Applied Linguistics, **2024, XX, 1–19**
https://doi.org/10.1093/applin/amae017
Advance access publication 21 March 2024
Article

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

Achieving Joint Attention and Understanding of Task Responsibilities in Synchronous Hybrid L2 Classroom Group Work

Teppo Jakonen[1,*], Heidi Jauni[2], and Olcay Sert[3]

[1]School of Languages and Translation Studies, University of Turku, Arcanuminkuja 1, FI-20014, Turku, Finland
[2]Language Centre, Tampere University, Tampere, Finland
[3]School of Education, Culture and Communication, Mälardalen University, Västerås, Sweden

[*]Corresponding author: Teppo Jakonen. Email: teppo.jakonen@utu.fi

This article addresses the need to better understand interactional asymmetries, challenges, and solutions in implementing synchronous hybrid language teaching. We investigate video-recorded peer interactions in a higher education language teaching context in which a student uses a telepresence robot, a remotely moveable videoconferencing tool, to participate in small-group task work in L2 English together with students who are physically located in the language classroom. Drawing on multimodal conversation analysis, we examine how the geographically dispersed peer group achieves, maintains, and repairs their joint attention on task-relevant learning materials as they are accomplishing a task, and how this kind of referential interactional work enables their co-operation as a group. Based on the analysis, we argue that in synchronous hybrid learning there is a need to reflexively adjust interactional practices to secure an intersubjective understanding of learning tasks and their progressivity. The findings also suggest that sensory and interactional asymmetries should be taken into account when developing and implementing synchronous hybrid learning environments that aim at equality of opportunities regardless of the participation mode.

1. Introduction

Video-mediated language teaching has increased significantly during the past few years. On the one hand, video-conferencing tools and platforms have become more suitable for educational needs because of technological innovations and features such as screen-sharing and collaborative text editing tools. At the same time, the recent coronavirus disease 2019 (COVID-19) pandemic forced many institutions and teachers to find ways to provide video-mediated 'emergency remote teaching' (Hodges et al. 2020) and design new instructional approaches. This has spurred large-scale research interest in practices of remote teaching in the fields of Applied Linguistics and Educational Sciences (see e.g. studies collected in González-Lloret et al. 2021; Tao and Gao 2022).

© The Author(s) (2024). Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

While a considerable literature on videoconferencing-based remote teaching is rapidly accumulating (e.g. Bozbıyık and Morton 2022; Çimenli et al. 2022; Malabarba et al. 2022), less attention has been paid to so-called synchronous hybrid educational contexts in which some students remain physically present in a classroom and others use videoconferencing to participate remotely in teaching in real-time. A recent systematic review of studies published in English identified only 47 studies reporting on synchronous hybrid teaching across different educational contexts and domains (Raes et al. 2020). The fact that research on synchronous hybrid instruction is still very much emerging is also visible in the existence of other concurrent labels used to refer to the phenomenon, such as *Here or There (HOT) instruction* (Zydney et al. 2019) and *hybrid-flexible* or *HyFlex teaching* (Beatty 2019). Thus far, the handful of studies that have investigated synchronous hybrid language education have tended to focus on participants' experiences using self-reports and interview data (e.g. Kohnke and Moorhouse 2021; Ulla and Espique 2022), leaving largely unaddressed the question what kinds of interactional challenges relate to the approach (but see Wang and Luo 2021; Neubauer 2022).

In this article, we address the apparent research gap around interactional practices in synchronous hybrid instruction, which is particularly acute in the context of language teaching, and respond to the call by Raes et al. (2020) for more multimodal investigations of actual instructional practices. As a field, Applied Linguistics is characterized by an interest in the relationship between language(s) and real-life concerns, and a drive to contribute societally by identifying solutions to such concerns. Aligning with this, the purpose of this article is to demonstrate some interactional challenges involved in synchronous hybrid instruction and to explore how participants adapt their interactional practices to the affordances and constraints of new kinds of technological platforms and tools. We address these concerns by reporting our observations on the use of telepresence robots as a tool to offer casual and individualized possibilities for remote participation in modern language classrooms in a higher education context in Finland.

In brief, telepresence robots are mobility-supplemented and remotely controlled videoconferencing tools that are used for hybrid meetings in offices, classrooms, and health care. These robots are typically equipped with cameras, screens, speakers and microphones, and wheels that enable their movement. When used in a classroom context, a remote student can operate the robot via an internet connection, move it around the classroom, control what is visible via the robot camera, and interact with locally present classroom participants. In comparison to other technologies for hybrid teaching, the mobility of a telepresence robot thus reconfigures the 'remote gaze' by enhancing the remote student's autonomy to decide and control what she sees and attends to in the physical classroom space. The purpose of this article is to investigate how geographically distributed participants sustain joint attention on task-relevant materials and the on-going task itself in this form of robot-mediated communication. By drawing on the methodological framework of multimodal conversation analysis (CA), we explore students' observable and interactional 'attention work' (Kunitz 2018) in a synchronous hybrid task context. Using single case analysis (e.g. Waring 2009), we analyse in detail an extended episode that shows how a group of four co-present students and one remote student make sense of group task instructions displayed on the classroom whiteboard in a higher education English as a foreign language classroom in Finland. Our study is guided by the following research questions:

[1] How do participants in synchronous hybrid education achieve joint attention on a textual artefact that is only located in the physical classroom?
[2] How do classroom-based participants make sense of the remote student's technology-mediated bodily-visual conduct as displays of task engagement and understanding?

2. Background

2.1 Joint attention and its relevance in pedagogical interaction

As Kidwell and Zimmerman (2007) argue, social interaction requires from participants an ability to coordinate their focus of attention with each other. This includes not only monitoring and

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

making sense of what the other participants are orienting their attention to during interaction, but also establishing mutual gaze 'between two or more co-present parties, or jointly to some entity in the local environment' (Kidwell and Zimmermann 2007: 609) when it becomes relevant. In this sense, joint attention is not only (or even primarily) a brain-internal phenomenon but also a rather practical task and an interactional and collaborative achievement that provides a basis for cooperative action and shared understanding (i.e. intersubjectivity).

Joint attention can be defined as a process of sharing experiences of observing an object or event (Tomasello 1995); in joint attention 'two or more people simultaneously attend to a single stimulus, together, each conscious that the experience is shared' (Enfield 2007: 400). In face-to-face interaction, gaze is a key resource for establishing and maintaining joint attention on environmental objects, structures, and space (e.g. Goodwin 1981). Participants routinely monitor each other's gaze direction on the basis of their head orientation, treating their gaze as a 'visible index of attention' (Hoey 2020) and engagement. By providing clues about whether or not a participant is momentarily engaged in the on-going interactional activity, gaze behaviour has also been found to configure turn organization, sequence organization, and turn-taking (e.g. Goodwin 1980; Kidwell and Zimmerman 2007; Auer 2021).

Many teaching and learning activities rely crucially on the coordination of participants' visual attention, be the focus of attention on specific participants, a learning material, or an instructional object. A growing number of interactional studies have explored such attention-organizing practices in pre-schools and primary schools (e.g. Lowi 2013; Nussbaum 2021; Strid and Cekaite 2022), secondary education (Kääntä 2014), and university contexts (Kunitz 2018). One strand of such CA work has explored the role of gaze in managing the participation framework and students' displays of engagement. These studies have shown that mutual gaze configures the way teachers allocate turns to students (e.g. Lauzon and Berger 2015), and that a student's gaze direction can be treated as a display of willingness (Mortensen 2008; Evnitskaya and Berger 2017; Ro and Burch 2020) or unwillingness (Sert 2013; Ishino 2022) to participate in instructional interaction. In one of these studies, Waring and Carpenter (2019) investigated how a teacher used gaze shifts during or immediately after accepting a student's contribution to manage students' attention and recipiency in an adult English language classroom. The authors argued that the teacher's gaze ensured that the larger pedagogical project of learning was implemented with efficacy (p. 137). In another study, Sert (2019) focused on sequences during classroom interaction in which students solicit the teacher's gaze to establish mutual orientation and maintain their participation. The study reveals how the teacher's responsive gaze becomes consequential in promoting student participation, as it is through this mutual gaze that student engagement is maintained. Similarly, in a study involving interactions between a tutor and a child, Carpenter (2016) investigated patterns of mutual orientation and gaze and showed that gaze can be used to direct a co-participant's attention. In short, these studies suggest that teachers and students use gaze direction as a resource for (re)configuring the classroom participation framework.

Another area of classroom-based CA research that informs our work has explored bodily-visual practices of scrutinizing and making sense of different kinds of learning materials as participants work on emergent *learnables* (Eskildsen and Majlesi, 2018). Ways of looking at materials can work as a resource for managing epistemic asymmetries (Kääntä 2014; Jakonen 2015) and be part of the co-construction of task-specific understanding in pedagogical activities. A study by Kääntä (2014) showed how students deploy gaze shifts and 'intense gaze' to mobilize co-participants' attention on task items shown on the board to initiate correction sequences. In another study, Strid and Cekaite (2022) investigated young children's peer interactions in a Swedish pre-school setting and showed that by using multimodal resources, children direct each other's attention and invite recipient gaze on material referents in the shared space. Thus, gaze direction can be a significant resource for mutual monitoring of co-participants' attention.

A particularly relevant notion to our study is what Kunitz (2018) has termed 'collaborative attention work' (see also Sert and Amri 2021). She documented how a group of university students working together on a classroom task mobilized other participants' attention to linguistic

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

forms visible on a computer screen in their shared environment. Kunitz (2018) argued that through collaborative attention work 'participants discover, co-construct, and act on emergent and student-selected objects of learning' (p. 64). In our data, students' attention work is mediated by videoconferencing, which configures participants' visual practices.

2.2 Gaze and visual asymmetries in video-mediated interaction and mobile robotic telepresence

Video-calling involves interaction between participants with 'non-mutual realities' (Ruhleder and Jordan 2001). Their realities differ not only with respect to how transmission delay, or 'lag', shapes turn-taking and sequence organization (Olbertz-Siitonen 2015), but also the ways in which visual and embodied conduct gets transmitted via cameras and screens. Compared to face-to-face settings, this complicates the use of gaze and gaze direction as a resource for interaction in video-mediated environments. As Luff et al. (2003: 55) observe, a characteristic feature of video-mediated interaction is that conduct is 'fractured from the environment in which it is produced and fractured from the environment it is received'. This can be seen in how different kinds of referential and deictic practices—which often involve embodied actions such as pointing—can be difficult to accomplish during a video call in exactly the same manner as in face-to-face situations. Thus, when a participant makes a pointing gesture during a video call, the other recipients do not necessarily see the target of the gesture if it is situated outside the view that the video camera transmits from the speaker's location. Identifying a remote co-participant's gaze direction can be equally cumbersome because of the mediating effect of the video camera, which provides the appearance of eye contact only when a participant is looking at their camera directly (e.g. Seuren et al. 2021).

An even more consequential asymmetry relates to the fact that in many videoconferencing set-ups, we are confined to seeing the other participant's environment either through the lens of a fixed camera (e.g. some classrooms equipped for hybrid education) or through the way the recipient of our video call positions their laptop or phone during interaction. Such a visual asymmetry can cause problems of intersubjectivity when participants need to sustain joint attention on objects. However, CA studies have shown that there are ways to tackle such constraints. Joint visual attention to objects can be secured through various kinds of showing practices (Licoppe and Morel 2014; Licoppe 2017) or visibility checks that enquire whether the recipient sees an object (Jakonen and Jauni 2021). Screen-based objects can also be spatially 'mapped' on computer screens through combinations of embodied actions and linguistic elements (Hjulstad 2016) and virtually pointed at with the mouse when using a shared digital workspace (Olbertz-Siitonen and Piirainen-Marsh 2021). Importantly, videoconferencing set-ups also differ with regard to who controls what each participant sees. In a video call between two participants using a phone or a laptop, each participant controls the other participant's visual field by way of positioning their device (and its camera) to show a particular 'slice' of their environment. Typically, participants tend to arrange themselves in a configuration of 'talking heads' (Licoppe and Morel 2012), showing each other their head and torso. However, visual control plays out very differently when using remotely moveable videoconferencing tools such as the telepresence robot that we investigate in this article. The participant connecting to a telepresence robot can remotely move the robot, which may be located for example in a healthcare facility (Due 2021), at home (Boudouraki et al. 2022) or in a classroom (Jakonen and Jauni 2021, 2022). The ability for movement enables one to 'look around' (Boudouraki et al. 2022) remotely and to reconfigure one's point of view with much more autonomy over the visual field than traditional videoconferencing tools allow. In contrast, the participants who interact with the user of the robot do not have such ability and instead typically see the robot-mediated remote participant's 'talking head'. In this sense, the telepresence robot can at least alleviate—and potentially even reverse—some of the visual asymmetries between classroom-based and remote students, as we will also seek to show in this study.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

3. Data and methods

3.1 Data and analytical focus

Our analysis draws on a broader database of approximately 15 hrs of video-recorded L2 teaching (English, Swedish, German, and Finnish) in the Finnish higher education context in which 1–2 students participate remotely in classroom-based teaching with the help of a Double 2 or Double 3 telepresence robot (as illustrated in Figure 1). During the lessons, the Double robot was in the classroom, and either an individual student or a pair of students connected to the device via the internet and participated in the lesson using the robot's videoconferencing features (camera, mic, speakers, wheels). The two robot models (2 and 3) have slightly different features and capabilities in terms of navigation and cameras. For our focus in this article, the central difference is the zoom function in Double 3, a feature that was not yet available in Double 2, which the participants use in the data that we report here. For the present analysis, a key technological feature is the wheels, which allow the remote student to move the robot and thereby 'move around' in the classroom physical space. The mobility thus provides the remote student the possibility to control their visual field that the robot's camera shows them on the computer screen in some remote space.

Our data can be seen as a series of qualitative naturalistic experiments (Heath and Luff 2018) in the sense that we were interested in how participants would use a new technology in a real-life educational context. We made the robot available to interested teachers and students and instructed them on how to use it, but, apart from that, let the participants decide how to include it in the lessons. One practical reason for this was also our wish to see if such a robot could be used as a tool for relatively casual and *ad hoc* synchronous hybrid instruction that would not require much advance preparation from the teacher. Our data contain lessons in which students participate remotely because of a temporary injury, travel, or a study abroad period, but also

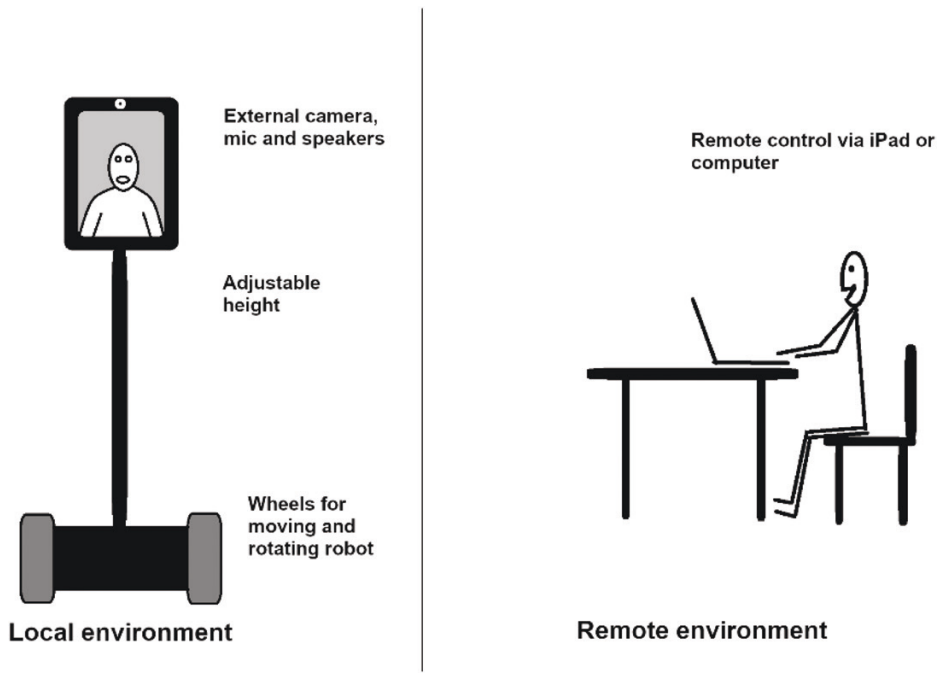


Figure 1. A double telepresence robot and its remote user (reproduced from Jakonen and Jauni 2023, CC BY 4.0 International License).

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

lessons during which the teachers and students take turns to experiment robot-mediated participation from another room on campus.

Going through the broader dataset, we became interested in how participants can achieve a joint visual attention on the material environment, although they do not share the same physical space (classroom) and had very asymmetrical sensory access to it. For the purposes of this study, we began to investigate situations in which the 'local' classroom-based participants observably orient to the remote student's visual attention, his 'looking at' and 'seeing' objects in the classroom space. In this article, we demonstrate our key findings by analysing a single extended case of small-group interaction during which joint attention on the task instruction becomes relevant and consequential for task accomplishment. Our main analytical interest relates to how the participants make sense of the remote student's visual behaviour by orienting to aspects of the robot's positioning and movement.

3.2 Single case analysis in CA

Conversation analytic studies can be collection-based (e.g. Çimenli et al. 2022; Jakonen and Jauni 2022), longitudinal (e.g. Pekarek Doehler and Berger 2018), and micro-longitudinal (Greer 2016; Sert 2017; Ishino 2018). They can also draw on single cases: a method within CA research referred to as single case analysis (SCA, e.g. Waring 2009; Balaman and Sert 2017). SCA illuminates interactional phenomena within a relatively short sequence of interaction through a very detailed investigation of an activity, and one of its purposes is to uncover an aspect of interaction which had not previously been noticed but is institutionally salient (Waring 2009). SCA allows analysts to 'uncover the orderliness of a single sequence of talk in order to make a case for the phenomenon under scrutiny' (Filipi and Chuang 2023: 244).

Like in other ways of carrying out conversation analytic research, single case analysis starts out with an 'unmotivated looking' of data to identify emergent phenomena for further examination (ten Have 2007). This is how we located the extended episode which we analyse in detail in Section 4. Following repeated viewing of the episode, presenting it at data sessions, and annotating further multimodal, prosodic, and linguistic details in the transcription, we observed a number of analytical phenomena that relate to visual challenges in synchronous hybrid task-based interactions. While our single case illuminates a very particular interactional and technological setting, we argue that the mediated nature of gaze, which is an inevitable aspect of any kind of video-mediated teaching, has relevance for understanding practices for achieving joint attention and understanding of tasks in other kinds of remote and hybrid environments. Because of this and the exploratory nature of the study, we chose SCA to present a rich description of the intricacies of interaction in the present context, with the hope that it can also lead to building larger collections in the future.

3.3 The task

The episode we analyse comes from an English lesson for technology students at a Finnish university, from a course that is required for a Bachelor's degree and that focuses on improving students' oral skills relevant to their field of study. The episode is part of a larger task that extends over five 90-minute lessons called the *Expert Panel*. In this task, groups of 4–5 students prepare to conduct a 20-minute panel discussion on their chosen field-relevant topic. The episode comes from the third lesson, during which the focal group is supposed to prepare their expert panel roles, discuss what they know about the topic, determine each member's area of expertise in the panel, predict possible audience questions, and prepare answers to those questions.

We knew that students would be working in small groups for the whole lesson and used that information to position cameras in the classroom space (see Kimura et al. 2018), placing one camera at the front and another at the back of the classroom. The two cameras aimed to capture what happened in the classroom space, including participants' multimodal actions and relevant material artefacts such as the whiteboard. We have used CA conventions to transcribe talk (Jefferson 2004) and embodied conduct (Mondada 2018), protecting participants' identities

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

through pseudonymization of names, removing references to the educational institution from the transcript, and by treating transcript images with pencil sketch filters (see online Supplementary Appendix).

4. Analysis: beginning a group task in a synchronous hybrid setting

In this section, we show how a geographically distributed student group mobilizes and negotiates joint attention on a task-relevant material artefact in a situation in which the very sharedness of physical space is problematic as one participant's presence in it – and gaze on its features – is video-mediated. The artefact is a task instruction slide shown on the classroom whiteboard (Figure 2). The extracts demonstrate how participants' joint attention on the slide serves the purposes of establishing understanding (Amri and Sert 2022) of the on-going pedagogical task and for sustaining its progressivity. We focus the analysis on how 'looking' (direction of gaze), 'seeing' (individual perception that can nevertheless be interactionally oriented to) and participants' joint attention are treated as a relevant part of task work through talk and technologically mediated embodied actions.

The group of students consists of five people. Four students (Ben, Heikki, Ville, and Jani) are physically present in the classroom, and Kimmo participates remotely via the telepresence robot, as illustrated in Figure 2. Before Extract 1, the group of students have selected Ville as the group's panel leader, and Ben has instructed him about what Ville's role entails in the eventual panel discussion. In Extract 1, we see how the group's focus gradually shifts from a previous joking sequence (not shown) back on task as Ben orients to the instructions on the whiteboard (lines 2–3) and begins to organize the group's preparation for the panel discussion by assigning task responsibilities.

Ben's turn at lines 2–3 breaks a silence, during which his gaze has been directed towards the whiteboard (transcript image 1.1a) for several seconds and he has been biting his lips. The turn provides the recipients a 'noticing' (Schegloff 2007) in the sense that Ben explicitly announces that he has registered something in the environment, identifies its location, and provides an embodied demonstration of intensively scrutinizing the environment (see Kääntä 2014). After a brief silence, Ben proceeds to read aloud one of the 'points' that he has just announced, a task

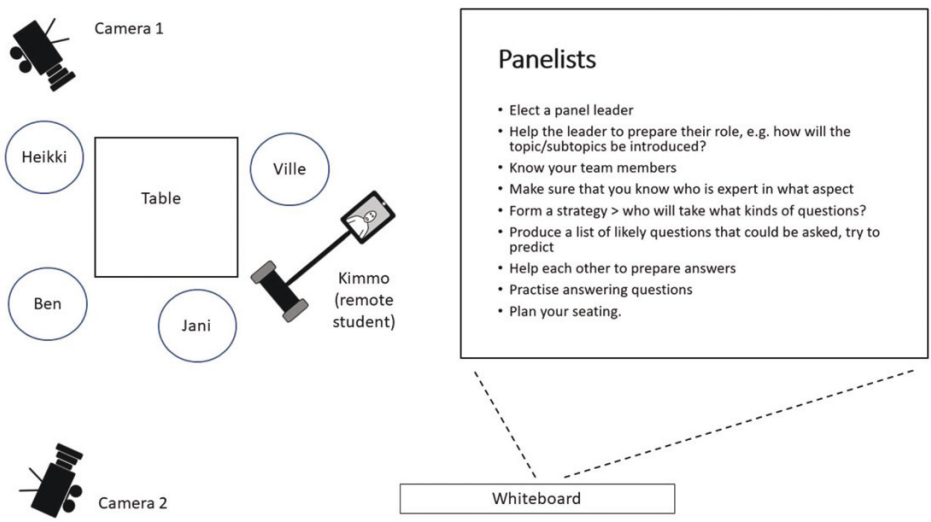


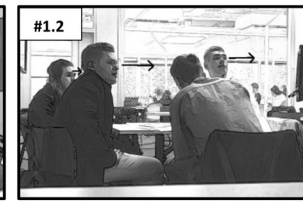
Figure 2: The classroom layout and task instruction slide.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

(1) Making joint attention and visual space relevant.

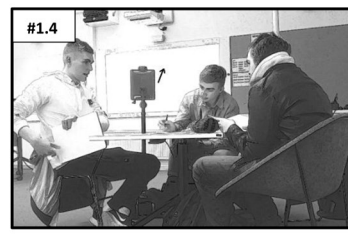
```
01          (3.0)
    ben     >>gz twd whiteboard, bites lips->
02  BEN       *ye::p,# (0.3) ther- there is some- °(kind)° err (0.4)
            ->*gz twd whiteboard->
    fig             #1.1a/b
03          points on the (.) whiteboard.
04          (0.7)^¤
    vil           ^turns gz and upper body twd board->
    hei            ¤gz twd board->
05  BEN     err "produce a #list of likely questions that could be asked,=
    fig                     #1.2
```






```
06          =t- ^try ¤to predict,"*
    vil       ->^
    hei             ->¤glances at ville, gz to the desk->
    ben                  ->*
07          (0.8)&(1.6)*(0.2)*
    kim           &turns the robot twd board->
    ben                *glances at heikki*
08  BEN     ^can (.) #(someone-)*^ (0.5) do you: Ville* have paper and a pen.&#
    vil     ^looks at robot------^
    ben                          *gz twd own bag-------*gz to desk->
    kim                                                          ->&
    fig               #1.3                                          #1.4
```






```
09          (0.3)
10  BEN     can you type &(it).
    kim                  &readjusts robot->
11          (0.8)&(0.3)*(0.8)    *(0.2)
    kim       ->&
    ben                *gz-ville*
12  VIL     of [course. ]
```

item displayed on the whiteboard (lines 5–6, see also Figure 2). In effect, such reading aloud is a practice for transitioning from one task item to another (Tůma 2022).

When noticing's occur in a sequence-initial position, such as here, they tend to mobilize the recipient's attention on the noticed feature and make the sharing of the noticing interactionally relevant (Schegloff 2007: 74, 219). Here, two on-site students, Ville and Heikki, soon align with Ben's noticing by turning their gaze from the desk towards the screen (Image 1.2) at the end of

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

the silence at line 5, just before Ben begins to read aloud the text on the whiteboard. The fourth on-site student, Jani, continues writing and maintains his gaze on the desk. The remote student, Kimmo, reacts to Ben's noticing later during the silence at line 7, after Ben's read-aloud has come to completion. Kimmo's reaction is a technology-mediated embodied response as he turns the robot around anticlockwise so that by the end of line 8 it faces the whiteboard and allows Kimmo to visually orient to the 'noticed' feature of the classroom environment (image 1.4). The turning robot is also visually monitored by Ville (image 1.3)

Ben proceeds with organizing the group's responsibilities in relation to the focal task item. Ben quickly glances at Heikki before beginning a turn that appears to request a volunteer (*can someone*). Shifting his gaze toward his own bag on the floor and then on the desk, he cuts it off and self-repairs the turn so that it queries whether Ville has a pen and paper. Without waiting for a response, Ben then requests Ville to act as the group's note-taker (line 10). Both turns by Ben are thus directly occasioned by the earlier noticing and the formulation of the noticed task item because producing a list requires writing equipment. Being in possession of a pen and paper can provide an accountable reason for being selected as a note-taker, and, conversely, a glance at one's own bag can display that Ben does not have a pen.

All in all, Extract 1 illustrates the kinds of technological contingencies that hybrid meetings have for visual conduct. While local participants can react to noticings and topicalizations of an environmental feature through embodied acknowledgments such as a quick glance that conveys a sense of having a shared and unproblematic visual access to the feature, the mobilization and interactional display of video-mediated attention can be more cumbersome and time-consuming. Here, Kimmo's response to the noticing takes the form of turning the robot and comes considerably later than the reactions by Ville and Heikki. Yet, turning the robot in this particular direction does constitute a competent and autonomous technologically mediated action. However, such different temporal dimensions in practices of looking at objects can also give rise to action misalignments, as Extract 1 illustrates. Extract 2 shows how the misalignment continues as Kimmo disengages from the group interaction and moves the robot closer to the whiteboard, and how the other students orient to such conduct.

As Ben, Jan, and Ville coordinate note-taking responsibilities at lines 12-16, the remote student is no longer in a 'face-to-screenface formation' (Due 2021) with any of the on-site students. With the robot screen facing away from the group, the remote student makes two separate movements with the robot: he first drives it forward a bit (lines 13–14) and then turns it slightly anti-clockwise (lines 16–17). Immediately after that, he drives the robot forward until it is less than two metres away from the whiteboard (in a position shown in image 2.2b), stopping the movement at the end of line 23.

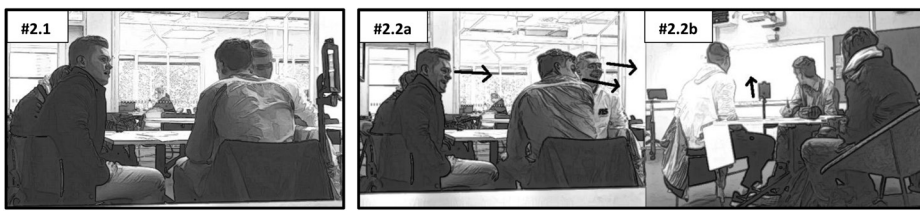
The classroom-based students all orient to the robot's movement away (and Kimmo's disengagement) from the group. Ben monitors the robot visually already at line 16 (see image 2.1), as Ville is responding to Ben's teasing reformulation of his earlier request to act as the note-taker (*can you type it*, line 10, vs. *have you the ability to type it*, line 13). Ben first peeks at the robot on the left side of Jan, who is seated next to him, and then on his right side. By summoning the remote student (line 18), Ben also opens a parallel line of talk with the negotiation about note-taking, which Ville continues in overlap by way of verbalizing his search activity (line 19). Jan looks over his shoulder at the robot who is moving away from him, and Ben, whose gaze has been fixed on the robot, breaks into a smile (line 19). Ville and Heikki's faces are also oriented towards the robot and the whiteboard (as shown in image 2.2a). After a 1.5s silence during which Ben smilingly looks at the robot, his smile transforms into audible laughter (line 21), which is soon joined by Ville (line 23). Jan also summons the remote student for the second time (line 22) but does not secure Kimmo's recipiency, who instead continues to drive the robot towards the board.

Turning and driving the robot so that its camera can relay an object that is being talked about is recognizable as a practice of moving to see something. In order to look in detail, the remote student needs to leave the group because the robot camera cannot be zoomed in on the object (in Double 2). The accountability of driving closer to the board is in that it indicates

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

(2) Seeking visual access by going (remotely) to the board.

```
12 VIL   of [course. ]
13 BEN      [have you] the (.) ability &to type (°it°).
   kim                                  &drives robot forward->
14       (0.8)&
   kim      ->&
15 JAN   *[can you (.)]* write, (0.4) *[on paper.]
16 VIL    [(need pen.)]                [I  can  w]rite on &paper.
   ben   *gz-robot-----*              *gz-robot->
   kim                                                 &turns robot->
17       #(1.0)&(0.4)
   kim       ->&drives robot twd board->
   fig   #2.1
18 BEN   [Kimmo,]
19 VIL   [let's ] see if I have, (1.1) +a *<pencil> here,
   jan                                  +turns to look at robot->
   ben                                     *smiles, looks at robot->
20       (1.5)
21 BEN   hhh hehehehe[he? [.hhhh
22 JAN               [Kimmo hhh
23 VIL                    [hehehehe#&
   kim                              ->&
   fig                               #2.2a/b
```



to the co-participants that turning the robot has not yet guaranteed visual access to the focal textual object. Performing such close scrutiny is a task-relevant action, contingent on the earlier noticing, but it has interactional implications in that it requires Kimmo to momentarily leave the group. As Extract 3 shows, the on-site students orient to this as their shared laughter is followed by jokes that categorize the robot's movement away from the group as an embodied display of non-task-related action.

The visual attention to the robot and the joint laughter in Extract 2 pave way for a series of playful comments on, and possible scenarios for, the robot movement. First, Heikki's comment at line 24 attributes a human intention to what the movement 'looks like' and is treated as a joint laughing matter by Ben, Ville, and Jan. This occasions another joke as Ben shifts footing and produces a turn as if he were speaking as Kimmo (I must go, line 28). Such a turn is an instance of hypothetical enactment (Holt 2007) and is hearable as an allusion to going to the bathroom. As Holt (2007) has shown, hypothetical enactments are typically used for joking or teasing and are often made recognizable as such by the use of laughter tokens, as is also the case here. The third jocular reference to the moving robot comes again from Ben, who produces an imagined scenario through laughter at lines 33, 36.

While all three comments categorize the robot movement as noticeable and deviant conduct, it is noteworthy that the comments operate with very different referential practices regarding who—or what—is being treated as the actor. The word 'it' used by Heikki at line 24 can be heard as a reference to the metal-bodied telepresence robot. In contrast, Ben uses an address term ('Kimmo'), which treats the remote student as the agent who is driving 'it' (line 36). Moreover,

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

(3) Making sense of robot movement.

```
24 HEI   look like (.) like it (would leave me.)
25 Ss    ((laughter))
26       (0.9)
27 VIL   hhh hehe[hehehe
28 BEN           [I m(h)ust go hh
29 ?     hehe[he
30 BEN       [(for-)
31       (1.3
32 VIL   (or [so:)
33 BEN       [if I open the door Kimmo will get the fuck [out of here.  ]
34 VIL                                                   [do you have pa]per=
35       =fo[r me.
36 BEN       [(to) d(h)rive it [around ((university name)) [hehehehe?       ]
37 VIL                         [I don't ha- (.)            [I don't have any=]
38       *&=paper w- (.) with me.
   ben   *looks at robot->
   kim    &turns robot->
```

Ben's turn is linguistically designed as if Kimmo were physically in the same room with the others, even if in reality he is already 'out of here' (line 33). Through such referential practices, the co-present students treat the remote co-participant as a hybrid assemblage consisting of a robot and a human (see also Due 2021): Kimmo's presence as well as the robot's agency are oriented to as ambiguous.

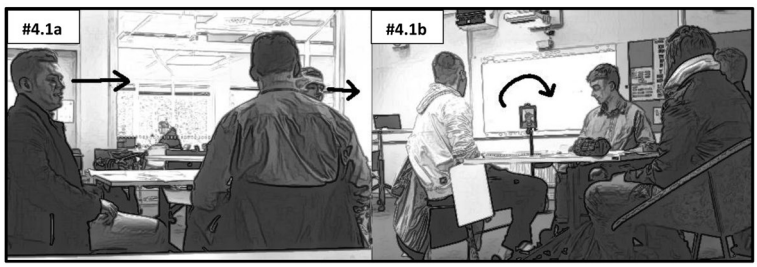
In Extract 4, we see how the participants resume their work of ensuring the visual availability of the task instructions on the whiteboard after the brief joking sequence in the previous extract. At line 38 of Extract 3, Ben disengages from the joking sequence by shifting his gaze towards the robot, which is still positioned in front of the screen but is beginning to turn around. As Extract 4 begins, Ben seeks Kimmo's attention and provides a series of referential directives to enable Kimmo to locate the instructions written on the whiteboard and to look at them.

Monitoring the robot's movement visually (image 4.1a), Ben summons Kimmo at line 41 and repeats the summons (line 43) shortly afterwards. Kimmo responds to the second summons as he has just turned the robot screen towards the group and is beginning to drive it back to the group's desk (line 45, image 4.1b). Now that Kimmo's recipiency is secured, Ben directs him to 'check' a task item on the whiteboard (lines 46–47, 49). The turn contains multiple pauses and instances of self-repair as well as a pointing gesture towards the screen. In overlap, Ville produces self-talk as he is searching his backpack, which may occasion some of Ben's disfluencies. The pauses and restarts may also relate to the very specific referential work Ben is doing by combining verbal turn elements to indicate the location of the task item ('fourth point from the bottom of the page'). This is done with a pointing posture that is visually maximally noticeable, as he extends his right arm at the level of his head (image 4.2). His arm positioned this way, Ben does a series of subtle gestures with his index finger: he fully extends his index finger (image 4.3), 'taps' downwards in the air during the first hesitation token 'err' at line 46 and both further 'errs' at line 47 (image 4.4), as if depicting the target of a word search ('bottom'). When Ben eventually utters 'bottom of the page', he stretches out his index finger so that it points slightly upwards (image 4.5), a trajectory which may be seen as a depiction of the direction from which one ought to be counting the bullet points. Altogether, Ben's multimodal directive turn treats the recipient as someone who has not registered the referred-to textual object when Ben read it aloud earlier (in Extract 1) The directive

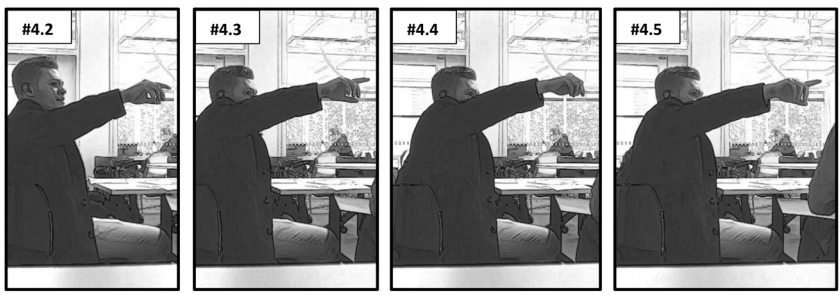
Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

(4) Instructing looking at the task instruction.

```
39 JAN   uh (.) do you (.) want some paper.=
   ben   >>looks at robot->
   kim   >>turns robot->
40 VIL   =I want some pa[per.]
41 BEN                  [ KIM]*mo,*
   ben                      ->*lifts rh mid-air, gz-robot*lowers rh->
42       (0.8)
43 BEN   Kimmo,*
   ben       ->*
44       (1.0)
45 KIM     &yeah#
   kim   ->&drives robot back to group->
   fig          #4.1a/b
```




```
46 BEN   *check *the, (.)# *err- (0.8) fourth, (0.8) err *&point# from=
   ben   *.....*points----*taps twice, then holds point-*extends finger->
   kim                                                    ->&
   fig                  #4.2                                  #4.3
47       =the- *(.) err [(0.7) the err::# (0.4)] *bottom of the (.) #page (0.3)
48 VIL                  [(err I have a ( ))    ]
   ben       ->*multiple taps--------------------*extends finger->
   fig                                  #4.4                    #4.5
```




```
49 BEN   on the *whiteboard *
   ben        ->*lowers hand*
50       (1.5)
51 BEN   check,=*
   ben          *points->
52 KIM   &=sorry?= &
   kim   &leans fwd&
53 BEN   =the* (.) four- fourth point# from the bottom of the w- err
   ben   ....*points and taps repeatedly->
   fig                              #fig4.6
```

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

(4) Continued

```
54        s- &sheet on the whiteboard=do you* see it.
     kim      &reverses the robot, makes a U turn->
     ben                                  ->*holds point->
55        (1.8)
56   BEN    *the four-  *[(.) fo- (.)*]the *#fourth* point,*
57   KIM                 [(wait little]
     ben   ->*extends arm*retracts arm*.....*points *,,,,,,,*
     fig                                    #4.7
```






```
58   BEN     &.hhh pro- produce a list of &likely questions.
     kim   ->&                            &drives robot fwd to the board->
59        (2.2)&(3.5)&(1.5) &(0.6)&(2.2)
     kim     ->&     &drives&     &turns around->
```

juxtaposes the remote student with the on-site students, whose visual access to the task item it takes for granted and as already established.

No response to Ben's directive is forthcoming by Kimmo (line 50), and Ben begins a turn that is hearable as a repetition of his directive (line 51). At the same time, Kimmo initiates an open class repair (line 52), simultaneously with which he can be seen on the robot screen to be poking his head forward. This indicates that the repair addresses a problem of hearing. Ben brings the emerging turn to completion at lines 53–54. Similar to his earlier directive, Ben points towards the whiteboard throughout the turn. In contrast, however, the turn is less disfluent and incremented with a latched-on visibility check (see Jakonen and Jauni 2021) at line 54.

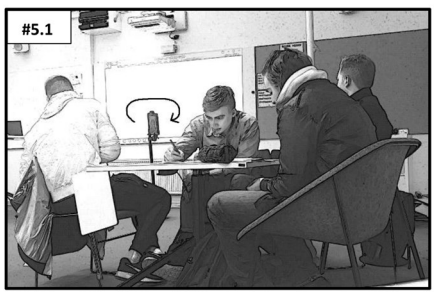
Kimmo begins to reverse the robot away from the group while Ben's turn is still in progress (line 54), demonstrating that he understands the directive to concern the whiteboard. In one continuous trajectory, he first reverses straight and then reverses and turns the robot anticlockwise until its screen faces the whiteboard, as illustrated by the difference in the robot's positioning between images 4.6 and 4.7. When Ben asks whether Kimmo 'sees it' (line 54), the robot screen is still facing the group, which does not yet enable visibility of the whiteboard for Kimmo. Both Ben and Kimmo orient to the project of manoeuvring the robot and the kind of delay that it presents for achieving visibility of the task item on the whiteboard. When Ben once more repeats the location of the item (line 56) to a retreating robot, he quickly extends and retracts his pointing arm, as if displaying hesitation through the lack of sustained gesture. In overlap, Kimmo appears to initiate repair to suspend the progressivity of the sequence. When Ben begins to read aloud the task item (line 58), the robot is already facing the whiteboard, and as the turn approaches its end, Kimmo drives the robot closer to the board. By not taking a turn during line 59, Ben is observably 'doing waiting' while the remote student manoeuvres the robot into a position that allows him visual access to the text on the whiteboard.

Towards the end of the silence (line 59), the remote student begins to turn the robot around towards the group. Extract 5 shows how he provides a sequentially relevant response to Ben's directive and visibility check by claiming that he can see the referred-to task item.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

(5) Claiming 'seeing' and realigning as a group.

```
60 KIM   yeah I can see it.#
   kim   >>turns robot around->
   fig                  #5.1
61 BEN   ye:s.
62       you can- (.) can you &come up (what) with &some (.) err questions.
   kim                   ->&                  &drives to group->
63       (2.7)&
   kim      ->&maneuvers robot by the desk->
64 KIM   about <our,> (0.5) [(presentation).
65 BEN                    y[e:s
66 BEN   ye::s (0.3) some (.) likely questions would (be) asked.
67       (2.4)
68 BEN   you can make a- &a (.) list, (0.4) on- err over &there if you (.) want
   kim               ->&turns robot-------------------&maneuvers robot->
69       if you have paper or something.&#
   kim                                ->&
   fig                                   #5.2
```






```
70       (1.3)&(0.4)
   kim      ->&reverses->
71 BEN   or &you can just type it &with your cell phone and send it to us=&
   kim    ->&                  &turns robot----------------------------&
72 JAN   =so do you need some (.) &paper Kimmo here [(.) you can& have (this)
73 Ss                                              [((laughter))
   kim                          &turns robot-----------------&
74 KIM   now/no (.) >thank you<
```

Turning the robot around in front of the whiteboard (image 5.1), Kimmo claims that he 'can see it' (line 60). As an action, this enables a re-alignment of the group so that Kimmo's participation in the task is secured. This is visible in how Ben uses the turn as an occasion to recruit Kimmo's effort (line 62), initially by assigning it as a task to him ('you can') but then self-repairing the turn into a request ('can you'). In addition to such ascertained intersubjectivity that group members are working on the same task item, Kimmo's re-alignment is also an embodied one, as he moves the robot back to the group's desk and thereby restores the group-based embodied participation framework (visible in image 5.2). Kimmo not only claims 'seeing' the relevant task item but also demonstrates it by requesting confirmation to a candidate understanding that he is expected to make questions about the group's presentation (line 64). This interpretation gets a confirmation from Ben (line 66).

All in all, Kimmo's claim and demonstration that he has seen the focal task item is pivotal for achieving a sense of working as a group and aligning to collaborate on the task of planning the group's future panel discussion. Via the robot, Kimmo is present as a group member whose contribution to the task may take some extra creativity from group members, as observable in

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

Ben's suggestions for how to make the list (lines 68–69, 71). At the same time, his presence is qualitatively different from the on-site students, to the degree that it can be turned into shared jokes (lines 72–74). While these jokes may appear as off-task behaviour, they are made possible by an intersubjectively shared understanding of what moving on with the task entails for the participants in this hybrid interactional learning environment.

5. Discussion and conclusions

In this article, we have used multimodal conversation analysis to investigate the details of 'collaborative attention work' (Kunitz 2018) as a hybrid student group is accomplishing an itemized language learning task, the instructions for which are displayed on a whiteboard in the physical language classroom. Instructional materials are important resources that language learners make relevant in talk and attend to in embodied ways as they navigate tasks (e.g. Hellermann and Pekarek Doehler 2010). However, in synchronous hybrid teaching, some task materials can also be asymmetrically available to a group of geographically dispersed students, as in the episode we have investigated. Reaching joint attention (Kidwell and Zimmermann 2007) can become procedurally complex when some participants have direct first-hand experience of the language classroom and its resources, and others see, hear, and move in the classroom remotely and in a mediated manner. In our case, the 'mediatedness' of video-mediated participation makes the material conditions of 'looking' and 'seeing' considerably different between classroom-based students and the robot-mediated student. The lengthy focal episode not only illustrates such complexities but also showcases how the students make sense of the situation and its asymmetries by adjusting their interactional practices in a reflexive manner.

In the present case, the task instructions displayed on the whiteboard are not the focus of students' continuous attention in the form of long-term sustained gaze on the screen. Rather, the instructions are made relevant and topicalized by the students as they progress through the itemized task (Extract 1). Bringing the instructions into the focus of attention requires referential work, and the students use talk, gaze, and gestures (Extracts 1, 4) to direct each other's visual attention to the whiteboard. These references allow the hybrid group to maintain its alignment in task work by making transitions from one task item to the other visible and by providing a context for further negotiation of task-related division of labour (Extracts 1, 2, 4, and 5). What is at stake for student participation when Ben makes the task instructions relevant in Extract 1 is therefore not so much scrutinizing their details together, but rather displaying to others that the instructions have been seen and registered so that the group can proceed to the next phase of the task. In the present case, 'seeing' requires just a quick glance towards the board from the classroom-based students, but from the robot-mediated remote student it takes a considerably longer, partly peer-instructed movement trajectory closer to the screen, which at the same time involves disengagement from the face-to-screenface formation (Due 2021). Therefore, even with the enhanced autonomy for remote gaze that the telepresence robot's mobility provides, for remote students, 'seeing' can still require a considerably more complex practice that may jeopardize their involvement and presence in a group. As we argued in Section 2, gaze and looking are central to participation and engagement in classroom interaction (see Lauzon and Berger 2015; Sert 2019), and through our multimodal sequential analysis, we have shown that remote engagement in a robot-mediated synchronous hybrid teaching may require mediation from participants who are physically present in the classroom. In this sense, joint attention can be a prerequisite of task progressivity, and without participants' mutual orientation to visual spaces, an intersubjective understanding of the task may not necessarily be established.

Our extracts have also shown that the classroom-based students orient to the robot's position and movement in the classroom space both in talk and through their embodied behaviour (gaze and laughter). Some of the jokes and laughter (Extract 3) seem to exhibit how participants encounter a new technology and make sense of how it potentially re-configures their social world and norms. However, the classroom-based students use the robot's position and movement trajectories as well as

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

its screen orientation as an index of what is likely to be within the remote student's attentional field (cf. Hoey 2020). In this sense, the robot screen is akin to a co-present participant's head orientation that provides clues about their moment-by-moment visual behaviour and attention. On the other hand, the robot's ability for movement is not only a way for the remote participant to independently reconfigure their point of view, but it can also be used to move from one place to the other or to transition from whole-class to group-based interaction (Jakonen and Jauni 2022). For classroom-based students, the robot's movement therefore poses a "Why this now?' question as they interpret the interactional purpose and consequences of the movement, such as in Extract 3 where the students (jokingly) treat the movement as something else than an attempt to see the whiteboard better.

Research on synchronous hybrid language education is largely uncharted territory, and our study has addressed the need for more research on multimodal instructional processes (Raes et al. 2020). Despite obvious technological challenges in this type of robot-mediated hybrid instruction, the geographically distributed focal participants demonstrate that they are in the end able to sustain their group-ness, ensure the recognizability of key task artefacts, and proceed with the collaborative task. As such, this calls for specific L2 interactional competences (e.g. Pekarek Doehler and Berger 2018) in designing one's actions according to the participation status (i.e. local vs. mediated) of the recipient. For instance, Ben's conduct (Extracts 4–5) demonstrates an awareness of a need to support the remote student's participation by specifying task instructions, directing the remote student's attention, and ensuring he has (visual) access to what is needed to adequately collaborate with the classroom-based students. Micro-level actions such as these help maintain a shared understanding of the task and its progression, and demonstrate that participants adapt their task work practices to the affordances and constraints of the socio-technological ecology of action. These actions are deeply multimodal, emerging through the situated and effective deployment of talk, gaze, gesture, and ways of navigating the material environment. Talk could surely be used to relay text-based task instructions, but it is not always enough. This also means that a key aspect in the design of meaningful synchronous hybrid learning environments is to try to alleviate asymmetries between face-to-face and remote participation by acknowledging that participation is a multimodal, social, and collaborative phenomenon.

Supplementary data

Supplementary material is available at *Applied Linguistics* online.

Funding

This work was supported by the Research Council of Finland [grant number 343480].

Notes on contributors

Teppo Jakonen works as Academy of Finland Research Fellow at the School of Languages and Translation Studies at University of Turku, Finland. His research explores the role of technology and materiality in social interaction in language teaching and learning activities. Jakonen's recent publications have appeared in journals such as *Modern Language Journal*, *Journal of Pragmatics*, *Computer Assisted Language Learning*, and *Classroom Discourse*.

Heidi Jauni works as a Development Manager at the Language Centre of Tampere University, Finland. Her research interests include conversation analysis, especially interaction in educational settings and mediated interaction. Jauni's recent work has appeared in *Classroom Discourse* and *Computer Assisted Language Learning*.

Olcay Sert is Professor of (English) language education in the School of Education, Culture and Communication at Mälardalen University. His research deals with classroom discourse, L2 interaction, and language teacher education. Sert's recent publications have appeared in journals such as *Cambridge Journal of Education*, *Modern Language Journal*, and *Linguistics and Education*.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

References

Amri, M. and O. Sert. 2022. 'Establishing understanding during student-initiated between-desk instructions in project work,' *Cambridge Journal of Education* 52/6: 667–89.

Auer, P. 2021. 'Turn-allocation and gaze: A multimodal revision of the 'current-speaker-selects-next' rule of the turn-taking system of conversation analysis,' *Discourse Studies* 23/2: 117–40.

Balaman, U. and O. Sert. 2017. 'The coordination of online L2 interaction and orientations to task interface for epistemic progression,' *Journal of Pragmatics* 115: 115–29.

Beatty, B. 2019. *Hybrid-Flexible Course Design*. EdTech Books.

Boudouraki, A., S. Reeves, J. E. Fischer and S. Rintel. 2022. 'Mediated visits: Longitudinal domestic dwelling with mobile robotic telepresence' in Proceedings of the 2022 CHI conference on human factors in computing systems, pp. 1–16.

Bozbıyık, M. and T. Morton. 2022. 'Transitioning between 'outside' and 'inside' knowledge in an online university EMI chemistry course,' *Applied Linguistics (Advance Access)* 44: 265–86.

Carpenter, L. 2016. 'Managing the participation of a young learner: A multimodal teacher practice,' *Working Papers in TESOL & Applied Linguistics* 16/2: 24–8.

Çimenli, B., O. Sert, and C. J. Jenks. 2022. 'Topic maintenance in video-mediated virtual exchanges: 'Rolling the Ball Back' in L2 interactions,' *System* 108: 1–21.

Due, B. L. 2021. 'RoboDoc: Semiotic resources for achieving face-to-screenface formation with a telepresence robot,' *Semiotica* 2021/238: 253–78.

Enfield, N. J. 2007. 'Social consequences of common ground' in N. J. Enfield and S. J. Levinson (eds): *Roots of Human Sociality: Culture, Cognition and Interaction*. Berg, pp. 399–430.

Eskildsen, S. W. and A. R. Majlesi. 2018. 'Learnables and teachables in second language talk: Advancing a social reconceptualization of central SLA tenets. Introduction to the special issue,' *Modern Language Journal* 102: 3–10.

Evnitskaya, N. and E. Berger. 2017. 'Learners' multimodal displays of willingness to participate in classroom interaction in the L2 and CLIL contexts,' *Classroom Discourse* 8/1: 71–94.

Filipi, A. and M. -S. K. Chuang. 2023. 'Chinese whispers: International Chinese students' language practices in an anglophone Higher Education context,' *Classroom Discourse* 14/3: 238–257.

González-Lloret, M., L. Canals, and J. E. Pineda. 2021. 'Role of technology in language teaching and learning amid the crisis generated by the covid-19 pandemic,' *Íkala, Revista de Lenguaje y Cultura* 26/3: 477–82.

Goodwin, C. 1981. *Conversational Organization: Interaction between Speakers and Hearers*. Academic Press.

Goodwin, C. 1980. 'Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning,' *Sociological Inquiry* 50/3–4: 272–302.

Greer, T. 2016. 'Learner initiative in action: Post-expansion sequences in a novice ESL survey interview task,' *Linguistics and Education* 35: 78–87.

Heath, C. and P. Luff. 2018. 'The naturalistic experiment: Video and organizational interaction,' *Organizational Research Methods* 21/2: 466–88.

Hellermann, J. and S. Pekarek Doehler. 2010. 'On the contingent nature of language-learning tasks,' *Classroom Discourse* 1/1: 25–45.

Hjulstad, J. 2016. 'Practices of organizing built space in videoconference-mediated interactions,' *Research on Language and Social Interaction* 49/4: 325–41.

Hodges, C. B., *et al*. 2020. 'The difference between emergency remote teaching and online learning,' *Educause*.

Hoey, E. 2020. 'Gaze' in J Stanlaw (ed.): *The International Encyclopedia of Linguistic Anthropology*. John Wiley & Sons.

Holt, E. 2007. '"I'm eyeing your chop up mind": Reporting and enacting' in E. Holt and R. Clift (eds): *Reporting Talk: Reported Speech in Interaction*. Cambridge University Press, pp. 47–80.

Ishino, M. 2018. 'Micro-longitudinal conversation analysis in examining co-teachers' reflection-in-action,' *System* 78: 130–47.

Ishino, M. 2022. 'Teachers' embodied mitigation against allocating turns to unwilling students,' *Classroom Discourse* 13/4: 343–64.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

Jakonen, T. 2015. 'Handling knowledge: Using classroom materials to construct and interpret information requests,' *Journal of Pragmatics* 89: 100–12.

Jakonen, T. and H. Jauni. 2022. 'Managing activity transitions in robot-mediated hybrid language classrooms,' *Computer Assisted Language Learning (latest articles)* 1–24. https://doi.org/10.1080/09588221.2022.2059518

Jakonen, T. and H. Jauni. 2021. 'Mediated learning materials: Visibility checks in telepresence robot mediated classroom interaction,' *Classroom Discourse* 12/1–2: 121–45.

Jakonen, T. and H. Jauni. 2023. 'Telepresent agency: Remote participation in hybrid language classrooms via a telepresence robot' in J. Ennser-Kananen & T. Saarinen (eds): *New Materialist Explorations into Language Education*. Springer, pp. 21–38.

Jefferson, G. 2004. 'Glossary of transcript symbols with an introduction' in G. Lerner (ed.): *Conversation Analysis: Studies from the First Generation*. John Benjamins, pp. 13–31.

Kääntä, L. 2014. 'From noticing to initiating correction: Students' epistemic displays in instructional interaction,' *Journal of Pragmatics* 66: 86–105.

Kidwell, M. and D. H. Zimmerman. 2007. 'Joint attention as action,' *Journal of Pragmatics* 39/3: 592–611.

Kimura, D., T. Malabarba, and J. K. Hall. 2018. 'Data collection considerations for classroom interaction research: A conversation analytic perspective,' *Classroom Discourse* 9/3: 185–204.

Kohnke, L. and B. L. Moorhouse. 2021. 'Adopting HyFlex in higher education in response to COVID-19: Students' perspectives,' *Open Learning: The Journal of Open, Distance and e-Learning* 36/3: 231–44.

Kunitz, S. 2018. 'Collaborative attention work on gender agreement in Italian as a foreign language,' *Modern Language Journal* 102: 64–81.

Lauzon, V. F. and E. Berger. 2015. 'The multimodal organization of speaker selection in classroom interaction,' *Linguistics and Education* 31: 14–29.

Licoppe, C. 2017. 'Showing objects in Skype video-mediated conversations: From showing gestures to showing sequences,' *Journal of Pragmatics* 110: 63–82.

Licoppe, C. and J. Morel. 2014. 'Mundane video directors in interaction: Showing one's environment in Skype and mobile video calls' in M. Broth, E. Laurier and L. Mondada (eds): *Studies of Video Practices*. Routledge, pp. 143–68.

Licoppe, C. and J. Morel. 2012. 'Video-in-interaction: 'talking heads' and the multimodal organization of mobile and Skype video calls,' *Research on Language and Social Interaction* 45/4: 399–429.

Lowi, R. 2013. 'Building understanding: The construction of joint attention in preschool,' *Issues in Applied Linguistics* 19: 85–117.

Luff, P., *et al*. 2003. 'Fractured ecologies: Creating environments for collaboration,' *Human–Computer Interaction* 18/1–2: 51–84.

Malabarba, T., A. C. O. Mendes, and J. de Souza. 2022. 'Multimodal resolution of overlapping talk in video-mediated L2 instruction,' *Languages* 7/2: 154.

Mondada, L. 2018. ''Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality,' *Research on Language and Social Interaction* 51/1: 85–106.

Mortensen, K. 2008. 'Selecting next speaker in the second language classroom: How to find a willing next speaker in planned activities,' *Journal of Applied Linguistics and Professional Practice* 5/1: 55–79.

Neubauer, D. 2022. *Chinese language instruction with novice learners: Target language topic development, engagement, and comprehension in online and hybrid classrooms*. Doctoral dissertation, University of Iowa.

Nussbaum, L. 2021. 'Showing attention: Student's turn co-construction and other-initiated other-repair' in D. Masats and L. Nussbaum (eds): *Plurilingual Classroom Practices and Participation*. Routledge, pp. 66–81.

Olbertz-Siitonen, M. 2015. 'Transmission delay in technology-mediated interaction at work,' *PsychNology Journal* 13/2–3: 203–34.

Olbertz-Siitonen, M. and A. Piirainen-Marsh. 2021. 'Coordinating action in technology-supported shared tasks: Virtual pointing as a situated practice for mobilizing a response,' *Language and Communication* 79: 1–21.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025

Pekarek Doehler, S. and E. Berger. 2018. 'L2 interactional competence as increased ability for context-sensitive conduct: A longitudinal study of story-openings,' *Applied Linguistics* 39/4: 555–78.

Raes, A., *et al*. 2020. 'A systematic literature review on synchronous hybrid learning: Gaps identified,' *Learning Environments Research* 23: 269–90.

Ro, E. and A. R. Burch. 2020. 'Willingness to communicate/participate'in action: A case study of changes in a recipient's practices in an L2 book club,' *Linguistics and Education* 58: 100821.

Ruhleder, K. and B. Jordan. 2001. 'Co-constructing non-mutual realities: Delay-generated trouble in distributed interaction,' *Computer Supported Cooperative Work (CSCW)* 10: 113–38.

Schegloff, E. A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge University Press.

Sert, O. 2013. '"Epistemic status check" as an interactional phenomenon in instructed learning settings,' *Journal of Pragmatics* 45/1: 13–28.

Sert, O. 2017. 'Creating opportunities for L2 learning in a prediction activity,' *System* 70: 14–25.

Sert, O. 2019. 'Mutual gaze, embodied go-aheads and their interactional consequences in second language classrooms' in J. K. Hall and S. D. Looney (eds): *The Embodied Work of Teaching*. Multilingual Matters.

Sert, O. and M. Amri. 2021. 'Learning potentials afforded by a film in task-based language classroom interactions,' *Modern Language Journal* 105/S1: 126–41.

Seuren, L. M., *et al*. 2021. 'Whose turn is it anyway? Latency and the organization of turn-taking in video-mediated interaction,' *Journal of Pragmatics* 172: 63–78.

Strid, E. and A. Cekaite. 2022. 'Calibrating joint attention and affective stances in young children's peer interactions,' *Journal of Pragmatics* 198: 29–42.

Tao, J. and X. A. Gao. 2022. 'Teaching and learning languages online: Challenges and responses,' *System* 107: 102819.

ten Have, P. 2007. *Doing Conversation Analysis*. Sage.

Tomasello, M. 1995. 'Joint attention as social cognition' in J. Bruner (ed.): *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum.

Tůma, F. 2022. "Okay, so, moving on to question two': Achieving transitions from one item to another in paired EFL speaking tasks,' *Slovo a slovesnost* 83/3: 163–87.

Ulla, M. B. and F. P. Espique. 2022. 'Hybrid teaching and the hybridization of education: Thai university teachers' perspectives, practices, challenges,' *Journal of Interactive Media in Education* 2022/1: 1–13.

Wang, C. and Y. Luo. 2021. 'Selected poster presentations from the American Association of Applied Linguistics conference 2021: Analysis of interactions in a synchronous hybrid English class,' *Language Teaching* 54/4: 578–82.

Waring, H. Z. 2009. 'Moving out of IRF (Initiation-Response-Feedback): A single case analysis,' *Language Learning* 59/4: 796–824.

Waring, H. Z. and L. B. Carpenter. 2019. 'Gaze shifts as a resource for managing attention and recipiency' in J. K. Hall and S. D. Looney (eds): *The Embodied Work of Teaching*. Multilingual Matters.

Zydney, J. M., *et al*. 2019. 'Here or there instruction: Lessons learned in implementing innovative approaches to blended synchronous learning,' *TechTrends* 63: 123–32.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae017/7633364 by guest on 21 January 2025