Social Interaction

Video-Based Studies of Human Sociality

_____

Sustained Pointing Gestures in Instructions and Questions: How the Temporal Extent of a Gesture Matters in Interaction

_____

Mojenn Schubert

Leibniz Institute for the German Language (IDS), Mannheim

Abstract

Gestures can be brief and compact in their execution, but also elaborate and extended. One way to utilise this kinetic flexibility is to extend one's gesture in time by holding it in its stroke position. This study explores the interactional function of gestural holds by investigating pointing gestures that are sustained beyond a sequence-initiating turn and into the responsive space following it. The study draws on video data from naturally occurring conversations in German and focuses on held pointing gestures after instructions and questions. It is shown that in both action environments, participants delay gestural closure to indicate that they still consider the addressee's response to be insufficient.

Keywords: pointing gestures, gestural hold, TRP, questions, instructions

1. Introduction

Pointing is a primary resource for communication in face-to-face interaction (Kita, 2003): Being one of the first gestures learned in early childhood (Butterworth, 2003), it provides for a basic communicative tool to direct the attention of others to objects, places or directions in the local environment of the participants. Maintaining a pointing gesture beyond the end of a turn and the moment when joint orientation has been established goes against expectations informed by previous research (Kendon, 2004; Stukenbrock, 2015). This paper examines such sustained pointings in the context of instructions and questions and argues that participants employ the temporal extension of their gesture to meet interactional demands that arise out of the locally situated context of conversation.

Sacks et al. (1974) state that not only turn-allocation but also turn-size is "locally managed, party-administered, and interactionally controlled" (p. 727). The duration or temporal extent of a gesture could be oriented towards and shaped by these interactional concerns in a similar way. To capture this potential interactional orientation in natural conversational data, this study analyzes instances in which the trajectory of a particular gesture – pointing – is extended by the integration of a hold (Kendon, 2004). The sequential contexts under investigation are first, sequence-initiating actions (Schegloff, 2007) that strongly call for a response (instructions and questions). This positional restriction allows us to explore the use of sustained pointings in structurally comparable environments where progress is highly dependent on interactional cooperation. It is shown that in sustaining gestures, participants often orient to what is necessary from an interactional point of view rather than dealing with intrapersonal processes of gesture-speech coordination.

2. Previous Research

2.1 Pointing gestures as part of actions in conversation

Generally speaking, pointing gestures are used to indicate an object, a direction or a location in space by projecting a vector from an extended body part (Kendon, 2004, p. 200; Kita, 2003, p. 1; McNeill, 1992, p. 28; Stukenbrock, 2015, p. 21). In its most prototypical use,[1] the pointing's target is physically present in the immediate surroundings and therefore perceivable by the interlocutors. Pointing is an effective way of drawing the attention of others – especially their

---

[1] The functional spectrum of pointing is of course not limited to *concrete pointing* (McNeill, 1992, pp. 18, p. 173) in this sense. *Abstract pointing* (McNeill et al., 1993) is used in interaction as well, as are cases of *Deixis am Phantasma* (Bühler, 1934, pp. 124).

gaze – to this target. The establishment of joint attention is, however, vulnerable to failing due to misalignments between the interactants. The pointing participant, on the one hand, needs to ensure that both gesture and target are perceivable to co-participants. This is done by monitoring the addressees' visual orientation, especially their gaze and bodily orientation (Stukenbrock, 2015, pp. 313). The addressee, on the other hand, must first locate the target in space and then come to an adequate interpretation of the intended referent (Clark, 1996, pp. 168; Clark, 2003, p. 246; Streeck, 2017, p. 125; Stukenbrock, 2015, p. 296). This process of reciprocal adjustment results – ideally – in perceptual alignment. Therefore, pointing gestures play a crucial role in referential practices (Erikson, 2009; Hindmarsh & Heath, 2000; Sidnell, 2005) and in building personal common ground (Clark, 1996, p. 113). To facilitate the disambiguation of potentially ambiguous referents, pointing gestures are often coordinated with verbal expressions such as demonstratives (Cooperrider, 2016), together forming an exophoric reference (Goodwin 1981a). These "composite signals" (Clark, 1996, p. 168) can become quite dense as they integrate not only gesture and speech, but also gaze and other embodied resources displaying one's spatial orientation (*multimodale Verdichtungsräume/places of multimodal densification;* Stukenbrock, 2015, pp. 32).

Since establishing joint reference is a fundamental starting point for conversational progression, pointing is a prevalent resource in a variety of social actions that either constitute or build upon recognitional acts. For instructions and questions, the social actions under investigation here, the status of locally shared reference, and thus the status of pointing, certainly differs in detail. But at the same time, they both represent prototypical sequence-initiating actions (Schegloff, 2007) that pose similar demands on their addressee, namely, to respond with a complying second action.[2] This provides a good opportunity to study whether and how participants show sensitivity to interlocutors' verbal and embodied conduct when holding a pointing gesture.

Instructions are understood here as social actions that are "designed to get someone to do something" (Lindwall et al., 2015, p. 145). Instructive actions are seen as embodying a stable deontic asymmetry (Stevanovic & Peräkylä, 2014, pp. 190) that holds for a local participation framework (Goodwin, 1981b, p. 137) with one participant claiming their deontic authority over the actions of another. Instructions can be considered a type of requests for action (Curl & Drew, 2008; Drew & Couper-Kuhlen, 2014), whereby the linguistic format often conveys more

---

[2] Of course, this also applies to other sequence-initiating actions such as (first) assessments, suggestions or offers, which I also encountered when exploring data for this study. However, I decided to focus on instructions and questions for various reasons: (i) Both actions are well understood in terms of the sequential expectations they build which makes it not only possible to account for action-specific usages but also to detect more general functions of gestural holds. This systematic comparative analysis would not have been possible with a more heterogeneous collection. (ii) They occurred most frequently and thus provided a sufficiently large database.

specific information on how the instructed action is to be performed (Deppermann, 2018a, p. 280). Pointing gestures are often employed to support direction-related instructions or to indicate objects (e.g. tools) that are central to the implementation of an instructed action: Rossi (2014, p. 309) shows that participants use pointings as a nonverbal request to perform a task at a specific location or the transfer of an object. In addition, pointings support recruiting moves (Floyd et al., 2020) and are commonly used as directional devices in the context of driving instructions (De Stefani, 2014; Deppermann, 2018a, 2021).

Question-answer sequences are prototypical adjacency pairs (Schegloff, 2007) aimed at requesting information to resolve an epistemic asymmetry between questioner and addressee (de Ruiter, 2012). Depending on their formatting, questions make different types of responses relevant (Heritage & Raymond, 2012; Raymond, 2003; Steensig & Heinemann, 2013; Stivers & Hayashi, 2010). Questions are understood here as invoking an epistemic asymmetry tilted towards the recipient and making, based on this asymmetry, (dis)confirmation, correction, or the delivery of information relevant in second position. How pointing gestures are employed in the context of questions has not yet been studied systematically, apart from scattered observations (Blythe et al., 2018, p. 160; Clark, 2012, p. 92). From what has been noted there, pointings can serve questions that work towards shared reference, but also function as an addressing device in multi-party interaction.

2.2 Gestural holds from an interactional perspective

The structure of gestural production has been extensively described in previous gesture research (Kendon, 1980, 2004; Kita et al., 1998; McNeill, 1992; Schegloff, 1984): The trajectory of a gestural unit (Kendon, 2004, pp. 111) minimally consists of a preparation of the articulators, a stroke-phase and a retraction towards a rest position. McNeill and Kendon stress that speakers constantly work towards gesture-speech-synchrony, which describes the temporal coordination of the stroke with that part of the co-occurring speech that semantically complements the gesture in a meaningful way (McNeill, 1992, pp. 25; Kendon, 2004, pp. 119). Integrating pre- and post-stroke holds that freeze a gesture in place (Kita et al., 1998, p. 26) allows for temporal manipulations that ensure this synchrony even in the case of gesture-speech discoordination: If speech lags behind, participants can hold their gesture in stroke position, prolonging it until the semantically associated utterance catches up (Kendon, 2004, pp. 136; Kita et al. 1998, p. 26). According to this reasoning, pointing gestures should normally be resolved as soon as the corresponding

lexical affiliate (Schegloff, 1984) has been uttered.[3] Indeed, a recent experimental study (Cooperrider et al., 2021, p. 10) supports the claim that pointings only rarely extend beyond the speech unit they are semantically associated with. Maintaining a pointing gesture beyond the end of a turn therefore defies expectations that were informed by previous gesture research. By approaching this phenomenon from a CA-perspective, this study aims to contribute to a better understanding of why participants sustain their pointing gesture. Because gesture research has focused strongly on intrapersonal, cognitive processes of gesture-speech coordination, interactional aspects that might account for holding a gesture have been largely disregarded. Close observation of the locally situated context of a gesture may instead reveal interpersonal factors that motivate on-line adaptations of a gesture's production.

In the case of pointings, extensions might be required to ensure that the recipient can successfully follow the gesture's direction by gaze and identify its target. Stukenbrock (2015) demonstrates that a held pointing gesture can provide a visual guideline for as long as the addressee needs it, irrespective of the boundaries of a turn: "The duration of the gestural apex is adaptively oriented to the addressees' behaviour *[own translation]*" (Stukenbrock, 2015, p. 69). Here, the participant clearly prioritises the establishment of joint attention over gesture-speech-synchrony. Occasional observations show that sustained pointing gestures can be withdrawn in response to an embodied shift of attention (Stukenbrock, 2015, p. 69) or a verbal turn that displays attention to or recognition of the target referent (Clark, 2005, p. 511; Sidnell, 2005, pp. 72). Providing perceptual support thus seems to be one way in which held pointing gestures address interactional needs. However, research on gestural holds in general suggests a much broader spectrum of potential uses for temporally extended pointings.

During sequences in which shared understanding is negotiated, holding an iconic gesture across a TRP can be used to mobilize response (Sikveland & Ogden, 2012). Similarly, when children prolong their pointing gestures, this leads to an increased responsiveness by their parents (Andrén, 2011). Going into detail about complex gesture production, Enfield (2009) argues that gestural holds serve a crucial function in two-handed gestures: Here, the stationary freezing of one gesturing hand allows for pragmatic backgrounding of its semantic content while the other, dominant hand continues to gesture and deliver the new, focal information (Enfield, 2009, p. 114). Gestural holds therefore function as "a kind of enchronic glue, inhabiting a structural position [...] across moves in a discourse trajectory" (Enfield, 2009, p. 146). Sustained pointing gestures operate on the level of interactional organization as well: By extending a pointing gesture

---

[3] Research on the temporal relationship between iconic gestures and speech suggests that gesture onset (Kendon, 1980; Schegloff, 1984) and even stroke onset (ter Bekke et al., 2020) slightly precede the respective lexical affiliate.

across its host turn, participants display their continuous engagement over the course of subsequent turns (Mondada, 2007, pp. 215).

All of the observations reviewed so far hint at one fundamental quality of post-stroke holds: They display continuity in a visually perceivable way. In spontaneous conversation, freezing a gesture indicates "the continuing relevance of the talk, turn or interactional project with which [it is] associated" (Sidnell, 2005, p. 75) and sustained pointings in particular convey the communicative signal 'I continue to want you to attend to this' (Clark, 2005, p. 511). However, what is indexed as 'not yet completed' could manifest differently depending on the local requirements of the social action in progress. Vice versa, withdrawing a held gesture is often coordinated with a verbal or embodied contribution that marks sequence closure (Mondada, 2007, pp. 215; Mondada, 2015, pp. 301; Sidnell, 2005, pp. 72; Sikveland & Ogden, 2012, p. 174; Stukenbrock, 2015, p. 69).

A careful look at previous research also shows that the phenomenon of holding a gesture has not been treated as the central object of investigation in its own right. Gesture studies conceive of holds as an epiphenomenon of the constant striving for synchrony between gesture and speech. Existing CA-studies give valuable insights into the interactional force of holds, but their data does not allow more general statements on the function of gestural holds as such (i.e. independent from the heterogenity of gestural form and sequential contexts). The design of the present study was chosen to counteract both shortcomings: By focusing on one particular gesture, its deliberate prolongation can be evaluated against a unified gestural form.[4] And focusing on a compact set of sequence-initiating actions allows us to analyse the gestural hold in a basic sequential position that promises strong interactional potential while still attending to the specific circumstances of the distinct action environments.

3. Data and Method

This study is based on video-recorded mundane and institutional talk-in-interaction in German. Data come from FOLK, the "Research and Teaching Corpus of Spoken German" which is hosted at the Leibniz-Institute for the German Language in Mannheim, Germany. From all persons recorded, FOLK obtained written informed consent and permission to use their data – including unmasked images – for scientific purposes (Schmidt, 2016).[5] From a total of 12

---

[4] This consistency of gesture type was not present in the study conducted by Sikveland & Ogden (2012): Their data consist of a group of iconic gestures which certainly take very different forms due to their idiosyncraticity (McNeill 1992, p. 41).

[5] To protect the identity of the individuals, names have been pseudonymized in the transcripts and masked in the audio. The corpus is published through the "Database

recordings covering institutional settings (e.g. driving school lessons and sales interactions) and mundane settings (e.g. board game sessions and home renovation), 8½ hours were selected and annotated for pointing gestures accompanying a sequence-initiating question or instruction.[6] All questions and instructions made transition relevant at the end of their turn (Ford & Thompson, 1996; Selting, 2000). Building on the definition by Kita (2003, p. 1), pointing gestures were defined as linear body movements articulated through an arm or hand. The configuration of the hand was not restricted to index finger pointing, so open hand pointing (Kendon, 2004, p. 206) was included as well. Coding was done in ELAN (Wittenburg et al., 2006) and included the notation of gesture phases (Kendon, 2004; Kita et al., 1998) according to the frame-by-frame method introduced by Seyfeddinipur (2006). Also, the temporal relationship between gesture and speech in first position and the responsive behaviour in second position was coded. Only pointing gestures that incorporated a post-stroke hold (Kita et al., 1998, p. 26) after reaching their apex and extended beyond the end of their host turn were included in the collection. In the analysis, close attention was paid to the moment at which joint attention was established: Pointing gestures that extended beyond this moment were considered to contribute more to the interaction than facilitating reference recognition.

In sum, the analysis builds on a collection of 55 cases, consisting of 29 instructions and 26 questions. The main method used is CA (Sacks et al., 1974; Sidnell, 2013) and multimodal interaction analysis (Deppermann, 2008, 2018b; Deppermann & Streeck, 2018; Mondada, 2013, 2016). The extracts presented in the paper were transcribed using the Jeffersonian system (Jefferson, 2004)[7] and conventions for multimodal transcription (Mondada, 2019).

---

for Spoken German" and is accessible online for scholars after registering at http://dgd.ids-mannheim.de.

[6] The selection was guided mostly by functional criteria (see 2.1) and permitted the inclusion of diverse turn-formats: Since in German, instructions can carry imperative (Deppermann, 2021; Zinken & Deppermann, 2017) as well as interrogative (Gubina, 2021) and declarative (Deppermann, 2018a) morpho-syntax, all of these formats were included. Similarly, for questions, both turns with interrogative (Stivers, 2010) and declarative (B-event statements, Labov & Fanshell, 1977) morpho-syntax were treated as potential candidates.

[7] Transcripts include an idiomatic English translation. Transcription lines of embodied behaviour are used as follows: -h for "hand", -rh for "right hand", -lh for "left hand", -g for "gaze", -f for "face", -b for "body", -d for "driving".

4. Instructions with Sustained Pointing Gestures

In this section, three examples of instructions show how sustained pointing is used in close coordination with the responsive behaviour of the adressee. It becomes evident that the trajectory of the gesture is produced in a way to reflect the temporality of the instructed action (Ex. 1, 2 and 3) as well as a corrective quality of the instruction itself (Ex. 2 and 3).

4.1 Continued attention throughout an activity

During a driving school lesson, the car is approaching an intersection in a residential area. As the junction comes into sight, the instructor RK informs his student EA that the rule 'right-before-left' applies in this case (l. 01-02), which means that they must give way to traffic coming from the right. He then gives a series of instructions about how to pass slowly and where to look.

Excerpt 1. *FOLK_E_00167_SE_01_T_01_DF_01 / c363-c375*

```
01  RK:   so  jetz ham wir schon wieder_n problem, da is schon wieder
           so now we have a problem again       there is again

02  RK:   rechts vor links, jetz trittst du ganz schnell die kupplung,
           right before left now you engage the clutch really quickly

03         (0.7)

04  RK:   <brem:st,>
           (you) brake

05         (0.3)

06  RK:   •wirst G+•A:#:NZ langsam;
           (you) become really slow
       rk-g •steering wheel•front window------------->
=>     rk-rh        +points ahead to street at right-> l.10
       Fig.          #1.1
```


Figure 1.1: RK points ahead while demanding that EA slow down; RK controls the steering wheel.

```
07         (1.5)
```

```
08  RK:  noch langsamer,
          even slower

09         •(1.0)
    rk-g ->•to student->

10  RK:  ~•jetzt läss_te die bremse $•+los, (.) l$o:•s?,
          now you let go of the brake     (.) (let) go
    ea-g >>front window------------$right------$front----------->
    ea-d ~car slowly approaches intersection--------------------->
    rk-g ->•front window------------•steering wheel•right window->
=>  rk-rh-------------------------->+moves pointing along right window->

11         $(0.9)
    ea-g ->$right window->


12  RK:  •jetzt guck+#ste~,
          now you look
    rk-g ->•to student--------------------->
=>  rk-rh---------->+points at right street-> l.14
    Fig.          #1.2
```


Figure 1.2: RK points to the right while controlling the steering wheel and monitoring EA.

```
13         ~(0.2)•(1.7)~
    rk-g ----->•right window------------->
    ea-d ->~reaches middle of intersection~

14  RK:  jetz •+lässte die +kupplung +*kommen-=
              now you release the clutch
         rk-g ---->•front window->>
     =>  rk-rh----->+snaps fingers+      +raises hand palm up-->
         rk-lh                            *raises hand palm up->

15  RK:  =und gibst et$was +*gas, und bist weg.
              and (you) accelerate a bit and (you) are gone
         ea-g            ->$front window>>
         rk-rh               -->+
         rk-lh               -->*
```

After giving some technical instructions to reduce speed (l. 02-04), RK begins to point ahead towards the intersecting street on their right (l. 06, fig. 1.1). Throughout the following talk, RK sustains his pointing gesture across several subsequent instructions. After telling his student to slow down a second time (l. 08), RK briefly looks at EA during a pause (l. 09) to observe her behaviour. He continues with another instruction regarding the brake and looks out the right window again (l. 10), which finally prompts EA to turn her gaze to the right as

well (l. 11). Thereby, she anticipates the upcoming relevance of the pointing target. Her attentional shift becomes visible to RK only in l. 12, when he looks back to EA. This second monitoring action is coordinated with the first instruction that explicitly refers to EA's visual orientation: *jetzt guckste* ('now you look', l. 12).[8] The pointing gesture complements the instruction by visually indicating the target of attention, the street at the right side of the car. This composite utterance (Clark, 1996) is produced precisely when the car has reached the intersection, so that both are now in a good position to monitor the traffic situation there. At this point, joint orientation has been mutually established: RK is able to see that EA is, in fact, paying attention to the street on the right (fig. 1.2). Instead of withdrawing, however, he maintains his pointing gesture while looking out the right window for another 1.9 seconds as they slowly enter the middle of the intersection (l. 13). Only then, when they have passed the major part of the junction, he does release the gesture and initiate a next instruction aimed at gaining speed again (l. 14-15).

The sustained pointing gesture in this extract contributes to the driving exercise in multiple ways. First of all, the early onset of the gesture enables the student to identify a location of relevance that becomes the focal point of attention during the course of driving. The gesture serves as a spatially and temporally stable point of orientation while the student has to manage multiple tasks at once (Haddington et al., 2014). Driving students need to be able to orient themselves in a highly dynamic environment, as the traffic situation on the roads is constantly changing. So, monitoring the local surroundings – and especially traffic having primary right of way – while operating the car becomes a crucial skill which is being instructed here with the help of the pointing gesture. The driving instructor not only needs to guide EAs attention to the intersecting street on a one-time basis, but he also must ensure that she keep her attention there for as long as traffic coming from the right could become relevant. This is why RK holds his pointing after l. 12 and up until l. 14: EA changed her gaze orientation, but that is only part of what is required in terms of attention. In sustaining the pointing gesture, RK indicates the continued relevance of the pointings' target and accordingly calls for continued attention towards it. With regard to the activity they are engaged in, the gestural cue creates a pragmatic background (Enfield, 2009, p. 114) against which the separate instructional steps to control the vehicle are to be interpreted.[9]

This extract is representative of a group of instructions (n=11) that explicitly call for the recipient's attention and are thus protoypical sites for pointing gestures. In contrast to what previously discussed literature suggests (Clark, 2005; Sidnell, 2005; Stukenbrock, 2015), a gaze change to the target or verbal display of recognition by the addressee only rarely leads to a resolution of the pointing

[8] Although no imperative design is used, this minimal, declarative format is common for instructions in spoken German (Deppermann, 2018a, p. 3).

[9] Special thanks to Pentti Haddington for pointing that out to me.

gesture in my collection (n=4). Instead, sustaining the pointing even after joint orientation has been muttually established is the default case. In these cases, the spatio-temporal contingencies of the local interaction impose additional demands on the participants' attention. Instead of communicating these demands through language, instructors signal the continuous relevance of attention by extending their gesture in time.

4.2 Increased attention to the implementation of a practical action

More often (n=18), instructions do not seek to direct the attention of others, but demand the implementation of a practical action. Here, too, the trajectory of the pointing gesture is closely coordinated with the progression of the addressee's response. In the following extract, which again comes from a driving lesson, TH practices reverse parking at the side of the street. As he is slowly reversing, he turns the steering wheel far to the right to guide the car closer to the sidewalk. His instructor HM regularly comments on the ongoing activity and initiates an instruction about the steering wheel in l. 02.

Excerpt 2. *FOLK_E_00416_SE_01_T_02_DF_01 / c32-c39*

```
01  HM:  ≠so. jetzt haste ≠die fünfunvierzig grad unge~≠fähr,
          PRT now you have about forty-five degree
      th-d                         >>car moves backwards~stationary>>
      th-h ≠turns steer. wheel right≠                #turns sw left->
02  HM:  •warte ma, +•jetzt ma≠chste $so zwei +$umdreh≠$ungen    zu+rück,+
          wait a moment now you´re making about two turns         back
      hm-g •side window•at steer. wheel------------------------------------>
      hm-h          +points at steer. wheel---+finger circles 2x left+,,,,+
      th-h  ------------------->≠               #turns sw left----->
      th-g                        $HMs hand-$at sw--$front window----->
03       (0.2)$(1.2)
      th-g ---->$at sw->
04  HM:  und +<halt zwei; es waren ≠schon drei; ≠.h
          and   hold on two that were already three .h
 =>   hm-h     +points at sw---------------------------->
      th-h ------------------------>≠               #turns sw right->
05       (0.5)#+(0.2)≠
 =>   hm-h ----->+lowers pointing hand->
      th-h ----------->≠
      Fig.      #2.1
06  TH:  wa$[s#: ]
          what
07  HM:    $[s#o.+]
              alright/like this
 =>   hm-h ------->+
      th-g ->$front window->
      Fig.      #2.2
```

Figure 2.1: HM points and looks at steering wheel while TH turns it back.

Figure 2.2: HM retracts pointing hand while uttering *so*.

```
08        (0.2)$(0.1)•+(0.1)
    th-g ---->$to HM->>
    hm-g ---------->•right window---------->
    hm-h              +palm vertical to front->

09  HM:  einfach nur die +räder +•grade; .h un jetzt fährste noch
         simply the wheels straight     .h and now you drive
    hm-h               -->+,,,,,,+
    hm-g                      -->•front window>>

10  HM:  n bisschen grad,
         straight for a bit
```

After the car has come to a halt (l. 01), HM tells her student to turn the steering wheel back to the left twice and complements her instruction gesturally with a pointing and circular movement of her index finger (l. 02). TH complies and turns the steering wheel to the left throughout the pause in l. 03 while looking at it. HM monitors TH's rotating movement and then suddenly intervenes to stop this ongoing action: She starts her turn with a progressivity-stopping *halt* ('hold on'), repeats the exact number of rotations required (*zwei*, 'two') and points out that TH already exceeded that number (*es waren schon drei*, 'that were already three') (l. 04). Although TH's attention is already focused on the steering wheel, she raises her hand to point at it a second time during this corrective instruction. This pointing, now, is maintained throughout TH's response (fig. 2.1): First, he stops his movement and then corrects the position of the steering wheel by turning it to the right again (l. 04). After watching this for another 0.5 seconds, HM slowly lowers her pointing hand and TH stops his manual action during this retraction (l. 05).

Regarding the function of the hold, its temporal extension again shows sensitivity to the progress of the respective response: HM sustains her gesture until the rotation has progressed to the point where it is deemed sufficient. Approval of TH's implementation is also given verbally, when she produces the sequence-closing German particle *so*[10] in coordination with retracting her hand

---

[10] As a deictic expression, *so* roughly corresponds to 'like this' and can confirm an embodied action as being 'exactly right' (Balantani, 2022). Additionally, it can serve

(l. 07, fig. 2.2). Also, TH himself treats the withdrawal of the pointing as a stopping signal for his ongoing action: Shortly after HM lowers her hand, he stops turning the steering wheel although verbally, his response has not yet been marked as sufficient. So here (l. 04-05), we can observe a reciprocal adjustment between the delivery of the instructive gesture and the delivery of the respective response. Not only the producer of the gesture, but also its adressee orients to the holding and withdrawal phase of the pointing as two separate communicative signals, with the first one indexing incompleteness and the need for continuation and the second one indexing completeness and closure.

While the turn design of the instruction shows sensitivity to what is already part of the common ground and what not, this cannot be said for the pointing accompanying it: Producing a second pointing gesture with the same target as before (l. 02) while both participants are clearly orienting to it already seems to promote conversational redundancy. Moreover, research on pointing gestures that re-introduce already known referents in conversation (Enfield et al., 2007) has shown that usually, subsequent pointing gestures are produced in a shorter and more reduced way than their initial counterparts. Here, however, the gesture is produced in a more elaborate way than before. Using a pointing gesture this way in this sequential position can therefore be considered marked. I argue that, instead of supporting referent identification, the upgraded format of the pointing gesture serves to display renewed, focused attention to a priorily established task under scrutiny: Since the first implementation of the instructed action did not meet her expectations, HM now shows increased attention towards the exact execution of the response. This additional quality of sustained pointing is strongly present in my collection: Often, sustained pointing gestures are used with intervening instructions that convey a higher attention towards the unfolding of the implementing action. Participants either correct an on-going practical action or adjust and specify the way it should to be done (n=13).

4.3 Combining temporal and attentional displays

Two functional dimensions of sustaining a pointing gesture have been identified so far: It serves to signal the continued relevance of the addresee's response and to display increased attention to its execution when it is in need of revision. That these are, in fact, separate dimensions that can overlap and yet be manifested in the production of the gesture will be demonstrated in the following extract. It is also representative of cases in which the initiated sequence unfolds in a dispreferred way so that pursuing response becomes necessary. The interaction in Excerpt 3 takes place during a theatre rehearsal. Throughout the

---

as a discourse marker that displays the closure of a sequence and transition to another (Barske & Golato, 2010).

session, the director PH combined pointing gestures with instructions several times to direct the actress's (SF) movements on stage or suggest changes to her performance. SF plays a scene in which she has to move through the audience and climb the stairs in front of the stage. The extract starts with her character ending a phone call (l. 02-03) while still standing at the bottom of the stairs and facing the audience, where PS is seated.

Excerpt 3. *FOLK_E_00266_SE_01_T_01_DF_01 / c929-c942*

```
01  SF:  NEIN ihr werdet auch nicht meinen exmann anrufen; der hat mit
          no you will not call my ex-husband either       he no longer

02        meinem laden (.) nichts mehr +zu tun; verstan≠den,+
          has anything to do with my shop understood
      sf-b   >>in front of stage towards audience-------≠turns around->
=>    ps-h                               +....................+

03        +•(0.4)          +#(0.5)≠(0.3)
      sf-b --------------------->≠foot on first step->
=>    ps-h +points at stage+holds-------------------->
      ps-g >>to SF•looks down----------------------->
      Fig.                 #3.1
```
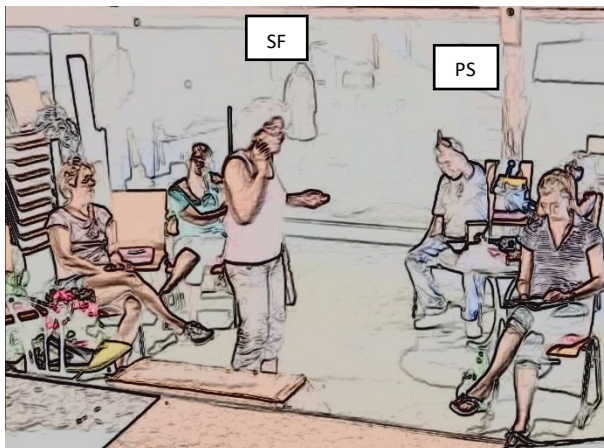

Figure 3.3: SF turns to stage as PS points in its direction.

Figure 3.2: PS looks at SF and points (shifted downwards) at her as he corrects her timing.

```
04  SF:  wehe die ru≠•fen meinen e•x an,
          woe betide them if they call my ex(-husband)
      sf-b ---------->≠climbs stairs------>
=>    ps-g ---------->•looks right-•to SF->

05        +(0.2)+
=>    ps-h +shifts pointing arm down+

06  PS:  +erst oben;
          not until (you are) on top
=>    ps-h +points at stage->
      Fig.    #3.2

07        ≠(1.0)          ≠(0.3)
      sf-b ->≠steps on stage≠step 2 on stage->

08  PS:  weh.
          woe
```

14

```
09        (0.4)≠
     sf-b  --->≠
10  PS:  ≠j[a,+]
            yes
11  SF:   [ w+]ehe die rufen mein+en ≠ex a•n,
             woe betide them if they call my ex(-husband)
     sf-b  ≠turns around to audience----≠
=>   ps-h  ---->+,,,,,,,,,,,,,,,,,,,+
     ps-g  --------------------------------->•to right->
12        (0.6)$•(0.2)
     sf-g        $inspects stage around her>>
     ps-g  ----->•to SF>>
13  SF:  donner≠wetter, (0.4)≠(0.3) nicht schlecht;
          gosh                      not bad
     sf-b       ≠hands on hips≠spins around>>
```

When PS raises his arm in preparation of a pointing (l. 02), this projects an upcoming directive action: Accordingly, SF anticipates that she needs to start moving and turns to the stage at the end of her script line (l. 02). The pointing gesture (fig. 3.1) is subsequently held throughout her climbing of the stairs until l. 11, which again reflects the duration of the expected responsive action. But unlike in the extracts before, the response is not delivered in a trouble-free manner: When SF climbs the first steps, she delivers the next line of the script ('woe betide them if they call my ex(-husband)', l. 04). This prompts PS to look up and give an instruction that corrects the exact timing of this line: SF should only say it when she has reached the top (l. 06). With this *on-the-fly* instruction (Krug & Schmidt, 2020, p. 267), he clarifies the spatial precondition that must be met in order for SF to proceed with her script. Similar to the formally upgraded pointing gesture in Excerpt 2, the corrective quality of the instruction is made visible through an additional, gestural cue: PS shifts his outstretched arm downwards (l. 05, fig. 3.2), so that the position of the gesture's apex is now different. Since this shift coincides with his correcting turn, I claim that it kinetically introduces an additional quality to the hold phase: While it was originally produced to mirror the progression of the instructed action, as in Excerpt 1, it now displays the gesturer's increased attention towards the revised implementation of it, as in Excerpt 2. Accordingly, PS closely monitors the spatio-temporal unfolding of SF's actions from this point on.

When the actress has finally reached the stage, continuing the script becomes relevant again, but she does not comply immediately. PS needs to pursue response: He provides a script prompt (l. 08: *weh*, 'woe') and, after another lack of uptake (l. 09), seeks a display of understanding (l. 10), when she finally turns around to the audience and recites the pending script line (l. 11). PS maintains his pointing gesture until SF says 'woe', when it becomes apparent that the script will now be performed in the right way, namely aligned with the spatial requirements of the scene. So in this extract, we can see how a slight adjustment of the gestural hold is used to display two orientations of the instructor at once. First, the hold indicates the temporal extent of the instructed action: 'Moving up

stage' is not a singular responsive event such as a gaze change, but rather a bodily action that extends over a longer period of time. Second, the timing of SF's script-body-coordination must be corrected: Now, the focal point of orientation for the gesture's completion is not only her movement, but also its temporal coordination with script progression.

In the context of instructions (n=29), sustained pointing gestures contribute to the local interaction in two main ways. Firstly, they provide a means of making visible the expected temporality of the instructed action. The gestural hold thus signals to the adressees to continue with what they are doing. The interactional reality of this is supported by cases in which the addressee also orients to the gesture's retraction as a stop signal (Ex. 1 and 2). Secondly, sustained pointing gestures are used to display close and renewed attention to the adressee's unfolding response. This upgraded attentional display through gesture is used with corrective or other secondary instructions that ask the adressee to revise their first attempt to perform an action (Ex. 2 and 3).

5. Questions with Sustained Pointing Gestures

The following section demonstrates the main uses of sustained pointing gestures during question-answer sequences. Similar to what was observed for instructions, speakers rarely orient to the recipient's gaze behaviour alone when sustaining their pointing gesture beyond a question (n=1 from 26 cases examined). Instead, the gesture is predominantly extended until further features of the recipient's responsive behaviour become transparent.

5.1 Continued relevancy of a selected addressee's turn

The following extract comes from a training course for first aid responders. Three trainees (PA and her assistants A5 and A6) are practising the treatment of a patient in an emergency situation. Shortly before the start of the transcript, PA, assuming the role of team leader, was asked to report the most important facts about the patient's condition to EP. Meanwhile, A5 and A6 are engaged with treating the patient.

Excerpt 4. *FOLK_E_00135_SE_01_T_01_DF_01 / c376-c384*

```
01  EP:  ja gut okay wie waren die pa[rame•ter,]
          well good okay what were the parameters

02  PA:                               [ hm_•hm,]
                                        uh huh
     pa-g                     >>looks to EP•to A6->

03        (0.3)

04  PA:  .h also sonst hat se eigentlich äh stabile werde[:, ]
              so apart from that she actually has uh stable values

05  A5:                                             [°wo]
                                                     where
         ~[sin denn  je]$[tz die na~deln;°]
          have the needles gone now

06  PA:  ~[äh ne freque]$[nz von hund~ert-][äh äh •$sättigung war~]
          uh frequency    of hundred         uhuh saturation was

07  A6:                  $[druck (.)  s~ech][zig   •$zu achtzisch,~]
                          pressure (.) sixty to eighty
     a5-b ~leans towards A6-----------~reorients to patient------~
     a6-g        >>to A5$to PA--------------------$to A5-------->
     pa-g -------------------------------------->•to EP--------->

08  PA:  ~≠neun[~zehn      ] ≠•neunund≠neunzig-
           nineteen           ninety-nine

09  A5:       [~°(gib mal)°]
                give (it)
     a5-b ~turns to A6~reaches for object in front of A6------------->
     a6-b  ≠leans towards A5---≠       ≠takes object, gives it to A5->
     pa-g -------------------->•to A6------------------------------->
10        ≠(0.52)
     a5-b #takes object from A6-->

11  PA:  un:d der druck +wa≠~r nothelferin_sechs;≠
          and the pressure was assistant six
     a5-b ------------------>~reorients to patient>>
     a6-b ---------------->≠,,,,,,,,,,,,,,,,,,,,,≠
=>   pa-h              +points to A6------------>

12  A6:  j#a (.) ~$sechzisch ach+tzi$ch [(rum/so);]
          yes (.) sixty eighty (roughly)

13  PA:                             [   sechz]+ig zu achtzig;
                                         sixty to eighty
     a5-b ------->~turns to floor>>
     a6-g -------->$to PA-----------$down>>
=>   pa-h -------------------->+,,,,,,,,,,,,,,,,,,,+
     Fig. #4

14  PA:  infu•sion simma grad dabei,
          infusion we are about to do right now
     pa-g --->•to patient>>
```

Figure 4: PA points at A6 who starts her responding turn with *ja* while she is still oriented to helping A5.

While PA lists several physical values of the patient (l. 04-08), A6 anticipates that she needs to contribute medical information to the ongoing reporting activity. In line 07, she temporarily disengages from aiding A5 and instead, directed at PA, gives the pressure value she obtained from the patient's earlier examination (l. 07: 'pressure (.) sixty to eighty'). After that, A6 attends to A5 again to help her recover something from the floor between them (l. 07-11). PA, who in parallel continues with her report without mentioning blood pressure (l. 06, 08), starts a turn that renders the value relevant again and invites a collaborative completion by A6: 'and the pressure was assistant six' (l. 11). This *fill-in-the-blank question* (Persson, 2017) is strategically cut off after the verb *war* ('was') and calls for the exact medical value as the missing complement. It is accompanied with a pointing gesture directed at A6, who is thus selected both verbally and gesturally as addressee of the question. Since A6 as target of the deictic gesture shows no referential link to the question, but a pragmatic one,[11] this pointing operates at the level of interactional organization rather than at the level of referential meaning-making. In my collection, these metacommunicative usages of pointing gestures[12] to select next speakers (Lerner, 2003) were found exclusively during questions.

PA holds her pointing gesture beyond the end of her question and into A6's answer in l. 12. A6, currently still involved with aiding her colleague and now being recruited for the reporting activity, first acknowledges her active recipiency with a turn-initial *ja* ('yes', fig. 4). Only then does she turn back to PA to deliver the requested value. Looking at the trajectory of the pointing, it is noticeable that the mere display of recipiency by A6 is not deemed sufficient for the gesture to be withdrawn. Only when the answer has progressed to the point where the

[11] Namely: Who has epistemic access to the information in question?

[12] In their study on embodied resources for selecting next speakers in multi-party settings, Blythe et al. (2018, pp. 160) mention a similar case of interpersonal pointing with a stick.

essential content of the requested information becomes transparent – the first value is produced in full and the second in part[13] – does PA retract her pointing (l. 13). The gestural hold reflects that the response is not yet considered sufficient and still needs to progress further from the point of view of the questioner. As soon as PA identifies the information she is seeking, she displays her understanding by repeating the combined value and continues further on the reporting activity (l. 14). In addition, the retraction of the held pointing gesture is again treated as a completion signal by the responding interlocutor, as in Excerpt 2: A6 shifts her gaze back to the patient as soon as the pointing hand withdraws.

5.2 Increased attention to the unfolding of the response

Asking a question provides an opportunity to pause the on-going line of action or talk in order to clarify epistemic preconditions before progressing further. This progressivity-stopping work is visible in cases of metacommunicative pointing as discussed above,[14] but also when participants ask for information about the target of the pointing gesture. Excerpt 5 shows an interaction in a perfume shop where the salesperson NR recommends different fragrances to his clients AR and CV. After having tested various perfume samples, the couple discusses their individual preferences. AR confirms with her partner CV that he likes the first fragrance (l. 01-04). After that, NR initiates a clarifying question that is accompanied by a sustained pointing gesture (l. 06-08).

Excerpt 5. *FOLK_E_00425_SE_01_T_01 / c961 - T_02 / c9*

```
01  AR:   hm aber w w [(.) äh dir dir gefällt i   ]m momen[t der h]ier;=ne,
            hm but w w   (.) uh you like this one at the moment don´t you

02  CV:               [ja ich (.) also es (meinste)]        [jaja;  ]
                       yes i (.) well it (you think)         yes yes

03  CV:   ge[nau der mit der $lime][tte $da] ja;
            exactly the one with the lime there yes

04  AR:     [glaub ich der $erste;][hm_$hm,]
              i think the first one   uh huh
      cv-g  >>looks at table$to AR------$down at table->

05         •(0.2)
      nr-g •looks up to CV->


06  NR:   f+ür $für [dich-]
```

---

[13] Experienced first aid responders can rely on candidate values that are frequently used to refer to blood pressure (e.g. 60, 80 or 85). Accordingly, PA can deduce from the uttered *acht* that the target value will be *achtzig* ('eighty') instead of *fünfundachtzig* ('eighty-five') for example.

[14] There, the reporting activity can only be continued once the blood pressure value has been retrieved.

```
                    for for   [you  ]
07  CV:             [ja;  ] $%nee [&#für] [$%si•+e,  ]
                    [yes  ]  no  [for  ] [her       ]
08  AR:                        [&#nee] [$%für m•+i]&ch-%#(0.3)+
                               [no  ] [for me     ]
=>  nr-h  +points to CV------------------------+,,,,,,,,,,,,,,,+
    nr-g  ----------------------------------->•to AR->
    cv-g  ---->$to NR-------$to AR--------$to NR>>
    cv-h                    %points to AR--%,,,,,,,,,,,,,%
    ar_h                         &points at herself &
    Fig.                         #5.1                    #5.2
```
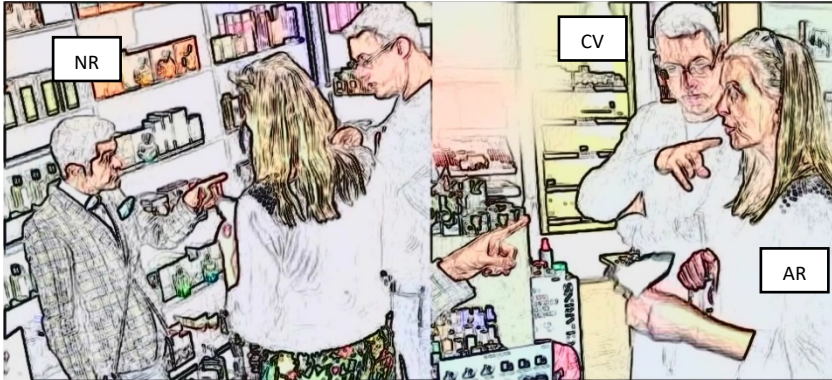

Figure 5.4: NR points and looks at CV during his response.
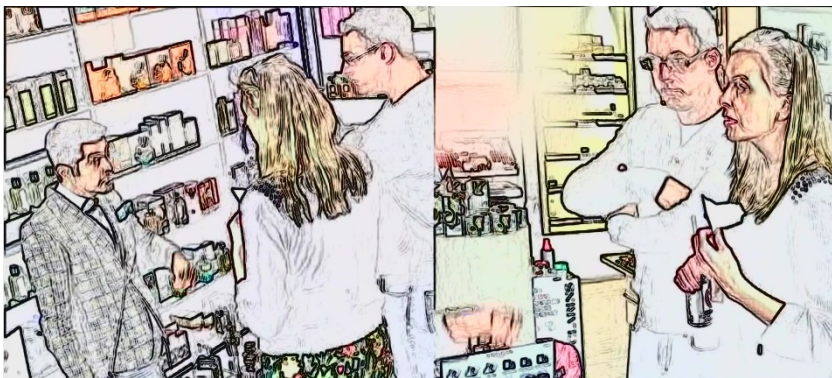
Figure 5.5: NR releases his pointing and reorients to AR.

```
09  NR:  oh; •der e (.) •≠äh der ↑erst[e?]≠
         oh the f (.) uh the first one

10  AR:                                  [ H]≠M_hm,
                                           uh huh
    nr-g ---->•to CV-----•at sample tissue->
    nr-b                 ≠takes sample tissue≠

11  NR:  ≠hm,• (0.7)≠
    nr-g --->•
    nr-b ≠smells sample≠

12  NR:  ja das is so_n ((makes a face, 0.6s.)) ((tongue click))
         yeah this is one such

13       (0.7)

14  AR:  aber [ich hab (den)]
         but I have (it)

15  NR:       [persÖnlich   ] würd ich den (0.5) eher für dich empfehlen-
              personally      I would (0.5) rather recommend this one to you
```

```
16  NR:   nicht für sie.
           not to her
```

After both CV and AR have agreed on the fragrance's high rating (l. 01-04), this projects that a concluding evaluation of the seller about his customers' choice is due next. But up until this point, neither CV nor AR has stated clearly for whom the perfume – a unisex product – is deemed most fitting, so NR needs to check his understanding at this crucial point in time. He initiates repair due to a problem of reference (Couper-Kuhlen & Selting, 2018, pp. 162) with the prepositional phrase *für dich* ('for you', l. 06) while looking and pointing at CV. Since this phrase syntactically extends AR's prior turn (l. 01) and provides a candidate understanding, it constitutes an *appendor question* (Sacks, 1992, p. 652; see also Lerner 2004). The question invites a simple confirming response and thereby indexes a high epistemic status k+ (Heritage & Raymond, 2012), which corresponds to NR's social role as a knowledgeable professional perfumer. His pointing gesture extends beyond the responsive space following his question (fig. 5.1). CV begins with a disconfirming *no* followed by a correction: Mirroring the multimodal format of NRs turn, he says *für sie* ('for her', l. 07) while pointing at AR. She herself gives a nearly identical answer ('no for me', l. 08) in combination with a short pointing at her chest as well. Again, the disconfirming token *no*, projecting a departure from a minimal response, does not suffice to withdraw the pointing gesture yet. Only when the corrected personal reference becomes clear – 'for her' is produced in full, 'for me' in part – does NR retract his gesture (fig. 5.2) and mark the receipt of new information with a change-of-state token (*oh*, l. 09) (Heritage, 1984).

The negotiation of reference is inherently coupled with the trajectory of the extended pointing gesture: It persists until the person for whom the fit of the product is assessed is intersubjectively identified. After that, NR displays surprise by repeating the fragrance in question with a jump in pitch on ↑*erste?* and high-rising intonation (l. 09). The assessment of his clients clearly contradicts his expectation, so instead of giving his approval, he treats it as being in need of further negotiation. He smells the fragrance again (l. 10-11), re-engaging with the product sensorially to warrant his upcoming assessment (Mondada, 2021), and argues that the perfume would be the better choice for CV instead (l. 15-16).

The sustained pointing gesture again signals that the answer in progress hasn't reached the focal information yet. Using a deictic gesture in this position can be seen as an overexploitation of referential resources: NR clearly already disambiguates the indexical pronoun ('for you', l. 06) by facing CV. I suggest that by mobilizing more communicative signals than needed, NR displays his increased attention to how the response plays out in detail. That NR is indeed closely observing the answer's progression is also evident in his other embodied behaviour: As soon as CV mentions his partner AR (l. 07), NR changes his gaze

to her. This coordination of gestural withdrawal and change of gaze resembles HM's behaviour in Excerpt 2.

When asking a question, sustaining a pointing gesture can serve two basic functions that show similarities to those in the context of instructions. First of all, it displays that the answer has not yet progressed so far that it is deemed sufficient from the questioner's point of view. Even though here, unlike with instructions, the expected response mostly consists of a verbal turn, the trajectory of the gesture is again closely tied to its progression. The gesture's hold is resolved as soon as the turn has unfolded to a point where the focal information is produced (Ex. 5) or inferable (Ex. 4). Secondly, sustained pointing gestures occur when there is little or no need for a gestural cue to support referent recognition. In busy environments, where multiple lines of actions are simultaneously relevant, this can be employed to attract and steer the attention of others to a specific, focused interaction (Ex. 4). It also serves as an attentional display, for instance when the information asked is crucial for the questioner's understanding of the current course of action or interactional progression in general (Ex. 4, 5).

6. Discussion and Conclusion

The analysis has revealed that participants sustain their pointing gestures beyond a first, response-inviting turn to attend to demands that arise out of the locally situated interaction rather than to production-related concerns about gesture-speech coordination. Although pointing gestures may initially be produced to facilitate the recognition of a target in space, their functional potential becomes much richer when they are extended in time. This functional shift becomes evident when examining the gesture's relation to what was assumed in previous research to be the crucial moment for a pointing's retraction: The establishment of joint attention to the target. Across the whole collection, only 5 cases (4 instructions and 1 question) were observed in which a gaze shift or verbal display of recognition by the addressee led to a withdrawal of sustained pointing. This strongly supports the claim that prolonged pointing gestures contribute more to interaction than supporting reference identification. It has been stated that "[s]ome indications locate things primarily in space and others primarily in time" (Clark, 1996, p. 165). In the case of sustained pointing gestures, these two dimensions of locating actually seem to converge and create a functionally rich communicative resource.

By maintaining their pointing gesture across a TRP and into the responsive space, participants display their orientation to how the response is carried out in detail. This holds for responses that take the form of practical action (after instructions) as well as the form of verbal turns (after questions). Since gestural

withdrawal is a well-established signal of completion (Mondada, 2007; Sidnell, 2005; Sikveland & Ogden, 2012; Stukenbrock, 2015), delaying this gestural closure gains a communicative quality in itself: Holding a gesture indicates that something is still in progress and therefore does not yet justify closure (of a multimodal contribution, of a sequence, of a joint project). The interactional reality of this mechanism is evident through cases where responding interlocutors treat the retraction of a held pointing gesture as a signal to stop with what they are doing (Ex. 2) or as permission to disengage from the focused interaction with the gesturer (Ex. 4).

Participants hold certain expectations with regard to how the addressee's response is to be carried out. In the case of instructions, the temporal extension of the gesture reflects the expected temporality of the instructed action itself. In the data analysed here, instructions mainly call for an embodied response that is not completed by a singular responsive event, but naturally extends over a longer period of time. Sustaining a gesture provides for a publicly visible tool to signal these temporal requirements without having to resort to speech. This division of labour across communicative resources allows participants to simultaneously meet other demands through verbal means, such as giving multiple subsequent instructional steps (Ex.1). For questions, the temporal extent of the pointing gesture reflects the progress of the answer to the point when the focal information can be obtained (Ex. 4 and 5). Questioners closely monitor the development of the responsive turn and mark it as not yet informative by maintaining the gesture associated with their initial question. Further research on other sequence-initiating actions[15] is needed to further test the robustness of this basic response-oriented function of gestural holds.

Furthermore, sustained pointing gestures are often used in contexts where, from an economy point of view, there is really no need for an additional gestural cue to support referent recognition. In these cases, the respective target and referent are already part of the common ground prior to the pointing and thus salient for all interlocutors (Ex. 2, 3 and 5). Producing a pointing gesture in this position and in this extended manner can be seen as a communicative upgrade whose function is to visibly renew joint attention rather than to refer to something deictically. Held pointing gestures serve as a spatially and temporally stable point of orientation that recipients can notice even in highly complex and dynamically changing environments. This attention-related function is visible in both instruction and question contexts: Often, the attention resources of the addressee are occupied by other, simultaneously relevant tasks or lines of actions (e.g. operating a car in Ex. 1, performing a scene in Ex. 3, treating a patient in Ex. 4). In multi-party contexts, where the participation framework may

---

[15] Since assessments, suggestions and offers were also observed in an earlier stage of this study, these first actions are promising candidates for follow-up studies.

frequently change, a strong gestural cue can also serve to temporarily establish a dyadically focused interaction between gesturer and adressee (Ex. 4).

Even though different social dimensions are of primary importance for instructions and questions – instructions relying more on deontics and questions more on epistemics – sustaining a gesture into and across the responsive space that follows them can achieve similar things. Instructors as well as questioners make use of the temporal extent of their gesture to signal their expectations for the next step of the sequence. This encompasses temporal as well as qualitative characteristics of the response, regardless of whether it is done by embodied or verbal means. Sustaining a gesture introduces an upgraded gestural cue that is designed to not only attract attention, but also to show increased attention to how the response is playing out. In general, moments of disruption seem to attract the use of held pointings, which resonates well with previous findings on held iconic gestures during negotiations of joint understanding (Sikveland & Ogden, 2012).

This study provides a first systematic account of gestural holds as an interactional ressource. It builds on earlier work (see chapter 2.2) which suggested that holds indicate continuation, but goes well beyond the scope of these sporadic observations: Analyzing a broad collection of 55 cases showed how and for what action-specific purposes participants use gestural holds in naturally occurring conversation. Cross-action comparisons also revealed functional commonalities and thus laid the foundation for a more general description of how gestural holds are used. Also, the methodological approach taken contributes to a more comprehensive understanding of the phenomenon: According to the conversation analytical perspective, which emphasises the situatedness of action and the temporal unfolding of interaction, sustained pointing gestures are to be understood as a communicative resource anchored within the local activity the participants are engaged in. Taking into account the interactional embeddedness of sustained pointings requires detailed sequential analyses such as those carried out here.

Gesture as a modality can be employed not only to express iconic, deictic or pragmatic meaning, but also to create an overarching layer of signalling that underlies different temporal constraints than speech. Participants utilize this affordance of gesture to display orientations and expectations that transcend the boundaries of their own turns. Paying attention to details of gesture production – in particular, the temporal extent of a gesture – can therefore contribute to a better understanding of social interaction itself.

Acknowledgements

References

Andrén, M. (2011). The organization of children's pointing stroke endpoints. In G. Stam Gale & M. Ishino (Eds.), *Integrating Gestures: The interdisciplinary nature of Gesture* (pp. 153–162). John Benjamin Publishing.

Balantani, A. (2022). Coordinating individual actions in joint activities: The case of the German deictic so. *Lingua, 277.* https://doi.org/10.1016/j.lingua.2022.103402

Barske, T., & Golato, A. (2010). German so: managing sequence and action. *Text & Talk*, *30*(3), 245–266. https://doi.org/10.1515/text.2010.013

Blythe, J., Gardner, R., Mushin, I., & Stirling, L. (2018). Tools of Engagement: Selecting a Next Speaker in Australian Aboriginal Multiparty Conversations. *ROLSI, 51*(2), 145–170. https://doi.org/10.1080/08351813.2018.1449441

Bühler, K. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Gustav Fischer Verlag.

Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9–33). Erlbaum.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 243–268). Erlbaum.

Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, *7*(4–5), 507–525. https://doi.org/10.1177/1461445605054404

Clark, H. H. (2012). Wordless questions, wordless answers. In J. de Ruiter (Ed.), *Questions: formal, functional and interactional perspectives* (pp. 81–100). Cambridge University Press.

Cooperrider, K. (2016). The Co-Organization of Demonstratives and Pointing Gestures. *Discourse Processes, 53*(8), 632–656. https://doi.org/10.1080/0163853X.2015.1094280

Cooperrider, K., Fenlon, J., Keane, J., Brentari, D., & Goldin-Meadow, S. (2021). How Pointing is Integrated into Language: Evidence From Speakers and Signers. *Frontiers in Communication* 6:567774. https://doi.org/10.3389/fcomm.2021.567774.

Couper-Kuhlen, E., & Selting, M. (2018). *Interactional Linguistics: Studying Language in Social Interaction.* Cambridge University Press

Curl, T. S. & Drew, P. (2008). Contingency and Action: A Comparison of Two Forms of Requesting. *ROLSI, 41*(2), 129–153.

Deppermann, A. (2008). *Gespräche analysieren: Eine Einführung.* Verlag für Sozialwissenschaften.

Deppermann, A. (2018a). Instruction practices in German driving lessons: Differential uses of declaratives and imperatives. *International Journal of Applied Linguistics*, *28*(2), 265–282.

Deppermann, A. (2018b). Sprache in der multimodalen Interaktion. In A. Deppermann & S. Reineke (Eds.), *Sprache im kommunikativen, interaktiven und kulturellen Kontext* (pp. 51–86). De Gruyter.

Deppermann, A. (2021). Imperative im Deutschen: Konstruktionen, Praktiken oder social action formats? In B. Weidner, K. König, W. Imo, & L. Wegner (Eds.), *Verfestigungen in der Interaktion. Konstruktionen, sequenzielle Muster, kommunikative Gattungen* (pp. 195–229). De Gruyter.

Deppermann, A. & Streeck, J. (2018). The body in interaction: Its multiple modalities and temporalities. In A. Deppermann & J. Streeck (Eds.), *Time in Embodied Interaction: Synchronicity and Sequentiality of Multimodal Resources* (pp. 1–29). John Benjamins.

De Stefani, E. (2014). Rearranging (in) space: On mobility and its relevance for the study of face-to-face interaction. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in Language and Linguistics: Geographical, Interactional, and Cognitive Perspectives* (pp. 434–463). De Gruyter. https://doi.org/10.1515/9783110312027.434

De Ruiter, J. P. (Ed.) (2012). *Questions: formal, functional and interactional perspectives.* Cambridge University Press.

Drew, P. & Couper-Kuhlen, E. (Eds.) (2014). *Requesting in Social Interaction*. John Benjamins.

Enfield, N. J. (2009). *The Anatomy of Meaning: Speech, Gesture, and Composite Utterances* (Language Culture and Cognition). Cambridge University Press. https://doi.org/10.1017/CBO9780511576737

Enfield, N. J., Kita, S., & de Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *JoP, 39*(10), 1722–1741. https://doi.org/10.1016/j.pragma.2007.03.001.

Erikson, M. (2009). Referring as interaction: On the interplay between linguistic and bodily practices. *JoP, 41*, 240–262.

Floyd, S., Rossi, G., & Enfield, N. J. (Eds.) (2020). *Getting others to do things. A pragmatic typology of recruitments.* Language Science Press.

Ford, C. E., & Thompson, S. A. (1996). Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 134–184). Cambridge University Press.

Goodwin, C. (1981a). Exophoric Reference as an Interactive Resource. In J. N. Deely & M. D. Lenhart (Eds.), *Semiotics* (pp. 119–128.). Plenum Press.

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

Gubina, A. (2021). Availability, Grammar, and Action Formation: On Simple and Modal Interrogative Request Formats in Spoken German. In O. Ehmer, H. Helmer, F. Oloff, & S. Reineke (Eds.), *'How to get things done' - Aufforderungen und Instruktionen in der multimodalen Interaktion* [Special Issue]. Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion, *22*, 272–303.

Haddington, P., Keisanen, T., Mondada, L., & Nevile, M. (2014) (Eds.). *Multiactivity in Social Interaction: Beyond Multitasking.* John Benjamins.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J.M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 299–345). Cambridge University Press.

Heritage, J. & Raymond, G. (2012). Navigating epistemic landscapes: acquiesence, agency and resistance in responses to polar questions. In J. P. de Ruiter (Ed), *Questions: Formal, Functional and Interactional Perspectives* (pp. 179–192). Cambridge University Press.

Hindmarsh, J. & Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *JoP, 32*(12), 1855–1878.

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

Kendon, A. (1980). Gesticulation and Speech: Two Aspects of the Process of Utterance. In M. R. Key (Ed.), *The Relationship of Verbal and Nonverbal Communication* (pp. 207–228). De Gruyter. https://doi.org/10.1515/9783110813098.207

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press.

Kita, S. (Ed.) (2003). *Pointing: Where language, culture and cognition meet*. Erlbaum.

Kita, S., Gijn, I.v., & Hulst, H.v.d. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction: international gesture workshop Bielefeld, Germany, September 17–19, 1997: proceedings* (pp. 23–35). Springer.

Krug, M. & Schmidt, A. (2020). Zwischenfazit: Sukzessive und simultane Verzahnung von Spiel- und Besprechungsaktivitäten – eine Instruktionsmatrix für Proben. *Gesprächsforschung Online*, *21*, 264–277.

Labov, W., & Fanshel, D. (1977). *Therapeutic discourse: Psychotherapy as conversation*. Academic Press.

Lerner, G. H. (2003). Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society, 32*(2), 177–201.

Lerner, G. H. (2004). On the place of linguistic resources in the organization of talk-in-interaction: Grammar as action in prompting a speaker to elaborate. *Research on Language and Social Interaction*, 37(2), 151–185.

Lindwall, O., Lymer, G., Greiffenhagen, C. (2015). The sequential analysis of instruction. In N. Markee (Ed.), *The handbook of classroom discourse and interaction* (pp. 142–157). Wiley.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

McNeill, D., Cassell, J., & Levy, E. T. (1993). Abstract deixis. *Semiotica, 95*(1–2), 5–19.

Mondada, L. (2007). Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies, 9*(2), 194–225.

Mondada, L. (2013). Conversation analysis: Talk and bodily resources for the organization of social interaction. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & S. Tessendorf (Eds*.), Body - Language - Communication: An international Handbook on Multimodality in Human Interaction (Vol. 1)* (pp. 218–226). De Gruyter.

Mondada, L. (2015). Multimodal completions. In A. Deppermann & S. Günthner (Eds.)*, Temporality in Interaction* (pp. 267–308.). John Benjamins.

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics, 20*(3), 336–366.

Mondada, L. (2019). *Conventions for multimodal transcription.* https://www.lorenzamondada.net/multimodal-transcription.

Mondada, L. (2021). *Sensing in Social Interaction*. Cambridge University Press.

Persson, R. (2017). Fill-in-the-Blank Questions in Interaction: Incomplete Utterances as a Resource for Doing Inquiries. *ROLSI, 50*(3), 227–248. https://doi.org/10.1080/08351813.2017.1340698

Raymond, G. (2003). Grammar and social organization: yes/no interrogatives and the structure of responding. *American Sociological Review*, *68*(6), 939–967.

Rossi, G. (2014). When do people not use language to make requests? In P. Drew & E. Couper-Kuhlen (Eds.), *Requesting in Social Interaction* (pp. 303–334). John Benjamins.

Sacks, H. (1992). *Lectures on Conversation. Volume II*. Edited by Gail Jefferson. Blackwell Publishing.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation*. Language, 50*(4/1), 696–735.

Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 266–295). Cambridge University Press.

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

Schmidt, T. (2016). Good practices in the compilation of FOLK, the Research and Teaching Corpus of Spoken German. In: J. M. Kirk & G. Andersen (Eds.), *Compilation, transcription, markup and annotation of spoken*

corpora. Special Issue of the International Journal of Corpus Linguistics [IJCL 21:3], (pp. 396–418).

Selting, M. (2000). The construction of units in conversational talk. *Language in Society,* 29(4), 477–517.

Seyfeddinipur, M. (2006). *Disfluency: Interrupting speech and gesture.* [Doctoral dissertation, Radboud University Nijmegen].

Sidnell, J. (2005). Gesture in the pursuit and display of recognition: A Caribbean case study. *Semiotica, 156*, 55–87.

Sidnell, J. (2013). Basic conversation analytic methods, In J. Sidnell & T. Stivers (Eds.)*, The Handbook of Conversation Analysis* (pp. 77–99). Wiley-Blackwell.

Sikveland, R.O., & Ogden, R. A. (2012). Holding gestures across turns: Moments to generate shared understanding. *Gesture, 12*(2), 166–199.

Steensig, J., & Heinemann, T. (2013). When "yes" is not enough – as an answer to a yes/no question. In B. Szczepek Reed & G. Raymond (Eds.), *Units of Talk – Units of Action* (pp. 207–242). John Benjamins.

Stevanovic, M. & Peräkylä, A. (2014). Three orders in the organization of human action: On the interface between knowledge, power, and emotion in interaction and social relations. *Language in Society, 43*(2), 185–207.

Stivers, T. (2010). An overview of the question–response system in American English conversation, *JoP, 42*(10), 2772–2781.

Stivers, T., & Hayashi, M. (2010). Transformative answers: One way to resist a question's constraints. *Language in Society, 39*, 1–25. doi:10.1017/S0047404509990637.

Streeck, J. (2017). *Self-Making Man: A Day of Action, Life, and Language.* Cambridge University Press.

Stukenbrock, A. (2015). *Deixis in der face-to-face-Interaktion.* De Gruyter.

ter Bekke, M., Drijvers, L., & Holler, J. (2020). The predictive potential of hand gestures during conversation: An investigation of the timing of gestures in relation to speech. In *Proceedings of the 7th GESPIN - Gesture and Speech in Interaction Conference. Stockholm: KTH Royal Institute of Technology.*

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A Professional Framework for Multimodality Research. In

Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation.

Zinken, J., & Deppermann, A. (2017). A cline of visible commitment in the situated design of imperative turns. Evidence from German and Polish. In M.-L. Sorjonen, L. Raevaara & E. Couper-Kuhlen (Eds.), *Imperative turns at talk: The design of directives in action* (pp. 27–63). Benjamins.