

1 **Participation and co-presence in the virtual world of**
2 **Second Life**
3 **Transitioning from a gathering to an encounter**

4 *Laura Kohonen-Aho & Anna Vatanen*

5 **1 Introduction**

6 Goffman (1963) suggests that there are two principal ways in
7 which social situations are organized, that is, how people are
8 present with someone in a shared space. The division
9 between ‘encounters’ and ‘gatherings’ – the two types of co-
10 presence – is made on the basis of the behavioral obligations
11 that pertain the given situation. In both situations, individuals
12 are physically close enough to perceive others and to sense to
13 be similarly perceived by them. In gatherings, such as when
14 strangers pass by one another on the street, the individuals do
15 not have a joint focus of attention and are not engaged in
16 being in focused interaction with others (Goffman 1963: 17,
17 88). In an encounter, on the contrary, the participants share a
18 joint orientation by, for instance, having a conversation,
19 which makes their interaction ‘focused’.

20 In computer-mediated communication, which is under
21 consideration in the present study, the division between
22 gatherings and encounters is an intriguing study area given
23 that the possibility for an unfocused gathering does not easily
24 exist: communication technologies are primarily developed
25 for connecting people for having focused interaction across
26 distances. Especially because of the lack of joint, surrounding
27 space where distributed individuals could just be co-present

with one another using virtual bodies but without having interactional obligations, situations of unfocused gatherings rarely emerge. For example, in video calls, people rarely just ‘hang around’ like they do in coffee houses or libraries, for instance. Instead, they are usually engaged in a focused encounter with one another during the entire time the mutual connection is open. Differentiating between a gathering and an encounter in computer-mediated communication is difficult also because of the lack of eye contact, subtle body movements and orientation to joint objects and surroundings, which in face-to-face settings are exploited as interactional devices and cues for differentiating gatherings and encounters and for transitioning from one to the other (Mondada 2009; De Stefani/Mondada 2018).

Both gatherings and encounters are, however, possible in computer-mediated communication as well, when three-dimensional virtual world (VW) technology is used. When compared to other communication technologies, VWs make mediated interactions more similar to face-to-face situations by the use of virtual embodiment, that is, avatar characters inside a joint space, and hence both gatherings and encounters may occur (Moore et al. 2006). However, the difference between a virtual gathering and a virtual encounter as well as the details of the interactional processes in transitioning between the two require further investigation. Mondada (2009) has investigated the transition process in face-to-face situations: between strangers passing each other on the street. However, for example due to the lack of eye contact between avatars, it is expected that this process is significantly different in VWs.

One way to approach gatherings and encounters in a VW context is to relate them to the concept of *presence*, and more specifically, to *co-presence* and *social presence*. Co-presence in a VW is mostly understood as the feeling of not being alone and as the awareness of other users in the joint space (i.e., being in a ‘gathering’). Social presence, on the other hand, refers to being with others in a joint virtual space but also experiencing psychological involvement and behavioral engagement with them (i.e., being in an ‘encounter’). (Biocca et al. 2003.) Despite the vast body of research on social presence in VWs (e.g., Mennecke et al. 2011; Schultze/Brooks 2019), relatively little is known about the detailed

70 interactional process of achieving social presence – that is, an
71 ‘encounter’ – between the co-present individuals. In the field
72 of computer-mediated communication, one reason for this
73 lack of understanding may be in the tradition of treating
74 social presence mainly as a product of the mind (see, e.g.,
75 Lombard/Ditton 1997), leaving the related interactional
76 practices for a lesser attention (Kohonen-Aho/Alin 2015;
77 Sivunen/Nordbäck 2015). As a distinction to the existing
78 research on social presence in VWs, we propose that this
79 concept should be investigated not merely as an individual’s
80 perception of others but rather as something that is
81 observable and negotiated in interaction – as is done in the
82 fields of interactional linguistics and conversation analysis, for
83 instance (see, e.g., Mondada 2009). Being present and
84 indicating presence are thus understood and investigated as
85 phenomena that involve not only linguistic but also – and
86 perhaps more importantly – bodily cues.

87 In the present study, we investigate how the VW
88 ‘gatherings’ turn into ‘encounters’. We analyze 40 transition
89 episodes in 12 video-recorded virtual team collaboration
90 sessions in the VW of Second Life. We apply the analytic
91 practices of multimodal conversation analysis (Goodwin
92 2000; Mondada 2016a) to capture the simultaneous
93 occurrence of talk and bodily action as they unfold moment
94 by moment. Our findings suggest that there are two main
95 processes used to transition from a gathering to an encounter.
96 Additionally, we examine how the VW encounters are
97 verbally opened (i.e., the participants’ first verbal turns after a
98 silence). Based on our findings, we discuss the role of virtually
99 embodied pre-beginnings in indicating what we call
100 “encounter-readiness” with co-participants in a VW. Our
101 study also contributes to the research on openings of
102 interactions from a conversation analytic perspective by
103 concentrating on multimodal practices in openings in a
104 specific technology-mediated setting. It also provides insight
105 on (social) presence as a behaviorally displayed entity in
106 VWs.

107 2 Background

108 2.1 Transitions between types of situations and activities

109 When co-present individuals are in a gathering, they can any
110 time transform the situation into a mutual encounter. This
111 transition usually occurs when one participant initiates a
112 conversation, or when mutual attention is created
113 nonverbally such as by establishing eye contact, using
114 gestures to get the other's attention, or signaling an intent for
115 an encounter with one's body orientation (Goffman 1963: 33–
116 37, 88–89). Mondada (2009) examined the establishment of
117 an interactional space between individuals in a public place.
118 According to her study, multimodal practices such as gaze
119 and body orientation have an essential role in this transition
120 before the first verbal turn is uttered. The studied setting
121 included strangers passing one another on a street, and one of
122 them, the “itinerary seeker,” opening an encounter to ask for
123 directions to a specific landmark. The focus was on what
124 happened before the first verbal turn and on how the
125 individuals spatially organized themselves and the
126 forthcoming encounter.

127 According to Mondada (2009), the formation of an
128 interactional space involves a three-step process (see Figure
129 1). When persons approach one another on a street, the
130 situation is a gathering, and one of them is still unaware of the
131 upcoming encounter. Then the one who has decided to open
132 an encounter starts to engage in what Mondada calls a pre-
133 opening, or an “embodied pre-beginning.” During this phase,
134 the future co-participants gradually engage in mutual
135 identification and recognition using multimodal resources
136 that can be divided in three steps. First, the initiator uses
137 gaze, glancing to identify and orient to a possible future co-
138 participant and to secure his/her reciprocity. Second, after the
139 first glances, the participants gradually adjust their body
140 movements and in conjunction coordinate slowing down
141 their walking. Finally, they both slowly stop moving and
142 stabilize the participation framework with their body
143 postures and orientation. Only after this embodied pre-
144 beginning, the first verbal turn takes place. In the studied
145 setting, the opening started with a pre-beginning particle such
146 as “uh(m),” followed by a phrase such as “pardon me”. These

pre-conditions for an encounter “are visibly and publicly assembled in time, within the progressive establishment of a mutual focus of attention and a common interactional space” (Mondada 2009: 1977; on the initiation of institutional encounters, see Mortensen/Hazel 2014).

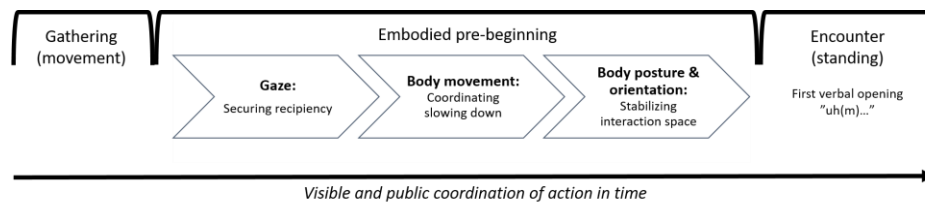


Figure 1: Transitioning from a gathering to an encounter (adapted from Mondada, 2009).

In their recent study, De Stefani and Mondada (2018) examined how both strangers and acquainted persons entered into casual encounters in public spaces. What mainly differentiated strangers and acquainted persons in encounters was their joint history or the lack of thereof, which leads to either the *identification* of an unknown person or the *recognition* of a familiar person as the potential interaction partner. Acquainted persons in De Stefani and Mondada’s (2018) study emphasized the opening of their encounter with the use of embodied resources. They produced emphatic greetings, which provided an additional display of mutual recognition.

Transitions from a gathering to an encounter may also occur when the participants have already been involved in an encounter but the continuous conversation has then lapsed. According to Hoey (2018), participants have three basic alternatives as to how to continue after a lapse – that is, after a silent moment between sequences of talk: they may move to end the interaction, continue with prior talk, or start something new. The participants practically achieve this transition from silence to talk not only by simply saying something, but subtle embodied behaviors are often exploited as well (Vatanen 2018; frthc). In the current study, however, the focal activity that precedes the imminent transitions has not been talking but engaging in an individual activity, and hence the silent ‘gatherings’ cannot directly be described as lapses (even though they involve no talk).

Rather, what is more at stake is a transition between two different activities.

Previous research has found that when participants have a clearly available common activity at hand, they frequently transition between talking and being engaged in that activity. The local devices for achieving such transitions have been described for activities such as students doing group work in classrooms (Szymanski 1999), friends playing video games (Mondada 2012), and families engaged in foraging activities (Keisanen et al. 2017). Displaying availability for interaction is crucially related not only to the participants' speech but also to their embodied behavior: certain body movements as well as gaze behavior function to elicit speech from the co-participant (Heath 1984). Directions and movements of the body and its parts are used to display participants' involvement in different activities and participation frameworks (Goffman 1981; Goodwin 1984; Schegloff 1998; Kamunen 2019). The participants' bodies in the given material environment create specific "contextual configurations", which frame and constitute the participants' actions (Goodwin 2000). The interactional space is created flexibly by arranging the participants' bodies in relation to the ongoing activity and the local environment – for instance, in situations where participants transition from one activity to another (Mondada 2013). The participants' body movements and spatial configurations are essential in making a transition from one activity to the next, for instance, when (trans)forming the participation framework of a group of people (Broth/Keevallik 2014; Råman 2018).

Analyzing the participants' embodied behavior during phases when they (possibly) transition from one activity to the next is crucial also in the present study. We are especially interested in how a participant makes oneself available for interaction and how they then jointly achieve the transition. That is, we investigate the ways in which individuals show their availability for interaction in a VW. Previous literature on interactions in VWs will be reviewed next.

2.2 Gatherings and encounters in a virtual world

VWs such as *Second Life* are persistent three-dimensional online environments developed for social interaction

(Schroeder 2008). Second Life, for example, includes various types of spaces and places where users can spend time and interact. VWs include several channels for interaction, such as text-based communication in open and private chats and within virtual artifacts (e.g., virtual whiteboards), audio connection, and above all, a customizable avatar character capable of movements (e.g., walking, jumping, flying) and gestures (e.g., waving, smiling, nodding). Avatar as a virtual body provides users a sense of presence in a virtual space, co-existence with others, and interaction with other avatars and virtual objects.

The shared space as well as the virtually embodied co-presence using avatars are unique features of VWs in comparison with other communication technologies where participants do not usually share a joint location with bodily representations. Thus, unlike other communication technologies, VWs provide for the possibility for spatially oriented interaction (Benford/Fahlen 1993), and, importantly, the possibility for both gatherings and encounters to occur. In general, communication technologies are not designed for unfocused gatherings where people often have at least peripheral (unconscious) social awareness of the co-present others (Goffman 1963: 83). In a VW, embodied users in a shared virtual space are continuously visible to one another, conveying the presence and location of the user in the virtual environment with cues about its body position as well (Schultze 2010). Thus, a shared space in VWs supports the process of seeing out of the corner of one's eye as well as glancing and overhearing, which are helpful for people in managing and coordinating their own activities as well as in predicting the activities of others in the shared space (Benford/Fahlen 1993).

In addition to focused encounters, previous research has also attended to some gathering-like situations in VWs. For example, Bennersted and Ivarsson (2010) observed that between different phases in online games, players engage in “waiting activities” such as jumping with their avatars. Jumping in the presence of other players was not intended as focused interaction but rather as a signal to others that the player was still active in the game while waiting. Online games commonly also include “idle animations”, different kinds of little activities that the avatars can do when a player

wishes not to be interrupted or leaves the game for a while. By using an idle animation such as reading a book with one's avatar, the player can remain in the presence of others in an unfocused manner. This type of behaviors are reminiscent of what has been shown to occur also in real-life situations when people wait. Waiting does not mean just standing still and/or doing nothing; instead, specific embodied resources are systematically employed to signal to other participants that waiting is taking place (Svinhufvud 2018; Ayaß 2020). Waiting can thus quite often be characterized as a type of a gathering.

The division between a gathering and an encounter in VWs relates to the research area of co-presence and social presence. Mennecke et al. (2011) introduced *the theory of embodied social presence in virtual worlds*. According to this theory, being in a VW and using its contents easily evoke the sense of presence in the virtual space, eventually leading to the sense of co-presence (a 'gathering'), followed by the sense of social presence (an 'encounter') with other users. However, this theory does not discuss the details on when, how and why the users engage in using the VW contents and avatars for achieving social presence, and the detailed process of transitioning from co-presence (gathering) to the state of social presence (encounter) has not been properly attended to. As seen above, previous research has given hints about the significance of spatially defined interactions as well as the use of avatars, but the detailed practices the users engage in to signal availability as well as to open an encounter after being in a gathering are not yet properly understood – a research gap that we in the current paper aim to fill.

Another reason for why the transition process itself has not gained much attention might lie in the difficulty of detecting the subtle interactional transition cues that people use in face-to-face situations in the context of VWs. Despite the increasing visual realism of VWs and the abilities of avatars to convey various types of social information, the users' ongoing activities are still far less obvious for others to detect than they are in real life. In face-to-face situations, participants use detailed observational (verbal and nonverbal) information when monitoring others in order to interpret their actions and to design appropriate responses to them. Accountability, projectability, and coordination of action crucially depend on

this observational information, which include the unfolding of turn taking in real time and the observability of gaze and other embodied activities – features that are still under development in VWs. Especially since avatars do not unintentionally “give off” cues (Goffman 1959: 2) about their users’ activities as human bodies do, information about ongoing activities in VWs needs to be explicitly communicated (Moore et al. 2006). Otherwise, the avatars just stand still and do nothing, possibly giving a false sense of availability.

Even though avatars are less accountable for their actions than real human bodies are, we suggest that since VWs enable the existence of both gatherings and encounters, also the transition process between the two can be investigated – even though it may be less sophisticated than in face-to-face situations (see Mondada 2009). This is what we attempt to do in the remainder of this paper.

3 Empirical study

3.1 Participants and the interaction setting

The research data are collected from a setting where 12 virtual teams collaborated in Second Life. The participants (N=36) in this study were recruited among students from two universities in Finland as well as among friends and colleagues, and they were randomly divided into three-member teams. To ensure that each team would be in equal position, the team members did not know one another beforehand and met face-to-face for the first time only after the session. Furthermore, the participants did not receive information about the researchers’ interests at all.

Second Life consists of spaces for social interaction and collaboration built on virtual islands. The space in this study was built on an island that was surrounded by a transparent wall that prevented team members from leaving the area. The space also included a virtual whiteboard that the teams could jointly use to complete the assigned tasks (see Figure 2). The teams used an audio connection to communicate verbally. Each team used the same three pre-selected avatars for navigating in the space and for using the whiteboard.

Although the teams comprised both male and female participants, the avatars had customary male appearances.

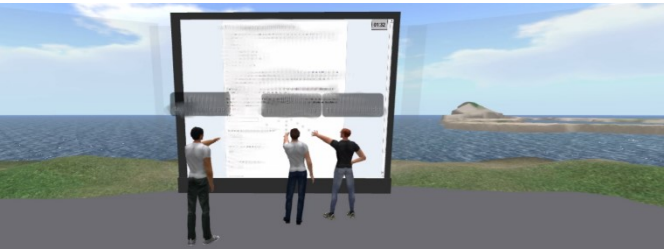


Figure 2: Collaboration space including avatars and a virtual whiteboard.

Each team participated in an assigned session that started with a brief orientation followed by collaborative and individual assignments. The orientation session aimed to familiarize the participants with the task types and Second Life’s functionalities. During the actual collaborative work, each team transitioned between two types of activities: collaborative tasks and individual questionnaires (see Figure 3). After each task, each team member was instructed to individually fill in a questionnaire featuring questions about how they perceived themselves, the other team members, and the joint interaction during the preceding task.

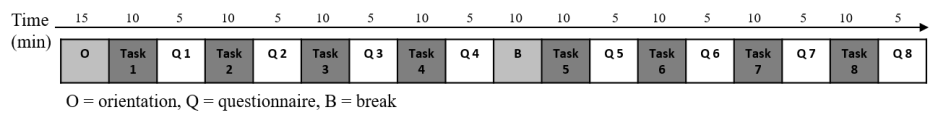


Figure 3: Structure of the team sessions.

3.2 Data collection

All teams in the collaboration setting were video-recorded. Each recorded session lasted 2.5 hours, resulting in 30 hours of video data. To capture the team members both in the virtual space as avatars and in their separate physical locations as “real bodies,” video cameras were embedded both in the virtual world (the “VW” videos) and in the physical real-life locations (the “RL” videos). The video data of each team comprises one VW video where the team members interact in the VW as avatars, and three separate RL videos that capture the team members in their physical locations.

374 3.3 Data analysis

375 In our analysis, we applied the principles of multimodal
376 conversation analysis (Goodwin 2000; Mondada 2016a). Our
377 analytic steps followed the basic conversation analytic
378 research process (Sidnell 2013). First, we viewed the video
379 recordings multiple times to identify the episodes that grasp
380 our phenomenon of interest: moments of transitions between
381 gatherings and encounters. Gatherings prevailed especially
382 during the questionnaire-filling time; at those moments, the
383 team members focused silently on individual rather than
384 collaborative work. The questionnaire filling moments can be
385 described as “allowable silences” (Hoey 2015), even “planned
386 silences” between the “planned encounters”, i.e., the
387 collaborative tasks.

388 All participants in the 12 teams filled altogether nine
389 questionnaires, one during the orientation and one after each
390 assigned task (108 questionnaire-filling episodes in total).
391 Most participants were ready sooner than expected, which
392 resulted in unexpected ‘surplus time’ until the next
393 collaborative task began. During this surplus time, the team
394 members either voluntarily initiated an encounter, or waited
395 for the next task in silence. In 40 of these episodes, there was
396 a transition from a gathering to an encounter, whereas in 66
397 of the episodes the surplus time was composed of a gathering
398 only. In these 66 gathering-only episodes, the team members
399 either just waited sitting still and silent (38 episodes), or
400 walked with their avatars in the joint space in silence, not
401 opening an encounter verbally (28 episodes). In addition, we
402 found two episodes where the team kept up the encounter
403 (i.e., a conversation) during the entire questionnaire-filling
404 episode, even though they were supposed to work
405 individually.

406 In our detailed analysis, we focused on the 40
407 gathering–encounter transitions. We transcribed the team
408 members’ verbal communication (using the Jeffersonian
409 conventions; see, e.g., Jefferson 2004) as well as their
410 embodied actions (applying the conventions developed by
411 Mondada (2016b)), including avatar movements and each
412 team member’s gaze direction, facial expressions, body
413 movements and hand movements on the keyboard (see
414 Appendix for the transcription symbols). Next, we analyzed

the sequential and temporal trajectories of the gathering-encounter transitions, described them in detail, and made comparisons across episodes. In what follows, we will present our findings.

4 Transitioning from a gathering to an encounter in a virtual world space

In general, the team members were rather cautious about initiating conversation immediately after they had finished their own questionnaires, possibly because there were not always clear signs of whether the others were ready with their questionnaires yet. Rather, the team members stayed silent for several seconds after finishing their own questionnaire. When interacting face-to-face, there are embodied strategies for dealing with “awkward silences” that may occur when continuous conversation lapses. For example, the participants may drink, eat, or engage in self-grooming or yawning (Hoey 2015; Vatanen submitted). In our data, the team members engaged in similar activities to ‘fill’ the silence or to pass the time when waiting in their separate physical locations. They drank water, yawned, stretched, or changed their body positions in their chairs. By carefully viewing RL videos of each participant, we also observed that some of them explored the virtual space without moving their avatars by using their mouse scroll to zoom and rotate their view of the virtual space. Zooming with a mouse scroll did not move the avatar but rather changed its perspective between a first-person perspective (seeing “with the avatar’s eyes”) and a third-person perspective (seeing from outside of the avatar). In addition to these private waiting activities, the team members also started to move their avatars in the virtual space when waiting (cf. Svinhufvud 2018; Ayaß 2020). When a team member finally initiated a conversation after having waited, the opening usually related to something other than directly asking whether the others were ready with their questionnaires.

Based on our analysis, the ways in which the participants transition from a gathering to an encounter in the 40 episodes fall into two main types. In Process 1 (12 episodes), the transition is accomplished using verbal means only: in these

cases, the encounter simply is verbally opened by one of the team members. In Process 2 (28 episodes), the transition process involves the use of avatars before the first verbal opening. In the following analysis, engaging in avatar movement before the first verbal opening will be called *a virtually embodied pre-beginning*. In this VW context, the virtually embodied pre-beginnings have different characteristics compared to the embodied pre-beginnings in face-to-face interaction described by Mondada (2009).

In addition to the two transition processes, there are differences in the first verbal opening turns. The openings include the social actions of *noticing*, *information-request*, *account*, and *proposal* that are related to the virtual space to which all team members had a joint visual access, or to something else such as the other team members or the tasks and questionnaires in the interaction setting. In Process 2, some of the openings are related to the avatar movement, which directly preceded the first verbal opening. The first opening is also dependent on the type of the preceding avatar movement during the virtually embodied pre-beginning (e.g., walking vs. jumping). Next, we analyze in more detail both transition processes as well as the types of the verbal openings and the virtually embodied pre-beginnings. In the following illustrations from our data, we call the team members seen in RL videos by their real (anonymized) names, and their avatars in the VW videos by “Name-A” (e.g., *Jaakko* and *Jaakko-A*).

4.1 Process 1: Transition directly to talk, no preceding virtually embodied behavior

Process 1 includes altogether 12 episodes where the transitioning from a silent gathering to an encounter takes place without any embodied preparation (see Figure 4). In these episodes, someone in the team just starts to talk at some point after finishing his/her own questionnaire.

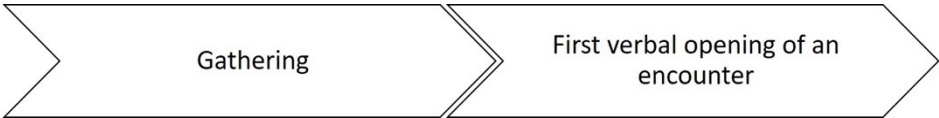


Figure 4: Process 1 of transitioning from a gathering to an encounter.

490 The 12 episodes in Process 1 can be divided into two
 491 categories according to the type of the first verbal turn that
 492 opens the encounter: noticing something (4 episodes) vs.
 493 requesting information (8 episodes). The content of the
 494 openings relates to either something that all team members
 495 have a joint visual access to in the virtual space or to
 496 something else, such as the interaction setting (questionnaires
 497 or collaborative tasks) or the other team members (e.g., their
 498 studies and free time). We illustrate this process by showing
 499 an excerpt of noticing something in the virtual space to which
 500 everyone has a joint visual access. Excerpt 1 begins when all
 501 team members, Juho, Susanna, and Jaakko, fill their
 502 questionnaires in silence. Juho is the first one to finish (line
 503 2).¹

504 **Excerpt (1)**²: Team 10, questionnaire 1, time 0:32:10.

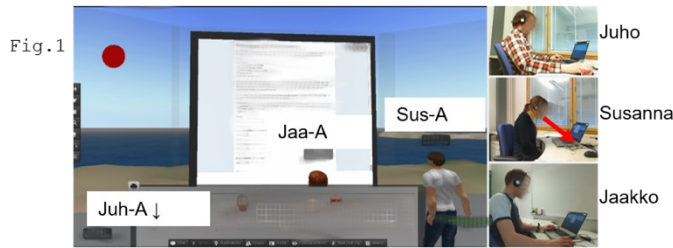
```

01      * (156.0)
    All:  *fill questionnaires-->
02      + (1.5) +V (5.2)
    Juh:  +finishes his questionnaire ((questionnaire on his screen turns grey))
    Juh:  +places his hand on keyboard and starts rotating his screen view-->
    Juh-A: Vdoes not move but stands still-->
03      ΔΔ (1.3) Δ (6.3) Δ (0.5) Δ (1.0) ΔΔ (18.6)
    Jaa:  Δfinishes his questionnaire
    Jaa-A: Δkeeps touching the whiteboard
    Jaa:  Δsits still, gaze towards his screen-->
    Jaa:  -->Δpresses a keyboard button
    Jaa-A: Δstops touching the whiteboard
    Jaa:  Δsits still, gaze towards his screen-->
    Jaa-A: Δstands still-->
04      %‡ (1.7) %Δ (2.4) %Δ (2.9)
    Sus:  %finishes her questionnaire and removes hands from the keyboard
    Sus-A: ‡stops touching the whiteboard
    Sus:  %sits still, gaze towards her screen-->
    Jaa:  -->Δglances to the camera
    Sus:  -->%stretches her shoulders-->
    Jaa:  Δdetaches hands from keyboard, adjusts headphones-->
05      Δ (12.0) %Δ (2.0)
    Jaa:  Δadjusts his chair
    Sus:  -->%leans abruptly towards her screen ((Fig. 1))
    Jaa:  Δplaces his hands back on the mouse and keyboard-->

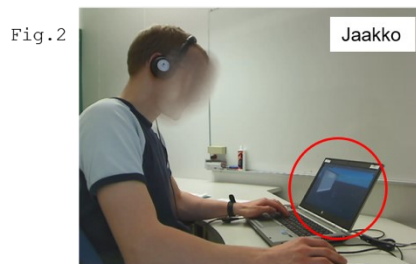
```

1 We detected the exact moment of finishing a questionnaire from the RL videos where we could see the questionnaire form turning grey in the team member's computer screen immediately after s/he submitted it.

2 The participants' embodied behavior in the data, both real and virtual, is transcribed applying Mondada's (2016b) conventions. However, the lines that are not in the focus in the analysis are transcribed more roughly, especially when it comes to the exact timing of the embodied behavior. The target lines are more precisely transcribed.



06 SUS: delfiini (.) oikealla.
dolphin (.) on the right.
 + \angle (1.0)
 07 Juh: -->+starts to turn his screen view to the right with his mouse-->
 Jaa: --> \angle starts to use his keyboard-->> ((Fig. 2))



08 % (1.3)
 Sus: %leans towards her screen
 09 JAA: .hh (.) \uparrow emminä nää Δ mitää delfiiniä.
 .hh (.) \uparrow I don't see Δ any dolphin.
 Jaa-A: --> Δ turns head to the right ((Fig. 3))



10 SUS: hh hh hh ((laughs)) %(.) tuolla skriinin takana,
 hh hh hh ((laughs)) %(.) **there behind the screen,**
 Sus: %places her hand on keyboard

505 After finishing his questionnaire, Juho starts to explore the
 506 virtual space by using his mouse scroll, which does not move
 507 his avatar (line 2). After 5.2 seconds, also Jaakko finishes his
 508 questionnaire but does not immediately engage in any
 509 detectable waiting activity but rather keeps sitting still, gaze
 510 towards his computer screen (line 3). After 6.3 seconds,
 511 Jaakko clicks his cursor off the whiteboard, which makes his
 512 avatar to stop touching the whiteboard with its arm but not to
 513 move otherwise. Jaakko sits still, gaze towards his screen for
 514 18.6 seconds while Juho keeps scrolling with his mouse and
 515 Susanna still fills her questionnaire. Then also Susanna
 516 finishes her questionnaire and immediately detaches her
 517 hands from her mouse and keyboard. Simultaneously her
 518 avatar stops touching the whiteboard (line 4). Susanna thus

disengages from the virtual world and starts to stretch her arms while waiting, while also Jaakko starts to adjust his headphones and chair. After stretching, Susanna suddenly leans towards her computer screen (line 5, Fig. 1). Simultaneously also Jaakko places his hands back on his mouse and keyboard, but does not use them. Then, Susanna makes the first verbal opening by informing others what she has just seen in the virtual space: a dolphin jumping in the virtual sea on her right, outside the collaboration area (line 6; see Bergmann (1990) and Hoey (2018) on using environment for generating talk). Juho and Jaakko do not provide immediate verbal responses, but they react to Susanna's opening bodily: Juho starts to rotate his screen view to the right and Jaakko starts to use his keyboard (line 7, Fig. 2). After altogether 2.3 seconds, Jaakko responds to Susanna, and only then, with a relatively long delay, Jaakko-A turns its head to his right (line 9, Fig. 3). As Jaakko has been unable to detect the dolphin, the encounter continues with Susanna providing a more detailed explanation of the dolphin's location (line 10).

In other words, here the transition from a gathering to an encounter is accomplished by only talking, without any (virtually) embodied preparations, and the first verbal turn is a noticing of something in the joint visual space. In addition, we found cases where the first verbal opening that is not preceded by any embodied behavior relates to requesting information about something that does not relate to the virtual space but something else, such as the activities the team members engage in during their VW interaction (e.g., the number of the remaining tasks or questionnaires), or the other team members (e.g., their previous experience with virtual environments). Next, we will analyze cases in the other transition process where the transition includes a virtually embodied pre-beginning before the verbal opening.

4.2 Process 2: Transition with a virtually embodied pre-beginning

In our data, the participants did not always remain waiting and managing the silence alone in their physical locations, but they used the waiting time to explore the joint virtual space by moving their avatars. We grouped these cases under transition process 2, which includes altogether 28 episodes

where the transitioning from a silent gathering to an encounter takes place with an embodied preparation phase (see Figure 5). This virtually embodied pre-beginning comprises different movements of the avatar.

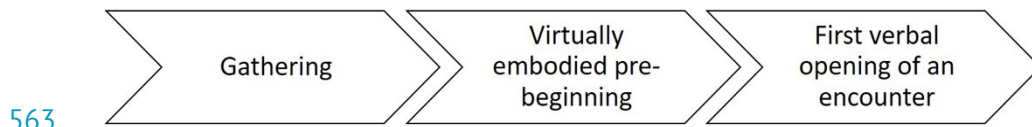


Figure 5: Process 2 of transitioning from a gathering to an encounter.

The 28 episodes in Process 2 fall into three categories according to the content-type of the first verbal turn that opens the encounter. The content of the first verbal opening turns relates either to the preceding avatar movement (12 episodes), to something else in the virtual space to which all have joint visual access (8 episodes), or to something else (8 episodes). As in Process 1, these openings include noticings and information-requests, but in addition, also accounts and proposals. In these episodes, it was also meaningful to make a difference as to who produces the first verbal opening: the one who engaged in the avatar movement preceding the first verbal opening, or someone other in the team.

The types of the virtually embodied pre-beginnings are quite different from the pre-beginnings in face-to-face situations described by Mondada (2009). Instead of subtle body orientations and eye contact with the interlocutor-to-be, virtually embodied pre-beginnings include walking around in the virtual space, jumping, and bumping into something (another avatar or the virtual whiteboard) with one's avatar. Let us now examine in more detail the three types of verbal openings and the preceding virtually embodied behaviors.

4.2.1 The opening relates to preceding avatar movement

In 12 episodes, the opening is either a *noticing* (8 episodes), an *information-request* (2 episodes), an *account* (1 episode) or a *proposal* (1 episode), all relating to the just-preceding avatar movement. In addition, these 12 episodes can be grouped according to who engages in moving the avatar and who then opens the encounter. In nine episodes (see Excerpt 2), the one who moves the avatar also opens the discussion. In three

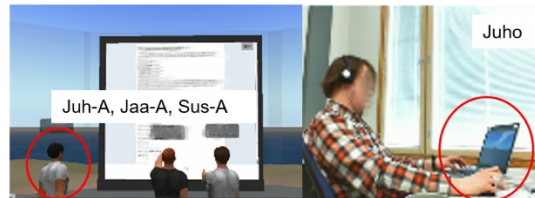
episodes (see Excerpt 3), the one who moves and the one who opens the encounter are different people.

In the following excerpt, Juho opens the encounter by requesting information about his own preceding avatar movement. The excerpt begins when all team members, Juho, Susanna, and Jaakko, fill their questionnaires in silence. Juho is the first to finish the questionnaire (line 2).

Excerpt (2): Team 10, questionnaire 5, time: 1:39:15.

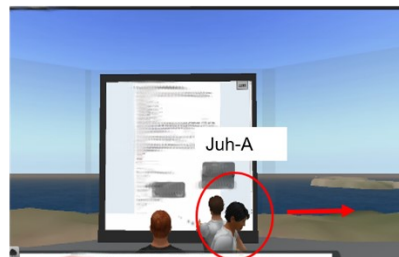
- 01 * (139.2)
All: *fill questionnaires-->
02 +V(3.8)+V(1.9)
Juh: +finishes his questionnaire ((questionnaire on his screen turns grey))
Juh-A: Vkeeps touching the whiteboard-->
Juh: +presses a keyboard button
Juh-A: -->Vstops touching the whiteboard
03 +V(25.6)
Juh: +starts to rotate his screen view with keyboard and mouse--> ((Fig. 1))
Juh-A: Vdoes not move but stands still--> ((Fig. 1))

Fig.1



- 04 ∠Δ+V(1.3) ∠Δ(1.5) ∠Δ(5.4) %‡(1.0) %+V(2.2)
Jaa: ∠finishes his questionnaire
Jaa-A: Δkeeps touching the whiteboard-->
Juh: -->+presses a keyboard button
Juh-A: -->Vturns head to the right
Jaa: ∠presses an arrow button a few times
Jaa-A: -->Δstops touching the whiteboard and walks few steps back-->
Jaa: ∠sits still, gaze towards his screen-->
Jaa-A: -->Δstands still-->
Sus: %finishes her questionnaire
Sus-A: ‡keeps touching the whiteboard-->
Sus: %sits still, gaze towards her screen-->
Juh: +clicks with his mouse few times
Juh-A: Vwalks past the space and stops ((Fig. 2))

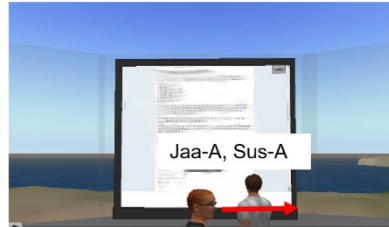
Fig.2



- 05 JUH: mihinköhän mä lähin oikee juoksee yhtäkki.
I wonder where I started to run all of a sudden.
06 SUS: hh hh hh hh ((laughs))
07 (1.0)

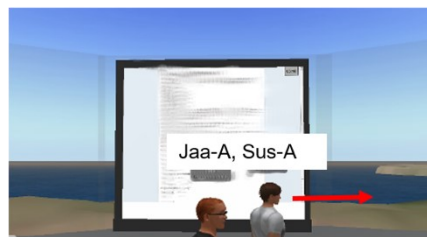
08 JUH: mikä tälle tuli.
what's the matter with this.
 09 JAA: \angle sua ei taas kiinnostanu tää homma Δ ilmeisesti.
 \angle you were not interested in this thing Δ apparently.
 Jaa: -- \angle starts to use his keyboard-->>
 Jaa-A: Δ turns towards Juh-A
 ((Fig. 3))

Fig.3



10 SUS: % \sharp hh hh hh ((laughs))
 Sus: -->%starts to use her keyboard-->>
 Sus-A: --> \sharp stops touching the whiteboard and turns towards Juh-A ((Fig. 4))

Fig. 4



11 JUH: .HHH hh hh ((laughs)) mä lähen litoo.
.HHH hh hh ((laughs)) I'll flee away.

603 Juho deactivates his whiteboard 3.8 seconds after finishing
 604 the questionnaire, which makes his avatar lower its arm from
 605 the whiteboard (line 2). Then he starts to explore the virtual
 606 space for 25.6 seconds. We see from the RL video that he
 607 uses only the ESC key and the mouse zoom that allow him to
 608 zoom and rotate the screen view without moving his avatar
 609 (line 3, Fig. 1). Jaakko finishes the questionnaire next. He
 610 deactivates his whiteboard as well, which makes his avatar
 611 take a few steps back; he does not, however, engage in
 612 further movements (line 4). Simultaneously, Juho presses a
 613 keyboard button that makes his avatar move a little: it turns
 614 its head to the right. After that, he remains sitting still, looking
 615 at his computer screen. Then, Susanna finishes her
 616 questionnaire, and one second later, Juho clicks his mouse,
 617 which makes his avatar walk across the virtual space, past
 618 Jaakko-A and Susanna-A (Fig. 2). Clicking somewhere in the
 619 virtual ground with his mouse seems to make Juho's avatar
 620 walk forward to the clicked location. This movement is
 621 apparently unintentional as Juho shares his surprise with
 622 others: "I wonder where I started to run all of a sudden" (line
 623 5), and "what's the matter with this" (line 8). Juho's questions
 624 about his avatar movement open the encounter, and Susanna

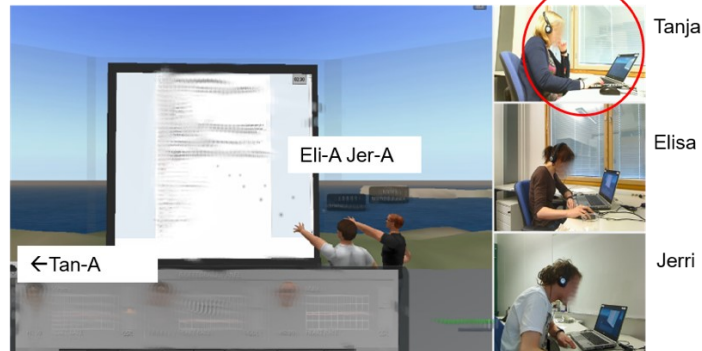
and Jaakko respond with laughter (line 6) and joking (9). They also start to use their avatars to turn to look at Juho-A (lines 9–10, Fig. 3, Fig. 4). Juho joins Jaakko's joking, and the encounter continues (line 11).

Also in the next excerpt, the encounter is opened by *requesting information* about something concerning the preceding avatar movement. In contrast with the previous excerpt, here the one who engages in the avatar movement (here: jumping) and the one who opens the encounter are different. The excerpt begins when all team members, Tanja, Elisa, and Jerri, fill their questionnaires in silence. Tanja is the first one to finish the questionnaire (line 2).

Excerpt (3): Team 9, questionnaire 1, time: 00:34:34.

01 * (139.0)
All: *fill questionnaires-->
02 +V(1.7) + (5.5)
Tan: +finishes her questionnaire ((questionnaire on her screen turns grey))
Tan-A: Vkeeps touching the whiteboard-->
Tan: +sits still, gaze towards her screen-->
03 TAN: +.hh
Tan: -->+leans back in her chair
04 +V(2.5)
Tan: +starts to use her keyboard--> ((Fig. 1))
Tan-A: ->Vstarts to walk around in the space--> ((Fig. 1))

Fig.1



05 ELI: hh (.) hh
06 (25.1)
07 JER: hhh (.) .hh hh (.) hhh
08 ∠Δ(1.0) ∠(6.0) V(2.0) ∠Δ(1.0) VΔ(6.1)
Eli: ∠finishes her questionnaire
Eli-A: Δkeeps touching the whiteboard-->
Eli: ∠sits still, gaze towards her screen
Tan-A: -->Vwalks in front of the whiteboard-->
Eli: ∠starts to use her keyboard-->
Eli-A: -->Δstops touching the whiteboard
Tan-A: -->Vwalks around in the space-->
Eli-A: Δturns around and starts to walk-->

09 (2.4) %Δ(2.3)
 Jer: %finishes his questionnaire
 Jer-A: ‡keeps touching the whiteboard-->
 Eli-A: -->Δwalks in front of the whiteboard-->
 10 JER: %hhh (.) %‡hh
 Jer: %leans back %clicks a keyboard button
 Jer-A: -->‡stops touching the whiteboard

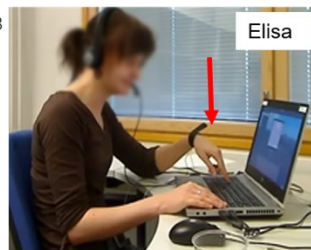
 11 %‡(6.0) %‡(3.0)
 Jer: %starts to use his keyboard-->
 Jer-A: ‡starts to walk-->
 Jer: -->%presses a keyboard button ((Fig. 2))
 Jer-A: -->‡jumps ((Fig. 2))

Fig.2



12 ELI: hh hh ((laughs)) fmiten sä ton teit.f
 hh hh ((laughs)) fhow did you do that.f
 13 JER: ee nappi.
 E button.
 14 Δ(1.0)
 Eli: -->Δpresses E button ((Fig. 3))
 Eli-A: -->Δjumps

Fig.3



15 ELI: ooh hahaha
 16 TAN: hh hh hh he he he he

638 For 5.5 seconds after finishing the questionnaire, Tanja sits
 639 still but then changes her body position and starts to use her
 640 keyboard and walk in the virtual space with her avatar (lines
 641 3–4). We see this movement only in the RL video on her
 642 screen, because her avatar has moved outside the view in the
 643 static VW video (line 4, Fig. 1). Elisa finishes next and remains
 644 sitting still, looking at her screen (line 8). After she has been
 645 silent for 6.0 seconds, Tanja-A appears in front of the
 646 whiteboard and walks near the other avatars. Although it is
 647 not certain whether Elisa notices Tanja's avatar, she starts to
 648 move her avatar as well right after Tanja-A appeared near her
 649 avatar. Jerri finishes his questionnaire as well (line 9) and

650 adjusts his body position while breathing loudly (lines 10). All
 651 team members move together in silence for 6.0 seconds (line
 652 11). Then Jerri suddenly jumps high up with his avatar (line 11,
 653 Fig. 2). Elisa reacts to this by laughing and asking how Jerri
 654 managed to jump (line 12). Once Jerri has responded that the
 655 E button in the keyboard makes the avatar jump (line 13),
 656 Elisa tries out jumping as well (line 14, Fig. 3), and the
 657 encounter continues.

658 In summary, in the cases above the transition to an
 659 encounter involved a virtually embodied pre-beginning, and
 660 the first verbal turn related to the preceding avatar
 661 movement. Next, we illustrate a transition process where the
 662 first verbal opening relates to the joint visual space or shared
 663 objects.

664 4.2.2 The opening relates to joint visual space or shared objects

665 In eight episodes, the opening is a *noticing* (5 episodes) or an
 666 *information-request* (3 episodes) about something in the
 667 virtual space that everyone has a joint visual access to. Thus,
 668 even though there is some preceding avatar movement, the
 669 first opening is not related to it, but the topic of the opening is
 670 something else in the joint virtual space. These eight episodes
 671 can be categorized according to the person who engages in
 672 moving the avatar and the person who opens the encounter.
 673 In five episodes (see Excerpt 4), the same person who moves
 674 the avatar also opens the encounter. In three episodes (see
 675 Excerpt 5), the team member who moves and the team
 676 member who opens the encounter are different.

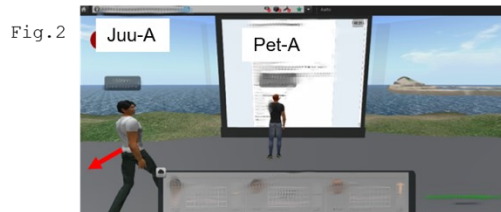
677 The following Excerpt (4) includes a *noticing* as the first
 678 verbal opening, and here it is the same team member who
 679 first walks with his avatar and then opens the encounter
 680 (even though here actually all participants move their avatars
 681 prior to the opening). This excerpt begins when all team
 682 members, Juuso, Filip, and Petra, fill their questionnaires in
 683 silence.

684 **Excerpt (4):** Team 3, questionnaire 7, time: 2:06:23.

01 * (67.0)
 All: *fill questionnaires-->
 02 +V(2.1)+V(1.5)+(2.3)+V(4.3)
 Fil: +finishes his questionnaire ((questionnaire on his screen turns grey))
 Fil-A: ∇ keeps touching the whiteboard-->
 Fil: +presses a keyboard button
 Fil-A: -->∇ stops touching the whiteboard
 Fil: +sits still, gaze towards his screen-->
 Fil: -->+starts to use his keyboard-->
 Fil-A: ∇ turns around and starts to run away from the
 whiteboard--> ((Fig. 1))



03 +V(6.3)∠Δ(1.5)
 Fil: -->+puts his hand on the mouse and starts to rotate his screen-->
 Fil-A: -->∇ stops moving and keeps standing still-->
 Juu: ∠ finishes his questionnaire
 Juu-A: Δ keeps touching the whiteboard-->
 04 JUU: .hhh
 05 ∠(2.4)
 Juu: ∠ sits still, gaze towards his screen-->
 06 JUU: ∠hhh
 Juu: -->∠ starts to use his keyboard-->
 Juu-A: -->Δ stops touching the whiteboard and starts to walk--> ((Fig. 2))

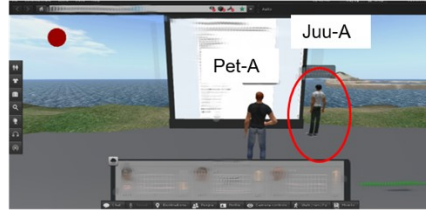


07 Δ(4.2)Δ(6.1)Δ(6.3)Δ(7.2)
 Juu-A: -->Δ walks past the whiteboard and stops
 Δ starts to walk around in the space-->
 -->Δ walks in front of the whiteboard and stops
 Δ turns around and starts to walk-->
 08 %‡(0.8)%(2.3)%‡(3.4)
 Pet: % finishes her questionnaire
 Pet-A: ‡ keeps touching the whiteboard-->
 Pet: % sits still, gaze towards her screen-->
 Pet: -->% starts to use her keyboard-->
 Pet-A: -->‡ stops touching whiteboard and starts to walk--> ((Fig. 3))



09 (2.5) ‡ (6.3) % ‡ (9.5) ∠ Δ (5.0)
 Pet-A: --> ‡ walks in front of the whiteboard-->
 Pet: --> % starts to zoom her screen view with the mouse-->
 Pet-A: --> ‡ stops moving and keeps standing still-->
 Juu: --> ∠ stops using the keyboard
 Juu-A: --> Δ stops walking, faces the sea ((Fig. 4))

Fig.4



10 JUU: ‡ hei tuol menee kala.
 ‡ hey there goes a fish.
 11 (1.2)
 12 JUU: [skriinin takana.]
 [behind the screen.]
 13 FIL: [(mitä)]
 [(what)]
 14 JUU: .hh hyppii kala.
 .hh ((there)) jumps a fish.
 15 (1.8)
 16 FIL: ‡ aa sä vasta nyt huomasit.
 ‡ oh only now you noticed.

685 Here all team members start to move with their avatars in the
 686 virtual space a few seconds after finishing their
 687 questionnaires. Filip finishes first (line 2, Fig. 1). He moves his
 688 avatar for 4.3 seconds, after which he switches to rotating his
 689 screen view with the mouse and keyboard, which stops his
 690 avatar movement (line 3). Then Juuso finishes his
 691 questionnaire and breathes loudly (lines 3–4). He sits still for
 692 2.4 seconds (line 5) and then again breathes loudly and starts
 693 to move his avatar (line 6, Fig 2). For altogether 23.8 seconds,
 694 Juuso is the only one to move his avatar (line 7). Finally also
 695 Petra finishes the questionnaire, and after 2.3 seconds she
 696 starts to move her avatar (line 8, Fig. 3). Juuso and Petra move
 697 simultaneously in silence for 8.8 seconds (line 9). The pace
 698 and direction of their movements seem to be rather random
 699 and they do not take any noticeable bodily contact with one
 700 another. Then, Petra stops moving her avatar and switches to
 701 just zooming her view with the mouse scroll. Juuso continues
 702 moving about in silence for another 9.5 seconds. Then he
 703 stops his avatar next to the whiteboard, the avatar facing
 704 towards the sea behind the collaboration space (Fig. 4). After
 705 5 seconds, Juuso announces a discovery he has made in the
 706 sea (line 10): a jumping fish. Filip reacts to Juuso's opening
 707 immediately by initiating repair with the question word *mitä*
 708 'what' (line 12). After Juuso partially repeats his preceding

709 turn, Filip responds by pointing out that this discovery is not
 710 very novel to him (see line 16).

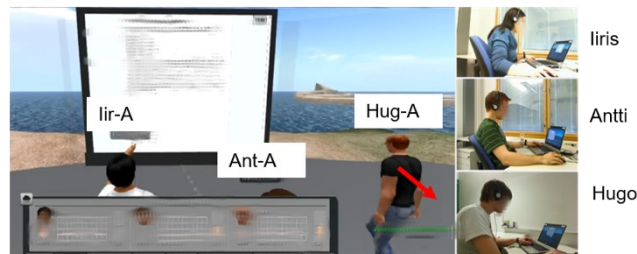
711 The following Excerpt (5) includes again a *noticing* as the
 712 first verbal opening. In contrast to Excerpt 4, here it is a
 713 different team member who engages in the preceding avatar
 714 movement (walking) and who opens the encounter. The
 715 excerpt begins when all team members, Iiris, Antti, and Hugo,
 716 fill their questionnaires in silence.

717 **Excerpt (5):** Team 2, questionnaire 1, time: 00:25:34.

01 * (163.0)
 All: *fill questionnaires-->
 02 + (1.0)
 Hug: +finishes his questionnaire ((questionnaire on his screen turns grey))
 03 HUG: +V.hh ∆∆jes.
 +V.hh ∆∆yes.
 Hug: +presses a keyboard button
 Hug-A: Vstops touching the whiteboard
 Ant: ∆finishes his questionnaire
 Ant-A: ∆keeps touching the whiteboard-->

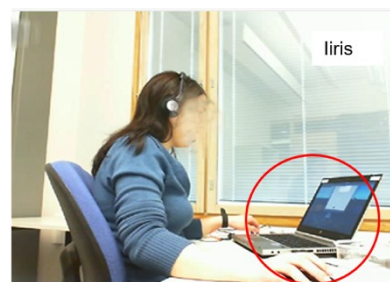
04 +V∆ (9.0)
 Hug: +starts to use his keyboard--> ((Fig. 1))
 Hug-A: Vstarts to walk--> ((Fig. 1))
 Ant: ∆sits still, gaze towards his screen-->

Fig.1



05 %# (0.6) %#V (1.6) %#V (14.5)
 Iir: %finishes her questionnaire
 Iir-A: ‡keeps touching the whiteboard-->
 Iir: %scrolls with her mouse-->
 Iir-A: -->‡stops touching the whiteboard
 Hug-A: -->Vwalks towards the whiteboard
 Iir: -->%zooms her screen view with her mouse-->> ((Fig. 2))
 Iir-A: ‡stands still-->>
 Hug-A: Vwalks around in the space-->

Fig.2



```

06      +V(0.9)+(3.0)
    Hug:-->+detaches hands from the keyboard
    Hug-A:-->Vstops moving
    Hug:      +sits still, gaze towards the screen-->>
07 HUG:  hhh
08 IIR:  ↑ei mulla oo kyllä mitään kalaa tuolla taustalla(h).
        ↑I really don't have any fish there in the back(h).
09 IIR:  [he he he he ]
10 HUG:  [shh hh hh ((laughs))]
11 IIR:  Enä odotin sit(h)ä mut ei tule.f
        fI was waiting for it(h) but it's not coming.f
12 HUG:  Enä ei mul on tuolla.f
        fwell no I have ((it)) there.f
13 ANT:  hh hh hh hh ((laughs))

```

Hugo is the first one to finish the questionnaire filling and produces an utterance “hh, yes” that signals his readiness with the assignment (lines 2–3). The others do not respond to this utterance but seem to treat it as self-talk (see Keevallik 2018). Simultaneously with Hugo’s utterance, Antti also finishes his questionnaire (line 3). Antti remains sitting still, gazing towards his computer screen and still having his avatar touching the whiteboard even though he is finished. Then, Hugo starts to move around with his avatar in the virtual space (line 4, Fig. 1). After 9 seconds, also Iiris finishes the questionnaire and starts to zoom her screen view with her mouse without moving her avatar (line 5, Fig. 2). Hugo stops moving his avatar and detaches his hands from the mouse and keyboard (line 6). After altogether 3.9 seconds, he sighs (line 7), which is followed by Iiris’s noticing that opens the encounter (line 8). Here it seems that Hugo’s avatar movement did not directly evoke the opening, but instead Iiris uses an (invisible) object in the virtual space to open the discussion: it is again the fish in the sea. Iiris continues to explain that she has been waiting for the fish to appear (line 11), which may indicate that she has been looking at the sea rather than Hugo-A’s movement. In any case, Iiris’s opening occurs only after Hugo-A stops moving, and it is also preceded by an audible outbreath by a co-participant. Both of these may have affected the timing of the verbal opening turn.

This excerpt is one of the six episodes where the first verbal opening refers to a topic that has already been discussed during the team session. The form of Iiris’s opening, “I really don’t have any fish there”, reveals that the team has already talked about the topic. Indeed, just before the team started to fill this questionnaire, Hugo asked the others if they have detected a fish in the sea. Antti had, but

751 Iris had not. Thus, here Iris uses the time after finishing her
752 questionnaire to look for the fish, and opens an encounter
753 after noticing that her search is not successful.

754 In summary, the cases above illustrated a transition process
755 involving a virtually embodied pre-beginning, and the first
756 opening turn here relates to the joint visual space or shared
757 objects. Next, we analyze a transition again preceded by
758 virtually embodied behavior, with other topics as the first
759 opening.

760 4.2.3 Other topics as the opening

761 In eight episodes, the verbal opening is either a *noticing* (2
762 episodes) or an *information-request* (6 episodes) on
763 something that does not relate to the preceding avatar
764 movement nor the virtual space. Rather, the content of the
765 opening has to do with the interaction setting (e.g.,
766 questionnaires or collaborative tasks) or the co-participants
767 (e.g., their studies and free time). These eight episodes can be
768 further categorized according to who engages in moving the
769 avatar and who opens the encounter. In six episodes, the
770 same person who moves the avatar also opens the encounter.
771 In two episodes, the person who moves and the person who
772 opens the encounter are different, as happens in Excerpt 6.
773 This excerpt involves an opening where the speaker *requests*
774 *information* about something not related to the joint space.
775 The excerpt begins when all team members, Paula, Selena,
776 and Oliver, fill their questionnaires in silence.

777 **Excerpt (6):** Team 1, questionnaire 6, time: 1:58:28.

```

01      * (121.0)
    All:  *fill questionnaires-->
02      +V(0.5)+V(1.0)+(1.0)
    Oli:  +finishes his questionnaire ((questionnaire on his screen turns grey))
    Oli-A:  ∇keeps touching the whiteboard-->
    Oli:    +clicks with his mouse
    Oli-A:  --> ∇stops touching the whiteboard and stands still-->
    Oli:    +sits still, gaze towards his screen-->
03      ∠Δ(1.2)∠(4.8)
    Sel:  ∠finishes her questionnaire
    Sel-A:  Δkeeps touching the whiteboard-->>
    Sel:    ∠leans back and stretches her arms-->
04  SEL:  ∠hhh
    Sel:  -->∠leans to the table, gaze away from her screen-->
05      %‡(1.3)%+V(0.2)
    Pau:  %finishes her questionnaire
    Pau-A:  ‡keeps touching the whiteboard-->>
    Pau:    %sits still, gaze towards her screen-->>
    Oli:    -->+starts to use his keyboard-->
    Oli-A:  -->∇starts to walk-->> ((Fig. 1))

```

Fig.1



```

06      ∠(3.0)
    Sel:  -->∠turns her gaze back to her screen-->>
07  SEL:  .hhh hhh
08  PAU:  tuli muuten mieleen siit deitti (.)
          it came to my mind by the way about that dating (.)
09  PAU:  s:ovellusta että siis oliko se silleen et,
          a:pplication so was it so that,
          ((the question continues and is responded to))

```

778 Oliver is the first one to finish his questionnaire, and he does
 779 not immediately engage in moving his avatar other than
 780 detaching its arm from the whiteboard (line 2). Selena finishes
 781 next and withdraws from her computer to stretch her arms
 782 and lean towards her table (line 3). Then she sighs (line 4).
 783 Paula is the last one to finish, and 1.3 seconds later Oliver
 784 starts to use his keyboard to walk around with his avatar (line
 785 5, Fig. 1). Now Selena returns to gaze towards her computer
 786 screen, and 3 seconds later breaths noticeably in and out
 787 (lines 6–7). As in Excerpt 5, the verbal opening of the ensuing
 788 encounter, Paula's question on lines 8–9, is preceded by an

audible in- and outbreath by a co-participant. The opening itself relates to something else than the avatar movement or the virtual space: the team members' free time. This is another one of the six episodes where the opening turn does not initiate a completely new topic but relates to a previous one: Selena has relayed that she has used the dating application Paula talks about during one of their previous task assignments. Paula's opening is thus understood as addressed to Selena, even though it is not certain if the preceding outbreath can be identified as produced by Selena.

In this excerpt, Oliver's prior avatar movement did not seem to relate to Paula's first verbal opening either, which was after all addressed to Selena. It is, however, possible that witnessing someone moving with an avatar can be used as a cue for detecting that at least one team member is ready with her/his questionnaire and thus available for an encounter (the same occurred in Excerpts 4 and 5). In addition, the audible breath may have again affected the timing of the verbal opening turn (as in Excerpt 5).

In summary, the case above illustrated a transition process involving a virtually embodied pre-beginning, with the first opening turn not relating to the avatar movement, joint visual space nor the shared objects. Rather, it relates to something else, which in the rare case of Excerpt 6 is a prior discussion topic of the team members' free time. Next, we present the implications of the virtually embodied behavior in VWs as indicating what we call *encounter-readiness*, after which we provide a summary and discussion of our findings.

5 Virtually embodied behavior indicates encounter-readiness

Previous studies on interaction in VWs have not been unanimous in how and to what extent virtually embodied behavior is actually used alongside verbal communication. Researchers have claimed that although people interact in a VW as avatars surrounded by virtual objects, they still mainly use talk in their communication (Sivunen/Nordbäck 2015). On the other hand, some studies also suggest that the possibility to use simultaneously text, audio, objects, and the avatar body is the key to enriching discussions and structuring interaction (Antonijevic 2008), and to increasing the

awareness of others (Allmendinger 2010). Some studies suggest that avatar stillness may signal even the lack of presence of the user behind the avatar (Bennerstedt/Ivarsson 2010) or false availability of the user (Moore et al. 2006). The findings of the current study shed more light on the role of virtually embodied behavior in moments when people in the VW have been silent, i.e., in a ‘gathering’, and are about to (re-)start talking.

Our findings suggest that the role of virtually embodied behavior is to function as a possible pre-beginning of an encounter: it signals one’s readiness with the preceding individual task and the availability to move to an encounter, i.e., to interact with others. We call the engagement in the virtually embodied behavior *encounter-readiness*. Displaying encounter-readiness seems to be an important cue for co-participants in considering opening an encounter. There are certain reasons for this, and some of them seem to do with certain functionalities of the present VW, Second Life. When the team members operate with the whiteboard, their avatar’s arm points towards the whiteboard, and dots appear from the avatar’s hand to indicate the activation of the whiteboard. When a team member stops activating the whiteboard, also the avatar’s arm should be lowered. An avatar automatically lowers its arm and just stands still after certain, sometimes a relatively long amount of time not being operated by the user. However, as we saw in all excerpts above, it was rather common that although a team member had completed the questionnaire, even detached their hands from the mouse and keyboard (see Excerpt 6) the avatar kept touching the whiteboard. Thus, it seems that if avatars were not moved elsewhere after completing the questionnaire, the team members had very few cues about whether the others were ready or not (in case they did not directly ask about it, which occurred only once in our data).

Our findings on virtually embodied behavior are consistent with Bennerstedt and Ivarsson’s (2010) observation about the use of avatar jumping between phases in games. According to them, during waiting periods in games, the players used avatar jumping to signal to the others that they were still in the game and had not left their avatar hanging behind. Also our data involve cases of jumping as a waiting activity (see Excerpt 3).

Altogether, our findings suggest that opening an encounter is more often preceded by virtually embodied behavior (28 episodes in Process 2) than by the participants' avatars just remaining still (12 episodes in Process 1). However, as mentioned in Section 3.3 above, our data also included 66 episodes where the 'surplus time' consisted of a gathering only, and no transition to an encounter occurred. These gatherings were of two types: waiting privately in silence for the next task (38 episodes), or walking with the avatar in the joint space but not opening an encounter (28 episodes). Some of these "gatherings" resemble certain moments in face-to-face interaction where the participants do not talk or do anything bodily but are nonetheless committed to being together at that moment, i.e., to the co-presence (see Vatanen, submitted).

These types of moments actually challenge the idea of dividing social situations strictly into the two categories of 'gathering' and 'encounter.' Rather, participant behavior in social situations seems to be better described as a continuum of orientations (see Vatanen, submitted). Somewhere in between focused encounters and unfocused gatherings are situations where the participants do not sustain a joint focus of attention, such as a conversation or another mutually coordinated activity, but nevertheless are committed to being co-present and together in the shared space – such as here the participants' commitment to accomplishing the whole collaboration session as a team. At these moments, the participants are physically – or, virtually – in the same space where they have just previously had a focused encounter, when the nature of the situation transforms into one where the togetherness and joint focus are more frail than they are in an encounter but stronger than in a gathering. (For more discussion and an example of such a situation in face-to-face interaction, see Vatanen, submitted.) In face-to-face interaction, the participants' orientations to one another and the co-presence can be traced observing their bodily behaviors, whereas in mediated co-presence such as a VW, the orientations are more difficult to prove, both for the analyst and especially for the participants who only have access to the co-participants' avatars. This seems to be related to the significance that avatar behavior has for

opening an encounter, which was discussed in the current study.

Comparing the 28 episodes where virtually embodied behavior was followed by an opening of an encounter (in Process 2), and the 28 episodes of avatar(s) walking but remaining in a gathering-like situation (or in a situation that seems to be in between the two situation types), it seems that moving about with the avatar indeed signals only availability or possibility for an encounter but does not automatically lead to it. For this reason, we conclude that virtually embodied behavior works only as a signal about *readiness* to move to an encounter, and it is up to the participants whether then to open an encounter or not. Consequently, not all virtually embodied behaviors can be called encounter *pre-beginnings* either; rather, that particular characterization can be done in retrospect only.

6 Summary and discussion

This study focused on the details of how gatherings turn into encounters in a VW. We aimed to further the empirical investigations of social interactions in VWs, especially by treating what in the field of computer-mediated communication has been called ‘social presence’ (i.e., an ‘encounter’) as a behaviorally displayed entity. We applied multimodal conversation analysis to examine the interactional practices that participants use in the gathering-encounter transitions. In our data, two main types of transition processes occur. First, there are episodes where the transition happens by one participant directly starting to talk, without any preceding virtually embodied behavior (Process 1). However, it is more common that the transition includes a virtually embodied pre-beginning phase that is then followed by verbal interaction (Process 2).

Contrasting our findings to Mondada’s (2009) and Mondada and De Stefani’s (2018) suggests that there are certain differences in how gatherings turn into encounters in a VW compared to face-to-face situations. Unlike in face-to-face situations (ibid.), in our data the interactional space was not stabilized at the end of the embodied pre-beginning, before the first verbal opening. In our setting, the physical

participants are rather stable behind their screens, but their avatars may keep moving when transitioning from a gathering into an encounter (and even during an encounter). The situation and the use of embodied resources are very different in a VW, and consequently, stabilizing the setting does not play a similar role. Furthermore, space and the (avatar) body seem to have different roles for establishing interactional space in a VW: movement in space helps to signal availability, whereas standing still signals a potential lack of (psychological) presence of the participant behind the avatar.

The interaction setting we studied is different from Mondada's (2009) also in other respects. Unlike Mondada, we did not study the very first encounters between strangers, or even unplanned encounters between acquainted persons (De Stefani/Mondada 2018). Rather, even though the participants in our setting were strangers to one another at first, they little by little became more acquainted with one another. They were also present with one another as avatars in a joint location the whole time, not briefly passing one another on a street. De Stefani and Mondada (2018) showed that people engage in interaction in a different manner depending on whether they interact with acquainted or unacquainted persons, which also has organizational consequences for the openings. In our setting, it was intriguing that the incremental familiarization of members within the teams did not seem to affect the ways in which encounters were opened time after time. Even though each team had the possibility to transition from a gathering to an encounter eight times during the whole collaboration session, the team members remained careful not to open the encounter immediately after finishing their own questionnaires, but usually only after several seconds and only after exploring the virtual space privately or by moving their avatars. In addition, such openings that would reveal any emerging acquaintanceship between the participants were rare (e.g., asking about the study fields of others). Thus, the incremental familiarization of the team members during the sessions did not seem to have consequences for how the encounters were opened; probably they could not achieve a sufficient level of familiarization after only three hours of interacting in

comparison with the persons in De Stefani and Mondada's study.

Our study also contributes to the prior literature on how interaction is resumed after it has been ceased for a moment. Previous research has detected three ways to resume interaction after a lapse in talk: participants may move to end the interaction, continue with prior talk, or start something new (Hoey 2018). In our data, it was interesting that the opening turn usually did not continue any prior topic (this happened in only six of the 40 episodes). Most often the participants opened the encounter by starting something new, often using the environment as a resource for generating talk. The environment-based openings concerned the joint virtual space and objects to which all team members had access, and in Process 2, also the preceding avatar movement. The rest of the openings related usually to the interaction setting or other team members. Furthermore, the social actions of the opening turns in our data involve noticings, information-requests, accounts, and proposals – in other words, a wider variety of actions compared to those observed in Mondada (2009) and De Stefani and Mondada (2018).

Our study illustrates how different kinds of avatar movements, such as walking and jumping, have different consequences for the following interaction. *Walking* with the avatar seemed to help a team member to notice things in the virtual space and then to use this information to open an encounter. In addition, witnessing walking seemed to work as a hint for the co-participants that the walker is ready with their questionnaire and thus available for an encounter. However, witnessing walking did not lead into initiating conversation, whereas witnessing playful avatar behaviors such as *jumping* was reacted to by other team members and thus invited others to open an encounter. A reason for this may be that unlike jumping, walking was familiar for all participants as they were instructed on that during the orientation session. Thus, co-present team members treated walking with “civil inattention” (Goffman 1963: 84), as would be done in gatherings on the street, for example. Furthermore, both walking and jumping functioned as “waiting behaviors” in our data (cf. Svinhufvud 2018; Ayaß 2020).

Our findings also elaborate Mennecke et al.'s (2011) theory on embodied social presence which implies that VW affordances such as avatar body, virtual space, and virtual objects as well as verbal and nonverbal communication need to be used to transition from 'co-presence' to 'social presence,' i.e., from a gathering to an encounter. However, this theory does not explicitly focus on the ways in which the VW affordances are used to achieve this transition. Our study thus offers a major new insight in this topic. Future research could explore how familiarity with one's avatar and the space might influence the ways in which encounters are opened in VWs as well as in constant virtual teams where the members know one another well. In addition, the transition from a gathering to an encounter in "public" VW spaces where the participants resemble more the passers-by in Mondada's (2009) study should be studied.

Appendix: Transcription conventions

Verbal communication (based on Jefferson 2004):

.	falling intonation
,	level intonation
↑	rise in pitch
spe <u>a</u> k	emphasis
spea:k	sound lengthening
£speak£	smiley voice
.hh	audible inhalation
hh	audible exhalation
HH	loud exhalation
he he	laughter
sp(h)ea <u>k</u>	laughter within talk
[beginning of overlap
]	end of overlap
(.)	micropause (less than 0.2 seconds)
(0.6)	pause in seconds
(())	comment by transcriber

1050 Embodied behavior (adapted from Mondada 2016b):

- * Symbols (one per participant) indicate moments when the
- + described embodied actions begin (or occur, if they are momentary),
- Δ and they are synchronized with corresponding stretches
- of talk or silence. The described embodied action lasts until
- the next action by that participant is marked in the transcript.
- *---> The action described continues across subsequent lines
- >* until the same symbol is reached.
- >> The action described begins before the excerpt's beginning.
- 1051 --->> The action described continues after the excerpt's end.

1052 References

- 1053 Allmendinger, Katrin (2010). Social Presence in Synchronous
 1054 Virtual Learning Situations: The Role of Nonverbal Signals
 1055 Displayed by Avatars. *Educational Psychology Review*, 22,
 1056 41–56. DOI: 10.1007/s10648-010-9117-8.
- 1057 Antonijevic, Smiljana (2008). From Text to Gesture Online: A
 1058 microethnographic analysis of nonverbal communication
 1059 in the Second Life virtual environment. *Information,*
 1060 *Communication & Society*, 11(2), 221–238. DOI:
 1061 10.1080/13691180801937290.
- 1062 Ayaß, Ruth (2020). Doing Waiting: An Ethnomethodological
 1063 Analysis. *Journal of Contemporary Ethnography*. DOI:
 1064 10.1177/0891241619897413.
- 1065 Benford, Steve/Fahlén, Lennart (1993). A Spatial Model of
 1066 Interaction in Large Virtual Environments. *Proceedings of*
 1067 *the Third European Conference on Computer-Supported*
 1068 *Cooperative Work*, 109–124. DOI: 10.1007/978-94-011-
 1069 2094-4_8.
- 1070 Bennerstedt, Ulrika/Ivarsson, Jonas (2010). Knowing the Way.
 1071 Managing Epistemic Topologies in Virtual Game Worlds.
 1072 *Computer Supported Cooperative Work (CSCW)*, 19(2),
 1073 201–230. DOI: 10.1007/s10606-010-9109-8.
- 1074 Bergmann, Jörg R. (1990). On the local sensitivity of
 1075 conversation. In I. Marková & K. Foppa (Eds.), *The*
 1076 *Dynamics of Dialogue* (pp. 201–226). Hemel Hempstead:
 1077 Harvester Wheatsheaf.
- 1078 Biocca, Frank/Harms, Chad/Burgoon, Judee (2003). Toward a
 1079 more robust theory and measure of social presence:
 1080 Review and suggested criteria. *Presence: Teleoperators*

- and Virtual Environments, 12(5), 456–480. DOI: 10.1162/105474603322761270.
- Broth, Mathias/Keevallik, Leelo (2014). Getting ready to move as a couple: Accomplishing mobile formations in a dance class. *Space and Culture*, 17(2), 107–121. DOI: 10.1177/1206331213508483.
- De Stefani, Elwys/Mondada, Lorenza (2018). Encounters in Public Space: How Acquainted Versus Unacquainted Persons Establish Social and Spatial Arrangements. *Research on Language and Social Interaction*, 51(3), 248–270. DOI: 10.1080/08351813.2018.1485230.
- Goffman, Erving (1959). *The presentation of self in everyday life*. New York: Anchor Books.
- Goffman, Erving (1963). *Behavior in public places*. New York: The Free Press.
- Goodwin, Charles (1984). Notes on story structure and the organization of participation. In M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 225–246). Cambridge: Cambridge University Press.
- Goodwin, Charles (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489–1522. DOI: 10.1016/S0378-2166(99)00096-X.
- Heath, Christian (1984). Talk and reciprocity: sequential organization in speech and body movement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 247–265). Cambridge: Cambridge University Press.
- Hoey, Elliot (2015). Lapses: How People Arrive at, and Deal With, Discontinuities in Talk. *Research on Language and Social Interaction*, 48(4), 430–453. DOI: 10.1080/08351813.2015.1090116.
- Hoey, Elliot (2018). How Speakers Continue with Talk After a Lapse in Conversation. *Research on Language and Social Interaction*, 51(3), 329–346. DOI: 10.1080/08351813.2018.1485234.
- Jefferson, Gail (2004). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). Amsterdam: John Benjamins.
- Kamunen, Antti (2019). How to Disengage: Suspension, Body Torque, and Repair. *Research on Language and Social Interaction*, 1–21. DOI: 10.1080/08351813.2019.1657287.

- 1122 Keevallik, Leelo (2018). Sequence Initiation or Self-Talk?
 1123 Commenting on the Surroundings While Mucking out a
 1124 Sheep Stable. *Research on Language and Social*
 1125 *Interaction*, 51(3), 313–328. DOI:
 1126 10.1080/08351813.2018.1485233.
- 1127 Keisanen, Tiina/Rauniomaa, Mirka/Siitonen, Pauliina (2017).
 1128 Transitions as sites of socialization in family interaction
 1129 outdoors. *Learning, Culture and Social Interaction* 14, 24–
 1130 37. DOI: 10.1016/j.lcsi.2017.05.001.
- 1131 Kohonen-Aho, Laura/Alin, Pauli (2015). Introducing a video-
 1132 based strategy for theorizing social presence emergence in
 1133 3D virtual environments. *Presence: Teleoperators and*
 1134 *Virtual Environments*, 24(2), 113–131. DOI:
 1135 10.1162/PRES_a_00222.
- 1136 Lombard, Matthew/Ditton, Theresa (1997). At the Heart of It
 1137 All: The Concept of Presence. *Journal of Computer-*
 1138 *Mediated Communication*, 3(2). DOI: 10.1111/j.1083-
 1139 6101.1997.tb00072.x.
- 1140 Mennecke, Brian/Triplett, Janea/Hassall, Lesya/Conde,
 1141 Zayira/Heer, Rex (2011). An examination of a theory of
 1142 embodied social presence in virtual worlds. *Decision*
 1143 *Sciences*, 42(2), 413–450. DOI: 10.1111/j.1540-
 1144 5915.2011.00317.x.
- 1145 Mondada, Lorenza (2009). Emergent focused interactions in
 1146 public places: A systematic analysis of the multimodal
 1147 achievement of a common interactional space. *Journal of*
 1148 *Pragmatics*, 41(10), 1977–1997. DOI:
 1149 10.1016/j.pragma.2008.09.019.
- 1150 Mondada, Lorenza (2012). Coordinating action and talk-in-
 1151 interaction in and out of video games. In R. Ayaß & C.
 1152 Gerhardt (Eds.), *The Appropriation of Media in Everyday*
 1153 *Life* (pp. 231–270). Amsterdam: John Benjamins.
- 1154 Mondada, Lorenza (2013). Interactional space and the study
 1155 of embodied talk-in-interaction. In P. Auer, M. Hilpert, A.
 1156 Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in language*
 1157 *and linguistics* (pp. 247–275). Berlin: De Gruyter.
- 1158 Mondada, Lorenza (2016a). Challenges of multimodality:
 1159 Language and the body in social interaction. *Journal of*
 1160 *Sociolinguistics*, 20(3), 336–366. DOI: 10.1111/josl.1_12177.
- 1161 Mondada, Lorenza (2016b). Conventions for multimodal
 1162 transcription. Retrieved from

- 1163 https://franz.unibas.ch/fileadmin/franz/user_upload/redakt
 1164 [ion/Mondada_conv_multimodality.pdf](https://franz.unibas.ch/fileadmin/franz/user_upload/redakt).
 1165 Moore, Robert/Ducheneaut, Nicolas/Nickell, Eric (2006).
 1166 Doing Virtually Nothing: Awareness and Accountability in
 1167 Massively Multiplayer Online Worlds. *Computer Supported*
 1168 *Cooperative Work (CSCW)*, 16(3), 265–305. DOI:
 1169 10.1007/s10606-006-9021-4.
 1170 Mortensen, Kristian/Hazel, Spencer (2014). Moving into
 1171 interaction – Social practices for initiating encounters at a
 1172 help desk. *Journal of Pragmatics*, 62, 46–67. DOI:
 1173 10.1016/j.pragma.2013.11.009.
 1174 Råman, Joonas (2018). The Organization of Transitions
 1175 between Observing and Teaching in the Budo Class.
 1176 *Forum: Qualitative Social Research* 19(1), 5. DOI:
 1177 10.17169/fqs-19.1.2657.
 1178 Sacks, Harvey/Schegloff, Emanuel/Jefferson, Gail (1974). A
 1179 simplest systematics for the organization of turn-taking for
 1180 conversation. *Language*, 50(4), 696–735. DOI:
 1181 10.2307/412243.
 1182 Schegloff, Emanuel (1998). Body torque. *Social Research*, 65,
 1183 535–596.
 1184 Schroeder, Ralph (2008). Defining Virtual Worlds and Virtual
 1185 Environments. *Journal of Virtual Worlds Research*, 1(1), 1–
 1186 3. DOI: <http://dx.doi.org/10.4101/jvwr.v1i1.294>.
 1187 Schultze, Ulrike (2010). Embodiment and presence in virtual
 1188 worlds: A review. *Journal of Information Technology*,
 1189 25(4), 434–449. DOI: 10.1057/jit.2010.25.
 1190 Schultze, Ulrike/Brooks, Jo Ann (2019). An interactional view
 1191 of social presence: Making the virtual other “real.”
 1192 *Information Systems Journal*, 29(3), 707–737. DOI:
 1193 10.1111/isj.12230.
 1194 Sidnell, Jack (2013). Basic Conversation Analytic Methods. In
 1195 Jack Sidnell/Tanya Stivers (Eds.), *The Handbook of*
 1196 *Conversation Analysis* (pp. 77–99). Chichester: Blackwell
 1197 Publishing.
 1198 Sivunen, Anu/Nordbäck, Emma (2015). Social Presence as a
 1199 Multi-Dimensional Group Construct in 3D Virtual
 1200 Environments. *Journal of Computer-Mediated*
 1201 *Communication*, 20(1), 19–36. DOI: 10.1111/jcc4.12090.
 1202 Svinhufvud, Kimmo (2018). Waiting for the customer:
 1203 Multimodal analysis of waiting in service encounters.

- 1204 Journal of Pragmatics, 129, 48–75. DOI:
1205 10.1016/j.pragma.2018.03.002.
- 1206 Szymanski, Margaret H. (1999). Re-engaging and Dis-engaging
1207 Talk in Activity. *Language in Society*, 28(1), 1–23. DOI:
1208 10.1017/S0047404599001013.
- 1209 Vatanen, Anna (2018). Trajectories of embodied activities and
1210 the management of lapses. Paper presented at the 5th
1211 International Conference on Conversation Analysis
1212 (ICCA2018), Loughborough, UK. 11–15 July.
- 1213 Vatanen, Anna (Forthcoming). The interaction order of silent
1214 moments in everyday life: Lapses as joint embodied
1215 achievements. In Michael Buchholz/Aleksandar
1216 Dimitrijevic (Eds.), *Silence and silencing in psychoanalysis: Cultural, clinical, and research perspectives*. Routledge.
- 1217 Vatanen, Anna (Submitted). Exploring co-presence during
1218 lapses (in Finnish everyday interaction): On encounters,
1219 gatherings, and ways to inhabit silent moments. Submitted
1220 to Jan Lindström/Ritva Laury/Anssi Peräkylä/Marja-Leena
1221 Sorjonen (Eds.), *Intersubjectivity in action*.