

Chapter 2

Telepresent Agency: Remote Participation in Hybrid Language Classrooms via a Telepresence Robot



Teppo Jakonen  and Heidi Jauni 

Abstract Videoconferencing technologies have become increasingly common in different sectors of life as a means to enable real-time interaction between people who are located in different places. In this chapter, we explore interactional data from synchronous hybrid university-level foreign language classrooms in which one student participates via a telepresence robot, a remote-controlled videoconferencing tool. In contrast to many other forms of video-mediated interaction, the user of a telepresence robot can move the robot and thereby (re-)orient to the space, the other participants and material objects that might be outside his immediate video screen. We employ an ethnomethodological and conversation analytic (EMCA) perspective to explore Barad's (Quantum physics and the entanglement of matter and meaning. Durham: Duke University Press: 2007) notion of agency as a distributed phenomenon that emerges from assemblages of humans and materials. We demonstrate the complex nature of telepresent agency by investigating where agential cuts lie in three short episodes that involve mediated perception, touch and movement. Based on the analyses, we discuss how the telepresence technology configures learning environments by making new kinds of competences and forms of adaptation relevant for teachers and students.

Keywords Agency · Telepresence robot · Video-mediated interaction · Remote student · Classroom interaction · Computer-assisted language learning (CALL)

T. Jakonen (✉)

School of Languages and Translation Studies, University of Turku, Turku, Finland

e-mail: teppo.jakonen@utu.fi

H. Jauni

Language Centre, Tampere University, Tampere, Finland

e-mail: heidi.jauni@tuni.fi

© The Author(s) 2023

J. Ennser-Kananen, T. Saarinen (eds.), *New Materialist Explorations into Language Education*, https://doi.org/10.1007/978-3-031-13847-8_2

21

Introduction

Different communication technologies are increasingly commonplace at work, in education and in free time as a way to enable real-time interaction between physically dispersed people. In particular, videoconferencing tools such as Skype, Zoom, FaceTime, Google Hangouts and Adobe Connect are already part of the everyday life of many individuals in different corners of the world. At the time of writing this chapter (2020–2021), many educational and professional organizations were suddenly forced to drastically increase the use of videoconferencing in their daily operations as an attempt to contain and slow the spread of the coronavirus pandemic (Covid-19) through social (or, more accurately, physical) distancing. In many schools and universities, turning face-to-face teaching into virtual classes was by no means an easy task for teachers, despite extensive research literature on blended/hybrid learning (Gleason & Greenhow, 2017) and telecollaboration (Dooly & O’Dowd, 2018).

Videoconferencing challenges our understanding of what it means to be present in some social environment or activity: how is the experience of presence a material phenomenon, and what kinds of implications does its material nature have for the way we think about agency? Perhaps a relatively easy example to illustrate what we mean here is to consider how, whenever we make a video call, the camera and the computer screen mediate what we see of the environment that is remote to us. It is usually less than what we perceive of our own ‘local’ environment in which we are physically present, and, depending on the technology, we might not necessarily even have the ability to control what the camera shows us. The camera is thus a powerful yet often unnoticed material tool: as Luff et al. (2003) have shown, it can “fracture” the ecology of action in video-mediated interaction so that if we, for example, point at something during a video call, it is not self-evident that our interlocutor sees both the pointing gesture and what is being pointed at. This can have significant implications for how shared understanding of the on-going activity can be achieved.

In this chapter, we explore this kind of remote – or telepresent – agency in a complex assemblage of technology, people, materials and space in an educational context. Investigating how university students participate in otherwise ‘regular’ face-to-face language classes via a drivable telepresence robot, we attempt to consider how agency is a social, interactional and materially mediated achievement. In a nutshell, telepresence robots are videoconferencing tools that give a participant the ability to move the camera that shows them a remote location (such as a classroom) by driving the robot that is physically in that location. Existing interview and survey-based studies from educational contexts suggest that telepresence robots can augment the sense of agency, presence and social inclusion of remote students (Cha et al., 2017; Fitter et al., 2018; Newhart et al., 2016). However, much less is known about how agency emerges through, and is managed in, the micro-level interactional practices involving telepresence robots. This chapter thus aims to contribute to research on telepresence robots and, more broadly, to interactional

research on videoconferencing by exploring what kinds of consequences the material and technological features of telepresence robots have for remote agency.

Being Telepresent in a Material World

Telepresence can be defined as “the sense of being in another environment” (Kristofferson et al., 2013). As a concept, telepresence goes back to (at least) the beginning of 1980s when Marvin Minsky (1980) used the term to describe remote, robotically enabled presence in some location involving “high-quality sensory feedback”. Minsky predicted that in the future such robotic telepresence would “feel and work so much like our own hands that we won’t notice any significant difference” (Minsky, 1980, p. 47). He envisaged telepresence above all as a technology that could be used in material environments that are hazardous to humans – examples include the outer space, undersea mining, nuclear power plants, and so on. In Minsky’s view (1980), a key aspect and the biggest challenge of telepresence would be achieving a realistic “sense of ‘being there’”.

Minsky’s definition raises a question what exactly makes us feel that we are ‘there’. In many ways, humans experience the world and engage in social relations through their bodies (Meyer et al., 2017). Thus, a primitive form of telepresence, of being ‘there’, can be provided optically: looking through a microscope or following live TV allows us to follow events in a place other than the one in which we are physically located. However, our experience of the physical world is not limited to the visual sense, but it routinely also involves other senses, such as auditory and haptic channels as well as a sense of where the limits of our body are. We can touch things, sense being touched, sense where people around us are by judging from which direction their sound is approaching us, and so on. Initially, it might seem that technology such as videoconferencing is just a tool that mediates the experience of the material world to us. However, it is not always easy to tell where a (technological) tool ends and a human being begins. For example, a blind man’s stick becomes over time “an instrument with which he perceives [...] an extension of the bodily synthesis” (Merleau-Ponty, 1945/2002, p. 176) instead of an object. Similarly, some user reports indicate that technologies such as the telepresence robot can through time “become integrated with one’s sense of self and sense of one’s own capabilities” (Takayama, 2015, p. 162).

Telepresence constitutes a context for social action in which the human body is at times a problematic resource – and for this reason it can be challenging to conduct co-operative activities via videoconferencing in exactly the same manner as face-to-face. One way to conceptualise these challenges is through Maurice Merleau-Ponty’s (1945/2002) phenomenological philosophy. He argued that in typical circumstances the living human body functions as our ‘zero point’ for making sense of the world and for acting in it. However, when acting and interacting via a telepresence robot, one needs to coordinate not only one’s own physical body but also the remote metal body of the robot. In our classroom data, the telepresence robot is a

material object through which a remote participant acts in the classroom, but it is also an embodied participant that other classroom participants can orient to and use as a resource for interaction. In order for the remote participant to take part in classroom activities, they thus have to co-ordinate the actions and movements of two different bodies, those of the remote body (robot) and those of their own living body, in a way that parallels how video gamers manage the movement of their digital avatars on screen in order to construct game-relevant actions (Bennerstedt & Ivarsson, 2010). The way the robot adds a re-embodied and movable extension of the self can lead to a fracture between the acting self and the sensory self. By offering simultaneous sensory feedback from two different locations, telepresence can also blur the distinction between these locations and challenge what Neisser (1988) has termed as the ‘ecological self’ – i.e., knowledge about oneself with respect to one’s physical environment.

Agency and Telepresence

In broad and traditional terms, agency can be seen as the degree to which “an agent (whether human or nonhuman) can act in the world of its own accord” (Takayama, 2015, p. 161). However, agency is also situated – we do things in the context of specific activities, and our actions and competence are judged in relation to contextual frames of reference and requirements. Barad (2007, p. 33) argues that “agencies are only distinct in relation to their mutual entanglement; they don’t exist as individual elements”. Although Barad’s (2007) agential realism represents a radical (re-)conceptualisation of the ontology and ‘locus’ of agency, we find that it is in many respects compatible with the way agency has been conceived of in the ethnomethodological and conversation analytic (EMCA) tradition. From an EMCA perspective, human action and interaction have a fundamentally co-operative and material character (e.g. Goodwin, 2013) so that the agency of a person is situated in, and emerges from the sequential context of action, the material objects, technological tools and other participants in the setting. Such a view can perhaps best be illustrated with an example from Charles Goodwin’s extensive research on the situated interactional competencies of an aphasic man in conversation with his family members. Goodwin (e.g. 2004) has shown how a man whose vocabulary a stroke reduced to only three words (*yes*, *and*, and *no*) can in spite of this limitation be a competent participant in conversation. This is possible because of the ‘laminated’ (Goodwin, 2013) nature of human action, i.e. how participants in interaction routinely disassemble and reorganize layers of different kinds of semiotic materials. Thus, the aphasic man in Goodwin’s studies can use another speaker’s lexicon and syntax as a ‘substrate’ and transform it, for example, by means of prosody and embodied displays of stance and footing. In that way, he is able to concurrently produce actions that participants treat as belonging to him. In Goodwin’s (2013, p. 15) view, this illustrates how “human beings inhabit each other’s actions”, which resonates well with Barad’s (2007)

view that individual agencies do not precede their interaction, but rather “emerge through their intra-action“(p. 33). What this suggests is that EMCA can offer a powerful empirical lens to investigate sociomaterialism and agential cuts from an emic perspective through participants’ (changing) orientations to the agency of persons, tools and material objects (see also Thorne et al., 2021, p. 110).

It is one thing to view events at a distance (for example via a video) and another to act and interact remotely in an agentive manner. Luna Dolezal (2009) has investigated the phenomenology of agency in recent, increasingly more high-tech forms of telepresence such as telesurgery whereby surgical operations are performed by manipulating robotic arms at a distance. She draws on Gallagher’s (2000) distinction between a sense of agency and a sense of ownership of an action as two distinct aspects of how we experience action (Dolezal, 2009, p. 218). Typically, we experience both of these senses together: for example, if I throw a ball so that it hits a window, I sense that I have caused the window to break (causal agency) and that my hand has undergone a throwing movement (ownership of action). Such a perception can be seen as a particular kind of agential ‘cut’ (Barad, 2007), a linking together of objects, beings and doings. However, telepresent actions can be different. Even if a person might see that they are doing some action, they do not necessarily feel the action as theirs because an embodied sensation of ‘owning’ it is missing. Similarly, when making a video call, we might see that we are physically close to another person but we do not (necessarily) sense the same kind of physical intimacy as when we are copresent. In Dolezal’s (2009, p. 218) view, this kind of “[d]issociation [of agency] from ownership” also has ethical consequences. Perhaps this is clearest in military applications of telepresence such as the use of drones to fire missiles with a remote user interface that reminds video games (see also Parks & Kaplan, 2017).

In this chapter, we investigate agency in remote participation in a video-mediated, physically distributed assemblage of humans, interactional spaces, human-created technological tools (e.g. the robot, computers, whiteboards), and physical classroom artefacts (chairs, desks etc.). In such a context, agency can be seen as entangled in the sense that the remote student “lack[s] an independent, self-contained existence” (Barad, 2007, ix) in this system without the other elements of the assemblage. Robot-enabled interaction between a remote student and co-present classroom participants is also asymmetric because the remote student has a very different kind of sensory access to the classroom. However, this and other material-technological conditions do not limit the remote student’s agency in the classroom in a deterministic manner. Of interest to us are the ways in which participants orient to interactional asymmetries and co-operate with each other to support robot-mediated remote participation. Analogous to Goodwin’s examples of how the co-operative organization of human interaction enables the aphasic man to act with considerable agency by using available resources for building action, telepresent agency emerges through coordinated and materially-embedded actions.

Data and Method

Our data consist of video-recorded English, Swedish, Finnish and German language lessons, taught to students of technology as part of their degree studies at a Finnish university. In the lessons at least one student participates from another location via a telepresence robot. Altogether, we have circa 12 hours of video-recorded lessons with class camera footage and (in the case of our English and Swedish classroom data) screen capture from the remote student's laptop. For the purposes of this chapter, we have selected extracts from the English and German classroom data. These lessons showcase first-time users testing the telepresence technology so that students took turns to go to another location on campus to participate in the lesson by operating the robot. The telepresence robot used in our data is Double 2, a device developed by Double Robotics for remote work and education purposes.

Double 2 has a mobile robotic base equipped with an iPad, external video camera, microphone and speakers. As Fig. 2.1 shows, the appearance of the robot is very schematic: it is an iPad on a stick, equipped with wheels. The key feature of the robot is its movability. The remote participant can control the robot via an online interface or with an iPad application. Using a computer, the robot is controlled with arrow keys, with which it can be moved around the classroom. Its height can also be adjusted, which is an important feature when joining groups of people that are sitting or standing. These abovementioned features enable the distant participants to re-orient to the material environment and other participants in a way that traditional

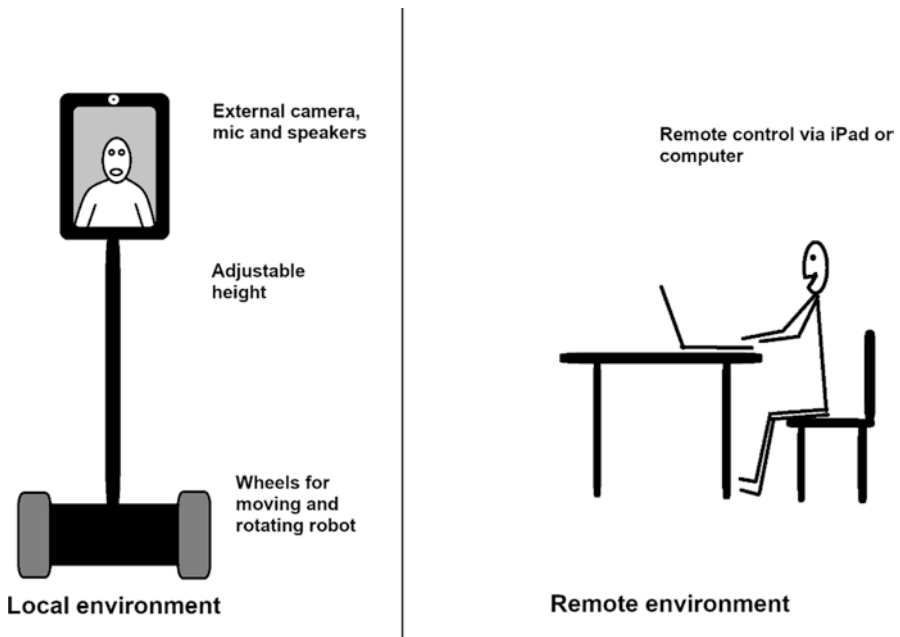


Fig. 2.1 A Double telepresence robot and its remote user

videoconferencing methods do not easily allow. However, Double 2 cannot be used to manipulate objects, and it also lacks the ability to pan or tilt the camera (these features are available in the newer version of the robot, Double 3).

Methodologically, we draw on conversation analysis (see Stivers & Sidnell, 2012). CA, which emerged in the 1960s in sociology (for in-depth accounts of CA origins, see Heritage, 2008; Psathas, 1995), has close connections to ethnomethodology (Garfinkel, 1967). It has since then spread beyond sociology into many other disciplines such as (applied) linguistics, psychology, medicine and anthropology. The sociological orientation is visible in an interest in understanding the organization of social actions and interaction, as well as explicating the kinds of resources that participants use to construct action and make sense of it. Analysing social interaction from a CA perspective usually proceeds through a bottom-up, inductive logic and an avoidance of pre-theorisation, in other words through ‘unmotivated looking’ (Psathas, 1995). From a CA perspective, interaction is viewed as an orderly and sequentially emerging phenomenon, and a key analytical strategy is investigating how participants treat each other’s actions in publicly observable ways in subsequent interactional turns – what Sacks, Schegloff and Jefferson (1974, p. 729) have referred to as a ‘next-turn proof procedure’. As Heritage (1984, 241–245) points out, in this way, CA conceptualizes interaction as structurally organized and individual turns-at-talk as both “context-shaped” (by the previous turn) and “context-renewing” (for some subsequent turn).

The transcription of interactional data follows standard CA conventions (Jefferson, 2004). In addition, we illustrate analytically relevant embodied phenomena by way of still images taken from the video. Their timing relative to talk is marked with hashtags (#) in the extracts.

Analysis

In this section, we discuss some ways in which, in the focal context, the agency of the remote student is a social, interactional and material accomplishment that emerges through participants’ coordinated and embodied actions. We do this by analyzing three examples, which illustrate telepresent agency in relation to seeing, touching and moving.

Agency and Perception

We begin by considering the sociomaterial assemblage with the help of two still images depicting the same moment in an EFL classroom. Figure 2.2 shows a frame grab from a video camera that was positioned at the back of the classroom. It shows a moment when a teacher is pointing at a whiteboard to show text written on it to two remote students who participate via a telepresence robot (the black object in



Fig. 2.2 Classroom view

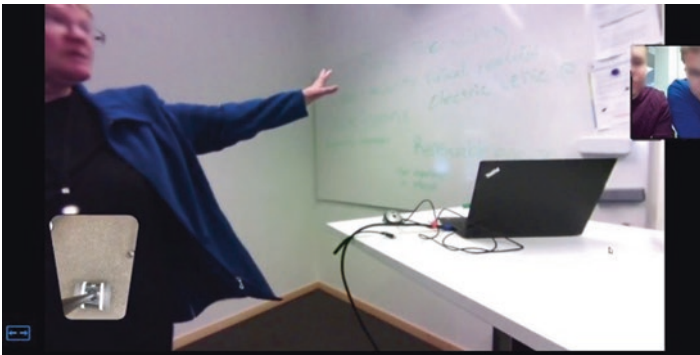


Fig. 2.3 Robot-mediated remote view into the classroom

front of the teacher). In contrast, Fig. 2.3 shows a frame grab from the two remote students' laptop screen at the same time, illustrating the remote students' visual access to the material environment of the classroom. The right-hand top corner shows the remote students' laptop camera recording, which is currently showing a half of each student's torso. This footage is streamed on the robot screen in the classroom and available to classroom participants.

Compared to the participants who are physically located in the classroom, this particular form of telepresence has some limitations with respect to sensing and experiencing the remote sociomaterial environment (the classroom). Some of the limitations relate to the properties of camera-mediated vision. Unlike the human eye, the robot camera offers no peripheral vision, which means that the visibility of objects is either 'on' or 'off', depending on whether they are within the frame perimeters or not. The camera cannot be zoomed or tilted in this version of the Double robot, which means that in order to see text on a whiteboard the remote students would need to drive the robot close enough to the board (as they are doing in Figs. 2.2 and 2.3). Similarly, viewing a paper document at a non-direct angle may

be more difficult than it is in the copresent condition (see also Jakonen & Jauni, 2021). In addition, while the robot can be remotely moved, turning the robot takes more time than it does for the average person to turn their head or body orientation. This kind of relative slowness in comparison to a human gaze shift could make it more challenging to follow talk between participants who are, for example, located in different corners of the classroom – or any other spoken exchange that involves rapid turn transitions. In our data, the classroom participants, especially teachers, orient to this asymmetry and conduct extra interactional work by way of checking, showing and guiding to ensure that classroom materials are visible to remote students (Jakonen & Jauni, 2021).

Seeing is a basic foundation of many kinds of interactions, something which has consequences for the accomplishment of other actions, such as moving from one place to another. For the remote participant, navigation in the classroom can be problematic because the video constitutes a 2D representation of a (familiar) 3D environment. Thus, navigation can require specific interactional practices from the participants, some of which we will discuss in more detail later in Extract 2.2.

Agency and Touch

Telepresence robots differ from each other with respect to the degree of anthropomorphism, i.e., to what extent their design includes human-like physical characteristics (Kristofferson et al., 2013; Li, 2015). Newhart et al. (2016) explored the use of telepresence robots by 6–16-year-old homebound students and found that anthropomorphism was a key factor in whether the classroom participants accepted and included the robot and its remote user as a regular member of the classroom. Interestingly, in one fifth-grade class, the teachers in the study had noticed that the students did not differentiate between the robot and the homebound student operating the robot, but referred to the robot with the student's name. Similar observations have also been made in workplace contexts: for example, Takayama (2015, p. 162) has noted that telepresence robots can through time “become invisible-in-use” and that they disappear “into the background of conscious attention”.

The Double 2 robot in our case has very few anthropomorphic qualities, and it is not specifically designed to look human. However, ‘seeing’ and ‘seeing as’ are not only psychological and optical phenomena; they are also situated and interpretative accomplishments (e.g. Goodwin & Goodwin, 1996; Nishizaka, 2017). As Goodwin (1994, p. 606) puts it, seeing is “lodged within endogenous communities of practice”. Thus, it is possible to see the Double 2 robot as a human body that has a head (the iPad that shows the remote participant's face), a neck/upper body (the pole on which the screen is attached) and a lower body (the wheels). This provides for a possibility to see the robot as the person who is interacting via it, perhaps more readily than in a situation where interaction is mediated by a tablet or a computer placed on a desk. Extract 2.1 illustrates this kind of orientation to the robot as an embodied human participant through an action that we call here, for the lack of a better term, as a mediated touch: a simulation of physical touch accomplished in

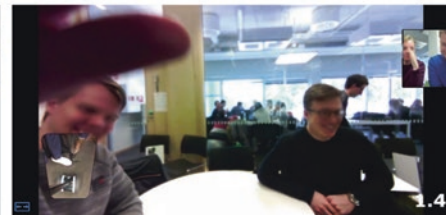
video-mediated interaction. The extract shows a peer group – two classroom students and two remote students (via one robot) – engaging in the parallel activity (Koole, 2007) of entertaining themselves while the teacher is asking others to write suggestions for group work topics on the whiteboard. The focal group jokingly treats the telepresence robot as if it were a human being by patting and stroking the robot’s head. This results in a largely non-verbal performance of social intimacy by way of peer-to-peer touch (see also Karvonen et al., 2018).

Extract 2.1 Mediated Touch and Physical Closeness

01 (5.3)#^{1.1}
 02 T so, (.) do you think we have enough topics:?
 03 (2.2)#^{1.2}



04 GREY? hh he he (.) ;he he
 05 (2.3)
 06 T great
 07 #^{1.3}(1.2)
 08 GREY hh (0.7) .hh he he
 09 (0.6)
 10 T let’s then start |voting.
 11 (2.6)#^{1.4}



12 GREY .hh he he [he he he .hhh he [he he
 13 BLACK [he he he he he he [he he
 14 T [the topics are,
 15 ((T continues list; DS turns robot towards T))

The group’s parallel activity takes place as the teacher is proceeding through a transition to a new activity phase (lines 2, 6, 10, 14–15). During this, the two remote students, who are visible in the top right-hand corner of image 1.1, drive the robot closer to the two classroom students, Grey (left in the image) and Black (right in the image). The two classroom students monitor the robot’s approach by gaze.

As image 1.2 above shows, Grey provides a ‘thumbs up’ gesture during the silence at line 3 to assess the movement and to signal that the robot has reached a suitable place close to the table. The bottom left-hand corner of image 1.2 illustrates how at this point the robot is already very close to Grey’s foot, considerably closer than is typical in human-robot interaction (Lauckner & Manzey, 2014). The

participants are now facing each other in what Kendon (1990) has termed as the F-formation, a basic spatial arrangement for human interaction in which parties have “equal, direct, and exclusive access” (p. 209) to the space between them. An F-formation can be achieved through a range of postural and group arrangements, such as when people are standing and chatting in a circle or seated side-by-side and work on a shared text, etc. F-formations are also formed by hybrid groups that consist of both co-present human participants and telepresence robots operated by a remote participant (Pathi et al., 2019), but their exact shape can depend on the material design of the robot (Kristofferson et al., 2013). To give an example, when a remote participant is visible to classroom members as a two-dimensional image on the screen, as in our data, a side-by-side spatial arrangement can be cumbersome because the remote participant’s field of view is narrower than that of a human eye.

The thumbs up gesture is followed by laughter and a silence (line 5), after which Grey pats the robot on the ‘head’ (top of the screen) as is visible in image 1.3. The patting is an instance of a mediated touch; the remote participants who operate the robot cannot feel the touch as a tactile sensory experience, but the participants can nevertheless use other embodied resources to simulate such an experience of touching and being touched. Here, the other resources include Grey’s posture (leaning head) and his facial expression (smile). The visibility of Grey’s hand in the top left-hand corner of the remote participants’ screen makes the action recognizable to them as a touch. Altogether, the lamination of these resources constructs the action as an instance of gentle patting, a form of affective touch (Cekaite & Kvist Holm, 2017) that demonstrates and builds social intimacy between the participants.

Grey’s patting gradually transforms into a stroking gesture by line 11, at which point one of the remote participants (Blue) pokes his head forward as if aligning with being patted and stroked (see the top left-hand corner of image 1.4). This kind of co-ordination of embodied actions by physically dispersed participants to achieve a simulation of human touch illustrates how both participants recognize the emergent action, its local sense and logic, and co-operate to accomplish it. Patting and stroking a peer’s head is socially a somewhat delicate action in many classroom contexts, perhaps even more so among adult students, and part of the situated humour around these actions comes from the unexpected nature of this kind of touch as a form of social intimacy in this setting. The shared joke is made possible by perceiving the materiality of the robot in such a way that it is seen as a human being, by finding equivalence between specific parts of the metal body of the robot and human body parts. The remote students agentively make this touch happen by driving the robot and by putting their head (Blue) into a position in which Grey can see it on the screen right under his hand.

Agency and Movement

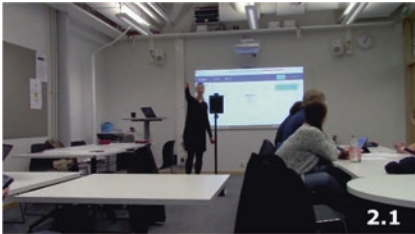
Extract 2.2 exemplifies how agentic movement by the remote student is collaboratively accomplished, and accommodated to, in the classroom. It shows how a German language teacher deals with a routine organisational task: assigning

students into small groups for an activity, here a quiz to be completed in groups. In the extract, the teacher's task is made more complex by the fact that the remote student (Timo) is part of a group with two classroom students (Lauri and Markus), who are seated at different ends of the classroom. The teacher thus needs to guide one classroom student (Lauri) and the remote student's robot to another desk for the activity.

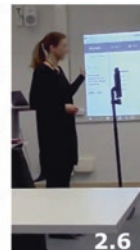
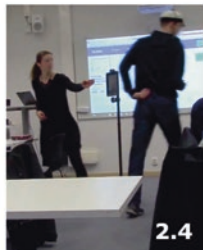
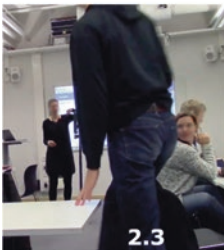
The extract shows how the remote student, who has positioned the robot in front of the classroom whiteboard (see image 2.1) follows and anticipates the teacher's instruction by beginning to move the robot. The teacher accommodates to this movement and supports the remote student's navigation of the robot into a group with an elaborate multimodal instruction (lines 6–7).

Extract 2.2 Changing Places

- 01 T los geht's ((clicks the quiz open))
'let's go'
 02 (0.6) ((reads groups from the whiteboard))
 03 T gut, (.) Timo sie sind mit Lauri und mit Markus^{#2.1} im <team>
'good, Timo you are with Lauri and Markus in a team'
 04 vielleicht ähm? (.) Lauri können sie mal hierher^{#2.2} gehen=
'perhaps ehm Lauri can you go here'



- 05 =wir nehm mal den anderen tisch diesmal
'we take the other table this time'
 06 also Timo^{#2.3} einmal, (0.5) einmal wir umdrehen, (0.4)^{#2.4}
'so Timo just we just turn around'
 07 zum Markus und zu Lauri.^{#2.5}
'to Markus and Lauri'
 08 (0.4)
 09 T genau.
'exactly'
 10 (1.2)
 11 T okay.^{#2.6}



- 12 (0.4)
 13 T helfen sie (ihn ein) bisschen (.) dass er sie findet,
'help him a bit so that he finds you'
 14 ((Remote student drives the robot to Markus))

The teacher assigns the remote student into a group by addressing him, announcing his group members (line 3), and by pointing at one of them (Lauri) to indicate his location in the classroom to the remote student, as shown in image 2.1. The teacher then implicates where the group ought to sit by requesting Lauri to go from the back of the room to another group of desks (where Markus is already seated, line 4). Image 2.2 illustrates how the teacher points towards Markus (on the right-hand side of the room) and how Lauri complies with the teacher's instruction by standing up and beginning to walk towards Markus's desk.

The remote student reacts to the teacher's turn at line 3 by beginning to turn the robot anticlockwise away from the whiteboard. The movement begins roughly when the teacher says 'ähm' (line 4) and stops at the end of line 4 into a position where the robot screen is facing the teacher (as it is in image 2.2). The movement is a demonstration of agency that shows that the remote student is able to anticipate what he should be doing next, even if the teacher has thus far merely named the remote student's group members.

The remote student continues to turn the robot roughly when the teacher says *diesmal* ('this time', line 5). This could be the beginning of a movement towards the assigned place (Markus's desk). Yet, the teacher provides a further instruction to the remote student, both verbally and in embodied ways (lines 6–7). The teacher makes a rotating gesture with her left hand (image 2.3) and points towards Markus's desk so that she continuously maintains herself in front of the screen of the turning robot (images 2.3–2.5). Doing this allows her to secure that her referential gestures will be visible to the remote student, whom she is directing to the desired location. The turning movement comes to a stop at the end of line 11 (*okay*), after which the remote student drives the robot straight ahead to Markus's desk (not shown here).

In this situation, it is noteworthy that the physical activity of moving oneself (or one's robot) to the appropriate place in the classroom is left to the remote student's task in much the same manner as the classroom student (Lauri). However, these two students are instructed and assisted by the teacher in a strikingly different manner. Whereas Lauri is 'just' verbally requested to go to Markus's desk (line 4), the instruction for the remote student is much more heavily supported by segmenting the requested action into turning around and moving straight ahead (lines 6–7) and what could be termed as hyper-iconic gestures. These instructional features display an orientation to the material constraints of telepresence and showcase a situated co-ordination of human and technological bodies, material environment and language in a fractured ecology of action in which referential practices are known to be complex (see e.g. Luff et al., 2003). In this sense, the instruction thus amounts to an embodied demonstration of professional competence by the teacher.

Conclusions

In this chapter, we have investigated issues related to agency in robot-mediated participation in language education. ‘Agency’ is itself a concept that is notoriously difficult to pin down, and here we have tried to explore its material and embodied nature by considering the nature of rather mundane senses (seeing and touching) and actions (moving) in video-mediated interaction (see also Muhonen & Vaarala, Chap. 4, this volume). Telepresence robots, such as the Double 2 robot in our data, are currently viewed as a potential technological tool for increasing the agency and social inclusion of vulnerable student groups relying on remote access to education (Cha et al., 2017; Fitter et al., 2018; Newhart et al., 2016). However, there is not much interactionally-oriented research examining the ways in which copresence and telepresence may be consequential for students’ possibilities for action, participation and agency in learning settings (but see Jakonen & Jauni, 2021; Liao et al., 2019).

From a conversation analytic perspective, a material tool such as the robot constitutes a resource for constructing and making sense of social action; the technology does not prescribe, a priori, any particular way to interact via it, even if such a way might have been envisioned by those who have developed the technology. Such a view has clear links to, for example, ecological perspectives that highlight the role of affordances for language learning (e.g. van Lier, 2000). From such a perspective, it can thus be difficult to assess any technological tool as inherently ‘good’ or ‘bad’, simply because human action can be constructed in a myriad of novel and unforeseen ways. This can be seen in how, despite the obvious technological limitations of the robot vis-à-vis copresent interaction – such as those related to the field of vision, speed and dexterity of movement, and the lack of haptic sensory feedback it affords – telepresent agency is still possible in social interactions that require seeing, touching or moving.

In all cases analysed in this chapter, the remote students are treated as agentic participants, but their agency is also co-operatively constructed and supported by classroom participants through practices of guiding, showing, and so on. The robot-mediated remote users are oriented to as needing particular kind of interactional support, which constructs these interactional situations as asymmetric. However, through the support, actions and participation become possible. This gives rise to a question where exactly agency is located in this kind of a sociomaterial assemblage (see also Guerretaz et al., 2021) involving telepresence, and in what sense are the remote student and the robot embodied participants in the classroom. For the remote student, the robot is a proxy or an extension of the self that mediates sensory information and provides a way to interact from a distance. The robot is also a material and agentic participant that classroom members orient to, and whose material and technological properties they must take into account as they design social actions addressed to the remote students: for example, by considering the arrangement of bodies in the classroom (Extract 2.2). Consequently, the ecological self (Neisser, 1988) and agency of the remote student are fundamentally dispersed across space,

existing in the remote location and the classroom, in this particular socio-technological assemblage.

In much of our data, remote students are given the primary responsibility to move the robot to relevant places within the classroom (e.g. Extract 2.2). However, at times remote navigation takes extensive teacher guidance and time. Perhaps paradoxically, extensive guidance constitutes an orientation to the asymmetric nature of robot-enabled hybrid teaching, but it increases the agency of the remote participant. Time-wise, a more effective means might be to just move the robot by carrying it from one place to another, similarly as one would move a laptop-mediated videoconferencing participant from one place to another. Yet, this does not happen, and part of the reason may be related to the way the robot can be seen as resembling a person: thus, lifting the robot by the pole would be akin to grabbing a human being by their neck.

The entanglement of agencies becomes visible through embodied actions that are addressed to, or that involve, the robot. The material shape of the robot seems to invite classroom students to treat it as an actual person for example by patting it on the head (Extract 2.1) or by giving high-fives. By touching the robot in a manner that resembles the way humans or animals are touched, classroom participants can treat it as an actor with agency. This agency does not necessarily stem from the robot's physical properties, but the situated role and meaning it has in the (distributed) ecology of action as the extension of the remote student's self, a kind of a 'stand-in' for an actual human being in an entanglement of materials and humans. This illustrates that "agential cut[s] between 'subject' and 'object'" (Barad, 2007, p. 140) can be complex, emergent and at times blurry in this kind of a sociomaterial assemblage.

In general, remote students – just like classroom-based students – participate in classroom interaction in a manner that demonstrates their understanding of the activities, the way they look for, and find, a local sense and order in the activities. Moreover, they participate in the unfolding of activities, and constitute those activities, by adapting their methods for accomplishing different actions to the interactional contingencies in a complex configuration of bodies, objects and technologies (see e.g. Girard-Groeber, 2018). In this way, the remote participants are taken as competent and agentic members of the classroom. Their sense-making is supported by knowledge of the kinds of practices, activities and roles that can be taken as typically relevant in this particular institutional setting. Adaptation is itself a demonstration of agency, and telepresent students' agency is enacted through the situated ways in which social order is co-operatively and repeatedly (re)produced in the setting.

Acknowledgements We are grateful to the anonymous reviewers, fellow contributors to this book, and the editors for their constructive feedback to earlier versions of this chapter. Any remaining errors and shortcomings are our own. This work was supported by the Academy of Finland [grant number 343480].

References

- Barad, K. (2007). *Meeting the universe halfway. Quantum physics and the entanglement of matter and meaning*. Duke University Press.
- Bennerstedt, U., & Ivarsson, J. (2010). Knowing the way. Managing epistemic topologies in virtual game worlds. *Computer Supported Cooperative Work, 19*(2), 201–230. <https://doi.org/10.1007/s10606-010-9109-8>
- Cekaite, A., & Kvist Holm, M. (2017). The comforting touch: Tactile intimacy and talk in managing Children's distress. *Research on Language and Social Interaction, 50*(2), 109–127. <https://doi.org/10.1080/08351813.2017.1301293>
- Cha, E., Chen, S., & Mataric, M. J. (2017). Designing telepresence robots for K-12 education. In *IEEE international symposium on robot and human interactive communication* (pp. 683–688). <https://doi.org/10.1109/ROMAN.2017.8172377>
- Dooly, M., & O'Dowd, R. (Eds.). (2018). *In this together: Teachers' experiences with transnational, telecollaborative language learning projects*. Peter Lang.
- Dolezal, L. (2009). The remote body: the phenomenology of telepresence and re-embodiment. *Human Technology, 5*(2), 208–226.
- Fitter, N. T., Chowdhury, Y., Cha, E., Takayama, L., & Mataric, M. J. (2018). Evaluating the effects of personalized appearance on Telepresence robots for education. *ACM/IEEE International Conference on Human-Robot Interaction, Part, F1351*, 109–110. <https://doi.org/10.1145/3173386.3177030>
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice-Hall.
- Girard-Grober, S. (2018). *Participation by hard-of-hearing students in integration classrooms: Facets of interactional competence*. Frank & Timme.
- Gleason, B. W., & Greenhow, C. (2017). Hybrid education: The potential of teaching and Learning with robot-mediated communication. *Online Learning, 21*(4), 159–176. <https://doi.org/10.24059/olj.v21i4.1276>
- Goodwin, C. (1994). Professional vision. *American Anthropologist, 96*(3), 606–633. <https://doi.org/10.1525/aa.1994.96.3.02a00100>
- Goodwin, C. (2004). A competent speaker who can't speak: The social life of aphasia. *Journal of Linguistic Anthropology, 14*(2), 151–170. <https://doi.org/10.1525/jlin.2004.14.2.151>
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse and Society, 18*(1), 53–73. <https://doi.org/10.1177/0957926507069457>
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics, 46*(1), 8–23. <https://doi.org/10.1016/j.pragma.2012.09.003>
- Goodwin, C., & Goodwin, M. H. (1996). Formulating planes: Seeing as a situated activity. In D. Middleton & Y. Engeström (Eds.), *Cognition and communication at work* (pp. 61–95). Cambridge University Press.
- Guerrettaz, A. M., Engman, M. M., & Matsumoto, Y. (2021). Empirically defining language learning and teaching materials in use through sociomaterial perspectives. *The Modern Language Journal, 105*(S1), 3–20.
- Heritage, J. (1984). *Garfinkel and ethnomethodology*. Polity Press.
- Heritage, J. (2008). Conversation analysis as social theory. In B. Turner (Ed.), *The new Blackwell companion to social theory* (pp. 300–320). Blackwell. <https://doi.org/10.1002/9781444304992.ch15>
- Jakonen, T., & Jauni, H. (2021). Mediated learning materials: Visibility checks in telepresence robot mediated classroom interaction. *Classroom Discourse, 12*(1–2), 121–145.
- Jefferson, G. (2004). *Glossary of transcript symbols with an introduction. Conversation Analysis: Studies from the First Generation*. Retrieved from http://www.liso.ucsb.edu/liso_archives/Jefferson/Transcript.pdf.

- Karvonen, U., Heinonen, P., & Tainio, L. (2018). Kosketuksen lukutaitoa. Konventionaalistuneet kosketukset oppilaiden välisissä vuorovaikutustilanteissa. In L. Lehti, P. Peltonen, S. Routarinne, V. Vaakanainen, & V. Virsu (Eds.), *Uusia lukutaitoja rakentamassa –Building new literacies (AFinLAN vuosikirja 2018)* (pp. 159–183). AFinLA.
- Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge University Press.
- Koole, T. (2007). Parallel activities in the classroom. *Language and Education*, 21(6), 487–501. <https://doi.org/10.2167/le713.0>
- Kristofferson, A., Coradeschi, S., & Loutfi, A. (2013). A review of mobile robotic telepresence. *Advances in Human-Computer Interaction*, 2013. <https://doi.org/10.1155/2013/902316>
- Kristoffersson, A., Severinon-Eklundh, K., & Loutfi, A. (2013). Measuring the quality of interaction in mobile robotic telepresence: A Pilot's perspective. *International Journal of Social Robotics*, 5, 89–101.
- Lauckner, M., & Manzey, D. (2014). 'Hey robot, please step back!' – Exploration of a spatial threshold of comfort for human-mechanoid spatial interaction in a hallway scenario. In *IEEE international symposium on robot and human interactive communication* (pp. 780–788). Edinburgh.
- Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human Computer Studies*, 77, 23–27. <https://doi.org/10.1016/j.ijhcs.2015.01.001>
- Liao, J., Lu, X., Masters, K., Dudek, J., & Zhou, Z. (2019). Telepresence-place-based foreign language learning and its design principles. *Computer Assisted Language Learning*, 1–26. <https://doi.org/10.1080/09588221.2019.1690527>
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured ecologies: Creating environments for collaboration. *Human-Computer Interaction*, 0024(776111462). <https://doi.org/10.1207/S15327051HCI1812>
- Merleau-Ponty, M. (1945/2002). *Phenomenology of perception*. Routledge.
- Meyer, C., Streeck, J., & Jordan, J. S. (2017). Intercorporeality. *Emerging socialities in interaction*.
- Minsky, M. (1980, June). Telepresence. *Omni Magazine*. Retrieved from <https://web.media.mit.edu/~minsky/papers/Telepresence.html>.
- Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical psychology*, 1(1), 35–59.
- Newhart, V. A., Warschauer, M., & Sender, L. S. (2016). Virtual inclusion via Telepresence robots in the classroom: An exploratory case study. *The International Journal of Technologies in Learning*, 23(4), 9–25. <https://doi.org/10.1103/PhysRevB.97.165108>
- Nishizaka, A. (2017). The perceived body and embodied vision in interaction. *Mind, Culture, and Activity*, 24(2), 110–128.
- Parks, L., & Kaplan, C. (Eds.). (2017). *Life in the age of drone warfare*. Duke University Press.
- Pathi, S. K., Kristoffersson, A., Kiselev, A., & Loutfi, A. (2019). F-formations for social interaction in simulation using virtual agents and Mobile robotic Telepresence systems. *Multimodal Technologies and Interaction*, 3(4), 69. <https://doi.org/10.3390/mti3040069>
- Psathas, G. (1995). *Conversation analysis: the study of talk-in-interaction*. Sage.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Stivers, T., & Sidnell, J. (Eds.). (2012). *The handbook of conversation analysis*. John Wiley & Sons.
- Takayama, L. (2015). Telepresence and apparent agency in human–robot interaction. In S. S. Sundar (Ed.), *The handbook of the psychology of communication technology* (pp. 160–175). Wiley Blackwell.
- Thorne, S. L., Hellermann, J., & Jakonen, T. (2021). Rewilding language education: Emergent assemblages and entangled actions. *The Modern Language Journal*, 105(S1), 106–125.
- van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245–259). Oxford University Press.

Teppo Jakonen, PhD, is Academy of Finland Research Fellow at the School of Languages and Translation Studies, University of Turku, Finland. His research primarily explores how language teaching and learning take place in and through social interaction, particularly in technology-rich settings.

Heidi Jauni, PhD, works as Development Manager at the Language Centre of Tampere University, Finland. Her research interests include conversation analysis, especially interaction in educational settings and technology-mediated interaction.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

