

Social Interaction. Video-Based Studies of Human Sociality.
2022 Vol. 5, Issue 1
ISBN: 2446-3620
DOI: 10.7146/si.v5i2.130876

Social Interaction

Video-Based Studies of Human Sociality

Talk and Embodied Conduct in Word Searches in Video-Mediated Interactions

Budimka Uskokovic & Carmen Talehgani-Nikazm
The Ohio State University

Abstract

This paper concerns ways L2 speakers utilize embodied practices when engaged in word searches in interactions via video-mediated interaction. Examining the role of embodiment in word search practices in semi-pedagogical conversations between native German speakers and German language learners, we demonstrate how interlocutors use the gesture of an upward extended index finger to manage extended word searches in the ongoing production of the turn. The gesture plus the request to wait (ein moment/"one moment") are used to indicate that the turn is temporarily put on hold, during which the L2 speaker orients to the screen to complete the ongoing word search. In addition, our paper scrutinizes how L2 speakers' embodied practices emerge in the context of video-mediated interaction.

Keywords: embodied practice, progressivity, video-mediated interaction, word search

1. Introduction

Using the analytical framework of conversation analysis and multimodality, we investigate talk and embodied practices of L2 speakers in conversations with L1 speakers in video-mediated interactions (VMI). By focusing on L2 speakers' word search sequences in the context of technology-mediated settings, we demonstrate how L2 speakers employ embodied resources to suspend the talk in progress in order to create space for a screen-based search. Word searches occur in cases where speakers temporarily stop their current course of action, and their turn in progress to search for a word or a phrase that is unavailable to them at the point in the conversation (Goodwin & Goodwin, 1986; Schegloff, 1979, 2007). Typically, speakers search for a word (e.g., a name of a person or an object) or larger structural units (e.g., a prepositional phrase or an adverbial expression of time). In the case of L2 speakers, entering a word search could mean either searching for a lexical item or grammatical item to which they momentarily have no access (i.e., lexical gap), or a lexical item that they have not yet acquired (Brouwer, 2003; Koshik & Seo, 2012; Kurhila, 2006; Skogmyr Marian & Pekarek Doehler, 2022/this issue).

Our analysis of L2 word searches in the digital environment of VMI reveals that when encountering trouble with the progressivity of their ongoing turn, L2 speakers frequently make use of the technologically enhanced setting and opt for a screen-based search for the next due element in their turn rather than inviting their L1 expert coparticipant to collaborate and provide a solution. In such cases, the L2 speakers suspend the talk for a moment and shift their orientation to the screen to initiate a screen-based activity until they find the sought-for item. Once they locate the lexical item, they integrate it into their turn, which was previously put on hold, with minimal delay. In the meantime, the L1 speaker orients to the L2 speaker's word search and withholds providing a potential solution. If L1 speakers have enough clues, or if they can draw inferences from the L2 speaker's prior turn, they may provide a candidate solution.

Our analysis also demonstrates another particular type of extended word search sequence which includes a verbal alert such as *ein moment* "one moment," accompanied by raising an extended index finger. We argue that the L2 speaker uses the gesture as an attention-getter to explicitly signal their commitment to completing the search and to create more space for their prolonged screen-based search. Thus, our analysis demonstrates how the embodied practice of showing an index finger (compared to one case in which the L2 speaker does not use a gesture) contributes to making the action of extended search using screen orientation recognizable for L1 recipients in the video-mediated interaction. In addition, we show, as revealed by the L1 speaker coparticipant's response, how they ascribe the action of request for more time to complete their screen-based search to their L2 speaker coparticipant's linguistic and bodily practice.

2. Gesture and talk in video-mediated interaction

Recent EM/CA-inspired research on VMI has shown how interactants' verbal and bodily practices may be shaped by the affordances and constraints that the particular technologically enhanced setting provides (Arminen et al., 2016; Mlynář et al., 2018). These findings have contributed to our understanding of how the organization and sequences of social actions, for example, opening video-mediated calls (Gan et al., 2020), showing objects (Licoppe, 2017; Rosenbaun & Licoppe, 2017), coordinating and organizing multiple temporal and sequential activities at the same time (Tuncer et al., 2020), and navigating online collaborative task-based activities (Pekarek Doehler & Balamán, 2021), have been adapted to VMIs. With respect to coparticipants' eye gaze direction, it has been argued that mutual gaze cannot be achieved due to camera- and screen-related obstructions (e.g., the distance between screen and camera) and the "fractured ecologies" (Luff et al., 2003, p. 5) of VMI, that is, coparticipants do not have shared access to each other's visual and screen domain (Heath & Luff, 1993).

Our multilingual VMIs are opportunities for L2 learners to meet with an L1 speaker and to have conversations on a variety of topics. There is no online task for them to collaboratively complete. In our data, we examine how L2 coparticipants accomplish word search actions in VMIs. In particular, we focus on the role of embodiment in L2 speaker word searches and how the L1 speakers orient to L2 speakers' embodied practices. Our findings contribute to the growing body of EM/CA research on how sequences of actions are accomplished (see Piirainen-Marsh et al., 2022/this issue) and the methods interactants use to organize their conduct in VMIs (see Mlynář et al., 2018 for a summary of EM/CA studies on VMIs).

3. Word search

Word searches are a type of self-initiated repair that are forward-oriented (Schegloff, 1979) and tend to put the progressivity of talk-in-interaction on hold (Sacks, 1992; Schegloff et al., 1977; Goodwin & Goodwin, 1986; Lerner, 1996). They become apparent when the participant experiences trouble in retrieving or finding a word needed to continue the turn (for a summary of repair, see Kitzinger, 2012). In such instances, speakers use a range of resources, such as sound stretches, cut-offs, and inter-turn pauses (Schegloff et al., 1977; Lerner, 1996), to signal to their interlocutors that they are having trouble accessing the next lexical item they need to continue their turn. Word searches may be designed to invite another speaker's participation and help in the search (Goodwin, 1981; Goodwin & Goodwin, 1986), and if they have enough clues, they may provide a candidate solution (Antaki, 2012). Speakers may also directly invite another speaker to provide a candidate solution using questions targeting a name (Goodwin, 1981).

In second language learner interactions, research has shown similar practices of word searches with some context-specific methods. Language learners may search for a range of language-related phenomena (e.g., words, word forms, syntactic structures, pronunciation) and present their solution try-marked for their coparticipant's confirmation (Koshik & Seo, 2012). When L1 speakers offer a candidate solution, it may be rejected by the L2 speaker insisting on bringing their turn constructional units into completion on their own (Theodórsdóttir, 2011). Language learners may use other resources to manage progressivity of L2 interaction. For instance, they may produce a candidate lexical item with rising intonation and gaze at their L1 recipient to seek their confirmation (Brouwer, 2004). L2 learners may also verbalize their lack of knowledge by posing self-addressed questions, such as "how do you say that" or "what is it" (Pekarek Doehler & Berger, 2019; Reichert & Liebscher, 2012; Svennevig, 2018).

In the past twenty years, there has been a growing body of research on the link between word searches and embodied practices in L2 interactions. These studies have shown that while searching for a word, L2 speakers typically employ embodied practices (Greer, 2013; Kurhila, 2006; Koshik & Seo, 2012; Skogmyr Marian & Pekarek Doehler, 2022/*this issue*), such as gaze shifts and indexical gestures. They may display a "thinking face" (Goodwin & Goodwin, 1986) or raise their eyebrows (Reichert & Liebscher, 2012). L2 speakers may also use a sharp head turn or tilt to the side with continued eye gaze toward the recipient to initiate repair (Seo & Koshik, 2010).

In the analysis section of the paper, we show that prior research findings on word searches (including the findings on embodied practices) in L1 and L2 interactions hold for the majority of the word searches in our data corpus. However, our analyses illustrate additional details, such as putting the progressivity of talk-in-interaction on hold and signaling to coparticipants that more time is needed to complete the word search by orienting to the screen and initiating a screen-based activity, which we argue are specific to the particularities of video-mediated interactions. We show that word searches that are long and complex practices are accompanied with a hand gesture that helps the current speaker secure space for their screen-based search. Below we provide a review of recent research on hand gestures.

3.1 Gestures in Word Searches

Previous research on L2 interaction has shown that, while engaging in a word search, L2 speakers routinely employ hand gestures and that it is a highly situated practice (Rydell, 2019). These studies have shown that L2 speakers use hand gestures to achieve understanding during a word search, and that the way in which L2 speakers employ their hand gestures depends on their environment (Rydell, 2019). Rydell's (2019) study on L2 speaker interaction shows that while engaging in a word search in analog test settings, L2 speakers employ hand

gestures to either invite the interlocutor to participate, or to signal the ongoing word search and still hold the turn. In addition, in a word search, L2 speakers may use hand gestures to describe and indicate to their recipient the word that is being sought (Egbert et al., 2004), and to index a particular domain of words, such as verbs or nouns (Hayashi, 2003).

In our paper, we focus on one particular gesture in word searches in video-mediated interaction. The gesture under investigation in this paper is the raised index finger (Kendon, 2004) accompanied by talk (see Figure 1).

Figure 1. *Raised Index Finger*



In this paper, we show how this raised index finger gesture is shaped by local circumstances of interaction, how L1 speakers orient to the multimodal activity and how, by employing such a gesture, L2 speakers hold the turn and signal to L1 speakers that they need to wait until the ongoing word search has been completed.

Even though hand gestures have been widely examined, predominantly in the pedagogical context (Lazaraton, 2004; Sert, 2017; Smotrova & Lantolf, 2013; Taleghani-Nikazm, 2008), the main goal of such studies was to investigate gesture use in instructional settings. None of these studies have looked at the L2 speaker's gestures and the coparticipant's orientation to them in the context of word searches in the digital environment of video-mediated interaction, when the recipient does not have full access to the current speaker's world, that is, when the interlocutors do not share the same physical space. Thus, our paper aims to examine the role of the raised index finger during word searches in VMIs when the coparticipants do not have access to each other's visual domains and screens. In our analysis, we demonstrate how such a "pragmatic gesture" — similar to a raised hand with a palm facing the interlocutor, which may signal to the interlocutors that they need to wait their turn (Streeck, 2009, p. 179) — is put to use for specific interactional practice, how it is made recognizable for its recipient, and how coparticipants orient themselves to the emerging activity (Oloff, 2021; Kamunen & Haddington, 2020). More specifically, we illustrate how this raised index finger gesture is used to get the recipient's attention, suspend talk and request the recipient to wait, and that it is understood as such by the L1 speaker coparticipant.

4. Data

The data examples illustrated here are from a corpus of video-recorded TalkAbroad interactions between German L1 speakers and German language learners (L2 speakers) collected by the first author in 2019. The video-based synchronous platform, TalkAbroad, provides a venue for real-time conversations about various language and cultural topics for L2 speakers with L1 speakers from 15 different countries. All L2 speakers of German in this study are students enrolled in a low intermediate German (third semester German) course at a large midwestern university in the United States. TalkAbroad partners are German L1 speakers who are typically students in their early twenties. In all instances, L1 and L2 speakers were informed about the video recordings and the aims of the project and gave their written consent for the use of data in videos and transcripts for research purposes.

The TalkAbroad conversations in our data corpus were recorded as integral class assignments in the third semester German course. The L2 speakers talked about different cultural topics, such as sports, transportation, and entertainment. The interlocutors in the examples presented in this paper met for the first time and took the time to get to know each other at the beginning of each conversation. The total recording time of the video data is 28 hours. A table with four different categories of word searches found in the corpus is presented below.

Table 1. Total number of word searches in the data corpus and in each category

Bodily and screen-based word searches without talk	Screen-based word searches with talk (<i>ein moment/one moment</i>) and raising index finger	Screen-based word searches with talk (<i>ein moment/one moment</i>) but without raising index finger	Screen-based word searches with talk (e.g., <i>ein moment/one moment</i>) and various gestures preceding talk	Total
19	8	5	18	50

In our collection, we have found 50 instances of word searches that we then grouped into four categories. The first category refers to searches carried out by bodily and screen-based activity without talk, that is, L2 speakers swiftly orient to the screen to look up a word without using the verbal alert *ein moment*. On the other hand, the remaining three categories refer to word searches in which L2 speakers orient to the screen to complete a word search and thereby signal to L1

speakers that they have initiated a screen-based word search by using a verbal alert. However, the three remaining categories differ from each other as follows: In the second category, L2 speakers raise their index finger to indicate that they need more time to complete their word search. In the third category, L2 speakers use a verbal alert to indicate an ongoing screen-based word search but they do not simultaneously raise their index finger. Finally, in the fourth category, L2 speakers use a verbal alert and various gestures, such as putting a finger on the lips or twirling hair. Out of the 50 word searches found in the corpus, we closely scrutinized the first three categories of word searches. In other words, the focus of this study is on word searches that include screen-based activity without verbal alert, with the L2 speaker's verbal alert such as *ein moment* "one moment" accompanied by an index finger extending upward to create space for an extended screen-based search, and with verbal alert but without gesture. We do not focus on the fourth category due to space limitations and our interest in the one particular gesture. The organization of embodied conduct and talk in the cases of the word searches that we analyze is not idiosyncratic, i.e., it is not used solely by one L2 speaker, but rather by multiple L2 speakers in the data.

Even though participants had mutual access to sound and image (screen) in real time, neither they nor the researchers had access to each other's screen. Since the participants' screen-based activities were not recorded and no screenshots of the L2 speakers' screens are available, we do not know the specifics regarding the kind of web-based dictionaries or translation tools the L2 speakers used during the screen-based word searches.

Our cases are analyzed by using the theory and methodology of conversation analysis (CA) (Sacks et al., 1974; Heritage, 1988; Hutchby & Woofitt, 1998). The method used is multimodal conversation analysis (Mondada, 2018). We provide one transcript for each excerpt. The top line is the original talk in German. The English translation is provided below each line in blue. The excerpts include anonymized screenshots of the participants at the indicated moment. The screenshots are numbered in red in the transcripts. The researchers used the free version of Bandicut, VSDC Free Video Editor, and Adobe Premiere Pro to edit and anonymize the videos. For the anonymization of screenshots, the researchers used the free version of the app Sketch Me!

5. Analysis

In the analysis section below, we demonstrate three different categories of word searches. In the first section (Excerpt 1), we present one example from the majority of L2 word search cases in our collection that shows how word searches are swiftly resolved by the trouble-source speaker's (L2 speaker) screen-based search, causing minimum delay in the progressivity of talk-in-interaction (Goodwin, 1989; Sacks, 1987; Schegloff, 2007; Streeck, 1995). In the second section, we present a selection of instances in which L2 speakers undertake long

and complex word search practices. We then show three cases in which an L2 speaker employs a raised index finger to suspend the ongoing talk. In the first case, the raised finger co-occurs with the verbal alert *eine momente*¹ “one moment” (Excerpt 2). In the second case, the raised index finger is followed by the verbal alert *ein moment* “one moment” (Excerpt 3). In the third and final case, the raised index finger remains “in the air” after the verbal alert has been uttered (Excerpt 4), which can be characterized as a gesture for “hold.” In doing so, we show that L1 speakers orient differently to the gesture and that not all L1 speakers recognize the screen-based activity as such immediately. Our analysis focuses on how L2 speakers initiate and resolve long and complex word search practices, how L1 and L2 speakers orient to them, and at which point in time L2 speakers use an index finger extended upward. We begin with examples in which L2 speakers search for a lexical item, L2-related information, and then for specific information. We end the analysis section with a word search with talk but without gesture.

5.1 Screen-based word search

Excerpt 1 illustrates an instance of word search in which an L2 speaker swiftly puts the conversation on hold to orient to the screen and conduct a screen-based search for three German words: *Mannschaft* “team,” *Ringen* “wrestling,” and *Hochschulsport* “high school sport” (lines 5-6, 9, and 12). This excerpt shows the smoothness of the word search when an L2 speaker shifts orientation to screen-based search without employing hand gestures. It is taken from a conversation between Mason (L2 speaker) and Kathrin (L1 speaker). Here, the coparticipants are talking about the kinds of sports they like and play. In lines 1 and 2, Kathrin asks Mason which sports he likes or plays.

>>Video<<

1 In German, only *ein moment* would be correct.

Excerpt 1. MAS + KAT

1 KAT: und welche sportArten magst du?
and which sports do you like?

2 oder MACHSt du↓
or DO you↓ do

3 MAS: ja ah-uh-ich uh wandern und klettern? [aber ich habe
yes ah-u-I uh hike and climb? [but I have
[

4 KAT: [mh

5 MAS: kei:n: uhm (0.2) tsk (.) %keine::% \$(0.5) \$mannschaft?
no: uhm (0.2) tsk (.) no:: (0.5) team?
%typing% \$reading\$

6 #u:h# sp#ort?# [ja. ja.= in highschool ich habe
u:h sport? [yes. yes.= in highschool I did
kat: #nods# #nods# [

7 KAT: [mhm

8 MAS: ^a::hm (0.3) tsk ^schwimmen, un:d
a:hm (0.3) swim, a:nd
^gaze up--->^

9 &a:hm#1&% (1.0)% \$rin #2&gen?&\$
a:hm (1.0) wre stling
&L hand %typing% \$reading &tilts h.&\$
&keyboard&



Mason (MAS)

Kathrin (KAT)

Figure #1

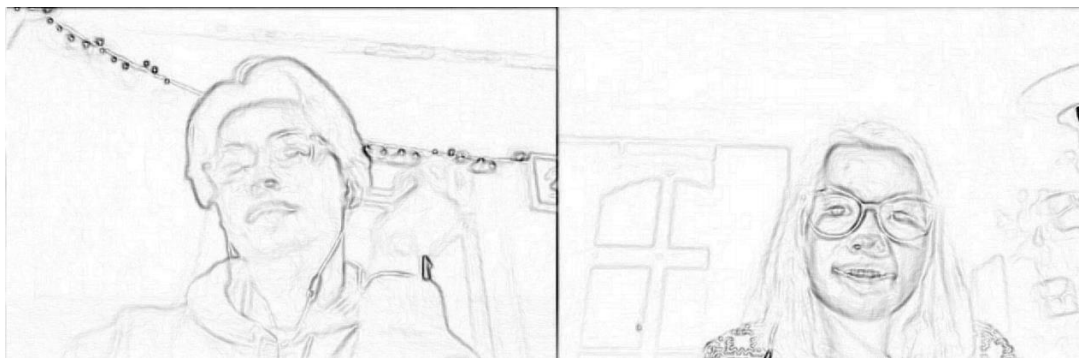


Figure #2

```

10 KAT:      *↑oh*           *↓okay*
             *moves back*   *nods up*

11 MAS:      ja das ist das aber ich bin im uni
             yes that is it but I am at college
12           und ich habe kein #3(0.2) %a::hm
             and I do not have any      a::hm
                                           %typing -->

```

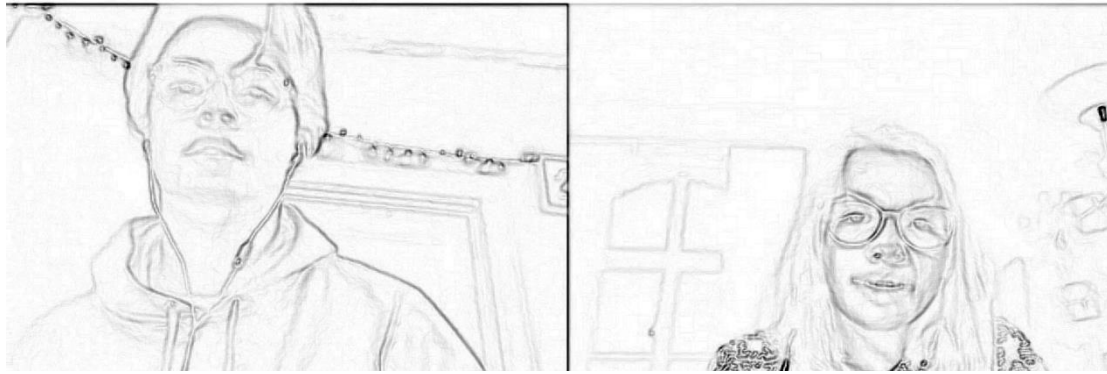


Figure #3

13 (0.3)
 mas: ----->
 14 KAT: also du machst% keinen **hochschulsport**
 well you do not do any college sports
 mas: ----->%
 15 MAS: ja, yes, yes hochschul- ja.ja.
 yES, yes†, yes† college- yes.yes.
 16 (0.2)
 17 MAS: [anyway] ja.ja.
 yes.yes.
 18 KAT: [okay]
 ((Kathrin continue to talk about her college sports activities.))

The first word search occurs in lines 5 and 6. After listing the sports he plays, Mason continues his turn with the conjunction *aber* "but," projecting delivering some contrastive information (lines 5 and 6). His turn exhibits sound-lengthening (*kein:n*: "no"), delay token *uhm*, intra-turn pause of 0.2, lip-smacking sound, and a micro pause indicating trouble finding the next element needed for completing his turn. It is here that Mason orients to the screen, shifts to typing, and immediately reads the noun *mannschaft* "team" from the screen and continues his turn. The shift to typing and reading the found word takes 0.5 seconds. Note Kathrin's nods (line 6) and *mhm* (line 7), displaying her orientation to the word search activity and signaling receipt of Mason's talk. Mason continues his response-turn by adding the different kinds of sports he played in high school, namely swimming and wrestling (lines 8 and 9). In line 8, we observe another word search displaying similar elements as in line 5: sound lengthening in speech perturbation plus intra-pause and lip-smacking, signaling trouble with the ongoing production of his turn. This word search is solved by Mason's producing the lexical item *schwimmen* "swim." Following this, his gaze shifts upward, projecting further problems with the ongoing production of his turn. Mason continues his turn with *un:d* "a.nd," which is also produced with sound lengthening and followed by vowel stretching *a:hm*. At this point, he moves his left hand from his face to

the keyboard, shifts his gaze to the keyboard and begins to type (Figure 1). After a 1.0-second pause, Mason reads *ringen* from the screen. Note that while Mason proposes the word *ringen*, he tilts his head (Figure 2), which may be indicating possible doubt of the word choice (Barrow, 2010). It is at this point that both speakers orient to the screen again. In return (line 10), Kathrin utters a high-pitched *oh*, an emotional change-of-state (Golato, 2012), signaling receipt of a new and surprising piece of information (wrestling may not have been what she expected to hear), latched to *okay*, whereby Kathrin displays her understanding of the new information. In both cases of screen-based word searches, Mason orients to the screen and reads the words (*mannschaft*, line 5, and *ringen*, line 9). These words do not seem to be part of his linguistic repertoire, or he cannot recall them at the moment. During Mason's screen-based activity, Kathrin gazes at the screen and does not provide a candidate solution, thereby displaying her orientation to Mason's solitary word search action.

In line 12, in the continuation of his response to Kathrin's question, Mason initiates the third word search, namely the word *hochschulsport*, "high school sports," which does not seem to be part of his linguistic repertoire. Again, we can observe a pause for 0.2 second, during which Mason shifts his gaze to the screen (Figure 3), a stretch of the vowel *a:hm*, and a 0.3-second gap (line 13), which are followed by a shift to the screen and visible typing. In line 14, probably based on Mason's turn in line 11 and his utterance *im uni* "at the college," Kathrin makes the inference *hochschulsport* "college sports" and offers a candidate solution. Mason accepts Kathrin's candidate solution and repeats part of it (line 15).

In sum, it is apparent that when encountering trouble progressing with the ongoing turn because the next item in the target language is not readily accessible, L2 speakers typically suspend the talk for a moment, orient to the screen, and initiate a screen-based activity until they find the sought-for item. Once they locate the lexical item, they integrate it into their turn. In our data collection, 19 word searches are resolved by the L2 speaker swiftly via screen-based activity without a verbal alert and gesture, as shown in Excerpt 1.

While the word searches presented in Excerpt 1 hold for the majority of our cases in the collection, we also noted cases revealing other features of repair operation. In these cases, the word search includes the utterance *ein moment* "one moment" accompanied by raising an extended index finger. These cases of word search are the focus of the next section.

5.2 Screen-based word searches accompanied with raised index finger

Excerpt 2 is from a conversation between Samuel (L2 speaker) and Herbert (L1 speaker). Here, upon multiple self-repairs involving replacement, Samuel engages in the action of word search (line 4). His word search involves the utterance *eine momente* "a moment," a hand gesture, and the screen-based activity of looking up the sought-for word. His gesture coincides exactly with the

utterance *eine momente* to alert his coparticipant of his incipient screen-based activity.

>>Video<<

Extract 2. 'brauchen'

1 HER: a:hm fängt man wenn man anfängt <pilot zu studieren>
when you start studying for a pilot

2 (0.2) direkt [an] ^SELBer zu fliegen
(0.2) immediately [to] ^fly on your own

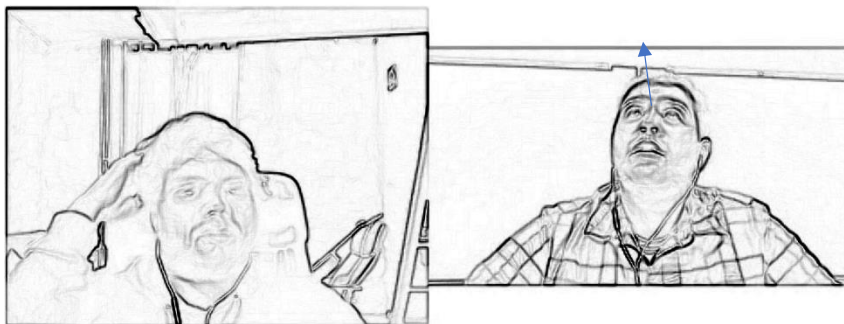
sam: ^gaze up - - - - ->

3 SAM: [.tsk] (1.0) a:hm (0.2) mhm ja↑ mei^nem ersten ^und^
[.tsk] (1.0) u:hm (0.2) uhm yes↑ m^y first ^and^

----->^ ^gaze up^

4 SAM: a:h (0.2) #1^ich-<how do you say need>
a:h (0.2) ^I-<

^gaze up----->



Herbert (HER)

Samuel (SAM)

Figure #1

5 ich^ &aha & ^ (0.3) ich dürft-^
I ^ &aha & ^ (0.3) I might-^

-->^ &head down& ^closed eyes----->^

6 #2&dürfte,&
might,
&tilts head&



Figure #2

7 #3hold on=



Figure #3

8 HER: =brauche?, vielleicht
=need?, maybe

9 SAM: #4&eine momente& %(0.5)% #5\$ich **brauche**\$
&one moment& %(0.5)% \$I **need**\$
&finger up-->& %typing% \$reading--->\$



Figure #4



Figure #5

10 #6 a:hm ^ein tausend fiv-fünf hundert^ (0.2) ^stunde^
 a:hm ^one thousand fiv-five hundred^ (0.2) ^hours^
 ^gaze up----->^ ^gaze up^



Figure #6

11 HER: eu hehehe
 wow hahahah
 12 SAM: flugzeit
 flight time
 13 HER: aha eu eu
 aha wow wow
 14 SAM: so (0.2)
 ((Samuel continues to explain how many hours he already had.))

Samuel's response to Herbert's question (lines 1-2) includes a word search. The word search is initiated by the vowel stretch of *a::h*, an intra-turn 0.2-second pause (line 3), and an upward gaze shift (Figure 1), followed by the "explicit word search marker" (Brouwer, 2003; Reichert & Liebscher, 2012) in English "how do you say need" (line 4). The upward gaze shift indicates that the reference "you"

is not a reference to Samuel's coparticipant. The self-talk "how do you say need" is a public display of his trouble with progressing the turn in German, namely not knowing the word "need" in German. In line 5, we have a restart *ich* and a downward gaze shift, a speech perturbation, a 0.3-second silence, and repetition of *ich* and the verb *dürft-* and self-repair *dürfte* (line 6) displaying Samuel's attempt at constructing a response to Herbert's question. Note that while Samuel proposes the word *dürfte* "might" in line 6, he tilts his head (Figure 2), which may be indicating possible doubt of the word choice (Barrow, 2010; Seo & Koshik, 2010).

Figure 3 shows that Samuel is looking at the screen which could be interpreted as creating an occasion for collaboration for Herbert. In line 7, Samuel shifts his gaze to the screen and produces "hold on," signaling to his coparticipant that he needs more time to complete his turn. Immediately following this, Herbert offers a candidate solution, which is not taken up by Samuel. In line 9, he utters *eine momente* "one moment," leans toward the screen, begins to type, reads from the screen (observable by his eye gaze movement) (Figure 5), and replaces the verb *dürfte* "might" with *brauche* "need." Samuel's German utterance *eine momente* is accompanied by a hand gesture (Figure 4), an extended forefinger ("index finger").

Samuel's gesture visually represents the temporal content of his utterance *eine momente* and together with talk explicitly indicates that he needs some time to engage in a screen-based search and thus marks Samuel's need to suspend the talk temporarily. In combination with *eine momente*, the speaker displays an active claim of the turn space and continued commitment to solving the trouble and turn progression. In doing so, Samuel directs Herbert's attention to his word search and underscores the transitioning (Kamunen & Haddington, 2020) to the emerging activity — namely searching for a word by orienting to the screen. In return, Herbert's gaze, silence, and body position do not signal any attempts to take the floor, and thereby he demonstrates his understanding of Samuel's gesture as a signal to wait and allow him to complete his screen-based search for the sought-for word. In addition, the raised index finger displays Samuel's speakership, it holds his turn, and exhibits his rights over the sequence (Mondada, 2007). Herbert orients to Samuel's request for a temporary disengagement from the talk: He maintains his gaze at screen and waits for his coparticipant to complete his search (line 10, Figure 6). Herbert receives Samuel's flight hours with surprise (lines 11 and 13), thereby displaying receipt and understanding of the new information and closes the sequence.

In sum, Excerpt 2 illustrates that word search is a *complex multimodal Gestalt* (Mondada, 2016) in that the L2 speaker utilizes all available resources in VMI to solve his forward-oriented repair (Schegloff, 1979) and resume the talk. He employs verbal and embodied resources simultaneously (line 9, Figure 4). Bodily, the L2 speaker leans toward and back from the screen, he shifts his gaze, and employs his index finger to direct L1's attention to the crucial moment for the

further development of the interaction. Verbally, the L2 speaker signals the delayed production of the lexical due item while reserving his right to the turn-in-progress (Schegloff, 1979) not only in his L1 (line 4) but also in his L2 (line 6).

In sum, the word searches in Excerpt 1 and Excerpt 2 are different sequentially. In Excerpt 1, the L2 speaker initiates his word search by gazing up (line 8), lengthening vowels (line 8), speech perturbations and pause (line 9). During the 1.0-second pause, he utilizes his keyboard to type and to thereby initiate a search-based activity. In the end, he reads from the screen and tilts his head to seek confirmation from Kathrin. On the other hand, in Excerpt 2, we observe an extended word search, which the L2 speaker again initiates by gazing up (line 3), speech perturbations and pause (line 4), and then an explicit word search marker “how do you say need,” which is then followed by an explicit request to wait “hold on” (line 7) in English and then utterance *ein moment* in German (line 9) asking to halt the progressivity of the ongoing talk for the sake of screen-based search for the item due. Similar to Excerpt 1, the side sequence (Jefferson, 1972) ends with the L2 speaker reading the result from the screen-based search. Thus, we argue that the other-addressed verbal alert (“one moment”) and raised index finger explicitly and publicly signal to the coparticipant that the word search that they have initiated requires the speaker to put the turn in progress temporarily on hold in favor of a search for the word which involves a screen-based activity. In doing so, the L2 speaker directs the coparticipant’s attention to their action of looking up a word, thereby discouraging their coparticipant’s involvement in their current action and securing the possibility of word search. The trouble in Excerpt 2 was a specific lexical item to which the L2 speaker did not have cognitive access. Excerpt 3 illustrates a word search in which the trouble is not a lexical item but the metric system, in other words, a fact.

The third Excerpt comes from the same category of phenomena — a gesture accompanied by the utterance *ein moment* “one moment.” However, in this instance, the gesture is followed by the utterance. The problem (that is, the word sought) is the speed limit in the U.S. in kilometers. After some talk during which the two coparticipants (Andy, AND, L2 speaker and Kathrin, KAT, L1 speaker) get to know each other, they move on to talk about cars, the German *Autobahn* “freeway” and driving speeds. In this context, Kathrin explains how fast her car drives, which is followed by an exchange of positive assessments of this information (lines 1 and 2). Following this, Andy produces an informing which contains the speed limit on the highways in the U.S. Our focus is on Andy’s word search in lines 3 and 4.

>>Video<<

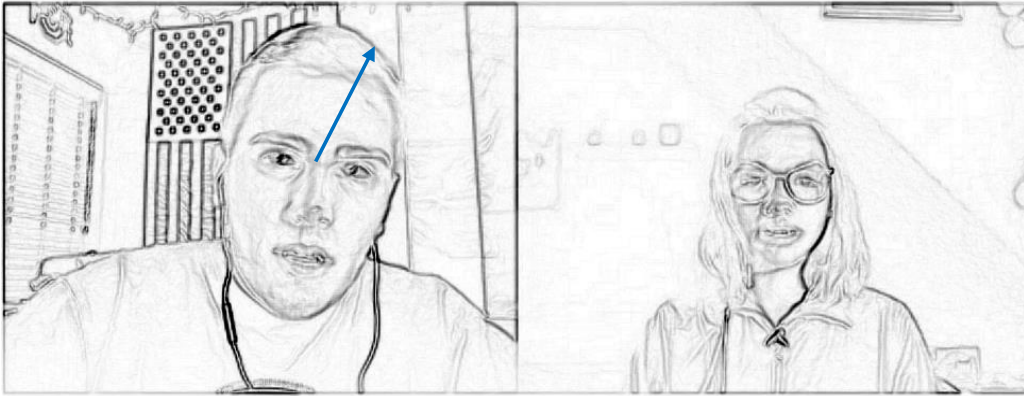


Figure #3

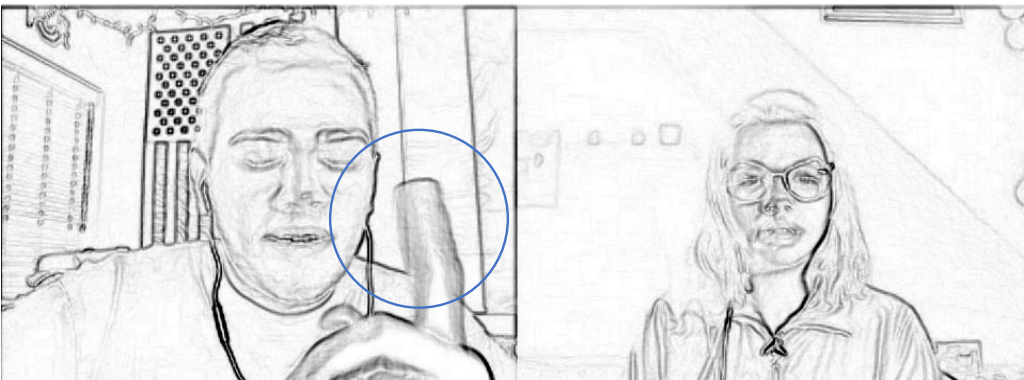


Figure #4

```

5      a:[hm]
      a:[hm]
and - - - - - >
6 KAT:  ^[hun]dert oder hundertzwanzig^ kaemha
        [hun]dred or hundred twenty kmh1
and - - - - - >
and ^gaze up, looking at the screen^
7      glaube ich
        I think
and - - - - - >
8 AND:  %a:hm es ist (.) einhundred a:hm (0.2)
        %a:hm it is (.) one hundred a:hm (0.2)
        - - ->%
9      einhundred dreizehn kilo-*kilometer* a:h
        one hundred thirteen kilo-kilometer a:h
        *nodding*

```

¹ The German abbreviation for kilometer per hour.

10 pro stunde ja=
 per hour yes=
 11 KAT: = hundert zehn ka[emha
 = hundred ten k[mh
 12 AND: [ja ja ahm [in den usa
 [yes yes ahm[in the states

 13 KAT: [okay [das
 [okay [it

((Andy continues asking about cars in Germany and whether each person in Germany should have a car.))

Following mentioning the location, *in den usa* “in the states,” we observe speech perturbations and sound-lengthening (*a:hm*, *a::h*) and a slight 0.3-second pause indicating Andy’s trouble with continuing constructing his turn. Note that Andy maintains his eye gaze at the screen (Figure 1) during the first adverbial phrase and shifts his eye gaze downward at the onset of the word search (Figure 2), while Kathrin maintains her gaze at the screen. Andy continues with the second component in his turn *das maximum ist* “the maximum is” (line 4), which is try-marked. This is grammatically continuous with the prior talk (before the speech perturbations and pause), projecting a complement (adverbial), namely the speed on a freeway in the U.S. (line 4). Andy shifts his gaze up (Figure 3) and produces the word *maximum* “the maximum,” indicating trouble accessing the next item due. He then raises his extended index finger, which is followed by the verbal alert *ein moment* “one moment” (line 4, Figure 4), signaling that the speaker requires further time for searching the sought-for word or information. Andy’s talk and gesture signal that further proceedings of talk will be temporarily put on hold. Immediately following this, Andy shifts his activity.

While Andy is gazing at his keyboard and displaying that he is still engaged in the screen-based activity, Kathrin offers a candidate solution with rising intonation (line 6) and completes her turn with *glaube ich* “I think” seeking confirmation (line 7). Andy, however, does not acknowledge the information. Rather, he continues to type until the onset of line 8. Andy’s raised index finger can be seen as him displaying publicly the need to refrain from conversation and holding the space to conduct the screen-based search, and not to confirm the try-marked candidate solution in line 6. In line 8, he provides a solution to his word search by reading from the screen (lines 8-10), thereby resuming the talk that had been suspended in favor of repair. In response, Kathrin repairs Andy’s utterance (line 11): Kathrin displays through her nodding that she understands Andy and rounds the number down to 110. In line 12, Andy acknowledges and aligns with Kathrin’s repair by using a confirming token *ja ja* “yes yes” to display that her prior utterance contains already known information and that Andy acknowledges and aligns with repair. In line 13, Kathrin returns to the main topic.

Similar to Excerpt 2, we can see that the gesture and the utterance *ein moment* occur in a context in which the word search involves a complex trouble source.

Prior to the talk and the gesture, the L2 speaker's turn displays several features of repair initiation projecting trouble with the progression of the ongoing turn. By using a gesture, the L2 speaker signals to the coparticipant that the word search requires more time and that they will disengage from further conversation in order to conduct a search using an outside resource. In combination with *ein moment*, the speaker displays an active claim of the turn space and continues commitment to solving the trouble and turn progression by themselves. However, not all L1 speakers orient to the gesture in like manner. In Extract 2, the L1 speaker orients to the gesture as a signal to wait and does so. In Extract 3, however, the L1 speaker provides her candidate solution during the L2 speaker's screen-based activity, thereby not orienting to her coparticipant's gesture nor complying with his verbal alert for some time. One of the reasons why Kathrin provides a candidate understanding and thereby ignores the request to wait might be due the nature of word search — it is not a straightforward word search, but rather a search for information (about speed limits and conversion from the American metric system into the German metric system). By providing a candidate understanding, Kathrin displays having enough cues and expertise to solve L2 speakers' word search. Note, however, that Andy does not accept Kathrin's candidate solution and that the word search is solved by the L2 speaker.

Excerpt 4 comes from the same category of phenomena — the gesture accompanied by the utterance *ein moment* “one moment.” However, in this instance, the L2 speaker (Sophia, SOP) holds her raised extended index finger even after uttering *ein moment* “one moment.” The word search instance of interest occurs in a context in which the L2 speaker attempts to correct the L1 speaker's (Herbert, HER) misunderstanding of the L2 speaker's prior talk. Here, the coparticipants have been talking about types of sports activities they do. Upon Sophia's explanation that she used to do track and cross-country running during her high school years (lines 3-8) and her self-correction in which she replaces the English word cross country in her previous turn (line 8) with its German equivalent *long streckenlauf*² (line 10), Herbert offers a candidate solution. Herbert's response turn demonstrates his inference from Sophia's prior talk, try-marked seeking confirmation. In response, Sophia disconfirms (*oh nein* “oh no”) Herbert's candidate solution (line 12) of her running a marathon and, following some laughter, continues correcting his misunderstanding (lines 12-17). It is in this context that she engages in a word search that includes *ein moment* “one moment” plus a gesture (line 15).

>>Video<<

2 Note that the German equivalent would be “der Langstreckenlauf”.

Excerpt 4. 'drei meilen'

1 HER: bist du schon ma:l (.) ski gefahren?
 have you ever been (.) skiing?

2 (0.2)

3 SOP: a::hm (0.2) ja das ist a:hm sehr gut
 a::hm (0.2) yes it is a:hm very good

4 (0.2)

5 HER: mhm

6 SOP: ja a:hm (0.3) ja. ich keine sport spiele a:hm
 yes a:hm (0.3) yes. I no sport games a:hm

7 in der high school a:h a:hm bin ich
 in the high school a:h a:hm I did

8 track und cross country? gelaufen
 run track and cross country?

9 HER: ah ja ja
 ah yes yes

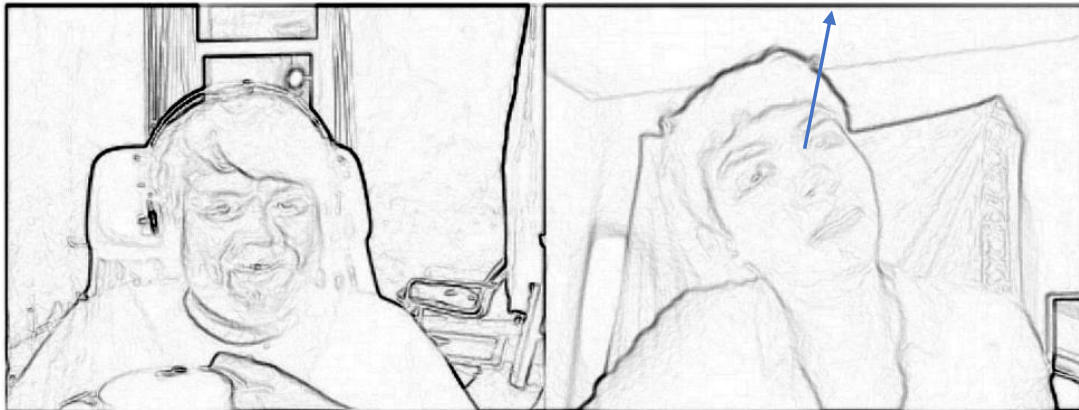
10 SOP: ja long strecken lauf
 yes long-distance running

11 HER: also marathon und so was?
 well marathon and something like that?

12 SOP: oh nein [laughter]
 oh no

13 HER: [laughter]

14 SOP: a:hm #1^(.) ^ bo- a::hm (0.2) .tsk
 ^gaze up ^



Herbert (HER)

Sophia (SOP)

Figure #1

```

15      #2^&ein moment    & &a:hm &.tsk #3%(0.3)%^
        one moment
        ^gaze down ----->^
        &finger up      & &exp. g.& %typing%
  
```

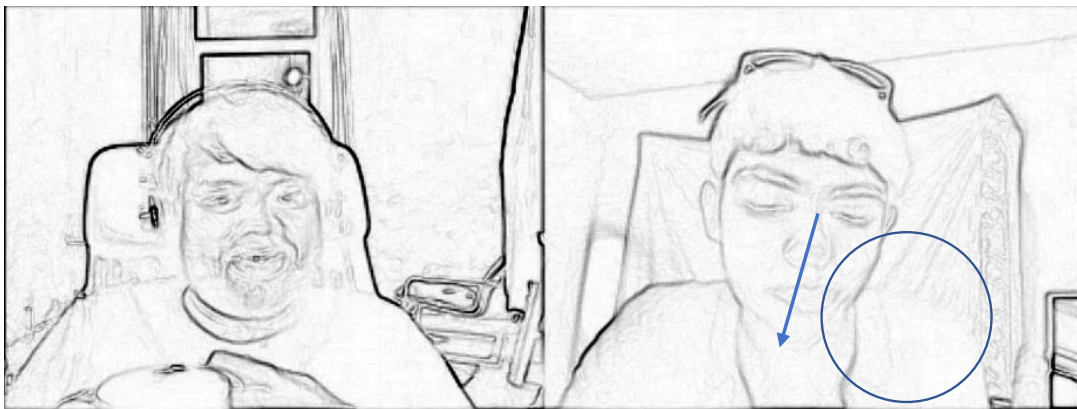


Figure #2

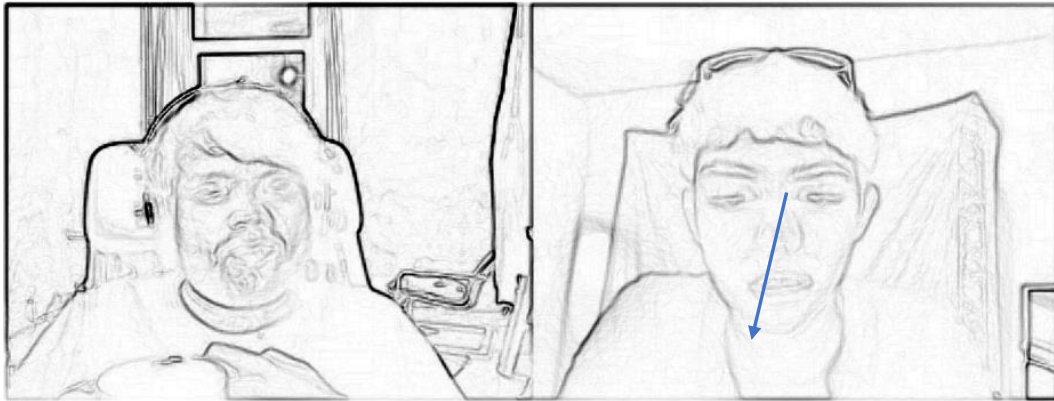


Figure #3

16 #4\$a:h three\$ or drei meilen
 \$a:h three\$ or three miles
 \$reading->\$

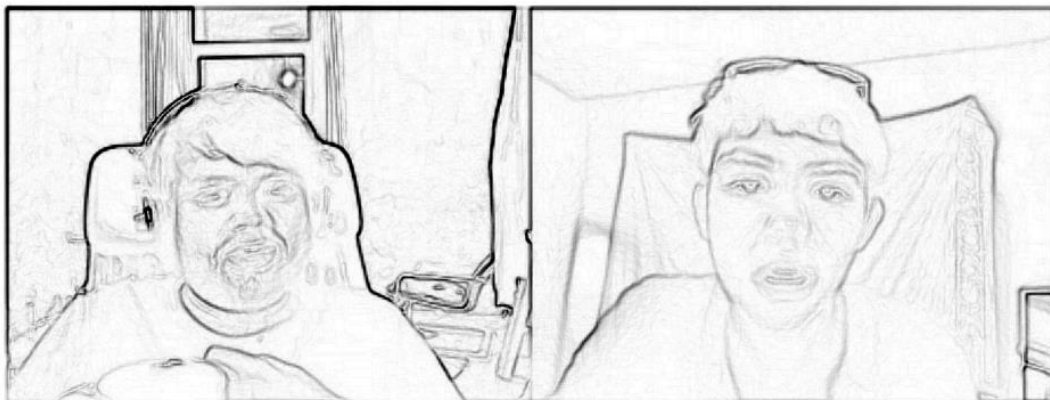


Figure #4

17 HER: @achso ja@ fünf kilometer oder so:?
 @I see @ five kilometers or so:?
 @gaze up@
 18 SOP: ja in cross country
 yes

While attempting to repair Herbert's inference that she has run a marathon, Sophia initiates a word search (line 15). The word search is initiated by stretching the vowel *a:hm* and by an upward gaze shift (Figure 1), which is accompanied by a short gap and followed by more speech perturbations and a 0.2-second gap (line 13). Following several indicators of problem with the progression of her ongoing turn, she utters *ein moment* "one moment," immediately shifts her gaze downward and shows her index finger (Figure 2). After another sound stretching and some explaining gestures, Sophia starts to type (line 15, Figure 3). Similar to the previous examples (see Excerpt 2 and 3), Sophia uses her gesture to keep her turn but also to publicly indicate that her word search might take some extended time because she needs to look up the sought-for word. She keeps holding her index finger even after the verbal alert, which also indicates that she is taking the time to "do thinking" before typing.

While Sophia is typing, Herbert continues to look at the screen and thereby displays his orientation to Sophia's screen-based activity and his understanding of her commitment to solving the word search herself. In addition, by waiting for Sophia to find a solution and by not offering another candidate solution, Herbert confers some epistemic authority to her, namely that he does not know what track and cross-country running is and that only Sophia can explain how many miles women usually race. In line 16, Sophia reads from the screen in English "three" (Figure 4) and then does a self-repair *drei meilen* "three miles," thereby completing the word search. In line 17, by uttering *achso* "I see," Herbert displays receipt of the new information and change of state and thereby displays that Sophia's repair resolution is successful and that they now share the same epistemic status (Heritage, 2012; Stivers et al., 2011). In addition, by converting the three miles into five kilometers, he offers a candidate understanding and seeks confirmation by try-marking it. In line 18, Sophia confirms Herbert's utterance.

This excerpt exemplifies another case of word search in which the L2 speaker uses the utterance *ein moment* "one moment" and an extended raised finger gesture. Compared to Excerpt 3, in this instance, the L1 speaker does orient to the gesture as a signal to hold on to his turn and as a request to not provide a candidate solution while the L2 speaker executes the screen-based activity. However, we argue that in word searches that involve a more complex trouble source, all three L2 speakers use "one moment" and an extended raised finger gesture to signal to their L1 coparticipant that the sought-for item is not readily accessible, that their search will require more time, and that the preceding business will be temporarily put on hold in favor of screen-based looking up of information. By asking for a moment, the speaker delays the progressivity of their turn to gain some time to "do thinking and searching" and index their continued commitment to the turn's production. Moreover, we argue that the practice of an extended raised finger gesture in combination with talk signals to the coparticipant that the word search in progress is not going to be a quick fix but rather an activity whose resolution requires temporary suspension of talk. The gesture in combination with talk (one moment) secures space for the extended search and holds the speaker's turn and exhibits their rights over the activity. Let us now look at another case in which an L2 speaker does utter *ein moment* ("one moment"), not accompanied by a gesture.

5.3 Word search accompanied by talk but no gesture

In all cases presented above, the L2 speakers have been able to complete their word searches by executing a screen-based activity to reestablish progressivity and understanding in talk-in-interaction. However, in one case (out of 5, see Table 1), the L2 speaker abandons his word search when prompted to answer a question. This case comes from a conversation between Herbert (HER), L1 speaker of German, and Samuel (SAM), L2 speaker, and precedes the

conversation in Excerpt 2. When Samuel attempts to say that two years are “left/remaining” until his college graduation, he displays problems in finding the German word *übrig* “remaining.” Our analysis focuses on the word search in line 4.

>>Video<<

Excerpt 5. ‘remaining’

```

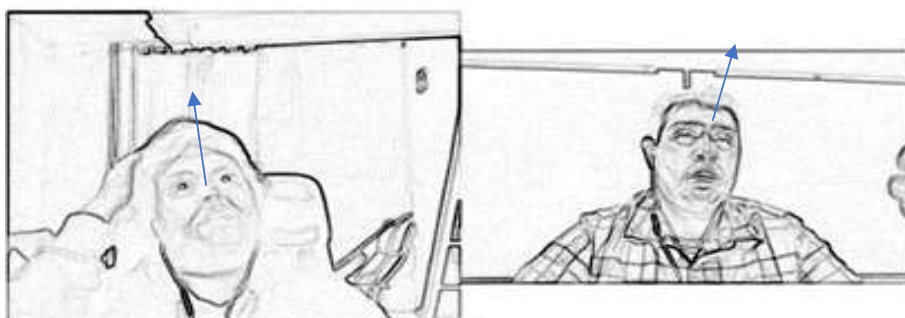
1 HER:      und ich weiß nicht wie la:nge studierst du schon?
             and I don't know how lo:ng have you been studying?
2           (0.5)
3 SAM:      ^a:hm ich ha:be studiert (.) für ^zwei jahre
             ^a:hm I ha:ve been studying (.) ^for two years
             ^gaze up ----->^
4           und ich *habe zwei jahre left*
             and I      *have two more years*
                       *left finger in the air*

5 HER:      @ja=
             @yes=
             @gaze up----->
6 SAM:      #1=^remaining^

her:      ----->

sam:      ^gaze up->^

```



Herbert (HER)

Samuel (SAM)

Figure #1

```

7 HER:      [und
             [and
her:      ----->
8 SAM:      #2^%[°how do you say°

```

her: ----->

sam: ^%gaze directed to the screen and typing ----->

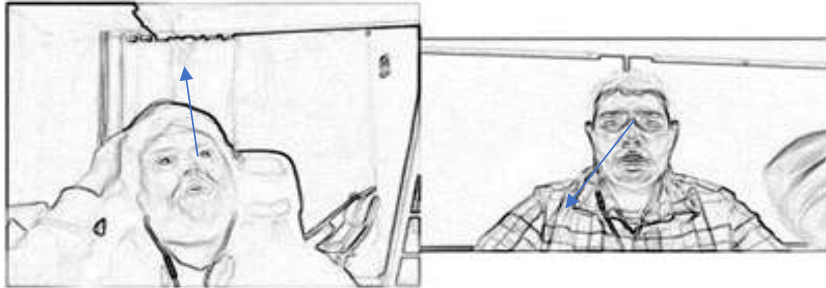


Figure #2

9 (0.2)

10 HER: #3@fängt man direkt an im flugzeug zu sitzen?%
@do you immediately start sitting alone on the plane?%

her --->@

sam ----->%



Figure #3

11 (1.0)

12 SAM: &eins noch?
&come again?
&moves twd screen----->
&r. hand twd right ear&

((Herbert is reformulating the question, see Excerpt 2.))

In response to Herbert's information-seeking question (line 1), after a 0.5-second delay (line 2), Samuel engages in a word search: he self-repairs by replacing "left" (line 4) with "remaining" (line 6). His eye gaze shifts upward (Figure 1) when conducting the self-repair of replacing the words. Note that both words are in English, signaling Samuel's lack of access to the German word *übrig*. Samuel continues by producing a self-addressed question (Hayashi, 2003) in low volume "how do you say" (line 8), thereby initiating a word search and explicitly and publicly displaying the kind of linguistic problem he is having, namely not knowing a lexical item in German. Samuel directs his gaze to the computer screen and the keyboard (Figure 2) to begin a screen-based activity. The gaze shift to the screen, the typing, and the verbal code-switch "how do you say" mark the beginning of Samuel's disengagement from the progressivity of talk-in-interaction and the onset of the search for the German word. In line 7, Herbert initiates a related follow-up question with the word *und* "and" (line 7) (cf. Heritage & Sorjonen, 1994; for and-prefaced questions), thereby orienting to the progressivity of interaction and not to Samuel's code switch. On the other hand, Samuel prioritizes the word search over progressivity.

After a 0.3-second pause (line 9), while Samuel is still engaged in typing, Herbert poses a follow-up question (line 10) — a yes-no polar interrogative question. At the beginning of this turn, he directs his gaze toward the screen. However, Samuel is still engaged in typing and searching for the German word (Figure 3). In line 10, Samuel stops typing, abandons his word search and looks at the screen. A one-second delay in line 11, followed by Samuel's initiation of repair in line 12 (*eins noch* "come again"), indicates that he has trouble with Herbert's question: Samuel either did not hear the question or did not understand it (Schegloff et al., 1977). In terms of embodiment, Samuel displays his trouble with Herbert's question as a hearing problem, note his head and torso moving toward the screen while making a cupping gesture (Mortensen, 2016) in line 12.

This example illustrates a case of an incomplete word search: The L2 speaker initiates a word search (Goodwin & Goodwin, 1986) without clearly underscoring the transition to the screen-based activity, abandons the screen-based search and shifts his gaze back to the screen. Samuel's utterance "how do you say" is not sufficient for Herbert to provide a candidate solution. Samuel's failure to find a word may be due to Herbert's question, which was produced midway through Samuel's screen-based search. At the same time, he misses the ongoing talk in Herbert's speech (Kozar, 2016) because his orientation was toward the screen and the search he was conducting. Our point is to show that the L1 speaker's orientation to the emerging activity is crucial for the L2 speaker's completion of their word searches. In other words, by raising the extended index finger, L2 speakers secure the opportunity to look up the lexical item due next and complete the word search themselves. By withholding a candidate solution, the L1 recipient demonstrates their understanding of the gesture talk combination as a request to withhold help and allow the L2 speaker to complete their screen-based search. In other words, we can say that L1 speakers ascribe the action of "hold" or "refrain

from help” to their L2 speaker’s talk plus upward extended raised index finger turn.

6. Discussion and conclusion

This study demonstrates the role of embodiment in word searches in VMIs. We demonstrated how the particular gesture of upward extended raised finger in combination with the utterance *ein moment* (“one moment”) is used by L2 speakers to suspend the activity in progress and secure the floor in the interest of creating a space to conduct a screen-based search, an emerging cognitive and manual activity. Unlike other gestures in word searches that elicit assistance from the coparticipants (Brouwer, 2003; Goodwin & Goodwin, 1986; Greer, 2013; Koshik & Seo, 2012; Kurhila, 2006; Reichert & Liebscher, 2012; Seo & Koshik, 2010), the L2 speaker’s embodied practices in our data restrain the coparticipant from providing a candidate solution, and instead secure the possibility of conducting a screen-based word search. Our analysis demonstrated how the particular embodied practice of showing the index finger contributes to making the action of an extended search using screen orientation (which involves prolonged suspension of the ongoing talk) recognizable for L1 recipients in the video-mediated interaction.

Our paper adds to the central argument found in previous studies on word searches that L2 speakers do not always rely on the expertise of L1 speakers and do not accept their offered candidate solution (Excerpts 2 and 3). Rather, the affordances of screen-based activities and the video-mediated environment enable them to complete their word searches and to enhance understanding and epistemic symmetry in L2 interaction, especially when the L2 speaker’s cultural expertise is required to complete a sequence (Excerpt 4). In doing so, the L2 speakers show their preference for self-repair (Schegloff, et al., 1977) and independent turn completion (see Theodórsdóttir, 2011). Moreover, our study demonstrates how word searches are highly situated practices (Goodwin, 1986; Rydell, 2019) and that the L2 speaker’s employment of semiotic resources such as gesture is connected to the affordances of the digital environment and the resources readily available for accomplishing an extended word search. Instead of remaining a solitary activity, L1 speakers orient to the explicit verbal alert “*ein moment*” and therefore the screen-based activity becomes a social activity (Pekarek Doehler & Balamán, 2021). Previous research showed that L2 speakers use verbal resources to announce their word searches (e.g., Greer, 2016; Reichert & Liebscher, 2012), and that word searches are social and collaborative activities (Goodwin & Goodwin, 1986; Matsumoto & Canagarajah, 2020). This study adds to the body of such research by examining embodied practices and the role of gesture in action formation and argues that shared orientation is necessary for collaboration in word searches. Even though the L2 speakers may not accept a candidate solution (Excerpt 2 and 3), by gazing at and letting the L2

speakers look up lexical items due, the L1 speakers acknowledge the L2 speaker's ongoing turn. The L2 speakers, on the other hand, actively claim and secure their right to keep the floor by employing the upward extended finger gesture and do not accept the L1 speakers' offered candidate solution if engaged in typing.

Our analysis contributes to our understanding of L2 interactions in the specific environment of video-mediated interaction and the verbal and embodied practices speakers utilize to maintain progressivity and establish understanding. It also suggests that future studies should further examine which additional embodied practices L2 speakers employ to manage progressivity and establish understanding in the particular interactional setting of VMIs.

References

- Antaki, C. (2012). Affiliative and disaffiliative candidate understandings. *Discourse Studies, 14*(5), 531–547.
- Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction, 49*(4), 290–309.
- Barrow, J. (2010). Electronic dictionary look-up practices of novice English learners. In T. Greer (Ed.), *Observing talk: Conversation analytic studies of second language interaction* (pp. 55–72). Tokyo: Pragmatics Special Interest Group of JALT.
- Brouwer, C. E. (2003). Word searches in NNS-NS interaction: Opportunities for language learning? *Modern Language Journal, 87*(4), 534–545.
- Brouwer, C. E. (2004). Doing pronunciation: A specific type of repair sequence. In E. Gardner & J. Wagner (Eds.), *Second language conversations* (pp. 93–113). Continuum.
- Egbert, M., Liebecker, L., & Rezzara, S. (2004). Inside first and second language speakers' trouble in understanding. In R. Gardner & J. Wagner (Eds.), *Second language conversations* (pp. 178–200). Continuum.
- Gan, Y., Greiffenhagen, C., & Licoppe, C. (2020). Orchestrated openings in video calls: Getting young left-behind children to greet their migrant parents. *Journal of Pragmatics, 170*, 364–380.
- Golato, A. (2012). German oh: Marking an emotional change of state. *Research on Language and Social Interaction, 45*(3), 245–268.

- Goodwin, C. (1981). Exophoric reference as an interactive resource. In J. N. Deely & M. D. Lehnart (Eds.), *Semiotics 1981* (pp.119–137). Plenum Press.
- Goodwin, C. (1986). Gesture as a resource for the organization of mutual orientation. *Semiotica*, 62, 29–49.
- Goodwin, C. (1989). Turn construction and conversational organization. In B. Dervin, I. Grossberg, B. O’Keefe, & E. Wartella (Eds.), *Rethinking communication: Paradigm Exemplars* (pp. 88–102). Sage Publications.
- Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62(1-2), 51–75.
- Greer, T. (2013). Word search sequences in bilingual interaction: Codeswitching and embodied orientation toward shifting participant constellations. *Journal of Pragmatic*, 57, 100–117.
- Greer, T. (2016). Multiple involvements in interactional repair: Using smartphones in peer culture to augment lingua franca English. In M. Theobald (Ed.), *Friendship and peer culture in multilingual settings* (pp. 197–229). Emerald.
- Hayashi, M. (2003). Language and the body as resources for collaborative action: A study of word searches in Japanese conversation. *Research on Language and Social Interaction*, 36(2), 109–141.
- Heath, C. C., & Luff, P. (1993). Disembodied conduct: Interactional asymmetries in video-mediated communication. In G. Button (Ed.), *Technology in working order: Studies of work, interaction, and technology* (pp. 35–54). Routledge.
- Heritage, J. (1988). Explanations as accounts: A conversation analytic perspective. In C. Antaki (Ed.), *Analyzing everyday explanation: A casebook of methods* (pp. 127–144). Sage Publications.
- Heritage, J. (2012). Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction*, 45(1), 1–29.
- Heritage, J., & Sorjonen, M. L. (1994). Constituting and maintaining activities across sequences: And-prefacing as a feature of question design. *Language in Society*, 23, 1–29.
- Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Polity Press.

- Jefferson, G. (1972). Side sequences. In D. N. Sudnow (Ed.), *Studies in social interaction* (pp. 294–233). Free Press.
- Kamunen, A., & Haddington, P. (2020). From monitoring to co-monitoring: Projecting and prompting activity transitions at the workplace. *Gespächsforschung-Online-Zeitschrift zur verbalen Interaktion*, 21, 82–122.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press
- Kitzinger, C. (2012). Repair. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 229–256). Wiley.
- Koshik, I., & Seo, M. S. (2012). Word (and other) search sequences initiated by language learners. *Text & Talk*, 32(2), 167–189.
- Kozar, O. (2016). Teachers' reaction to silence and teachers' wait time in video and audioconferencing English lessons: Do webcams make a difference? *System*, 62, 53–62.
- Kurhila, S. (2006). *Second language interaction*. John Benjamins.
- Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, 54(1), 79–117.
- Lerner, G. H. (1996). Finding face in the preference structures of talk-in-interaction. *Social Psychology Quarterly*, 59(4), 303–321.
- Licoppe, C. (2017). Showing objects in Skype video-mediated conversations: From showing gestures to showing sequences. *Journal of Pragmatics*, 110, 63–82.
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured ecologies: Creating environments for collaboration. *Human-Computer Interaction*, 18, 51–84.
- Matsumoto, Y., & Canagarajah, S. (2020). The use of gesture, gesture hold, and gaze in trouble-in-talk among multilingual interlocutors in an English as a lingua franca context. *Journal of Pragmatics*, 169, 245–267.
- Mlynář, J., González-Martínez, E., & Lalanne, D. (2018). Situated organization of video-mediated interaction: A review of ethnomethodological and conversation analytic studies. *Interacting with Computers*, 30(2), 73–84.
- Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, 9, 194–225.

- Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20(2), 2–32.
https://doi.org/10.1111/josl.1_12177
- 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. (2016). The body as a resource for other-initiation of repair: Cupping the hand behind the ear. *Research on Language and Social Interaction*, 49(1), 34–57.
- Oloff, F. (2021). Some systematic aspects of self-initiated mobile device use in face-to-face encounters. *Journal für Medienlinguistik*, 2(2), 195–235.
- Pekarek Doehler, S., & Balamán, U. (2021). The routinization of grammar as a social action format: A longitudinal study of video-mediated interactions. *Research on Language and Social Interaction*, 54(2), 183–202.
- Pekarek Doehler, S. & Berger, E. (2019). On the reflexive relation between developing L2 interactional competence and evolving social relationships: A longitudinal study of word-searches in the ‘wild’. In J. Hellermann, S. W. Eskildsen, S. Pekarek Doehler, & A. Piirainen-Marsh (Eds.), *Conversation analytic research on L2 interaction in the wild: The complex ecology of learning-in-action* (pp. 51–75). Springer.
- Piirainen-Marsh, A., Lilja, N., & Eskildsen, S. W. (2022/this issue). Bodily Practices in Action Formation and Ascription in Multilingual Interaction: Introduction to the Special Issue. *Social Interaction. Video-Based Studies of Human Sociality*, 5(1). <https://doi.org/10.7146/si.v5i2.130866>
- Reichert, T., & Liebscher, G. (2012). Positioning the expert: Word searches, expertise, and learning opportunities in peer interaction. *The Modern Language Journal*, 96(4), 599–609.
- Rosenbaun, L., & Licoppe, C. (2017). Showing ‘digital’ objects in web-based video chats as a collaborative achievement. *Pragmatics*, 27(3), 419–446.
- Rydell, M. (2019). Negotiating co-participation: Embodied word searching sequences in paired L2 speaking tests. *Journal of Pragmatics*, 149, 60–77.
- Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J. R. E. Lee (Eds.), *Talk and social organization* (pp. 54–69). Multilingual Matters.
- Sacks, H. (1992). *Lectures on conversation*. Blackwell.

- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. In T. Givon (Ed.), *Discourse and syntax* (pp. 261–286). Academic Press.
- Schegloff, E. A. (2007). *Sequence organization in interaction*. University Press.
- Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53(1–2), 361–382.
- Seo, M. S., & Koshik, I. (2010). A conversation analytic study of gestures that engender repair in ESL conversational tutoring. *Journal of Pragmatics*, 8, 2219–2239.
- Sert, O. (2017). Creating opportunities for L2 learning in a production activity. *System*, 70, 14–25.
- Skogmyr Marian, K., & Pekarek Doehler, S. (2022/this issue). Multimodal Word-search Trajectories in L2 Interaction: The Use of Gesture and how it Changes over Time. *Social Interaction. Video-Based Studies of Human Sociality*, 5(2). <https://doi.org/10.7146/si.v5i2.130867>
- Smotrova, T., & Lantolf, J. P. (2013). The function of gesture in lexically focused L2 instructional conversations. *Modern Language Journal*, 97(2), 397–416.
- Stivers, T., Mondada, L., & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 3-24). Cambridge University Press.
- Streeck, J. (1995). On projection. In E. Goody (Ed.), *Social intelligence and interaction* (pp. 87–110). Cambridge University Press.
- Streeck, J. (2009). *Gesturecraft. The manu-facture of meaning*. John Benjamins.
- Svennevig, J. (2018). What's called in Norwegian?: Acquiring L2 vocabulary items in the workplace. *Journal of Pragmatics*, 126, 68–77.
- Taleghani-Nikazm, C. (2008). Gestures in foreign language classroom: An empirical analysis of their organization and function. *Selected Proceedings of the 2007 Second Language Research Forum*, 229–238.
- Theodórsdóttir, G. (2011). Language learning activities in everyday situations: Insisting on TCU completion in second language talk. In G. Pallotti & J.

Wagner (Eds.), *L2 learning as a social practice: Conversation-analytic perspectives* (pp. 185–208). National Foreign Language Resource Center.

Tuncer, S., Lindwall, O., & Brown, B. (2020). Making time: Pausing to coordinate video interactions and practical tasks. *Symbolic Interaction*, 44(3), 603-631.