	O	
	M	
S		#5
		

005 P3: hm mh,

P3	F	
	L	
S		#6
		

006 TMR: oder (.) also mir ist egAl WER,
 or (.) so I don't care who

P3	F	
	M	
S		#7
		

007 jEder einen thermoMIX?
 each of you one Thermomix

TMR	F	
	O	
S		#8
		

EXAMPLE 1 (Continued)

008 P3: mhm

P3	M	↑...
P2	M	↑↑.↑.
S		# 9
		

009 P2: aHA,
oh

P2	A	↓ ↓
	F	✗ ✗
S		# 10
		

010 TMR: und ihr drückt auf kOchvorgang STARTen.
and you press start cooking

P3	F	 *
----	---	---

EXAMPLE 1

Start cooking (TM_2021_P2.005_Kochvorgang starten). Link video: <http://f.kanjo.de/pduOt>.

the cheese (step 5) and transfer the shredded cheese into a bowl (step 6). The next step is then to fill in oat flakes into the pot so that the machine can grind them. On the screen of the TM, the participants see the instruction (see [Supplementary material image 2](#)).

The machine directs the cook by means of a two-part written instruction: first there is the action to be performed (in den Mixtopf geben/*add to the pot*) and then the ingredient and the quantity (100 g Haferflocken, kernig/*100 g oat flakes, kernel*). On the left side of the screen, there is an image of the pot and a digital scale showing the users how much they have already put into the pot. The extract starts at the end of executing the instructed action. The participants are the Thermomix® representative (TM-R) and two participants of the CE-meeting (P2 and P3) ([Example 2](#)).

While P3 has already finished filling in the oats, P2 still performs the instructed action. At the same time, the TM-R monitors her and the TM's screen with the scale and comments on P2's execution of the instructed action with formulating the possibility to add a little bit more oats (kannst ruhig noch ein MÜH?! *you can feel free to add a bit more*, l.01) by using the modal verb “kannst”/ (you) can (see [Deppermann and Gubina, 2021](#) for more details). P2 interprets TM-R's hint as a directive and performs the new instructed action by adding another spoonful of oat flakes. The TM-R evaluates and

confirms the action (genau siehste super/ *exactly look perfect*, l.02–04). This prototypical instruction sequence between the human instructor (TM-R) and the instructee (P2) contains three parts ([Lindwall and Mondada, 2025](#)): an instruction (l.01), the instructed action, and the evaluation (l.02–04). Unlike other human-to-human instructions, this one is embedded into the machine-to-human instruction sequence and challenges the weight of the machine's directives and therefore the TM as instructional agent. The participants and the TM-R implicitly negotiate TM's instruction by dealing with the question how rigid one has to follow them and how wide the scope of deviation can be. The TM-R accounts for her instruction by explaining that slightly more or less oat flakes will not change the recipe and finally the dough (wenn da ein gramm mehr oder weniger machts das net/ *if there is one gram more or less it does not matter*, l.05–06). P2 confirms the explanation (nicht so schlimm/*not so bad*, l.08) and thus gives more weight to TM-R's instruction than to TM's instruction and positions the TM-R in the role of the leading instructor.

Even though the participants become more and more experienced in dealing with the TM (either through self-experiencing the machine or through observing the other participants cooking), they frequently orchestrate the step-by-step guided cooking with the explanations and instructions of the TM-R. The example shows that P3 not only waits



FIGURE 4
TM_2018_instruction (video still by the author).

to continue her cooking process until P2 has completed the two instructed actions (filling in oat flakes and adding a bit more) but even until the TM-R formulates the directive to press the “continue” button on the machine’s screen to proceed to the next step preparing the bread roles (so dann macht mal auf weiter/*so then press continue*, l.09).

The example furthermore shows that the TM users normally have to decide on their own whether they have successfully executed the current processing step and can proceed to the next one by pressing the “continue” button so that the TM displays the next instruction. Here, the machine-to-human instruction sequence differs from the human-to-human instructions where the TM-R as expert and instructor comments on and evaluates the instructed action. This makes clear that the TM-R not only positions herself spatially as an intermediary between the machine and the participants but also deontically and epistemically by commenting on the quality of the performed action, evaluating the progress and setting the moment when to move on to the next step.

Finally, the participants follow the directive to continue by tapping on the screen and touching the “continue” button – P2 with her right index finger (👉, l.09) and P3 with her left index finger (👈, l.09). The

sound of the TM (🔊, l.09) indicates that this action has been executed. The display screen changes and shows the next step within the machine-to-human instruction chain, which is to close the pot by replacing the lid (see [Supplementary material image 3](#), “Messbecher in den Mixtopfdeckel einsetzen/*insert the measuring cup into the mixing bowl lid*”). As the example shows, this directive and the way it

is written and visualized on the screen causes trouble and makes it relevant for the TM-R to formulate additional instructions and perform supporting actions.

The TM-R builds on her previous explanations what the measuring cup and the mixing bowl lid are and on the fact that the participants have already performed this action in step 4 displayed by her using the adverb “wieder”/ “*again*” (jetzt sagt er euch wieder den deckel drauf/*now it tells you again the lid on it*, l.10–11). Thus, she does not simply verbalize the instruction by reading out loud what is written on the screen but adds additional information about the participant’s epistemic status, attributing a higher epistemic stance based on the recurring action. P3 then performs the instructed action independently and even presses “continue” on her own (l.18) while the TM-R is involved with P2. P2, on the other hand, struggles with the lid and the measuring cup. The TM-R verifies that P2 already holds the lid and the measuring cup in her hands (l.13–16). She then formulates another verbal instruction that P2 only has to place the cup and the lid on the pot (so tust es drauf/*this is how you put it on*, l.17). Not until P2 still displays difficulties in executing the instructed action, the TM-R moves forward into the instruction space of P2’s cooking ensemble. She touches the lid with the fingers of her left hand (👉, l.18) and positively evaluates P2’s action (genau/*right*, l.19). Here, the TM-R assists with initiating a short, embodied repair sequence and helping P2 executing the instructed action. She finally evaluates this sequence (das ist alles gut/*that’s all right*, l.21), removes her hand

(👇, l.20) and steps back into her spatial position behind the participants. With this she leaves the cooking ensemble.

Although P3 has already pressed the “continue” button and therefore decided that she has successfully executed the instructed action, the TM-R addresses her directive to both participants (*jetzt macht ihr wieder weiter/now you go on with continue again*, l.22). P2 follows the instruction by touching the green “continue” button on the

TM’s screen with her left index finger (l.22). The sound of the TM again indicates that this action was performed, the display then shows the next step.

In this extract, the TM-R is monitoring the performed instructed actions of the participants and adds directives that are on the one hand educational in that they teach the participants not only how to execute the instructed action but also when the instructed action is completed. On the other hand, her directives structure the event in that they drive the CE-meeting forward. Consequently, the participants form a (temporary) instruction triangle which causes a more complex sequential order of the single instruction sequences and the whole coordination of the

two instruction chains (the steps in brackets are optional, but common):

- a. Start cooking process by tapping on the button.
- b. TM shows instruction on the display.
- c. Participants read the instruction (*aloud*).
- d. (*Participants negotiate TM’s instruction, request the TM-R to interpret the instruction*)
- e. (*TM-R gives new instruction*)
- f. Participants perform the instructed action.
- g. (*TM-R assists, corrects the instructed action*)
- h. (*Participants request the evaluation of the instructed action*)
- i. (*TM-R evaluates the instructed action and instructs to tap the button “continue”*)
- j. Participants tap the button “continue.”
- k. TM shows the next instruction on the display.
- l. Steps b to j are repeated until the last step of the cooking process was executed.

01 TM-R: kannst ruhig noch ein MÜH?
feel free to add a bit more
 02 geNAU,
exactly
 03 SIEHste; (--)
look
 04 SUpEr;
perfect
 05 wenn da ein gramm mEhr oder WEIniger;
if there's one gram more or less
 06 macht das NET;
it doesn't matter
 07 °hh
 08 P2: nicht so SCHLIMM;
not so bad
 09 TM-R: so (.) dAnn macht mal auf WEIter, (-----)
so then press continue



P3	F	
TM	Sd	
P2	F	
TM	SD	

10 jEtzte (.) sagt_er euch WIEder,
now it tells you again
 11 den [dEckel DRAUF,
the lid on it
 12 P2: [MESSbecher,
measuring cup
 13 TM-R: nEe des i_des is der MESSbescher;=
no this is this is the measuring cup
 14 =und der DEckel;
and the lid
 15 des PASST;
that fits
 16 geNAU;
right
 17 so tust_es DRAUF;
this is how you put it on

EXAMPLE 2 (Continued)

18

geNAU, (-----)
right

TMR	F'	
P3	F'	
TM	Sd	
S	#10	
		

19

(da PASsen;)
that fits

20

geNAU;
right

TMR	F'	
S	#11	
		

21

das is Alles GUT;
that's all right

22

jetzt macht ihr wieder WEIter, (--)
now you go on with continue again

P2	F'	
TM	Sd	

EXAMPLE 2

One gram more (TM_2021_P1.001_Ein Gramm mehr). Link video: <http://f.kanjo.de/lqJHp>.

This again shows that and how the human-to-human and the machine-to-human instruction sequences are intertwined, that the human-to-human instructions depend on the machine-to-human instructions and that both instructional agents (the TM-R and the TM) induce and shape the actions of the participants. As the example already demonstrated, this leads to troubles, causes negotiation sequences and affects the way the TM-R initiates instructions.

4.3 Instruction troubles

The next example illustrates that the instruction format of the TM-R further changes during the preparation process. The

participants move on in preparing the bread rolls and are now instructed by the machine to add half of the shredded cheese (see step 3–6, part 4.2 of this paper). The machine’s directive (Zugeben Hälfte des zerkleinerten Käses/add half of the shredded cheese) causes some trouble since it is not clear enough for the participants how to measure the weight of the cheese. Here, the TM-R comes in as an expert on using the machine and directing the execution of the instructed action (Example 3).

After completing the previous step, P2 sighs and comments on the CE-meeting and using the TM (l.01–02), before she independently

presses the “continue” button (, l.02). The TM then displays the next instruction on the screen, which P2 partly reads aloud (l.4). At

this point, the TM-R modifies the machine's instruction by suggesting that the participants add more cheese than the machine specifies. She justifies this by sharing her knowledge about the further course of the cooking process and her expertise on the whole recipe, namely that they need some cheese as topping for the bread rolls but not that much (l.07).

P2 then displays that she struggles with the machine's instruction because the machine does not indicate the weight of the already added cheese as it did in a previous step (see [Example 2](#), l.15–17 and 32–36). Furthermore, this causes troubles with the instruction of the TM-R. Nevertheless, she performs the instructed action but seeks confirmation from the TM-R regarding the adequacy of her actions, specifically asking whether she has added enough cheese (l.19). TM-R positively evaluates this (l.20–21). Subsequently, the host offers tips on how to effectively execute the instructed action without the TM measuring the weight. So, she proposes to remember how much cheese one has filled in the machine in step 3 of the recipe and calculate the half or three quarter of that individually (l.23–30). While the host and P2 further discuss this point and highlight the deficits of the TM at this point (*ich dachte das zeigt wieder an/I thought that it displays again*), the TM-R steps back and observes the actions of P2. Finally, P3 independently continues the cooking process by pressing the “continue” button (l.38), and P2 also goes further on, overlapping her action of touching the screen with the TM-R's instruction (l.37).

The extract illustrates how the participants, the TM-R, and the host negotiate the TM instructions by adding, requesting, and refining instructions as well as incorporating their own cooking knowledge to discuss and, if necessary, modify the instructions provided by the machine. These negotiations occur throughout the entire CE-meeting but more frequently when the participants have acquired a certain amount of knowledge about how to operate the machine. In that respect, the TM-R initiates educational instructions and strengthens her instructional role, during which she may physically reposition herself to become an active part in executing the machine's instructions.

5 Conclusion

The paper's aim was to shed light on the interactive accomplishment and management of following human-to-human instructions and machine-to-human instructions in the interactive context of Thermomix® Cooking-Experience meetings. The focus was especially on how the TM-R displays and claims her role as expert and as instructor. The analysis shows that she uses multiple multimodal resources such as their bodily-spatial orientation, manual gestures, other visual-bodily resources and verbal resources to hold herself available for the participants and become an active part of the cooking ensemble if necessary.

The examples showed that the process of receiving and following instructions by a human instructor as well as a machine instructor while using the Thermomix® as cooking tool appears as a socially organized and locally accomplished joined project of the participants of Cooking Experience meetings. The TM-R displays her understanding of her role as invited expert in that she structures and organizes the CE-meeting, demonstrates her own expertise in cooking

as well as in handling the food processor and holds herself available and relevant by providing educational as well as directive instructions during the event. Although she withdraws her position as focus person and positions herself spatially at the side or behind the cooking ensemble, she constantly observes the actions of the participants and has the right to step into the instruction space between the participants and the machine. Furthermore, she manipulates the machine and assists in executing the instructed action manually and provides additional verbal instructions. The participants in turn ratify the representative and her role as instructing agent in that they give her the permission to enter the cooking ensemble, making it an instruction triangle, prioritize her instructions over the instructions of the machine and align the cooking process with her directives.

During the CE-meetings, the participants learn how to interact with the smart machine and develop practices that Graf described for experienced users: the user “is given a choice of recipes coded by software engineers to work with this machine but improves or adapts it according to her own embodied knowledge and the preferences or needs of each family member for any given meal” (Graf, 2023, p. 55–56). This observation is in line with what [Brown and Laurier \(2012\)](#) describe for the use of GPS systems while driving a car (see section 2 of this paper). Although the device offers a step-by-step guidance and formulates directives, each time a new instruction appears, the users have to decide if and how they follow it and perform the instructed action. The same is true for using the TM. Additionally, the manufacturers proclaim a guarantee of success if the users strictly follow the instructions of the machine. This makes it even more relevant for the TM-R to display and maintain her role and for the participants to ascribe more weight on either the TM-R's or the TM's instructions. The examples show that this is a constant process throughout the CE-meeting and that the instructions of the TM-R are not automatically assigned with more authority than the instructions of the TM. In this respect, the TM-R as expert is made relevant by herself but also by the participants, regarding knowledge about the machine as well as cooking knowledge beyond using smart machines. While the participants do not ask on how to hold a knife or cut vegetables, the discussion is on the following questions: Do we have to follow the instructions of the TM (right now)? In which way exactly do we have to follow the instructions of the TM? How rigid do we have to follow the instructions of the TM (and still have a guarantee of success)? How is the scope of deviation when using the machine and following the instructions? What kind of ingredients do we have to use exactly?

Within the setting of the CE-meetings, the participants not only deal with different instruction giving agents (the TM-R and the TM) but also with different instruction formats and instruction purposes: The analysis showed that the instructions of the TM do not have an educational purpose that is, they do not aim to transform the epistemic status of the users from un-known to known when it comes to cooking. Moreover, the TM instructions are “designed to get someone else to do something” (Goodwin, 2006, p. 517) that is, they aim to structure, guide and coordinate the process of cooking. They are step-by-step directives. The instructions of the TM-R on the contrary also have an educational character. Although both TM-Rs initiate coordination instructions (e.g., directing the participants to press the “continue” button), they are more and more becoming

involved in interpreting the TM's instructions and give educational instructions that help the participants to (prospectively) use the machine on their own.

As using smart technology and AI in the kitchen and during cooking becomes more and more part of the daily lives of domestic cooks (Graf, 2023; Jaber et al., 2024), there is still relatively little known on how these technologies affect interactions when cooking together. This paper offered a glance into this field of research by

analysing a specific teaching and learning context with the help of interaction analytic methodology. With this it contributes to research on human-machine interaction and collaboration interests in social interaction. Further research is thus necessary to get more insights in how smart digital technologies shape social interaction (in domestic kitchens). This includes research on the use of smartphones, tablets, online videos in general, and all sorts of kitchen robots.

01 P2: HACH,
((P2 signs))
02 AUFregend. (--)
Exciting

P2	F		
TM	Sd		



03 P3: ((lacht))
((laughing))
04 P2: die HÄL[fte des zerkleinerten käse;
the half of the chopped cheese
05 TMR: [so (-) die HÄLfte d-
okay the half of
06 da könnt_er ruhig en müh MEHR drauf machen;
you can put a little bit more in it
07 weil (--) wir müssen dann noch en bisschen oben DRAUF
because we must add a little bit on top
machen,
08 also mÄcht ruhig ein dreiVIERtel;
so don't hesitate to add three-quarter
09 das LANGT;
that's enough
10 äh des is oKAY; (1.5)
äh that's okay
11 P2: äh::=
äh::
12 TMR: =da steht NICHTS da;
there is nothing there
13 [die HÄLfte (.)] genau;
half of it right
14 P2: [ACH d-
oh
15 ich wollt grad SAgen;=
I just wanted to say
16 =das wär SCHÖN wenn des jetzt,
it would be nice if it now
17 aber (-) GUT;
but okay
18 TMR: nee; (---)da da:s;=
no there that
19 P2: so?
like this
20 TMR: geNAU;
exactly
21 genau (.) des is oKAY;
right that's okay

EXAMPLE 3 (Continued)

22 P2: hm_mh,
 23 Hst: ja da KANN man ja sagen;=
okay one could say
 24 =man macht hundertfünfzich gramm Käse;
one takes 150 gram cheese
 25 we_ma SAGT (-) ne?
if one says right
 26 dreivIERtel;
three quarter
 27 TMR: ge[NAU (.) ja,
right yes
 28 P3: [so;
okay
 29 Hst: muss man (xx) sich MERken;=
one must remember
 30 =wie viel käse [man gemacht HAT;
how much cheese one prepared
 31 P2: [NE:-
no
 32 ich dacht das zeigt wieder AN;
I thought that it displays again
 33 Hst: ja [ne;;
yes no
 34 P2: [ne,
well
 35 aber das zeigt_s in DEM fall;
but in this case

P3	F	
TM	Sd	

36 jetzt mal NICH an; (1.5)
it doesn't display anything
 37 TMR: so dann [WEIter,
okay then continue
 38 P2: [WEIter,
continue

P2	F	
TM	Sd	

EXAMPLE 3

Half of the cheese (TM_2021_P1.001_Hälfte Käse). Link video: <http://f.kanjo.de/cqylA>.

Data availability statement

The datasets presented in this article are not readily available because of the data protection regulations regarding the consent forms of the study participants. Requests to access the datasets should be directed to CD, carolin.dix@uibk.ac.at.

Ethics statement

Ethical approval was not required for the studies involving humans as the studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this

study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

CD: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported

by the University of Innsbruck, Open Access Publication Fund and University of Innsbruck, Faculty of Language, Literature and Culture.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author declares that no Gen AI was used in the creation of this manuscript.

References

- Albert, S., and Hall, L. (2024). Distributed agency in smart homecare interactions: a conversation analytic case study. *Discourse Commun.* 18, 892–904. doi: 10.1177/17504813241267059
- Björklund-Flärd, D. (2024). Seeing together: the organization of looking and seeing in navigational driving instructions. *Lang. Commun.* 98, 45–59. doi: 10.1016/j.langcom.2024.06.003
- Brown, B., and Laurier, E. (2012). “The normal natural troubles of driving with GPS,” in Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY: ACM. 1621–1630.
- Couper-Kuhlen, E., and Selting, M. (2018). Interactional linguistics. Studying language in social interaction. Cambridge: Cambridge University Press.
- De Stefani, E., and Gazin, A.-D. (2014). Instructional sequences in driving lessons: Mobile participants and the temporal and sequential organization of actions. *J. Pragmat.* 65, 63–79. doi: 10.1016/j.pragma.2013.08.020
- Deppermann, A., and Gubina, A. (2021). Positionally-sensitive action-ascription. Uses of Kannst du X? ‘Can you X?’ In their sequential and multimodal context. *Int. Ling.* 1, 183–215. doi: 10.1075/il.21005.dep
- Dix, C. (2023). Transcribing facial gestures: combining Jefferson with the international SignWriting alphabet (ISWA). *Social interaction. Video-Based Stu. Hum. Soc.* 6:143071. doi: 10.7146/si.v6i3.143071
- Evans, B., and Lindwall, O. (2020). Show them or involve them? Two organizations of embodied instruction. *Res. Lang. Soc. Interact.* 53, 223–246. doi: 10.1080/08351813.2020.1741290
- Goffman, E. (1983). The interaction order: American Sociological Association, 1982 presidential address. *Am. Sociol. Rev.* 48, 1–17. doi: 10.2307/2095141
- Goodwin, M. (2006). Participation, affect, and trajectory in family directive/response sequences. *Text Talk.* 26, 515–543. doi: 10.1515/TEXT.2006.021
- Graf, K. (2023). Cyborg cooks: mothers and the anthropology of Smart kitchens. *Digital Culture Soc.* 9, 49–70. doi: 10.14361/dcs-2023-0104
- Habscheid, S., Hoffmann, D., Hector, T., and Waldecker, D. (2025). “Voice assistants in private homes. Introduction to the volume” in Voice assistants in private homes: Media, data and language in interaction and discourse. eds. S. Habscheid, T. Hector, D. Hoffmann and D. Waldecker (Bielefeld: Verlag), 9–30.
- Helmer, H., and Reineke, S. (2021). Instruktionen und Aufforderungen in Theorie und Praxis. Einparken im Fahrunterricht. *Gesprächsforschung – Online Zeitschrift zur verbalen Interaktion.* 22, 114–150.
- Jaber, R., Zhong, S., Kuoppamäki, S., Hosseini, A., Gessinger, I., Brumby, D. P., et al. (2024). “Cooking with agents: designing context-aware voice interaction,” in Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 1–13.
- Kendon, A. (1990). Conducting interaction. Patterns of behavior in focused encounters. Cambridge: Cambridge University Press.
- Krug, M. (2025). Coordinating multiactivity transitions in dance rehearsals: spatial configurations and instructional practices. *Front. Commun.* 10:1520005. doi: 10.3389/fcomm.2025.1520005
- Lilja, N., and Piirainen-Marsh, A. (2022). Recipient design by gestures. How depictive gestures embody actions in cooking instructions. *Soc. Int. Video-Based Stu. Hum. Soc.* 5:130874. doi: 10.7146/si.v5i2.130874
- Lindwall, O., and Ekström, A. (2012). Instruction-in-interaction. The teaching and learning of a manual skill. *Hum. Stud.* 35:2749. doi: 10.1007/s10746-012-9213-5
- Lindwall, O., and Lymer, G. and, Greiffenhagen, C. (2015). The sequential analysis of instruction. In N. Markee (Ed.) Handbook of classroom discourse and interaction. New York, NY: John Wiley & Sons 142–157.
- Lindwall, O., and Mondada, L. (2025). Sequence organization in the instruction of embodied activities. *Lang. Commun.* 100, 11–24. doi: 10.1016/j.langcom.2024.11.003
- Lynch, M., and Lindwall, O. (2024). “Introduction. Instructed and instructive actions” in Instructed and instructive actions. The situated production, reproduction, and subversion of social order. eds. M. Lynch and O. Lindwall (London: Routledge), 1–18.
- Messner, M. (2023). Die Interaktion zwischen Dirigent:in und Musiker:innen in Orchesterproben. Mehrsprachige und multimodale Interaktionsmuster zwischen Worten und Tönen, Händen und Füßen. Berlin: deGruyter.
- Mondada, L. (2014). “Cooking instructions and the shaping of things in the kitchen” in Interacting with objects. Language, materiality, and social activity. ed. M. Nevil (Amsterdam: John Benjamins), 199–226.
- Mondada, L. (2016). Challenges of multimodality: language and the body in social interaction. *J. Socioling.* 20, 336–366. doi: 10.1111/josl.1_12177
- Oloff, F. (2017). “Genau als redebeitragsinterne, responsive, sequenzschließende oder sequenzstrukturierende Bestätigungspartikel im Gespräch” in Diskursmarker im Deutschen. Reflexionen und Analysen. eds. H. Blühdorn, A. Deppermann, H. Helmer and T. Spranz-Fogasy (Göttingen: Verlag für Gesprächsforschung), 207–232.
- Ortner, H. (2023). Sprache, Bewegung, Instruktion. Sprache, Bewegung, Instruktion. Multimodales Anleiten in Texten, audiovisuellen Medien und direkter Interaktion. Berlin: deGruyter.
- Parkhurst, S., and Parkhurst, D. (2008). A cross-linguistic guide to SignWriting. A phonetic approach. Available online at: http://www.signwriting.org/archive/docs/sw0617_Cross_Linguistic_Guide_SignWriting_Parkhurst.pdf (Accessed February 5, 2025).
- Pelikan, H., Broth, M., and Kevallik, L. (2022). When a robot comes to life: the interactional achievement of agency as a transient phenomenon. *Soc. Int. Video-Based Stu. Hum. Soc.* 5:129915. doi: 10.7146/si.v5i3.129915
- Sanchez-Stockhammer, C. (2021). Multimodal cohesion through word formation: sublexical cohesive ties in online illustrated step-by-step cooking recipes. *Disc. Context Media.* 43:100536. doi: 10.1016/j.dcm.2021.100536
- Schmidt, A., and Deppermann, A. (2021). Instruieren in kreativen Settings. Wie Vorgaben der Regie durch Schauspielende ausgestaltet werden. *Gesprächsforschung – Online Zeitschrift zur verbalen Interaktion.* 22, 237–271.
- Schmidt, A., and Deppermann, A. (2023). Showing and telling. How directors combine embodied demonstrations and verbal descriptions to instruct in theater rehearsals. *Front. Commun.* 7:955583. doi: 10.3389/fcomm.2022.955583
- Schubert, M. (2024). Sustained pointing gestures in instructions and questions. How the temporal extent of a gesture matters in interaction. *Soc. Int. Video Stu. Hum. Soc.* 7:137058. doi: 10.7146/si.v7i2.137058
- Scollon, R., and Scollon, S. (2003). Discourses in place. Language in the material world. London: Routledge.
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J. R., Bergmann, P., and Birkner, K. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, 353–390.
- Smart, N., and Szczepek Reed, B. (2025). Co-creation of activity spaces in an amateur dance group: interactional construction of the teaching space. *Front. Commun.* 9:1517858. doi: 10.3389/fcomm.2024.1517858

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2025.1589894/full#supplementary-material>

Stukenbrock, A. (2014). Take the words out of my mouth: verbal instructions as embodied practices. *J. Pragmat.* 65, 80–102. doi: 10.1016/j.pragma.2013.08.017

Szczepek Reed, B. (2023). Go on keep going: the instruction of sustained embodied activities. *Discourse Stud.* 25, 692–717. doi: 10.1177/14614456231153578

Szczepek Reed, B. (2024). You don't need me shouting here: when instructors observe learners in silence. *Res. Lang. Soc. Interact.* 57, 169–192. doi: 10.1080/08351813.2024.2340406

Szczepek Reed, B. (2025). Horse-directed vocalizations: clicks, trills, and /ho:/. *Lang. Commun.* 100, 25–45. doi: 10.1016/j.langcom.2024.11.006