



Social Interaction. Video-Based Studies of Human Sociality.  
2020 Vol. 3, Issue 2  
ISBN: 2446-3620  
DOI: 10.7146/si.v3i2.114992

# **Social Interaction**

## **Video-Based Studies of Human Sociality**

---

**Multimodal displays of understanding in vocabulary-oriented sequences**

---

**Jenny Gudmundsen & Jan Svennevig**

*University of Oslo, MultiLing - Center for Research on Multilingualism in Society  
across the Lifespan, P.O.Box 1102, Oslo, Norway*

### *Abstract*

*How do participants use multimodal means to display joint understanding after dealing with vocabulary problems in second language interactions? In this article, we investigate the resolution phase of vocabulary-oriented sequences, more specifically word searches, word explanation sequences and vocabulary checks. We argue that understanding is displayed not only by change-of-state tokens, but typically by a more encompassing multimodal gestalt involving facial expression, gestures, and torso movements. A central feature of this gestalt is what we propose to call the change-of-state face, which consists in raised eyebrows and widened eyes and is produced while the head and/or the torso is raised upwards. Typically, this display of a change-of-state is followed by further demonstrations of understanding, such as verbal and gestural repeats. Finally, the solution of the problem is often 'celebrated' by means of reciprocal smiling. Not all vocabulary-oriented sequences are equally successful. Facial expression and gesture may also be used to resist and question a proposal of a candidate solution to a word search or a request for an explanation. In this way, multimodal means are crucial to establishing the success of vocabulary-oriented sequences and returning to the main business of the talk.*

Keywords: conversation analysis, multimodality, embodiment, understanding, change-of-state, iconic gesture, smiling, language café

## **1. Introduction**

Since Harvey Sacks described the fundamental distinction between *claiming* and *demonstrating* understanding (Sacks, 1992, p.141), researchers in the field of multimodal conversation analysis have pursued this line of thinking by specifying the sequential, situated, and embodied dimensions of displaying understanding. Many studies have focused on instructional settings such as archaeological apprenticeship (Goodwin, 1994), clinical dental training (Hindmarsh et al., 2011) and car instructions (Mondada, 2011). These studies show how different multimodal resources may be specialized for displaying various aspects of understanding: “While the talk may simply take the form of a claim of understanding, the body provides an exhibit to support or undermine the claim” (Hindmarsh et al., 2011, p. 500).

In this article, we investigate multimodal displays of understanding in vocabulary-oriented sequences in second language (L2) interactions, that is, word search sequences (Goodwin and Goodwin, 1986), vocabulary checks (Hosoda, 2006) and word explanation sequences (Mazeland and Zaman-Zadeh, 2004). Previous studies on embodied features in such sequences have mostly focused on the initiation phase, for instance, how speakers initiate word searches by averting their gaze from the interlocutor and producing a “thinking face” (Goodwin and Goodwin, 1986), or how they may provide hints to the interlocutor by producing iconic gestures (Hayashi 2003, Greer 2013). This study aims to describe embodied displays of understanding in the resolution phase of these sequences, with a special focus on a facial expression we propose to call the *change-of-state face*. Furthermore, we show how gestural repeats are used to demonstrate understanding. We argue that multimodal means are essential to establishing the success of vocabulary-oriented sequences.

## **2. Displays of understanding**

Verbal displays of understanding may take many forms. Understanding may be demonstrated by speakers formulating in their own words the gist or import of the interlocutor’s previous turn (Sacks, 1992, p. 141). The change-of-state token (‘oh’) may be used to propose that its producer has undergone some kind of change in his or her locally current state of knowledge, information, orientation or awareness

(Heritage, 1984). Among its many functions in different sequential environments, Heritage describes a practice where the change-of-state token prefaces a turn within which understanding is displayed:

*“The ‘oh’ functions as a realization claim which, in occurring prior to the subsequent turn components conveying the substance of the understanding achieved, proposes confidence in the adequacy of the understanding subsequently displayed” (Heritage, 1984, p. 323).*

Subsequent studies have described a range of different forms and functions of change-of-state tokens in various languages (cf. Heinemann and Koivisto, 2016). A function of special relevance here is to mark realization and restoration of understanding (‘now-understanding’) after a temporary lack of epistemic access or intelligibility (Koivisto, 2015, Heinemann 2016). For instance, the Finnish change-of-state token *aa* is observed to occur in situations where the speaker initiates repair involving an understanding problem. After the repair proper, it occurs as a sign of problem resolution and thereby invites sequence closure and resumption of the main activity (Koivisto, 2015). Furthermore, tokens marking *realization* may in many languages conventionally occur in collocation with a response word, such as *nå ja* (‘oh yes’) in Danish (Emmertsen and Heinemann, 2010) and *oh ja* in Dutch (Seuren, Huiskes and Koole, 2016). To date, no studies have been conducted on Norwegian change-of-state tokens, but the tokens found in our extracts, *åh* (‘oh’) and *åja* (‘oh yes’), seem to conform to the function described above, namely of claiming ‘now-understanding’ and realization.

Embodied displays of understanding have also been described by some researchers. Hindmarsh et al. (2011) showed how students in dental training accompanied verbal claims of understanding with gaze and posture, namely leaning into the patient’s mouth to ‘look with’ the demonstrator. The verbal claim is thus supported by visual evidence that the students were physically engaging in the collaborative viewing. In a study of driving lessons, Mondada (2011) observed that the change-of-state token “AH,” together with maximally raised eyebrows, projected a participant’s final demonstration

of understanding. Before the change-of-state token occurred, there were several embodied actions that displayed the customer's understanding, such as looking at objects in the car pointed to by the instructor. Fukuda and Burdelski (2019) have shown how visitors on a museum tour demonstrated understanding by interacting with visible, imagined and tactile objects, such as seeing and handling them, and thereby co-constructing the tours as institutionally organized activities.

### **3. Vocabulary-oriented sequences**

In recent years, language learning 'in the wild' — outside pedagogical settings (Hutchins, 1995) — has become a lively field of study within the conversation-analytic (CA) approach to second language acquisition (SLA) (Kurhila, 2006; Wagner, 2015; Theodórsdóttir and Eskildsen, 2011; Theodórsdóttir, 2018). The current study contributes to this field by investigating vocabulary-oriented sequences in so-called language cafés, where second language speaking immigrants meet first language speaking volunteers in order to provide the former with an opportunity to practice their language skills.

Vocabulary-oriented sequences have been thoroughly investigated within the CA-SLA tradition, since they constitute the “stuff of the L2 learning process” (Eskildsen, 2018, pp. 49). They are meta-linguistic and meta-communicative and constitute side sequences in relation to the overarching activity of the talk. In order for the participants to return to the main sequence and re-establish the progressivity of the talk, the participants need to establish joint identification or understanding of the vocabulary item being dealt with. In this study, we have included three types of sequences: word searches, vocabulary checks and word explanations. *Word search* sequences involve the speaker displaying problems in finding a word and appealing to the interlocutor for assistance (Goodwin and Goodwin, 1986). A similar practice, *vocabulary checks*, occurs when the L2 speaker produces a word and appeals to the interlocutor to confirm that it is correct, for instance, by producing it with rising intonation (Hosoda, 2006). The word searches and vocabulary checks analyzed in this data take the form of self-initiated other-repairs. The third type of repair sequence we focus on in this study is *word explanations*. They occur when the trouble source is the meaning of a word and

the interlocutor initiates repair by asking for clarification (Mazeland and Zaman-Zadeh, 2004). Common to all three sequence types is that there is a problem of identification or mutual understanding regarding a specific word. Thus, the resolution of the sequence involves the reestablishment of intersubjectivity, and this process is the subject of the current study.

A few studies have outlined the characteristics of the completion phase of vocabulary-oriented sequences. Theodórsdóttir (2018, p. 36) stated the importance of the L2 speaker's acceptance of a candidate solution to complete a word search, and Park (2007, p. 14) argued that both the speaker's and the recipient's satisfaction with the word is taken into account to finalize the search. Others have shown that speakers display acceptance of the repair in third position by repeating the candidate item (Svennevig, 2018a), repeating the gesture (Gullberg, 2011), producing acceptance tokens such as ("yes" and "right") and nodding their head (Hosoda, 2000). Park (2007) found that the acceptance may also be followed by a joint celebration, consisting of smiling faces and the use of higher volume. Moreover, participants tend to show more excitement after search sequences that become expanded and prolonged.

#### **4. Data**

The data material consists of video recordings of naturally occurring interactions at a language café in Oslo, Norway, collected by the first author in 2018 (Gudmundsen, 2019). Language cafés are an arena where immigrants can practice their second language skills with volunteers from the host country. This is an international phenomenon that has spread at an increasing pace in recent years (Gundersen, 2011). The attendees gather into small groups before starting the conversation. The conversations are informal and may be unstructured or semi-structured (Johnston and Audunson, 2017).

The total recording time of the video data is 5 hours. The conversations are among one or two L2 speakers and one L1 speaker. 85 vocabulary-oriented sequences were identified, consisting of word searches, vocabulary checks and word explanations. For this study, we have selected 4 extracts which represent the three different sequence

types and are particularly rich in multimodal displays of understanding, especially the characteristics of the change-of-state face.

Central to the analysis is the opportunity to observe minute details in the participants' facial expressions. We are therefore grateful for their permission to let us publish video extracts in non-anonymized form.

The method used is multimodal conversation analysis (Mondada, 2014). We provide one transcript for each excerpt. These excerpts include references to figures showing screen shots of the participants at the moment indicated.

## **5. Analysis**

This section includes analysis of the three different sequence types and concludes with a deviant case, in which the displays of understanding are not reciprocal. Due to space limitations, we provide only the resolution phase of the sequences and merely paraphrase how the sequence develops up to that point.

### **5.1 Vocabulary check**

This extract is from the conversation between the participants Omar (L2) and Tine (L1). The example shows the ending of a vocabulary check initiated by Omar. The participants have been talking about Omar's visit to the dentist. When Omar attempts to say that the dentist had to sew in Omar's mouth, he performs an iconic gesture (McNeill, 1992) – a sewing movement with his hand. He produces the word 'sew' incorrectly as "si:r" rather than "sy:r." He then asks Tine explicitly whether she recognizes the word. During the vocabulary check sequence, Tine has made several unsuccessful guesses of words with similar pronunciation. We will now see the resolution part of the sequence, where she finally identifies the correct Norwegian word.

## Excerpt 1: 'sew'

1 OMA: nei nei nei det va va-va va  
**no no no it wa wa-wa wa**

2 jeg #syr , ikke [ syr klær ]  
**I sew not sew clothes**

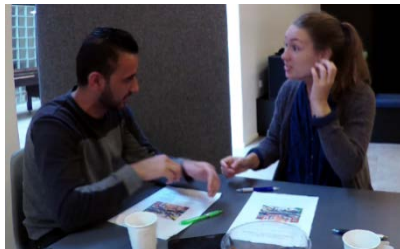
Fig: #1

3 →TIN: [ #↑åh #sy:r #ja:h ]  
**oh sew yes**

Fig: #2 #3 #4



Ex1-Fig1



Ex1-Fig2



Ex1-Fig3



Ex1-Fig4

4 TIN: [ .hh ]

5 OMA: [#syr]  
**sew**

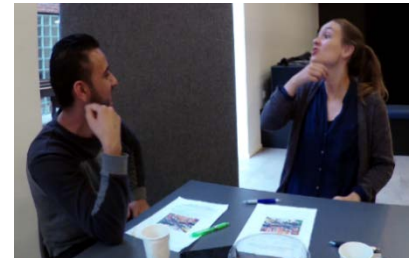
Fig: #5

6 TIN: #↑syr ?=  
**sew**

Fig: #6



Ex1-Fig5



Ex1-Fig6

7 OMA: =ja=  
**yes**

8 TIN: #=j↑a::h ,  
**yes**

Fig: #7



Ex1-Fig7

In line 1 and 2, the failure of the vocabulary check sequence is manifested as Omar explicitly rejects Tine's previous proposal and displays problems in initiating a new explanation. As a consequence, Tine scratches her head, having something similar to a "thinking face" (Goodwin and Goodwin, 1986) while she is gazing down at Omar's



gestures (fig. 1). After Omar has produced the word once more in line 2, this time with a more target-like pronunciation (“sy:r”), Tine utters an emphatic change-of-state token with high pitch and volume in line 3. Simultaneously, as she removes her hand from her head, she widens her eyes, raises her eyebrows, and raises her head (fig.2). This facial expression is what we propose to call a *change-of-state face*. In combination with the verbal change-of-state token, it marks the realization of what Omar had been trying to say and displays ‘now-understanding’ of his account.

After this first display of understanding, she repeats the target word “sew” (“syr”) with high pitch, followed by the confirmation “yeah” (“jah”), thereby demonstrating her identification of the word. In Norwegian, such repeats with final response particles claim understanding and acceptance of the previous contribution (Svennevig, 2004). While Tine says the word “jah” in line 3, she leans her head and upper torso backwards and starts smiling, thereby displaying a positive emotional stance (fig. 4). Immediately, Omar reciprocates the smile. At this moment, Omar starts to perform a sewing movement close to his mouth (fig. 4), and in the next turn in line 5, he repeats the target word while he continues to sew (fig. 5) – thereby confirming Tine’s identification of the word. Tine repeats the word once more and mirrors the iconic gesture in line 6 (fig. 6). By doing so, she does not just *claim* understanding, she also *demonstrates* it (Sacks, 1992). The ending of this vocabulary check shows an emphatic uptake with multiple verbal and gestural repeats, smiling and acceptance tokens indicating that the participants accept and celebrate the solution. The reason it is so pronounced may be that the vocabulary-oriented sequence had been expanded and prolonged (Park, 2007, p. 18).

In this example of a vocabulary check sequence, we see that displaying understanding takes the form of a *multimodal gestalt*, a combination of an emphatic verbal change-of-state token and a pronounced shift in facial expression and posture. Furthermore, we can observe that identification of the word is not just claimed, but demonstrated by repeating it and by mirroring the iconic gesture used to depict its meaning.

## 5.2 Word search sequence

The following extract shows a word search (Goodwin and Goodwin, 1986) involving the participants Omar (L2) and Tine (L1). This time, Omar has been describing a picture on an exercise sheet, depicting a man looking at a shop window. When he tries to say that the man might be using the window as a *mirror*, he displays problems in finding the right Norwegian word. A prolonged collaborative word search sequence follows, and after several unsuccessful attempts to find the word, Tine suddenly displays that she understands what he is trying to say. However, it still takes some time before she finds the word. We show the first part of the resolution here (2a), and the final part in the next excerpt (2b).

### Excerpt 2a: 'mirror'

1 OMA: (0.5) hva ska jeg #si::  
**what can I say**

Fig: #1

2 (1.3#)

Fig: #2

3 →TIN: †åh (#0.4) #e:h (#0.2)  
**oh uh**

Fig: #3 #4 #5



Ex2-Fig1



Ex2-Fig2



Ex2-Fig3



Ex2-Fig4



Ex2-Fig5

4 TIN: du #står eh , =  
**you stand uh**  
 #6

5 OMA: =ja  
**yes**



Ex2-Fig6

When Omar says ‘how can I say it’ he has a “thinking face” with averted gaze, and Tine frowns while leaning towards him in a listening position (fig. 1). Towards the end of the following pause, she suddenly changes her facial expression: she raises her eyebrows, widens her eyes, and raises her head (fig. 2). Immediately after, she utters the change-of-state token “åh” (line 3), drops the pen from her hand and starts leaning her torso backwards (fig. 3). At this point, her eyebrows are maximally raised (Mondada, 2011, p. 548). Omar orients to this multimodal gestalt by raising his head and looking at Tine, while he drops his hand from his chin (fig. 3). Thereby he treats it as displaying understanding and relieving him of the need to search further. In the following minimal pause, Tine starts to perform an iconic gesture: her hand is moving back and forth (fig. 5). This iconic gesture also recycles a similar gesture that Omar had used while trying to explain the word (not shown). Tine then says “you stand” (“du står”), initiating a contextual cue related to the concept of ‘looking at oneself in the mirror’. Both the gesture and the verbal utterance thus demonstrate her candidate understanding of the meaning of the word being searched for. As she initiates this explanation, Omar starts to smile (fig. 6). By doing so, he displays acceptance of her candidate understanding. Thus, mainly by non-verbal means, the participants display the first signs of mutual understanding of the *concept* Omar tries to convey. But the correct *word* has not yet been found. The final resolution of the word search is achieved only in the extract below, following directly after the previous one.

## Excerpt 2b: 'mirror'

6 TIN: når #du:  
**when you**

Fig: #7

7 (.)

8 OMA: j[ a ]  
**yes**

9 TIN: [skal] #[fi ]kse [deg]  
**are fixing yourself**

Fig: #8

10 [ja] [ ja ]  
**yes yes**

11 (0.3)

12 TIN: sp#tejl .  
**mirror**

Fig: #9

13 (.)

14 OMA: #£speil£ .  
**mirror**

Fig: #10

15 (.)

16 TIN: [ sp↑eil . ]  
**mirror**

17 OMA: [ £speil£.]ja .  
**mirror yes**



Ex2b-Fig7



Ex2b-Fig8



Ex2b-Fig9



Ex2b-Fig10

In lines 6 and 9, Tine elaborates her verbal cue as she continues to perform iconic gestures and looks straight ahead, as if she is looking at herself in a mirror. The first gesture seems to illustrate shaving (fig. 7) and the second, fixing her hair (fig. 8). Simultaneously, Omar smiles broadly, nods, and utters acknowledgement tokens — thus claiming recognition of the concept being explained and depicted. As noted, the demonstration only hints at the *meaning* of the word and may indicate that Tine is searching for the word herself. But in the end, she utters the candidate solution “mirror”

(“speil”) with high pitch (line 12) as she turns her gaze to Omar and points quickly towards him (fig. 9). This gesture highlights the candidate item, both by functioning as a ‘beat’ and by metaphorically depicting that she ‘spotted’ or ‘pinpointed’ the solution. Furthermore, she embodies the change-of-state from not remembering to remembering the word by leaning forward, raising her eyebrows, and widening her eyes. Omar repeats the candidate solution with a smiley voice in line 14. Simultaneously, he points quickly to the worksheet on the table and enhances the salience of the gesture by following it with his gaze (fig. 10) (cf. Streeck, 2009). This invokes the relevance of the word for the exercise. At this point, Tine smiles and thereby shows her satisfaction with the solution (Park, 2007). Afterwards, the participants repeat the word once more with smiling faces (in line 16 and 17) in a post-expansion rehearsal sequence (Svennevig 2018a), which also celebrates the solution and proposes to close the vocabulary-oriented sequence.

In this word search sequence, the multimodal display of a cognitive change-of-state occurs twice, first when Tine claims recognition of the concept and utters the change-of-state token (extract 2a), and then when she finds the word searched for and produces the candidate solution (extract 2b). In both cases, she raises her eyebrows and widens her eyes. The verbal component is different in the two cases, as they represent different types of cognitive shifts (understanding a concept and identifying a word form), but a common feature is the shift to high pitch and volume. The postures also vary, but they are both clear *shifts* in posture. The posture shift typical of displaying understanding is the first one, raising the head and leaning backwards. Leaning forward seems more to be associated with presenting (symbolically ‘handing over’) the solution to the interlocutor.

### **5.3 Word explanation sequence**

The next example shows the ending of a word explanation sequence (Mazeland and Zaman-Zadeh, 2004) between the participants Khalid (L2) and Sofia (L1). Sofia has been saying that one of her hobbies is rock climbing, when Khalid signals that he does not understand the Norwegian word for ‘climb’ - ‘klatre’. Sofia explains the word by

describing various aspects of the activity of climbing and simultaneously depicting them by means of gestures.

**Excerpt 3: ‘rock climbing’**

- 1 SOF: og så klatrer man opp  
**and then you climb up**
- 2 og har en sånn tau sik#ring.  
**and have this rope protection**



**Ex3-Fig1**

Fig: #1

- 3 (.)#



**Ex3-Fig2**

Fig: #2

- 4 → KHA: ah d-du liker cli-#climb,  
**oh y-you like cli-climb**



**Ex3-Fig3**

Fig: #3

- 5 SOF: #yeah,

Fig: #4

- 6 rock [climbing. ]

- 7 KHA: [(su) climb]

- 8 SOF: [#yeah.]



**Ex3-Fig4**

- 9 KHA: [#ja: ] du liker [det,]  
**yes you like that**

Fig: #5

- 10 SOF: [ ja. ]  
**yes**



**Ex3-Fig5**

In lines 1 and 2, Sofia continues to provide descriptions of climbing in short chunks (or ‘instalments’), leaving room in between for Khalid to provide a response (Svennevig, 2018b). In line 3, Khalid slightly raises his eyebrows and marginally widens his eyes. When his eyebrows are maximally raised, he utters the change-of-state ‘ah’ in line 4. Simultaneously, he raises his head slightly backwards (fig. 2). It should be noted that

Khalid's change-of-state face is not as pronounced as in the other examples. In the same turn, he also demonstrates understanding by initiating an understanding check, using the equivalent English word "climb." At the point when he utters this word, he points towards the wall with a vertical flat hand (fig. 3). The gesture is iconic in that it depicts the wall and the upwards direction of the climbing. Furthermore, it is a recycling of a gesture used by Sofia when explaining the word (not shown), thus referring anaphorically back to the explanation. This embodied action thus further demonstrates his understanding. In lines 5 and 6, Sofia confirms this suggestion by providing a more specific term ("yeah, rock climbing") while she mirrors Khalid's gesture (fig. 4) and follows his code-switching to English. In the next turn, Khalid resumes the main activity by switching back to Norwegian and producing a news receipt to Sofia's original report about her hobbies. At this point, both of them are smiling (fig. 5). Moreover, their hands return to home position (Kendon, 1975, 2004), thereby signaling closure of the side sequence.

In this example, we find the same multimodal gestalt, a verbal change-of-state token accompanied by a change-of-state face and a posture shift. Here it is used to mark the identification of the meaning of an unknown word in the L2, and this meaning is subsequently demonstrated both verbally and gesturally.

#### **5.4 A deviant case**

In all cases presented up to now, the solution to the vocabulary problem has been accepted by the interlocutors and an ensuing celebration has occurred. However, in one case, the interlocutors resist the solution proposed, and a central part of this resistance is performed by withholding the embodied displays of acceptance and positive emotional stance. The example shows part of an expanded word search sequence involving (from left to right) Abu (L2), Lena (L1) and Jamal (L2). In the repair initiation, Jamal asks Lena for the Norwegian word for the English word 'stamp'. However, this word is polysemous, and in Norwegian there are two possible translations: 'stempel' (an instrument for stamping a mark on a paper) and 'frimerke' ('a postage stamp'). At first, Lena hesitates but Abu seems to understand as he joins Jamal in subsequently trying to explain the meaning of the word to her. In doing so,

they use a range of multimodal resources available, such as iconic gestures and material objects (a paper sheet). Lena orients to this by gazing at their movements, slightly frowning. They seem to be depicting a postage stamp, and Abu even refers to a “small paper” (line 2), but, as we shall see, Lena suggests the word “stempel.”

### Excerpt 4a: ‘stamp’

1 JAM: hvis du [ skal sette #papir som ]  
**if you are to put paper like**

Fig: #1



Ex4-Fig1

2 ABU: [ liten papir sånn o:g, ]  
**small paper this and**



Ex4-Fig2

3 JAM: i bost ,  
**in post**

4 ABU: °de [ skriver noe å° ]  
**they write something and**

5 JAM: [ eh bost ]

6 →LEN: [ h ja: #↑åja ] stemp#el.  
**yes oh yes stamp**

Fig: #2 #3



Ex4-Fig3

7 (.)

8 JAM: stemp#el?  
**stamp**

Fig: #4



Ex4-Fig4

9 LEN: ja[ h , ]  
**yes**

10 ABU: [°st↑empel°]  
**stamp**

During Jamal’s turn, Lena raises her head and eyebrows, widens her eyes, and looks at Jamal while she utters the change-of-state token “åja” (‘oh yes’) in line 6 (fig. 2). By adding the response particle “ja” (“yes”), she marks the solution as something that she *only now* realizes and thereby takes the blame for not having understood before (perhaps related to the fact that the other interlocutor, Abu, seems to have understood



it a long time ago, cf. Emmertsen and Heinemann, 2010). Having signaled a change-of-state by means of this multimodal gestalt, she then proposes the candidate solution “stempel” with emphatic prosody, as she leans towards Jamal and starts to smile (fig. 3). This display of emotional stance indicates that she is rather confident in having found the correct word. Just like in the word search sequence above (extract 2b), the posture shift may iconically depict that she is offering the solution to the participant who initiated the sequence.

However, in this case the participants do not accept and celebrate the proposed solution as in the other vocabulary-oriented sequences. After a minimal pause, Jamal repeats the candidate with rising intonation in line 8 and makes a backwards head tilt while gazing at Lena (fig. 4), and thus inviting a response from her (Aoki, 2011, Rossano, 2012). Abu also repeats the word with try-marked intonation in line 10 (Sorjonen, 1996). These repetitions request a confirmation from the L1 speaker and thus display some problem with it. The participants also do not reciprocate Lena’s smile, and their facial expressions remain rather unaltered (fig. 4). And in the next excerpt (directly following the previous one), Jamal explicitly rejects the proposed solution:

#### Excerpt 4b: ‘stamp’



Ex4-Fig5

11 JAM: (---)(>ikke-s-#ikke nei sånn< .)  
**not I- not no like that**

Fig: #5

12 LEN: ja h , =  
**yes**

13 =det er- eh en #rektangel,  
**it is uh a rectangle**

Fig: #6



Ex4-Fig6

14 LEN: ja h ,  
**yes**

15 (.)

16 så vi ligge #på ark ,  
**that we put on a paper**

Fig: #7

17 (0.2)

18 >også vi sender til sånn #post .<=  
**and we send by this mail**

Fig: #8

19 LEN: #=<sup>↑</sup>mhm.

Fig: #9

20 (1.1)

21 ABU: >hva heter #det<  
**what's it called**

Fig: #10

22 (0.4)



Ex4-Fig7



Ex4-Fig8



Ex4-Fig9



Ex4-Fig10

In this extract, Jamal performs a third position repair (Schegloff, 1992). First, he rejects Lena's candidate solution by saying "ikke" and "ikke nei sånn" ("not" and "not like that") while moving his fist up and down as if stamping something (fig. 5). The gesture thus demonstrates his understanding of the candidate solution. Lena's response, a positive acknowledgement token ('jah' - yes) is not a type-conforming response to such a rejection and may indicate that she does not understand the utterance as a rejection (maybe because she is not gazing at Jamal's gesture). Then Jamal reinitiates his explanation of the meaning of the word in a neatly coordinated multimodal explanation sequence, divided up into instalments (lines 13-18) (Svennevig, 2018b). He produces each chunk with rising intonation and leaves a pause in between, providing space for Lena to respond. For every piece of information, he illustrates with iconic gestures: first drawing a rectangle on the paper (fig. 6), then pressing his palm down on it as if sticking a stamp onto an envelope (fig. 7), and finally stretching out his hand into the air as if sending something off (fig. 8). This explanation sequence clearly shows that the word he is searching for is 'frimerke' ('postage stamp').

In line 19 Lena provides a minimal response ("mhm") with high pitch while she nods and has a serious facial expression, which she keeps throughout the upcoming pause of 1.1 seconds in line 20. She thus does not seem to orient to Jamal's talk as providing

clues in order to elicit a new candidate solution. At the end of the pause, Jamal shifts his gaze to Abu and thereby invites further assistance from him. In line 21, Abu says “hva heter det” (“what’s it called”), thereby incrementing Jamal’s turn with a question making the activity of searching for a word explicit.

The search goes on for quite some time after this before it is solved. However, our point here is to show that L2 speakers do not always accept a solution proposed by an L1 speaker. Instead of accepting and registering the candidate proposal with a repeat with falling intonation (as in 2b), they initiate repair by repeating it with rising intonation and then upgrade to explicitly rejecting it. Their embodied behavior is central in this process of resisting a proposed solution to their word search. They do not reciprocate the smiling but instead retain an unaltered facial expression and posture.

## **6. Discussion and conclusion**

Vocabulary-oriented sequences have in common that there is a problem of understanding a concept being explained or identifying a word produced. Our study shows that finding a solution to such a problem is typically marked by uttering a change-of-state token with high pitch and volume, raising the eyebrows, and widening the eyes, and by raising the head upwards and the torso backwards. Typically, the facial expression changes from a thinking face (Goodwin and Goodwin, 1986) or frowning. The contrast in posture and facial expression is thus significant, as can be seen when comparing pictures of the speakers before and after the solution is found (Figure 1). The shift seems to be more pronounced the longer the repair sequence has lasted (thus explaining the less pronounced change in example 3).

**Example 1**



**Example 2**



**Example 3**



**Example 4**



*Figure 1: The change in posture and facial expression related to change-of-state.*

Previous studies have described the characteristics of a similar facial expression, with widened eyes and open mouth, combined with the sound 'oh', when describing surprise and astonishment (Darwin, 1872/2007; Ekman and Friesen, 1975; Heath et al., 2012). This indicates that this multimodal gestalt can express different meanings in different contexts, and that the communicative practice it represents depends on the local sequential context.

This study adds to the body of research on displays of changes-of-state in interaction by describing the embodied aspects of such displays and how they, in combination with verbal tokens, constitute a multimodal gestalt. Previous studies have focused on the verbal tokens used, and to some degree their prosodic characteristics (e.g. Koivisto, 2015 and Heinemann, 2016), but to our knowledge no studies have addressed systematically how a change-of-state is expressed by facial expressions and posture.

The multimodal displays of understanding occur in the third position of repair sequences and typically preface a turn displaying the understanding of the foregoing explanation. Thereby, the study shows the relevance of not only claiming, but also demonstrating understanding in the resolution of vocabulary-oriented sequences in L1/L2 conversations. The demonstrations often recycle iconic gestures from the problem phase, accompanied by verbal repeats and acceptance tokens. Previous studies on L2 learning have shown how participants use "return gestures" – when the recipient displays ongoing understanding of the embodied activities of the speaker – to achieve and maintain intersubjectivity and orient to language learning (Eskildsen and Wagner, 2013, 2015). And in word search sequences, participants repeat gestures and verbal utterances until both parties have reached mutual understanding (Gullberg, 2011, p. 141). This study corroborates these findings and underlines the importance of iconic gestures as a resource for demonstrating understanding in vocabulary-oriented sequences.

The excerpts also reveal some variation between the different types of vocabulary-oriented sequences. While the posture shift upwards and backwards seems to be

associated with displaying understanding of the *meaning* of a word being explained (as in 1, 2a, 3), the identification and presentation of a word *form* seems to be more associated with leaning forward towards the interlocutor, thus symbolically depicting the act of *offering* the solution to him or her (as in 2b and 4a).

Finally, our analysis shows how smiling functions as a display of acceptance and shared emotional stance toward the solution. The speaker offering the solution initiates the smiling, and the interlocutor reciprocates it simultaneously with or immediately after a claim or demonstration of understanding. That it constitutes a token of acceptance and appreciation of the solution is evident by the fact that it co-occurs with other types of positive stance displays, such as verbal evaluations and acknowledgement tokens. It is also demonstrated by the fact that, in the deviant case, the smile is not reciprocated by either of the participants.

Previous studies of language learning in the wild have pointed to the importance of both the speaker and the recipient displaying acceptance of a candidate solution for the search sequence to be completed and closed (Theodórsdóttir, 2018, Park, 2007). This study corroborates this finding by illustrating the importance of smiles, verbal repeats, and gestural demonstrations of understanding for establishing the success of a proposal. The deviant case also shows how the lack of such displays may be used to resist and reject a candidate solution.

## 7. References

Aoki, H. (2011). Some functions of speakers head nods. In: Streeck, J., Goodwin, C. and LeBaron, C. (Ed.), *Embodied Interaction: Language and Body in the Material World* (pp. 93-105). Cambridge: Cambridge University Press.

Darwin, C. (2007). *The expression of the emotions in man and animals*. Chicago: Filiquarian Publishing. (Original work published 1872).

Ekman, P. and Friesen, W. V. (1975) *Unmasking the Face: A guide to Recognizing Emotions from Facial Clues*. Englewood Cliffs, NJ: Prentice-Hall. Reprint edn, Palo Alto, CA: Consulting Psychologists Press, 1984.

- Emmertsen, S. and Heinemann, T. (2010). Realization as a Device for Remediating Problems of Affiliation in Interaction. *Research on Language and Social Interaction*, 43 (2), 109-132.
- Eskildsen, S. W. (2018). 'We're Learning a Lot of New Words' Encountering New L2 Vocabulary Outside of Class. *The Modern Language Journal*, 102 (supplement 2018), 46-63.
- Eskildsen, S. W. and Wagner, J. (2013). Recurring and shared gestures in the L2 classroom: Resources for teaching and learning. *European Journal of Applied Linguistics*, 1, 139-161.
- Eskildsen, S. W. and Wagner, J. (2015). Embodied L2 Construction Learning. *Language Learning*, 65, 268-297.
- Fukuda, C. and Burdelski, M. (2019). Multimodal Demonstrations of Understanding of visible, imagined and tactile objects in guided tours. *Research on Language and Social Interaction*, 52(1), 20-40.
- Goodwin, C. and Goodwin, M.H. (1986). Gestures and coparticipation in the activity of searching for a word. *Semiotica*, 62, 51-75.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606-633.
- Greer, T., (2013). Word search sequences in bilingual interaction: codeswitching and embodied orientation toward shifting participant constellations. *Journal of Pragmatics*, 57, 100-117.
- Gudmundsen, J. (2019). Felles forståelse på prøve. En multimodal samtaleanalyse av språkorienterte sekvenser i andrespråkssamtaler. [Mutual understanding at risk: A multimodal conversation analysis of vocabulary-oriented sequences in second language conversations] MA thesis. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.
- Gullberg, M. (2011). Multilingual Multimodality: Communicative Difficulties and Their Solutions in Second-Language Use. In: Streeck, J., Goodwin, C. and LeBaron, C. (Ed.), *Embodied Interaction: Language and Body in the Material World* (pp.137-151). Cambridge: Cambridge University Press.
- Gundersen, A. (2011). Norwegian libraries and special services for immigrant women Paper presented at World Library and Information Congress: 77. IFLA General Conference and Assembly, 12.-18. August 2011, San Juan, Puerto Rico.

- 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, 109-141.
- Heath, C., Vom Lehn, D., Cleverly, J. and Luff, P. (2012) Revealing Surprise. Peräkylä, A. and Sornojen, M. L. (Ed.) *Emotion in Interaction* (pp.212-234). Oxford: OUP.
- Heinemann, T. (2016). From 'looking' to 'seeing': Indexing delayed intelligibility of an object with the Danish change-of-state token  $n\uparrow$   $\text{\AA}\downarrow$ . *Journal of Pragmatics*, 104, 108-132.
- Heinemann, T., & Koivisto, A. L. (2016). Indicating a change-of-state in interaction: cross-linguistic explorations. *Journal of Pragmatics*, 104, 83–88.
- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In: Atkinson, J.M. and Heritage, J. (Ed.), *Structures of Social Action* (pp. 299-345). Cambridge: Cambridge University Press.
- Hindmarsh, J., Reynolds, P. and Dunne, S. (2011). Exhibiting understanding: The body in the apprenticeship. *Journal of Pragmatics*, 43, 489-503.
- Hosoda, Y., (2000). Other-Repair in Japanese Conversations Between Nonnative and Native Speakers. *Applied Linguistics*, 11, 39-63.
- Hosoda, Y. (2006). Repair and relevance of differential language expertise in second language conversations. *Applied Linguistics*, 27, 25-50.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge: MA: The MIT Press.
- Johnston, J. and Audunson, R. (2017). Supporting immigrant's political integration through discussion and debate in public libraries. *Journal of Librarianship and Information Science*, 00 (0), 1-15.
- Kendon, A. (1975). Gesticulation, speech and the gesture theory of language origins. *Sign Language Studies*, 9, 349-373.
- Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.



- Koivisto, A. (2015). Displaying now-understanding: The Finnish change-of-state token *aa*. *Discourse Processes*, 52, 111-148.
- Kurhila, S. (2006). *Second Language Interaction*. Amsterdam: John Benjamins.
- McNeill, D. (1992). *Hand and Mind*. Chicago: Chicago University Press.
- Mazeland H. and Zaman-Zadeh, M. (2004). The Logic of Clarification: Some Observations about Word-Clarification Repairs in Finnish-as-a-Lingua-Franca Interactions. In: Gardner, R. and Wagner, J. (Ed.), *Second language conversations* (pp. 132-156). New York, NY: Continuum.
- Mondada, L. (2007). Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse studies*, 9, 194-225.
- Mondada L. (2011). Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, 43, 542-552.
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137-156.
- Park, I. (2007). Co-construction of word search activities in native and non-native speaker interaction. *Working Papers in TESOL & Applied Linguistics*, 7(2), 1-23.
- Peräkylä, A. and Ruusuvuori, J. (2009). Facial and Verbal Expressions in Assessing Stories and Topics. *Research on language and social interaction*, 42, 377-394.
- Rossano, F. (2012). Gaze in Conversation. In: Sidnell, J. and Stivers, T. (Ed), *The Handbook of Conversation Analysis* (1th ed., pp. 229-256). Chichester: Wiley-Blackwell.
- Sacks, H. (1992). *Lectures on Conversation*. Vol 2. Jefferson, G., introduction by Schegloff, E.A (Ed.). Oxford: Basil Blackwell.
- Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 1295-1345.
- Seuren, L. M., Huiskes, M. & Koole, T. (2016). Remembering and understanding with oh-prefaced yes/no declaratives in Dutch. *Journal of Pragmatics*, 104, 180-192.
- Sorjonen, M. L. (1996). On repeats and responses in Finnish conversations. In:

- Ochs, E., Schegloff, E. A.; Thompson, S.A. (Ed.), *Interaction and Grammar* (pp. 277-327). Cambridge: Cambridge University Press.
- Streeck, J. (2009). Gesture as Communication 1: It's Coordination With Gaze and Speech. *Communication Monographs*, 60, 275-299.
- Svennevig, J. (2004). Other-repetition as display of hearing, understanding and emotional stance. *Discourse Studies*, 6, 489-516.
- Svennevig, J. (2018a). "What's it called in Norwegian?" Acquiring L2 vocabulary items in the workplace. *Journal of Pragmatics*, 126, 68-77.
- Svennevig, J. (2018b). Decomposing turns to enhance understanding by L2 speakers. *Research on language and social interaction*, 51, 398-416.
- Theodórsdóttir, G. (2018). A Closer Look at Correction and Explanation Practices in Everyday L2 Interaction. *The Modern Language Journal*, 102 (supplement 2018), 30-45.
- Theodórsdóttir, G. and Eskildsen, S.W. (2011). The use of English in everyday as a second language: Establishing intersubjectivity and doing learning. *Nordand*, 6, 59-85.
- Wagner, J. (2015). Designing for language learning in the wild. Creating social infrastructures for second language learning. In: Cadierno, T. and Eskildsen, S.W. (Ed.), *Usage based perspectives on second language learning* (pp.75-104). Berlin: de Gruyter Mouton.

---

<sup>i</sup> We are most grateful to the participants at the language café for their permission to let us record the video material and analyze and publish selected parts. We would also like to thank the three anonymous reviewers for extensive and useful comments to previous versions of this paper, and the editors for being tolerant with Jenny's transcription debut in a scientific journal. This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 223265.