Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion (ISSN 1617-1837)
Ausgabe 25 (2024), Seite 54-74 (www.gespraechsforschung-ozs.de)

Multimodal responses to questions: turn-initial *oh* coordinating with raised eyebrows in video-mediated interaction[1]

Marina Reis de Souza

English abstract
In video-mediated interaction, in addition to the vocal conduct, the face of participants usually comes in as a prominent communicative resource and can be employed for action formation. This study uses multimodal conversation analysis to investigate the role of facial gestures in coordination with the turn-initial particle *oh* in responsive actions. The analysis shows that raised eyebrows in turn-initial position are used in responses to polar questions that are treated by participants as inapposite. This embodied response marks issues of askability. Raised eyebrows co-occur with *oh* in the same sequential position to signal a problem with the question and its unexpectedness. The paper draws on a corpus of English as a lingua franca interactions over Skype. Hence, it contributes to the study of multimodality in technology-mediated environments.

Keywords: video-mediated interaction – Conversation Analysis – *oh* – facial gestures – multimodality.

German abstract
Neben dem verbalen und stimmlichen Verhalten tritt in videogestützten Interaktionen meist das Gesicht der Teilnehmer*innen als relevante leibliche Interaktionsressource hervor und es kann zur Bildung der Handlung eingesetzt werden. Der Schwerpunkt des vorliegenden Beitrags liegt auf der Rolle fazialer Gesten in Koordination mit der *turn-initialen* Partikel *oh* in responsiven Handlungen. Die Analyse zeigt, dass gehobene Augenbrauen zusammen mit *oh* in turn-initialer Position verwendet werden, um unangebrachte oder unpassende Entscheidungsfragen als in irgendeiner Weise problematisch (z. B. als inadäquat oder unvorhergesehen) zu kennzeichnen und damit deren *askability* zu bearbeiten. Die Studie stützt sich auf einen Korpus von Skype-Gesprächen, in denen Englisch als Lingua Franca verwendet wird. Sie trägt zur Untersuchung multimodaler Praktiken der Handlungskonstitution in computergestützter Kommunikation bei.

Keywords: videogestützte Interaktion – Konversationsanalyse – *oh* – Gesichtsausdruck und -gesten – Multimodalität.

[1] I would like to thank the editors and the two anonymous reviewers for their valuable comments and suggestions on this paper. I would also like to thank Prof. Friedrich Lenz, PD Dr. Maximiliane Frobenius, and Dr. Janet McLaughlin for their comments on the earlier version of this paper and the ViMELF team for giving me access to the corpus. An earlier version of this paper was presented as a panel contribution in the 18th International Pragmatics Conference (Université Libre de Bruxelles, July 2023).

1. Introduction

In social interaction, participants can mobilize different multimodal resources to accomplish actions. The deployment of multiple resources for these goals is possible because

> [f]ace-to-face interaction is, by definition, multimodal interaction in which participants encounter a steady stream of meaningful facial expressions, gestures, body postures, head movements, words, grammatical constructions, and prosodic contours (Stivers/Sidnell 2005:2).

The term *multimodal Gestalt* is commonly employed to describe the routinized organization of verbal and embodied conduct and can be defined as "a web of resources formatting an action" (Mondada 2014:139).

Participants are constantly monitoring their co-participants' turns and orienting towards them in a systematic and organized manner. An example of this systematicity appears in responses to questions since "[a]sking a question places significant constraints on what the recipient does next" (Stivers/Hayashi 2010:1). In the case of polar questions, for example, participants are usually limited to type-conforming *yes* or *no* responses (Raymond 2003; Stivers/Hayashi 2010), with positive answers being the preferred option (Schegloff 2007). However, there are other ways in which participants can respond to polar questions (see e.g., Enfield et al. 2019). Some of these formats include the type-conforming tokens *uh huh, yeah, nope* (Raymond 2003), transformative answers (Stivers/Hayashi 2010), and repetitions (Stivers 2005).

Moreover, "[e]ach question that a speaker poses does more than *ask a question*" (Stivers 2022:38). A polar question design can be employed to "initiate repair on another's talk, […] assess the weather, and request information" (ibid). Therefore, responses are not limited to affirming or confirming (see Gipper/Groß 2023) or *yes*, *no*, and similar formats.

In their responses, participants can challenge the *askability* of questions (Stivers 2011). In other words, participants orient to whether a question should have been asked or to a potential problem in the question. There are several ways in which participants can challenge askability (see Stivers 2011). One way in which this occurs is through answers prefaced by the turn-initial particle *oh*, that can mark inapposite questions (Heritage 1998). The term 'inapposite' describes a question that is treated as "problematic as to its relevance, presuppositions, or context" (ibid:296). That is, in the responsive turn, participants indicate some kind of problem regarding the adequacy of the inquiry. For example, participants can employ the turn-initial particle *oh* when questions are considered "self-evident by virtue of its physical context, or of persons' cultural or personal knowledge" (ibid:301).

Questions can be answered using vocal conduct only or embodied conduct (e.g., De Stefani 2021) and responses can also consist of multiple multimodal resources. Raised eyebrows, for example, can be employed to signal problematic issues in interaction (see e.g., Hömke 2018; Kaukomaa/Peräkylä/Ruusuvuori 2014).

The current study uses data from Skype, a video-mediated interaction (VMI) technology "that enables synchronous communication via a video link" (Due/ Licoppe 2020:2). In this setting, the face is one of the few communicative resources that participants have access to because the rest of the body is usually outside the camera's field of view.

This paper draws on the frameworks of multimodal conversation analysis to study two answer formats to respond to polar questions. First, the main finding is presented, which consists of cases in which participants treat questions as inapposite by responding with raised eyebrows and the turn-initial particle *oh*. The co-occurrence of *oh* and eyebrow movement forms a multimodal Gestalt (Mondada 2014). Part of the interactional force of this Gestalt is provided by *oh*, which is characterized as a "change-of-state token" (Heritage 1984) and signals, for example, unexpectedness (Heritage 1998). Second, cases which include eyebrow movements but the response is not *oh*-prefaced are also analyzed. These examples occur in repeated eliciting of informings, that is, when participants re-ask for information already previously discussed.

2. Multimodal practices of beginning responsive actions

The conversation analytic literature emphasizes the significance of the turn-initial position. Schegloff (1987:71) points to its importance "for the projection of the turn-shape or the turn-type of the turn that is being begun at that turn beginning". Turn-beginnings provide participants with a first indication to interpret what will come next in the interaction. This initial position is often occupied by turn-initial particles, such as *oh*, *so*, and *well* (Heritage 2018; Bolden 2006). Response tokens, such as *yes* and *no*, are also used in this initial position. These are defined as "brief, non-topical responses; responses that indicate that a piece of talk by one speaker has been heard and registered by the recipient of that talk" (Heinemann 2015:39). However, in responses to questions, response tokens differ from turn-initial particles. The latter can project action(s), such as preferred or dispreferred responses (see Heritage/Sorjonen 2018).

Particles positioned at the beginning of turns become relevant to the notion of *nextness*, a central aspect in conversation analysis (CA). Heritage (2013:333) explains that turn-initial particles

> can either project what we may term 'unmarked nextness' in which the current speaker continues a sequence of turns within some expected set of parameters (e.g., 'Are you going to the movies'/'Yes') or, alternatively, they can index some kind of departure from expectations for the subsequent turn (e.g., 'Are you going to the movies'/'Well I wasn't planning to. …').

Therefore, turn-initial particles can provide further details on the subsequent turns by anticipating expected or unexpected trajectories in interaction. Each turn-initial particle has particularities and specific functions, according to the sequential

position and sequence type in which they are employed. For example, the particle *well* often prefaces dispreferred responses (Heritage 2018; Schegloff/Lerner 2009).

The current work focuses on a specific particle, namely *oh* at turn-beginnings. The use of *oh* in English has been extensively investigated (see e.g., Bolden 2015; Heritage 1984, 1998, 2002, 2018; Norrick 2009; Schiffrin 1987). Generally, the particle *oh* is described as a *change-of-state token*, that is, "its producer has undergone some kind of change in his or her locally current state of knowledge, information, orientation or awareness" (Heritage 1984:299). Moreover, *oh* is described as a *backward-looking* particle (Heritage 2018; Schegloff/Lerner 2009). This indicates that the particle usually orients to an action from the previous turn.

In CA, particles' functions and meanings are analyzed according to their sequential placement (Heritage/Sorjonen 2018). *Oh* can be used in all sequential positions (see Heritage 1984, 1998, 2002, 2013; Bolden 2006). In the second position, *oh* is used to preface responses, assessments, and for information receipt (see e.g., Heritage 1998).

Focusing on responses with turn-initial *oh*, Heritage (2018:296) explains that the particle can signal a problem with the preceding question or "[f]oreshadow reluctance to advance the conversational topic invoked by the inquiry". In this position, the particle is backward-looking since it signals some type of problematic issue related to the previous question. The terms 'problematic' and 'inapposite' have been taken from Heritage (1998) for the purposes of this paper to describe *oh*-prefaced responsive actions that indicate a problem with the preceding inquiries.

However, the current paper does not investigate turn-initial particles in isolation, but rather how they co-occur with facial gestures. Turns and turn-beginnings are not only characterized by vocal conduct but also by embodied behavior. Mondada (2022a:292) explains that "[e]ven if language is a key resource for some types of activities, it is not a priori the most fundamental one". Consequently, the body, the face, and the vocal conduct of participants are interactionally relevant.

CA studies of face-to-face interaction have demonstrated the importance of analyzing how multimodal resources play a role in the formation of responsive actions. Regarding question-answer sequences, De Stefani (2021) shows that participants can respond to questions with silent nods. Kendrick and Holler (2017) observe that gaze aversion precedes dispreferred responses.

Regarding facial gestures, Hömke (2018:67) observes that eyebrow movements are the second most common facial movement in interaction. The study reveals that "the co-occurrence of eyebrow actions and verbal repair initiations suggests […] they may be co-expressive in signaling problems in hearing or understanding" (ibid:80). Previous studies on eyebrow movements present similar findings. Kaukomaa/Peräkylä/Ruusuvuori (2014) find that participants, in certain cases, produce turn-opening frowns in the transition between turns until they begin their vocal turn. This eyebrow movement signals that the following turn "will not perfectly fulfil the expectations aroused by the previous turn" (ibid:144). Dix and Groß (2023:19) describe, for example, how raised eyebrows are used as a "change-of-state marker" in responses. The embodied resource signals new or surprising information. These studies demonstrate that eyebrow movements provide important cues in social interaction and that their deployment is commonly related to issues arising from the previous turn.

Other than employing stand-alone verbal tokens or facial gestures, participants use coordinated multimodal resources to accomplish a responsive action. As previously discussed, this coordination of resources is known as a multimodal Gestalt (Mondada 2014). An example of a multimodal Gestalt in a responsive position is discussed in Pekarek Doehler et al. (2021) and Pekarek Doehler (2022). The authors demonstrate that the utterance 'I don't know/I dunno' coordinates with gaze aversion and is employed as a dispreferred answer to inquiries. Another example is presented by Helmer/Betz/Deppermann (2021). Their study of interactions in German reveals that the particle *okay*, when prefaced by change-of-state tokens such as *ah* and *achso,* co-occurs with nodding. These form a multimodal Gestalt to show that "speakers do not only accept information as sufficient, but also indicate that it has been new (or forgotten)" (ibid:380). De Stefani (2022:5) observes that multimodal Gestalts

> are, of course, not just recurrent, uniform reproductions of the "same" multimodal arrangement. They are sensitive to the local and sequential environment at hand, and they are highly adaptable – with regard to their temporal deployment, the coordination of the different modalities, and their manifestation in space. And they are recognizable, for interactants, as such.

In other words, multimodal Gestalts refer to the mobilization of embodied and linguistic resources that possess a local meaning and are interpreted regarding several contextual dimensions.

While these studies on face-to-face interaction have analyzed the vocal and embodied conduct of participants in responsive actions, facial gestures, in particular eyebrow movements, have only recently received increasing attention (see e.g., Dix/Groß 2023; Heller/Schönfelder/Robbins 2023; Stolle/Pfeiffer 2023). The multimodal resources which participants deploy in video-mediated settings are still under-researched. These settings are characterized by a habitual *talking heads* configuration, that is, only the heads of participants can be seen as these are positioned in front of the camera (Licoppe/Morel 2012). Moreover, Luff et al. (2003) describe video-mediated settings as *fracture ecologies* and explain that participants do not have full access to their co-participants environments. In VMI, the notion of *non-mutual realities* (Ruhleder/Jordan 2001) is also relevant. This notion describes how interaction is affected by delay in VMI (see also Ilomäki/Ruusuvuori/Laitinen 2021; Seuren et al. 2021).

3. Data and procedure

This study draws on the Corpus of Video-Mediated English as a Lingua Franca Conversations (ViMELF 2018, published under Creativ Commens License CC BY NC 4.0). The corpus is composed of audio and video recordings of dyadic conversations between university students of different European nationalities.

In total, the corpus includes 20 recordings of audio and video material. For the current study, approximately 6 hours and 35 minutes of recordings were analyzed. This time frame represents a sub-corpus of 11 videos in which one participant is German and the other is either Italian, Spanish, or Finnish. These recordings were selected because the remaining conversations are either audio-only or partial video recordings and are therefore not appropriate for the analysis of embodied behavior.

The participants' names were anonymized and an anonymization filter was applied to the original videos. The pronouns used throughout this paper follow the participants' gender self-identification present in the corpus' sociolinguistic background data.

The conversations are the participants' first encounters and the screen recordings were produced by the participants themselves. Participants received a topic prompt to start the conversation and the instruction to talk for at least thirty minutes. However, they did not have to only concentrate on the prompted topic (Brunner/ Diemer/Schmidt 2017). The transcription of the Skype calls is based on modified versions of Dressler and Kreuz's (2000) transcription conventions and the CASE Transcription Conventions (2017) for hesitation. The transcription of embodied actions follows Mondada's (2018, 2022b) conventions for transcribing multimodality (see Appendix for transcription conventions).

The analysis was conducted using multimodal conversation analysis (Mondada 2018). It includes four excerpts that are part of a larger collection of cases in which participants respond to different types of inquiries. Table 1 shows an overview of the corpus concerning polar questions, inapposite questions, and answers.

Polar questions			**Answers to inapposite questions**	
Question's main action	***Total***	***Inapposite questions***	***oh-prefaced + raised eyebrows***	***raised eyebrows***
Requests for new information	268	7	-	5
Requests for confirmation	65	12	1	7
Other initiation of repair	12	1	1	-
Directive-type question	15	2	2	-
Assessment	8	1	-	-
Total	**368**	**23**	**4**	**12**

Table 1: Overview of polar questions and answers in the sub-corpus of ViMELF (2018)

Firstly, instances of the multimodal Gestalt *oh* + raised eyebrows in responsive position were collected. It was observed that this coordination of turn-initial particle and embodied behavior was used to respond to polar questions that participants treated as inapposite (see Section 1). Four cases of this phenomenon were found. There were no cases in which the response was *oh*-prefaced but did not coordinate with raised eyebrows. Then, raised eyebrows in responsive position were collected. Twelve cases of raised eyebrows in responsive position were employed to respond to questions marked by participants as inapposite. These twelve cases were not prefaced by *oh*. The functions of these two types of responses were then compared and analyzed.

Secondly, the investigation was extended to understand what type of polar questions participants were responding to. All polar questions of the corpus that addressed a potential sensitive or problematic topic and could be deemed inapposite were collected. One-word newsmarks, such as *really* (Thompson/Fox/Couper-Kuhlen 2015) were excluded from the collection. In total, 23 inapposite questions were identified.

Thirdly, all polar questions of the corpus were collected and classified. The questions' main action was coded in accordance with Stivers' (2022:39) coding system (see also Stivers/Enfield 2010). Requests for new information aim at obtaining information, such as *have you ever been to an English-speaking country?* (this and the following examples are taken from ViMELF). Requests for confirmation aim at confirming previously discussed matters: *so it's also a long distance relationship?*. Other initiation of repair are questions in which the "[q]uestioner seeks confirmation about trouble in speaking, hearing, or understanding" (Stivers 2022:39; see also Schegloff 2000). An example of request for confirmation taken from the corpus is *they don't?*. Directive-type questions are those in which the participant asking a question wishes the receiver to perform an action, such as *can you u:h, I don't know, recommend something?*. In assessments, the participants asking the question wish to receive confirmation on an assessment: *and German as well isn't German like super difficult to learn as well?*. As seen in Table 1, most of the cases of questions marked as inapposite are either requests for new information or for confirmation.

To summarize, the table shows there are 23 questions which address potential inapposite matters. In 4 cases, participants respond to these questions with the turn-initial particle *oh* in coordination with raised eyebrows. In 12 cases, participants' responses coordinate with raised eyebrows. Although participants answered to all 23 questions, in most of them (16 cases or approximately 70%), raised eyebrows co-occurred with the responses. These 16 cases account for the responses with and without the particle *oh*.

4. Results

4.1 Prefacing responsive turns: turn-initial *oh* coordinating with raised eyebrows

Section 4 illustrates how participants employ the turn-initial particle *oh* with raised eyebrows at the beginning of responses to questions. It is argued that the responsive action signals a problem with the question concerning its appositeness. Excerpt (1) provides a first example. This passage describes the opening sequence of the conversation between SB51 and ST01.

Excerpt (1): Name [00:00:40]

```
01 SB51: (1.0) uhm,
02       so uhm,
03       ... ST01's name,
04       well that's: uh a kind of strange name to me,
05       .. cause u:hm,
06       yeah I'm not really familiar to it,
```

```
07        .. u:hm,
08        like- #can you tell me where the name comes from o:r,
   fig          #1a
```

Figure 1a

```
09 ST01: +oh+# it's a it's a region name,
         +RE+
   fig        #1b
```

Figure 1b

```
10        it's: similar to city,
11        in uh Spanish maybe that sounds more familiar to you?
12 SB51: ah yeah right yeah.
13 ST01: okay so,
14        ... it's a name from my- from my area.
15 SB51: ah okay alright.
16        cause I asked a friend of mine,
17        uhm she studies Spanish,
```

After an initial 'how-are-you' sequence (omitted from the transcript), SB51 produces a *so*-prefaced topic shift (line 2). She then comments on her interlocutor's name: *well that's: uh a kind of strange name to me, .. cause u:hm, yeah I'm not really familiar to it,* (lines 4-6). SB51 then asks a directive type question: *like- can you tell me where the name comes from o:r,* (line 8, Figure 1a).

The request for information posed by SB51 seems to be built around matters of entitlement and contingency (see Curl/Drew 2008). In their study of telephone calls to doctors, Curl and Drew (2008) have found that "[b]y making a request using model verbs and related forms, speakers treat their request as noncontingent […] and therefore their request as unproblematic" (ibid:147). Therefore, it might be the case that SB51 has not initially treated the request as unfitting. However, another

characteristic of SB51's request is that it presents a turn-final 'or'. Previous research on interactions in German reveals how '*oder*' (English 'or') in turn-final position is used to "weaken polar constraints" (Drake 2016:180).

Though in a different conversational context, the design of this question could indicate that SB51 recognizes that asking the meaning of a person's first name might be perceived as a sensitive topic. ST01's response is *oh*-prefaced: *oh it's a it's a region name* (line 9). The participant does not use the tokens 'yes' or 'no', but he provides the requested information, explaining the origin of his name. Co-occurring with the turn-initial particle *oh*, ST01 raises both eyebrows (transcribed as 'RE', Figure 1b). With the deployment of these multimodal resources, ST01 points to a problem which potentially relates to four aspects that concern issues of askability (see Stivers 2011). Firstly, the topic might be regarded as sensitive because SB51 starts their video call by providing an evaluative account of her interlocutor's first name, calling it a *strange name* (line 4). Secondly, the request for information is positioned at the very beginning of their conversation, just after the opening sequence, as they are still knowing each other. Thirdly, the question could also be perceived as unexpected. Fourthly, requests are among dispreferred first pair parts (Schegloff 2007). Taking the above-mentioned aspects into consideration, it is possible to conclude that the Gestalt *oh + raised eyebrows* in this responsive position is not only employed as a reaction to the request for information but also to the preceding evaluation of ST01's first name (lines 4-6).

Heritage (1998:313) states that *oh*-prefaced answers "project reluctance to talk about the topic raised by the inquiry" in different ways. These are characterized, for example, by "minimal, unelaborated or occasionally even dismissive [responses or when] producers unilaterally shift topic immediately after the *oh*-prefaced response, or shortly thereafter" (ibid:314). That is, the content of *oh*-prefaced responses usually reveals that the continuation of the ongoing conversation topic is undesirable. However, these characteristics are not found in Excerpt (1). ST01 provides further information about names somewhat similar to his (line 10) and asks: *in uh Spanish maybe that sounds more familiar to you?* (line 11). SB51 produces an affirmative answer: *ah yeah right yeah.* (line 12), displaying her understanding of ST01's name. As an attempt to conclude the topic, ST01 utters: *okay so, … it's a name from my- from my area.* (lines 13-14).

Excerpt (1) has shown how the multimodal Gestalt *oh + raised eyebrows* is used for the local management of a sensitive topic. However, Excerpt (2) demonstrates how the Gestalt is employed to manage another type of inapposite inquiry, that is, when there is a misunderstanding. In this example, HE19 and SB93 discuss how much they pay to study where they live.

Excerpt (2): University fees [00:22:00]

```
01 HE19: do you pay for university in Germany?
02 SB93: oh just a bit yeah.
03 HE19:        +[okay]+.
                +nods--+
04 SB93: I think [we pay],
05       <one hundred eighty euros>,
06       .h for six months.
07 HE19: o:h,
08       that's nothing.
```

```
09        yeah.=
10 SB93: =I know it’s nothing,
11        ye[ah].
12 HE19:    [it’s] free here for foreign students as well?
13 SB93: (1.7) uh in Finland?
14        or-,
15 HE19: .t +yeah+,
             +nods+
16        so university is free?#
   fig                          #2a
```

Figure 2a

```
17 SB93: and YOU PAY?
18 HE19: .. +>oh no# no+ so<,
             +RE, shakes head+
   fig                #2b
```

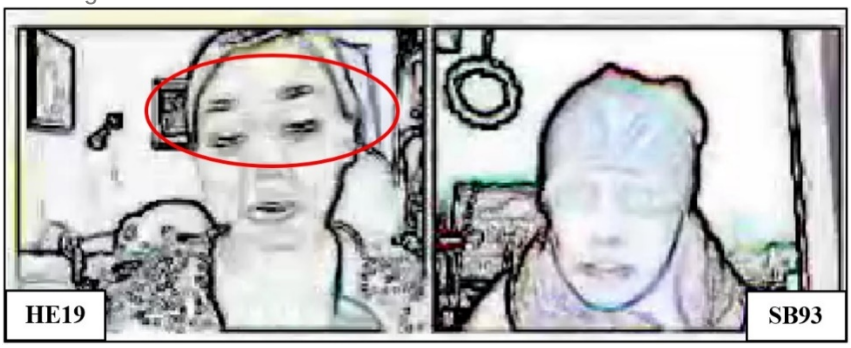

Figure 2b

```
19        e-,
20        for everyone,
21        yeah,
22 SB93: it’s COMPLETELY free?
23 HE19: >yeah<.
24        [yeah]
25 SB93: [oh okay].
```

The sequence is launched by HE19’s question *do you pay for university in Germany?* (line 1). After an affirmative answer followed by the acknowledgment token *okay* (line 3) (Couper-Kuhlen 2021), SB93 details the amount she pays (lines 4-6). Following this initial exchange, HE19 assesses the amount of money SB93 pays to study: *that’s nothing*. (line 8). SB93 produces a partial repeat (Stivers 2005), showing she agrees with her interlocutor. Then, HE19 self-selects and explains that foreign students do not have to pay for university in Finland (line 12). After a pause of 1.7 seconds, SB93 produces a follow-up question to check the information: *uh*

in Finland? or-, (lines 13-14). HE19 nods and produces an affirmative answer, stating that university is free (lines 15-16, Figure 2a).

SB93 launches a question that displays inferences from the previous talk: *and YOU PAY?* (line 17). This question is classified as an *other initiation of repair*. According to Stivers (2022:56-57), "[*a*]*nd* is associated with marking questions that follow on from a prior as part of a project and downplaying both the contingent nature of the question and the novelty of it". In other words, *and* functions to reestablish the connection between a previous turn and the current turn. HE19's response is formed by the turn-initial particle *oh* and repetitions of the negative token *no*. At the same time, HE19 raises both eyebrows and shakes her head (Figure 2b). SB93's question can be perceived as problematic because the information should have already been understood. Heritage (2013:331) also explains that turn-initial elements, such as *oh*, can "cluster around 'expectation canceling' functions", which, in this case, are used to resist to the assumption that HE19 pays to study at a university. The repetitions of the token *no* also support this interpretation. According to Stivers (2004:288), "[m]ultiple sayings function to display that the speaker finds the prior speaker's course of action problematic".

Concerning the facial gesture, a recent study shows that raised eyebrows can be used "for 'zooming in' to the surprising piece of information" (Dix/Groß 2023:20). In Excerpt 2, it can be observed that the raised eyebrows function as an embodied resource to emphasize the requested information and make clear that HE19 does not pay for university. The multimodal resources are employed in this position to indicate that the preceding question presents a problematic issue. Due to a misunderstanding, HE19 produces a request for confirmation.

HE19 then explains that university is free *for everyone*, (line 20). A request for reconfirmation (see Gipper/Groß 2023) is launched by SB93: *it's COMPLETELY free?* (line 22). HE19 responds with multiple affirmative tokens (Stivers 2004), which can also be interpreted as a minimal response and an unwillingness to advance the topic. Furthermore, it indicates that the information should, by now, be clear. The sequence ends with SB93's *[oh okay].* (line 25) as a sequence-closing third (Schegloff 2007), showing that she has understood the information.

In sum, participants' responses are characterized by the Gestalt *oh + raised eyebrows.* In Excerpts (1) and (2), the responsive action conveys that the requested information or confirmation is inapposite, either because it contains a sensitive matter or because the information was misunderstood. Also, in Excerpt 2, as noted previously, the interlocutor displays resistance and corrects the inference of the first turn.

4.2 Responsive turns with raised eyebrows: repeated eliciting of informings

The analysis of Skype calls reveals that raised eyebrows in turn-initial position, when coordinated with the particle *oh,* indicate a problem with the preceding polar question. Similarly, the facial gesture alone can be deployed to indicate a problem. However, differently from *oh* + raised eyebrows, raising eyebrows without the turn-initial *oh* relates to the repeated eliciting of an informing.

In their responsive actions, participants raise their eyebrows to indicate that the requested information should already be known because it was covered earlier in

the conversation. Excerpts (3) and (4) exemplify these cases. Excerpt (3) shows the initial part of the call between SB75 and HE01 while they are introducing themselves and explaining why they are participating in the recordings of ViMELF (2018).

Excerpt (3): Finland (00:01:45)

```
01 SB75: .h where are YOU from?
02 HE01: ... .h uh:m,
03       I'm from a small city close to Helsinki?
04 SB75:   [°oh°].
05 HE01: ..[in Finland?]
06 SB75: ..°nice yeah°.
   {HE01 talks about her major and participation in the study}
15 HE01: .. .h well I'm study:ing Economics?
16       at uh a:n a university in- in Helsinki?
17       an:d I have to take academic writing,
18       which is an English course so I can write my thesis in English.
19 SB75: ... OH that's interesting.
20       that's >very cool<°.
21       so are you from Finland?
22       or: u:hm are you from another country?#
   fig                                          #3a
```

Figure 3a

```
23 HE01: .. .h +no# I'm from Finland.+
                +RE and closes eyes---+
   fig             #3b
```

Figure 3b

```
24 SB75: .. oh okay so:.
```

In the first part of their exchange, SB75 asks where HE01 is from (line 1) and HE01 answers she comes from a city close to Helsinki, in Finland (lines 3 and 5). In lines 7 to 14, HE01 reports she is participating in the corpus study as a part of a course in academic writing (omitted from the transcript). In lines 15 to 18, HE01 provides

more information, stating she needs to take academic writing for her studies. In line 15 she explains: *.. .h well I'm study:ing Economics at uh a:n a university in- in Helsinki?.* SB75 produces an *oh*-prefaced assessment (lines 19 and 20) and asks a request for information designed as an alternative question (see Stivers/Enfield 2010): *so are you from Finland. or: u:hm are you from another country?* (lines 21 and 22). Figure 3a shows both participants as they look at their screens and no facial gestures could be observed.

After a short pause and an in-breath, HE01 produces an answer *.. .h no I'm from Finland.* (line 23). During her response, she raises both eyebrows and closes her eyes (Figure 3b). The embodied response relates to a question concerning something that should already be known. SB75's assessment provided in line 6: *..°nice yeah°.* does not seem to signal any problematic issue, that is, she does not seem to realize the same question had been asked twice. HE01 has, multiple times, given information about where she is from, stating she is from a city close to Helsinki (line 3), reaffirming its location: *.. [in Finland?]* (line 5), and affirming she studies at a university in Helsinki (line 16). Her eyebrows only return to a *home position* (Sacks/Schegloff 2002) at the end of the turn.

A similar example is shown in Excerpt (4), in which SB78 and HE04 talk about languages.

Excerpt (4): French (00:27:45)

```
01  SB78: our area here is VEry close to France?
02         .. [so],
03  HE04:     [okay].
04  SB78: .. we always start with with French in school.
{participants discuss other topics for around 7 minutes}
305 HE04: have you studied any# other +languages#?
    fig                          #4a                #4b
                                          +RE->
```

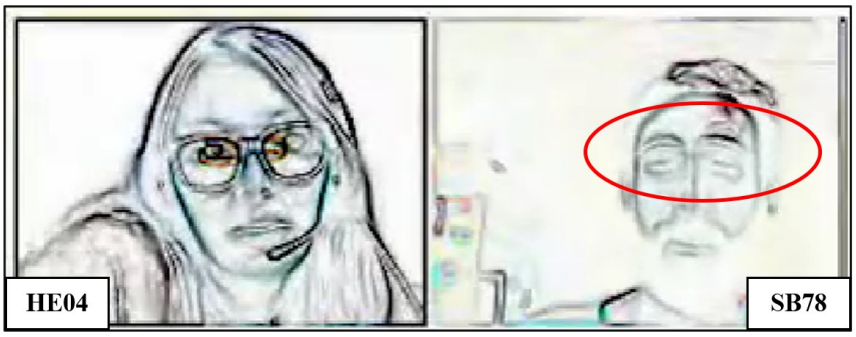

Figure 4a

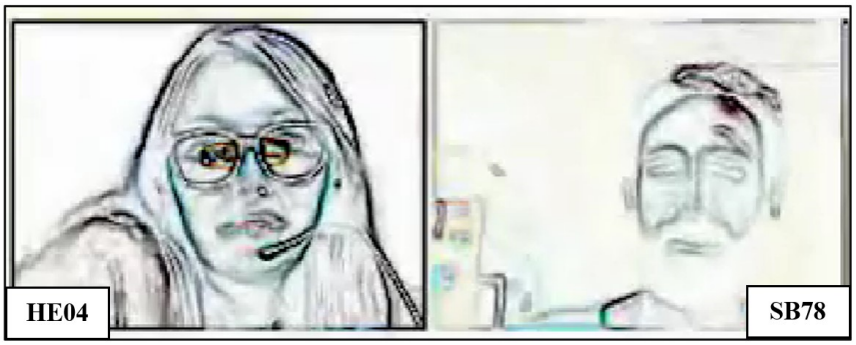

Figure 4b

```
306 SB78: .. yeah+ +French#+ of course.
                ->+
                    +E flash+
    fig                    #4c
```

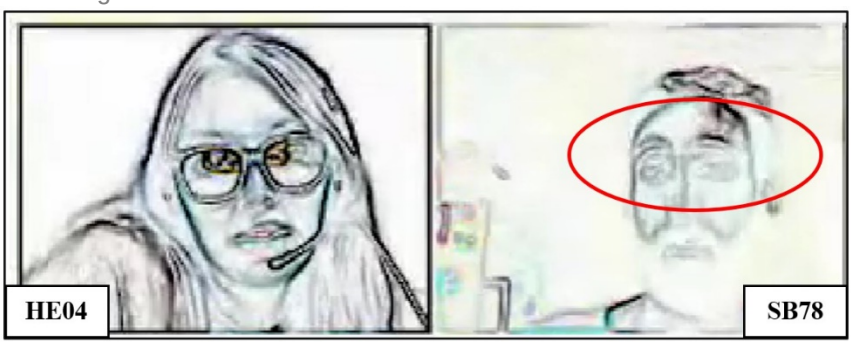

Figure 4c

```
307 SB78: .. because we have to here ((laugh)).
308 HE04: ... yeah?=
309 SB78: =because it's-
310       ... we- u:h this area,
311       .. of Germany used to be French,
312       .. at [some] time in history.
313 HE04:       [uh huh]?
314 SB78: so only wait let me,
315       .. not even sixty years ago,
316 HE04: ... okay,
317 SB78: ... it was French so we we still,
318       ... have to learn it,
319       bu- but I've forgotten °all of it°,
320       .. °I have to [admit]°.
321 HE04:               [yeah]?
```

In the initial part of the excerpt, SB78 explains that he lives close to France (line 1) and that, in this region, they study French at school (line 4). After discussing other topics for approximately seven minutes (omitted from the transcript), HE04 asks the polar question: *have you studied any other languages?* (line 305). Figure 4a shows that no facial gesture from either HE04 or SB78 is observable. When HE04 utters the noun *languages*, SB78 raises his eyebrows (Figure 4b). Both eyebrows remain raised throughout the initial part of his answer: *yeah French of course.* (line 306), that signals givenness of information.

It can be argued that the first instance of eyebrow lift, that is, when HE04 utters *languages* (line 305), is analytically unrelated to the request due to transmission delay. However, while he utters the word *French*, an eyebrow flash is noticed (transcribed as 'E flash', Figure 4c). Therefore, with the embodied response, SB78 orients to the inappositeness of the question. As line 4 shows, SB78 has previously affirmed that, where he lives, people learn French at school. This is also supported by *of course* and a further explanation given by SB78 in line 307: *.. because we have to here*. According to Stivers (2011:104), "when 'Of course' is preceded by 'Yes', confirmation […] is arguably prioritized, and the challenge to question's askability is secondary to that". Nonetheless, the facial gesture seems to orient to the repetition of the request for information from HE04 and it seems to indicate that HE04 has either forgotten or misheard SB78's previous statement.

HE04 then produces a confirmation check … *yeah?*= (line 308), which seems to index that she does not recall hearing this information before. SB78 then provides an explanation, stating that the region where he lives in Germany used to be a part of France (lines 309-312 and 314-316) and reaffirms the initial information (lines 317 and 318).

Excerpts (3) and (4) have shown responses to self-explanatory information, that is, questions that had been previously asked and answered in the conversation. These questions display incorrect assumptions and these are blocked by the interlocutors' responses. As an embodied display to resist to the wrong inference, participants raise both eyebrows and this facial gesture is coordinated with their vocal turn.

5. Summary and Discussion

This paper has shown how participants respond to questions with *oh* and raised eyebrows or just raised eyebrows in a video-mediated environment. The questions are marked by issues of askability (see Stivers 2011). *Oh* is commonly used to preface answers that signal a problematic issue with the preceding question (Heritage 1998). I have addressed how participants' responsive turns are *oh*-prefaced and the turn-initial particle is accompanied by the raising of both eyebrows. These examples were then compared with cases in which raised eyebrows coordinate with the responsive turn, but do not co-occur with the production of *oh.*

The present study has shown how the Gestalt *oh* + *raised eyebrows* is used to respond to directive-type questions, requests for new information, and requests for confirmation. However, it is important to note that these examples are part larger corpus with other types of inapposite questions.

In Excerpts (1) and (2), participants respond to questions and mark them as out of place or unexpected by employing *oh* + raised eyebrows. While Excerpt (1) is an example of a potentially sensitive question, addressing the origin of a participant's first name, Excerpt (2) is marked by a misunderstanding because a participant seeks to re-confirm an already stated fact. In Excerpts (3) and (4) participants also deploy raised eyebrows to respond to questions. However, they do not use the change-of-state token *oh.* Participants treat the questions as unexpected by raising both eyebrows as responses to questions about matters that were previously discussed and should be known.

The analysis points to two emerging practices. In the first practice, participants display the inadequacy of the inquires by challenging its askability with *oh* and raised eyebrows. According to Heinemann (2009:170), "by explicating exactly why a question should not have been asked, the answerer would be challenging the questioner's inapposite behavior in an overt manner". It is also explained that this "would only serve to further suspend the progressivity of the activity" (ibid). Therefore, though *oh*-prefaced responses can signal inappositeness, they do not hinder the interaction. Existing literature on *oh* focuses on epistemics, particularly in informing sequences (see e.g., Heritage 2012; Thompson/Fox/Couper-Kuhlen 2015). In such sequences, the particle *oh* serves to signal that the participant was informed about something. However, in the case of responses to inapposite questions participants assume that the co-participants have *epistemic access* (see Thompson/Fox/Couper-Kuhlen 2015) to the response. In contrast, the backward-

looking property of *oh* placed in a responsive position shows how the particle orients to the preceding questions by signaling a trouble and "that the speaker's attention has been drawn towards something unexpected" (Heritage 2018:165).

In the second practice, participants also respond to inapposite questions with raised eyebrows that co-occur with the vocal turn (in Excerpts 3 and 4) but do employ the turn-initial particle *oh*. Eyebrow movements such as frowning (Kaukomaa/Peräkylä/Ruusuvuori 2014) and raised eyebrows (Hömke 2018) are related to interactional trouble. While the embodied conduct of participants marked by raised eyebrows relates to strong and problematic assessments and inferences in all four analyzed cases, the latter cases do not display the notion of unexpectedness and change-of-state signaled by *oh*.

There are several implications that emerge in the study of embodied actions in VMI. For example, multimodal turns in video-mediated interaction occur in a complex environment in which recognizing facial gestures is a challenging task. Latency (see e.g., Seuren et al. 2021) and problems with video quality are among the potential issues of video-mediated settings, which can interfere with the recognition of facial gestures by co-participants and by the analyst. Furthermore, connection issues might also influence the vocal conduct of participants and, consequently, turn-taking behavior. A further example concerns the data collection, for example, whether the video was recorded by one participant or the other. This can affect the analysis since the same interaction might have been recorded differently, for instance one participant may experience some sort of delay while the other does not (see e.g., Ruhleder/Jordan 2001).

My findings contribute to the emerging field of research in gesture deployment in VMI by addressing how participants make their facial gestures visible in this type of setting, given its potential constraints. Further research needs to address whether other types of embodied behavior co-occur with the turn-initial particle *oh* in this sequential position. Furthermore, studies need to be conducted to uncover other embodied resources participants employ to respond to inapposite questions and whether these practices are exclusive to a video-mediated environment.

6. References

Bolden, Galina B. (2006): Little Words That Matter: Discourse Markers 'So' and 'Oh' and the Doing of Other-Attentiveness in Social Interaction. In: Journal of Communication 56 (4), 661-688.

Bolden, Galina B. (2015): Discourse Markers. In: Tracy, Karen (ed.): The International Encyclopedia of Language and Social Interaction. John Wiley & Sons, 1-7.

Brunner, Marie-Louise / Diemer, Stefan / Schmidt, Selina (2017): '... Okay so Good Luck with That ((Laughing))?' – Managing Rich Data in a Corpus of Skype Conversations." In: Hiltunen, Turo / McVeigh, Joe / Säily, Tanja (eds.): Studies in Variation, Contacts and Change in English 19 (Big and Rich Data in English Corpus Linguistics: Methods and explorations).

CASE transcription conventions (2017): Birkenfeld: Trier University of Applied Sciences. http://umwelt-campus.de/case-conventions. Accessed: 02.08.2023.

Couper-Kuhlen, Elizabeth (2021): Language over Time: Some Old and New Uses of OKAY in American English. In: Interactional Linguistics 1 (1), 33-63.

Curl, Traci S. / Drew, Paul (2008): Contingency and Action: A Comparison of Two Forms of Requesting. In: Research on Language and Social Interaction 41 (2), 129-153.

De Stefani, Elwys (2021): Embodied Responses to Questions-in-Progress: Silent Nods as Affirmative Answers. In: Discourse Processes 58 (4), 353-371.

De Stefani, Elwys (2022): On Gestalts and Their Analytical Corollaries: A Commentary to the Special Issue. In: Social Interaction. Video-Based Studies of Human Sociality 5 (1).

Dix, Carolin / Groß, Alexandra (2023): Surprise about or just registering new Information? Moving and holding both eyebrows as visual change-of-state markers. In: Social Interaction. Video-Based Studies of Human Sociality 6 (3).

Drake, Veronika (2016): German questions and turn-final oder. In: Gesprächsforschung 17, 168-195.

Dressler, Richard A. / Kreuz, Roger J. (2000): Transcribing Oral Discourse: A Survey and a Model System. In: Discourse Processes 29 (1), 25-36.

Due, Brian Lystgaard / Licoppe, Christian (2020): Video-Mediated Interaction (VMI): Introduction to a Special Issue on the Multimodal Accomplishment of VMI Institutional Activities. In: Social Interaction. Video-Based Studies of Human Sociality 3 (3).

Enfield, Nick J. et al. (2019): Polar answers. In: Journal of Linguistics 55 (2), 277-304.

Gipper, Sonja / Groß, Alexandra (2023): Less than Confirming and Doing More than That: Comparing Responses to Requests to Confirmation in German and Yurakaré. In: Contrastive Pragmatics, 1-45.

Heinemann, Trine (2009): Two Answers to inapposite inquires. In: Sidnell, Jack (ed.): Conversation Analysis: Comparative Perspectives. Cambridge University Press, 159-186.

Heinemann, Trine (2015): Negation in Interaction in Danish conversation. In: Skrifter Om Samtalegrammatik, 2 (12).

Heller, Vivien / Schönfelder, Nora / Robbins, Denise (2023): Displaying a critical stance. Eyebrow contractions in children's multimodal oppositional actions. In: Social Interaction. Video-Based Studies of Human Sociality 6 (3).

Helmer, Henrike / Betz, Emma / Deppermann, Arnulf (2021): Coordination of OKAY, Nods, and Gaze in Claiming Understanding and Closing Topics. In: Betz, Emma / Deppermann, Arnulf / Mondada, Lorenza / Sorjonen, Marja-Leena (eds.): OKAY across Languages: Toward a comparative approach to its use in talk-in-interaction. Amsterdam: John Benjamins, 363-393.

Heritage, John (1984): A Change-of-State Token and Aspects of Its Sequential Placement. In: Aktinson, J. Maxwell / Heritage, John (eds.): Structures of Social Action: Studies in Conversation Analysis. Cambridge, U.K., Cambridge University Press, 299-345.

Heritage, John (1998): Oh-Prefaced Responses to Inquiry. In: Language in Society 27 (3), 291-334.

Heritage, John (2002): Oh-Prefaced Responses to Assessments: A Method of Modifying Agreement/Disagreement. In: Ford, Cecilia E. / Fox, Barbara A. / Thompson, Sandra A. (eds.): The Language of Turn and Sequence. Oxford University Press, 196-224.

Heritage, John (2012): The Epistemic Engine: Sequence Organization and Territories of Knowledge. In: Research on Language and Social Interaction 45(1), 30-52.

Heritage, John (2013): Turn-Initial Position and Some of Its Occupants. In: Journal of Pragmatics 57, 331-337.

Heritage, John (2018): Turn-Initial Particle in English: The Cases of Oh and Well. In: Heritage, John / Sorjonen, Marja-Leena (eds.): Between Turn and Sequence: Turn-Initial Particles across Languages. Amsterdam: John Benjamins, 155-190.

Heritage, John / Sorjonen, Marja-Leena (2018): Analyzing Turn-Initial Particles: Chapter 1. Introduction. In: Heritage, John / Sorjonen, Marja-Leena (eds.): Between Turn and Sequence: Turn-Initial Particles across Languages. Amsterdam: John Benjamins, 1-22.

Hömke, Paul (2018): The Face in Face-to-Face Communication: Signals of Understanding and Non-Understanding. PhD thesis, Nijmegen: Radboud University Nijmegen.
https://repository.ubn.ru.nl/handle/2066/200670 Accessed 01.08.2023.

Ilomäki, Saraki / Ruusuvuori, Johanna / Laitinen, Jaana (2021): Effects of Transmission Delay on Client Participation in Video-Mediated Group Health Counseling. In: Qualitative Health Research 31 (12), 2328-2339.

Kaukomaa, Timo / Peräkylä, Anssi / Ruusuvuori, Johanna (2014): Foreshadowing a Problem: Turn-Opening Frowns in Conversation. In: Journal of Pragmatics 71, 132-147.

Kendrick, Kobin H. / Holler, Judith (2017): Gaze Direction Signals Response Preference in Conversation. In: Research on Language and Social Interaction 50 (1), 12-32.

Licoppe, Christian / Morel, Julien (2012): Video-in-Interaction: 'Talking Heads' and the Multimodal Organization of Mobile and Skype Video Calls. In: Research on Language & Social Interaction 45 (4), 399-429.

Luff, Paul / Heath, Christian / Kuzuoka, Hideaki / Hindmarsh, Jon / Keiichi, Yamazaki / Oyama, Shinya (2003): Fractured Ecologies: Creating Environments for Collaboration. In: Human–Computer Interaction 18 (1-2), 51-84.

Mondada, Lorenza (2014): The Local Constitution of Multimodal Resources for Social Interaction. In: Journal of Pragmatics 65, 137-156.

Mondada, Lorenza (2018): Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. In: Research on Language and Social Interaction 51 (1), 85-106.

Mondada, Lorenza (2022a): The Situated and Methodic Production of Accountable Action: The Challenges of Multimodality. In: Maynard, Douglas W. / Heritage, John (eds.): The Ethnomethodology Program: Legacies and Prospects. New York: Oxford University Press, 289-321.

Mondada, Lorenza (2022b): Conventions for multimodal transcription. (Version 6.0.1, Dec. 2022). https://www.lorenzamondada.net/multimodal-transcription. Accessed: 02.08.2023

Norrick / Neal R. (2009): Interjections as pragmatic markers. In: Journal of Pragmatics 41 (5), 866-891.

Pekarek Doehler, Simona (2022): Multimodal Action Formats for Managing Preference: Chais Pas 'Dunno' plus Gaze Conduct in Dispreferred Responses to Questions. In: Journal of Pragmatics 197, 81-99.

Pekarek Doehler, Simona / Polak-Yitzhaki, Hilla / Li, Xiaoting / Stoenica, Ioana Maria / Havlík, Martin / Keevallik, Leelo (2021): Multimodal Assemblies for Prefacing a Dispreferred Response: A Cross-Linguistic Analysis. In: Frontiers in Psychology 12, 1-24.

Raymond, Geoffrey (2003): Grammar and Social Organization: Yes/No Interrogatives and the Structure of Responding. In: American Sociological Review 68 (6), 939-967.

Ruhleder, Karen / Jordan, Brigitte (2001): Co-Constructing Non-Mutual Realities: Delay-Generated Trouble in Distributed Interaction. In: Computer Supported Cooperative Work 10, 113-138.

Sacks, Harvey / Schegloff, Emanuel A. (2002): Home Position. In: Gesture 2 (2), 133-146.

Schegloff, Emanuel A. (1987): Recycled Turn Beginnings: A Precise Repair Mechanism in Conversation's Turn-Taking Organization. In: Button, Graham / Lee, John R. E. (eds.): Talk and Social Organisation. Clevedon, England: Multilingual Matter, 70-85.

Schegloff, Emanuel A. (2000): When "others" initiate repair. In: Applied Linguistics 21 (2), 205-243.

Schegloff, Emanuel A. (2007): Sequence Organization in Interaction: A Primer in Conversation Analysis 1. Cambridge: Cambridge University Press.

Schegloff, Emanuel A. / Lerner, Gene H. (2009): Beginning to Respond: Well-prefaced Responses to Wh-Questions. In: Research on Language and Social Interaction 42 (2), 91-115.

Schiffrin, Deborah (1987): Discourse markers. Cambridge: Cambridge University Press.

Seuren, Lucas M. / Wherton, Joseph / Greenhalgh, Trisha / Shaw, Sara E. (2021): Whose Turn Is It Anyway? Latency and the Organization of Turn-Taking in Video-Mediated Interaction. In: Journal of Pragmatics 172, 63-78.

Stivers, Tanya (2004): 'No No No' and Other Types of Multiple Sayings in Social Interaction. In: Human Communication Research 30 (2), 260-293.

Stivers, Tanya (2005): Modified Repeats: One Method for Asserting Primary Rights From Second Position. In: Research on Language & Social Interaction 38 (2), 131-158.

Stivers, Tanya (2011): Morality and question design: "of course" as contesting a presupposition of akability. In: Stivers, Tanya / Mondada, Lorenza / Steensig, Jakob (eds.). The Morality of Knowledge in Conversation. Cambridge University Press, 82-106.

Stivers, Tanya (2022): The Book of Answers: Alignment, Autonomy, and Affiliation in Social Interaction. Oxford, New York: Oxford University Press.

Stivers, Tanya / Enfield, Nick (2010): A Coding Scheme for Question–Response Sequences in Conversation. In: Journal of Pragmatics 42 (10), 2620-2626.

Stivers, Tanya / Hayashi, Makoto (2010): Transformative Answers: One Way to Resist a Question's Constraints. In: Language in Society 39 (1), 1-25.

Stivers, Tanya / Sidnell, Jack (2005): Introduction: Multimodal interaction. In: Semiotica 156, 1-20.

Stolle, Sarah / Pfeiffer, Martin (2023): Stand-alone facial gestures as other-initiations of repair. Social Interaction. In: Social Interaction. Video-Based Studies of Human Sociality 6 (3).

Thompson, Sandra A. / Fox, Barbara A / Couper-Kuhlen, Elisabeth (2015): Responses in Informing Sequences. In: Grammar in Everyday Talk: Building Responsive Actions. Cambridge: Cambridge University Press, 50-118.

ViMELF (2018): Corpus of Video-Mediated English as a Lingua Franca Conversations. Version 1.9. The CASE Project. http://umwelt-campus.de/case/.

7. Transcription conventions

Transcription conventions for speech (Dressler/Kreuz 2000, modified)

?	Rising intonation
.	Falling intonation
,	Continuing intonation
TEXT	Stress or loud speech
(0.6)	Pause, tenths of a second
..	Pauses of one-half a second or less
…	Pauses of more than a half-second
<text>	Spoken slowly
>text<	Spoken rapidly
:	Lengthened syllable
-	Word cut-off
=	Latched talk (lack of temporal gap between two speakers)
[ ]	Overlapping speech
°text°	Spoken softly
H	Audible breathing
.h	In-breath
h	Out-breath
( )	Unclear or unintelligible speech
“ “	Shift in speaker’s voice
((laugh))	Laughter
word	Word replaced for privacy reasons
{ }	Analyst’s comments

Hesitation (CASE Transcription conventions 2017, modified)

mhm	Closed mouth, agreement or understanding
m/m:/mg/m:h	Closed mouth, single sound
uh huh	Open mouth
uh/u:h	Open mouth, single sound, often with hesitations
e:r	Open mouth, clear pronunciation or /r/, often with hesitations
uhm/u:hm	Open mouth that is closed at the end, also with hesitations

Multimodal transcription (Mondada 2018, 2022b)

* *	Descriptions of embodied actions are delimited between
+ +	two identical symbols (one symbol per participant and per type of action)
Δ Δ	that are synchronized with correspondent stretches of talk or time indications.
*--->	The action described continues across subsequent lines
---->*	until the same symbol is reached.
>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
.....	Action's preparation.
----	Action's apex is reached and maintained.
,,,,,	Action's retraction.
ric	Participant doing the embodied action is identified in small caps in the margin.
fig	The exact moment at which a screen shot has been taken
#	is indicated with a sign (#) showing its position within the turn/a time measure

Marina Reis de Souza
Institut für englische Sprache und Literatur
Stiftung Universität Hildesheim
Universitätsplatz 1
31141 Hildesheim

reis@uni-hildesheim.de

Veröffentlicht am 6.4.2024