



Micro-Sequential Coordination in Early Responses

Arnulf Deppermann  and Axel Schmidt

Pragmatics-Department, Leibniz-Institut für Deutsche Sprache

ABSTRACT

Our study deals with early bodily responses to directives (requests and instructions, i.e., second pair parts [SPPs]) produced before the first pair part (FPP) is complete. We show how early bodily SPPs build on the properties of an emerging FPP. Our focus is on the successive incremental coordination of components of the FPP with components of the SPP. We show different kinds of micro-sequential relationships between FPP and SPP: successive specification of the SPP building on the resources that the FPP makes available, the readjustment or repair of the SPP in response to the emerging FPP, and reflexive micro-sequential adaptations of the FPP to an early SPP. This article contributes to our understanding of the origins of projection in interaction and of the relationship between sequentially and simultaneity in interaction. Data are video-recordings from interaction in German.

Introduction

Directives are “designed to get someone else to do something” (Goodwin, 2006, p. 515; Searle, 1976). Primary instances of directive actions are requests (Drew & Couper-Kuhlen, 2014) and instructions (Lindwall et al., 2015). Our focus is on requests for objects and actions and on instructions for actions that are to be complied with immediately. Requests and instructions and their responses are primary examples of adjacency pairs: They are ordered in the sense that the first pair part (FPP) makes the second pair part (SPP) conditionally relevant (Schegloff, 1968, 2007). The SPP is due as soon as the FPP has reached a transition relevance place (Sacks et al., 1974), that is, when it is possibly complete. Yet, as the articles in this special issue show, conditionally relevant responses can come early, that is, before the FPP is possibly complete.

Verbal early responses provide evidence that participants do not always avoid overlap in producing a projected next action in interaction, as predicted by the well-known model of turn-taking by Sacks et al. (1974). Instead, sometimes respondents orient to producing a response as soon as it is recognizable (see Jefferson, 1973) for them which type of action is made relevant by a turn that is still *in fieri*. Early response onset is particularly common in bodily, nonverbal responses to requests and instructions that, by virtue of their embodied nature, do not make for a turn that overlaps the FPP.

But how are participants in an interaction able to produce an early response and, moreover, to do this in a way which is astonishingly seldom subject to misunderstanding and failure? The key to an answer to this question lies in the mechanism of projection (Auer, 2005; Streeck, 1995; Streeck & Jordan, 2009): In order to understand how early responses can be produced, we must inquire into the sources of projection that participants use to anticipate which action is due next. Sources of projection are located on different temporal levels:

- (Shared) interactional histories and social learning make action sequences and conditionally relevant actions predictable and lead to the establishment of routines. This kind of knowledge is

CONTACT Arnulf Deppermann  deppermann@ids-mannheim.de  Institut Für Deutsche Sprache, R5, 6-13Mannheim 68161, Germany.

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

considered as “personal common ground” (Clark, 1996, pp. 334–336) and accumulates over time in communities of practice (Deppermann & Schmidt, *in press*; Hazel, 2018).

- The larger context of the interaction, for example, genre- and activity-type related expectations and series of sequences, can make whole sequences of actions as next sequences expectable. This kind of knowledge is considered as “communal common ground” (Clark, 1996, pp. 332–334). Actions such as queuing up at a market stall (Dausendschön-Gay & Krafft, 2009) or placing objects on a counter in the context of a sales interaction (Clark 2005) project typical next steps in a routine activity (such as ordering products). If next actions are highly expectable within an ongoing project, requests may be produced without using language (Rossi, 2014).
- The properties of an emerging FPP create projections, which can be used for early response generation.

Of course, these three orders are not independent from each other. Indeed, prior interactional experience is crucial for being able to immediately identify components of an emerging turn that can be used to anticipate which next action is due.

In this article, we study early bodily responses to requests and instructions. We focus on the third temporal source of projection outlined above, that is, the properties of the emerging FPP. Projection here concerns at least two interrelated issues:

- Which action will be implemented by the turn *in fieri*?
- Which next action is made conditionally relevant by the turn *in fieri*?

These two general issues are linked to a number of more specific projective properties that are incrementally made available by the turn *in fieri*, such as Who is addressed? Is an object involved in the action to be performed? Which (kind of) object will be requested? Where is the object located?

In order to study these questions empirically, we have to attend to the precise temporal relationships of the coordination between the emerging FPP and the emerging SPP. These relationships lie at the intersection between simultaneity and sequentiality: While early responses are simultaneous in the sense that they start while the FPP is still *in fieri*, they are yet micro-sequentially coordinated with the emerging FPP. Micro-sequential coordination means that the early response builds on components of the FPP that have already been produced and that create projections for the type of action underway and the next action that is due. Alternatively, projections may only concern specific properties of the actions, such as which object is involved, who should perform the next action, and which segment of the surrounding space will be relevant for the next action.

Micro-sequential relationships between an FPP and an early response are particularly complex if not only the onset of the early response is coordinated with components of the FPP under production but if there is a successive incremental coordination of several components of the emerging FPP with several components of the emerging SPP. In these cases, the SPP is not produced *en bloc*, but its internal temporal structure is adapted to the internal temporal structure of the FPP, making use of the projections emanating from the components of the FPP that have already been produced. In our article, we are specifically dealing with early responses that are micro-sequentially coordinated in this stepwise manner.

In the following sections, we first briefly review studies and concepts from prior research on which the study of early responses can build and then present our data. The main body of the article is devoted to the study of micro-sequential coordination of early bodily responses. As a background, we briefly present a contrasting case of an early response that is coordinated with the FPP only concerning its onset. We then turn to cases in which several components or phases of the SPP are coordinated with the projections emanating from components of the FPP that are successively produced. We finally analyze cases in which the process of the production of the FPP *in fieri* is affected by an early response, that is, reciprocal micro-sequential adaptations of FPP and SPP. In the Discussion, we summarize our findings and discuss their implications for our understanding of the relationship between simultaneity

and sequentiality in response generation and concerning the role and nature of projection and anticipation in social interaction.

Theoretical background

Although early responses have not become a systematic focus of research yet, there are several investigations that offer important contributions to their study. Experimental studies in psycholinguistics have found that a shared visual environment supports the fast and errorless accomplishment of joint activities (Clark & Krych, 2004; Clark & Marshall, 1981). If instructor and instructee have visual access to each other's actions and relevant spatial configurations, they can continuously monitor the current state of the task, the activity of the instructee, and the understanding of instructions given. Moreover, fast coordination of instructions and implementing actions, even in overlap, as well as modifications of speakers' utterance midway are possible (Clark & Krych, 2004; Gergle et al., 2004). Following the principle of least collaborative effort (Clark & Wilkes-Gibbs, 1986), instead of producing response tokens (such as *okay*) or formulating actions (e.g., *I put the red block on the green block*), people just carry out the instructed action, while checking if the instructor accepts their move. Pointing to or placing things facilitate the cooperation: "Coordinating and executing joint activities are not just linked. They are often achieved with the same behavior" (Clark 2005, p. 524). In the context of mutual visual access, "cross-timing" (Clark & Krych, 2004, p. 69) of verbal instructions and physical implementation is common. Psycholinguistics studies on language comprehension (e.g., Sauppe, 2016) have shown that linguistic input is incrementally integrated by hearers and that language comprehension is anticipatory in principle. Listeners use the semantic information provided by verbs to anticipate object referents before they are mentioned (Altmann & Kamide, 1999). While these studies show that speaking and listening are collaborative processes that build on incremental understanding and quasi-simultaneous feedback loops, there is no research yet on the micro-sequential organization of the verbal and embodied resources used in FPPs and SPPs that account for projections that are used to bring about early responses.

Conversation analytic studies have repeatedly observed early responses and highlighted the role of projection for their occurrence. Charles Goodwin (1979, 1981) and Marjorie Harness Goodwin (1980) have shown how simultaneous responses to an emerging turn are taken into account by the turn-producer in the emerging continuation of that turn. Jefferson (1973) has pointed out that once a turn has reached its recognition point, a fitting response may already be produced in overlap. Jefferson (1985) has shown that speaker's laughter even can be designed to elicit recipient's laughter already while the speaker's turn is still under way. Especially studies on learning and instructing (Deppermann, 2018; Koschman, 2011; Lindwall et al., 2015) as well as workplace studies (Heath & Luff, 2000) have shown that participants monitor their partners' behavior in order to anticipate next actions and to respond early. Early responses can be designed to demonstrate independent knowledge (e.g., by a student [Hindmarsh et al., 2011] or a patient [Hindmarsh, 2010] in dentistry). Early responses are one variant of possible temporal relationships between FPPs and SPPs. In several articles, Mondada (2013, 2015, 2018a, 2018b) has studied in detail different temporal relationships between FPPs and SPPs, mainly focusing on instructions and bodily instructed action as we do as well in this article. Mondada (2014) shows how surgeons' bodily adjustments prior to a request provide a resource for assistants to preempt or anticipate the instruction. Especially in routine activities (such as carrying objects or furniture together), manual actions can even start before being instructed (see also Mondada, this volume). "Thus, the verbal directive confirms this action, rather than initiating it" (Mondada, 2018a, p. 99, 2018b). In fact, in the theoretical discussion of her conventions for multi-modal transcription she points out that there may be micro-sequential relationships between emerging turns, which have to be captured by a transcription that allows representing the precise time-courses of actions and their interrelations. Early compliant actions in response to requests and instructions are cases of projected actions.

Projection is a key mechanism both for sequence organization (Schegloff, 2007) and for the grammar of turn-construction (Auer, 2005). On both levels first elements (first actions, first grammatical items) project next elements (responsive actions, required and matching next grammatical units), which allow the recipient to anticipate how a turn will (possibly or probably) develop, which action it will implement and which response is expected from the recipient. Anticipation, thus, systematically builds on projection. The sooner a respondent can anticipate what is expected from them, the sooner they are able to produce a fitting response. Levinson (2013) discusses how early projection of the action type (see also Deppermann, 2013) allows for early response generation. Certain grammatical formats of turn-construction (like clauses with a verb in first position) provide for early projection of the design of the turn, its completion and the action-type it implements, thereby also projecting early the expected response.

Methods

Our study draws on the analysis of 41 video-recorded instances of embodied responses to requests and instructions recorded between 2010 and 2019. Thirty-eight instances come from the German reference corpus of talk-in-interaction FOLK, which is hosted by the Leibniz-Institut für Deutsche Sprache (transcript-aligned video data are available online via dgd.ids-mannheim.de including additional information on meta-data of the corpora). The cases come from a diverse range of activity types, such as emergency drills, driving lessons, room redecoration, bathroom renovation, cooking, joint video games, and physiotherapy. Three additional instances come from theater rehearsals. We selected cases in which an embodied response starts early with regard to a verbal directive. Verbal transcripts are transcribed according to GAT2 (Selting et al., 2011), multimodal transcripts follow the conventions by Mondada (2018a).

Results

Relationships of micro-sequential coordination between FPPs and SPPs in early responses

In this section, we study different ways in which the time course of early responses is coordinated with the time course of a FPP. The study of these questions involves a detailed multimodal analysis of the projective potentials of the different bodily and verbal resources of the FPP on which the emerging SPP builds.

We are interested in the following issues:

- *Sources of projection*: Which components of the emerging FPP are used for the production of an early response? Which projections emanate from the individual components of the FPP and how do they inform individual components or phases of the early response?
- *Temporal ordering*: Can we observe patterns of the temporal order of verbal and bodily resources in FPPs on which patterns of the successive unfolding of early responding SPPs build?
- *Micro-sequential complexity*: Does only the onset of the early response depend on properties of the emerging FPP, or is there a more complex, successive micro-sequential coordination between components of the FPP and the emerging SPP?
- *Reciprocal coordination*: Is the temporal unfolding of the FPP itself affected by the occurrence of the early response?

We first present a case of an early response that is coordinated with the FPP *only in its onset*. The, we devote the next section to more complex cases of micro-sequential coordination in which the early response is incrementally produced, successively specified, or repaired in coordination with the projections that the emerging FPP make available step by step. Finally, we analyze cases in which the FPP is itself reflexively adapted to early response by abandoning before completion or self-repair.

Early response without micro-sequential adaptation

In contrast to the other extracts in this article, our first case represents an early response without micro-sequential adaptation of an SPP to an emerging FPP. In Extract 1 from a family breakfast, father (F) asks daughter (D) to give him the butter. D's early response is coordinated with the FPP *only in its onset*:

*Extract 1: butter/family breakfast (FOLK_E_00355_SE_01_T_01_DF_01 - 9:44-9:54)*¹

```

01          $$& (0.5)      $&(0.7)          &(0.8) $&(0.2) &(0.15) $(0.15)$

f-g  $head movements/looks for butter on table-----$

d-g   $at plate----$at bowl-----$at butter----$at plate-->

d-h   &at plate----&reaching left-&poising&bowl-&butter-->

02  F    $gibst      $du mir  $BITte $&die butter.
      give.2SG   you me.dat please the butter
      can you please give me the butter

f-g   $at D--->>

d-g   >>at plate--$at butter-$at F--$at butter--->>

d-h   touches butter-----&passes butter to F--->>

03          (0.44)

```

During the 2.5 seconds silence before his turn in line 02, F looks for the butter by moving his head slightly in different directions. D is still involved in her own project, looking down and turning from using the butter to using something else (the chocolate cream, as turns out later) without paying attention to F. She poises her hand over two items (a little bowl and the butter), ending up touching the butter, which is part of D's individual project and not yet part of the upcoming request sequence. D's actions become responsive when F asks D to give him the butter addressing her by gaze (line 02). With the following gaze shift from her plate to the butter, D starts her responding action of passing the butter to F. D complies early, passing the butter before the requested object (article and noun: *die butter*, "the butter") is mentioned. However, there is no micro-sequential coordination between FPP and SPP. D draws on gaze (address) and linguistic cues of F's emerging FPP: verb-first second-person interrogative *gibst* "give" + second-person subject pronoun *du* "you" + dative object pronoun *mir* "me" strongly project a request in German. The requested object (*the butter*) is inferred from contextual information: Immediately before F's turn starts, D has already touched the butter. She gazes at the object already after F's turn-initial item *gibst*. Before F produces the object phrase, D shifts her gaze shortly to F, when she starts to lift and pass the butter to F, without further orientation to the emerging FPP. Instead, she follows the passing of the butter with her gaze until F takes it.

While the response in Extract 1 starts early but then develops independently from the emerging FPP, the further course of other early responses adapts micro-sequentially to the course of the emerging FPP. Since SPPs are not produced en bloc in these cases, it becomes clearer on which specific resources used in the FPP the producer of a SPP relies.

Micro-sequential adaptation of SPP

In this section, we analyze cases in which early responses are incrementally produced, successively specified or repaired in coordination with the projections that the emerging FPP makes available step by step.

Successive specification of FPP and SPP

We start with a less complex case, in which the SPP producer can only rely on verbal cues, because participants have no mutual visual contact. In Extract 2 from a physiotherapy session, the patient (P), lying on a table (Figure 1), is instructed by the therapist (T), standing behind him, to stretch his left leg:

Extract 2: *stretch leg*/physiotherapy (FOLK_E_00360_SE_01_T_01_DF_01 - 14:53-15:00)

```

01          (4.14)

02  T      mal #das rech ^LINke bein          ^auch ^# ausstreck#en;
          PRT ART right left leg              also   expand.INF
          stretch the ri left leg as well

          P-P          ^contracts leg muscle ^lifts^stretches leg--->
          fig          #fig.1                      #fig.2      #fig.3

03          (0.94) #^
          P-P  -----^
          fig          #fig.4

```

Like in the case before, the response is produced early: P stretches his leg before the action (“*stretch*”) is mentioned. In contrast to Extract 1, the SPP is not simply produced en bloc. P adapts his emerging SPP, stretching of his left leg (the right leg is already stretched), incrementally to the verbal information given by the emerging FPP. When T produces *linke bein*; “left leg”, P contracts the muscles of his left leg (and of his left arm and shoulder as well). The projected linguistic format here is a deontic infinitive (Deppermann, 2006): It starts with naming the object, while the verb is only produced afterward. This object-first format in German projects an upcoming deontic action, which in this context can be expected to be a request. When T says *auch* (“also”), P begins to lift his left leg upward (Figure 2), obviously, because the adverb *auch* (“also”) in this turn-position projects that his left leg be moved into the same position that his right leg already inhabits. P finally starts to extend his leg when T says *aus* (“ex”, Figure 3), thus again anticipating the precise kind of action to be performed, before the verb *ausstrecken* (“expand”) is fully completed (Figure 4).

In Extract 2, a SPP is incrementally produced only relying on the emerging verbal formulation of the FPP (see Mondada, this volume for a similar case (Ex. 1)). In contrast to Extract 1, where reaching the butter is an action produced in one go once the requestee has recognized which action the FPP is implementing, moving the leg in Extract 2 consists of three discrete movements: contracting, lifting, and finally stretching, which are incrementally adapted to the production of an FPP. Since T has visual access to the actions of P, moving his leg is taken by P as a demonstration of understanding (Hindmarsh et al., 2011). No verbal acknowledgment by P is necessary (Clark & Krych, 2004).

Extract 3 is a more complex case of micro-sequential adaptation of the SPP to the FPP. The early response builds on both verbal and embodied properties of the emerging FPP, which are used to specify the response successively. A couple is renovating the bathroom of their home. The husband (H) asks his wife (W), who assists him, to hand him a screwdriver.



Figure 1–4. Figure 1. Body positions at the beginning; **Figure 2.** Upward move of leg; **Figure 3.** Leg stretching starts; **Figure 4.** Leg is stretched out.

Extract 3: bathroom renovation (FOLK_E_00299_SE_01_T_01_DF_01 - 2:49-2:58)

```

01  H    äh #+*kannst$ du# %$mir#* schon $ma$ $da #den$ &den# SCHRAUbenzieher+$ ge%&#ben;
      PRT  can.2SG you me.dat PRT    PRT  there the the screwdriver give.INF
      uh can you already give me there the the screwdriver

      h-b    +turns to object, gets up-----+
      h-g    *at object-----*
      h-h    $. . . . . $points object--$, , , , , , , $
      w-b    %turns to object-----%
      w-g    $. . . . $at object-----$
      w-h    &reaches for object-----&grasps--->
      fig    #fig.5    #fig.6 #fig.7    #fig.8    #fig.9    #fig.10

02    (.) dann kann ich NÄMlich (0.2) schon mal die-
      then can.1SG I PRT    already PRT the
      then I can already

03    (0.2) &(1.3)
      w-h    ----&gives object to H--->>

04  H    dinger Lösen hier;
      things loosen.INF here
      loosen those things here
  
```

H asks W to give him a screwdriver (which is lying on the window sill; see red marking in [Figure 5](#)). W already starts to reach for the object before H finishes his turn in line 01. We now look more closely at how the production of the FPP, the request for the screwdriver, and the incipient response, W's turning to the screwdriver, are synchronized. Immediately after starting his request with a turn-initial particle, H turns his body slightly in the direction of the object ([Figure 5](#)). At the same time, he begins to get up and produces the second person modal verb construction *kannst du* ("can you"). The syntactic verb-first position projects an interrogative, which is routinely used for making a request, thus opening a new interactional sequence. Together with the interrogative, H gazes to the object of



Figure 5–7. **Figure 5.** H starts getting up and turns body; **Figure 6.** W orients toward domain of scrutiny; **Figure 7.** H points to requested object.

the upcoming request, the screwdriver (**Figure 6**). On the production of the modal verb *kannst* (“can”), H starts to point toward the object. The gesture reaches its apex when he says *mir* (“me”; **Figure 7**). After the subject *du* (‘you’), the second lexical item after the modal verb, has been produced, W begins to turn her body into the direction of the object (**Figure 6**). At this point, neither the object nor the requested action have been mentioned. W thus responds solely to the orientation of H’s body, his pointing gesture and the emerging syntactic request format. W follows H’s gesture with her gaze (see Mondada, 2015; Stukenbrock, 2015, 2018).

As H says *da* (“there”), she is gazing at the requested object (**Figure 8**), which still has not been mentioned (neither has the requested action). Now a shared focus of attention is established: H and W are orienting to the same referential target, the screwdriver. W starts to reach for the object immediately after *da* (“there”), that is, as soon as she has spotted it (**Figure 9**). She grasps the object before the end of H’s turn upon the first syllable of the infinite verb *geben* in clause-final position (‘give’; **Figure 10**), which closes the turn and finally denotes the action W is instructed to perform.

In Extract 3, first phases of the SPP are produced while a FPP is still underway and even while the precise nature of the expected response, that is, both the requested action and the requested object, has not yet been named. There is a micro-sequential coordination between the production of the emerging FPP and the emerging SPP. The respondent builds a fitted response by using just those resources which the emerging FPP affords as soon as they are available for building an emerging response immediately. In particular, at the beginning the bodily display of direction is used to identify the direction of the movement that matters for the requested action and thus the domain of scrutiny (see Goodwin, 2003; Stukenbrock, 2015). Afterward the



Figure 8–10. **Figure 8.** W gazes at domain of scrutiny; **Figure 9.** W reaches toward object; **Figure 10.** W grasps object.

encoding of the agent (who is to perform an action: *du* “you”) and of the beneficiary (*mir* “me”) and—a bit later—the gestural reference (singling out the referential target) are produced consecutively in the FPP and used to build the emerging SPP well before the referent and the expected action are verbalized by the FPP.

The step-by-step-orientation of an SPP to an emerging FPP becomes particularly clear if requested objects or actions cannot be as easily inferred from contextual information as is possible in Extract 3 for the screwdriver. In Extract 4, the SPP-producer orients early to a not yet specified object and refines her response as soon as the object reference is delivered. Mother (M) is cooking together with daughter (D). M prompts D to taste the noodle water:

Extract 4: noodle water/cooking (FOLK_E_00327_SE_01_T_01_DF_01 – 21:00-21:05)

01 (2.25)

02 M *\$weißt=e # \$was probierst=e mal bitte#\$+*
 know.2SG=you try.2SG=you PRT
 You know what will you please try a little bit

m-g *\$at daughter--->*

d-g *\$at mother-----\$at stove--->>*

d-b *+turns tow stove--->*

fig *#fig.11 #fig.12*

03 M *en *\$bisschen +vom NÜdel#wass\$er,*
 a little bit of-the noodle water
 of the noodle water

m-g *--->\$at stove-----\$at bowl in front of her--->>*

m-b **leans slightly forward--->*

d-b *+moves tow stove--->>*

fig *#fig.13*

04 M *ich glaub mir is das SA: #*1z*
 i think.1SG me.dat is.1SG the salt
 i think the salt is

fig *#fig.14*

m-b *-----**

M’s FPP starts when she establishes mutual gaze with D (Figure 11) and frames her turn by a misplacement marker (*weißte was*, “you know what”), displaying that the upcoming turn is unrelated

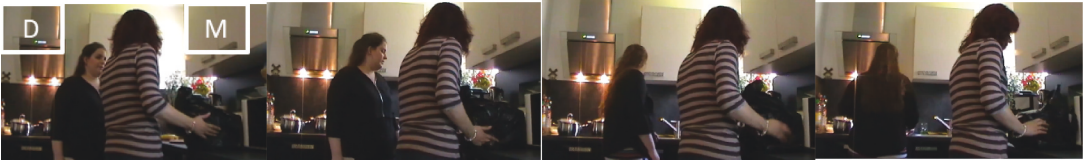


Figure 11–14. **Figure 11.** Mutual gaze; **Figure 12.** M gazes at stove, D turns toward stove; **Figure 13.** D continues turn toward stove; **Figure 14.** D orients to object.

to the previous line of action and therefore unexpected. Nevertheless, after M has produced the verb–first second-person interrogative *probierste mal* (“will you try”), D immediately withdraws her gaze and starts to turn toward the stove where there are several pots, not waiting for the object that is projected by the transitive verb *probierste*. D infers from *probierst du* that a request is underway. The semantics of the verb *probierste* (“try/taste”) projects that the object is edible. In addition, when M produces *probierste*, she leans slightly forward and gazes in the direction of the stove (Figure 12). Both the verbal cues and the concurrent embodied conduct by M allow D to orient early to the domain of scrutiny, the stove, where the pots are located, and to infer the kind of action expected from her.² At this point, however, she does not yet know what she is supposed to taste as there are several pots on the stove. While M carries on with *bisschen was vom* (“a little bit from”), D continues to move toward the stove (Figure 13).

Only when M produces the object to be tasted (*nudelwasser*, “noodle water”), D completes her complying action by finally orienting to the requested object (Figure 14, during M’s account *ich glaub mir ist das salz*, “i think the salt is”) and by subsequently tasting the water (not part of the transcript). The verbalization of the object here is crucial for adjusting the response to the precise referent that is targeted by the requester.

In Extract 4, there are possible alternative referents in the domain of scrutiny to which the requester could be referring. This requires the respondent to wait for the naming of the object in the FPP, until she can orient to the object of the request (see Hanna & Brennan, 2007 for comparable findings in experiments). Therefore, the SPP is adapted in its course to the emerging FPP: The respondent first only orients toward the domain of scrutiny which is available because of the combination of the semantics of the verb in the FPP, the requester’s bodily orientation, objects in the domain of scrutiny and the overall joint project of cooking, while the manipulation of the target only becomes (unambiguously) possible when it is named.

In Extracts 2 to 4, SPPs are incrementally adjusted to the stepwise production of an emerging FPP. By this, responses are not only early, they are micro-sequentially aligned with the projections the emerging FPP affords. Next, we discuss a case where a FPP producer not only projects what action they perform and what they expect as next action, but they are actively involved in the implementation of the SPP.

Reciprocal actions

Sometimes, FPP producers not only indicate by verbal and embodied cues what they expect others to do, but they start a complementary action designed to co-accomplish the projected SPP, for example, reaching for the object that is requested or pushing a body part that should be moved. In such cases, FPP and SPP are particularly tightly coordinated, forming one intercorporeal, joint bodily action together (see Heath & Luff, 2021 and in this volume on presenting and grasping tools in surgery; see; Meyer et al., 2017 on intercorporeality). In Extract 5 from an emergency drill, the officer-in-chief (P) asks the assistant (A) to hand him a paper. While producing his request, he starts to reach out his hand for the requested object.

Extract 5: documents/emergency drill (FOLK_E_00139_SE_01_T_01_DF_01 – 5:04-5:07)

```

01      *$(0.87) $#
      p-b      *turns body tw. A/documents
      p-g      $at A/documents--->
      a-g      $at As2--->
      fig      #fig.15

02 P      %geben#      sie&#      mir      $mal      die&%      SPAp#iere her.$
      give.2.SG      you      me.dat PRT      the      documents

      give me the documents

      p-g      -----$at doc.----->
      p-h      %reaches for doc.-----%grasps/takes doc.---->
      a-g      $at doc.-----$at P>>
      a-h      &passes doc.-----&hands over docs--->
      fig      #fig.16      #fig.17      #fig.18

04      &(0.27)      &
      a-h      &withdraws&

```

Before P produces his verbal request to hand over the documents (line 2), he straightens his body, turns to A and gazes at her (line 1), projecting the transition from the previous activity of talking to another assistant (AS2) to a new activity which involves A (Figure 15). The content of the documents had been a topic some seconds before in the interaction. At the beginning of Extract 5, the documents are in A's hand. A already starts passing the documents to P when he has produced the turn-initial second person verb + subject pronoun *geben sie* "give+you" construction, that is, well before the requested object is mentioned (Figure 17). This construction clearly implements an object request. A's early response here builds on three sources: the documents have been topical before, P already starts to stretch out his hand in the direction of the documents when he produces the first word of his request (Figure 16), and A already holds the documents in her hand. As in Extract 1, A does not need to reorient herself to another direction and place from which the requested object would have to be fetched (as was necessary in Extracts 3 and 4). When P produces *sie* ("you"), A starts passing the documents, while P's



Figure 15–18. **Figure 15.** P gazes at A; **Figure 16.** P starts reaching for the documents; **Figure 17.** A starts passing the documents; **Figure 18.** Transfer of the documents.

hand has already met her half way (Figure 17). When P says *papiere* (“papers”), both participants’ actions—grasping and passing—meet and co-accomplish the transfer of the papers (Figure 18).

In Extract 5, the production of an SPP is adjusted to an emerging FPP that not only projects the next requested action but that is itself involved in the accomplishment of the requested action by executing a reciprocal action, in this case reaching for a requested object. The incipient bodily co-construction enables a quite early response. In cases like this, where FPP and SPP accomplish a bodily joint action, their micro-sequential coordination concentrates on its intercorporeal accomplishment (cf. Heath/Luff, this volume). This is also indicated by gaze directions: Once A has started to pass the papers, she follows the course of her passing action with her gaze thereby, at the same time, monitoring the reciprocal action of P’s reaching for the papers in order to be able to conduct the handing over appropriately. Yet Extract 5 also shows how bodily and pragmatic preparedness of the respondent before the start of the request matters: The papers were interactionally salient slightly before and A is holding the object of the request.

So far we discussed cases in which the producer of an SPP draws successively on information provided by an emerging FPP. In all extracts, the production of an SPP was unproblematic and successful. In the next section, we discuss cases in which the production of an SPP is readjusted or repaired in coordination with the projections that the emerging FPP makes available step by step.

Micro-sequentially occasioned repair of SPP

In Extract 6, from a driving school lesson, the teacher (T) instructs the student (S) to look backward when reverse parking. Before T speaks, he deploys a pointing gesture indicating the direction where to look:

Extract 6. *Look back/driving school* (FOLK_E_00167_SE_01_T_01_DF_01 – 14:49–14:52).

```
01          $$ (0.6)                $% (0.3) $
t-g   $at S/steering wheel----$at rearview mirror-->
t-h                                     %.....
s-g   $looks forward-----$

02 T       je#tz  $%guck          nach #GAnSz #hi%ntn.$$#
      now      look.imp   to.PRP PRT   back
      now look fully back

t-g   -----$at S-----$at rearview mirror-->
t-h   .....%points backwards-----%,,,,,,
s-g   .....$backward-----$readjusts looking-->
fig   #fig.19                #fig.20#fig.21   #fig.22

03          (0.94) %
s-h   ,,,,,,%
```

S starts to look backward before T has finished the first word of his instruction *jetzt* (“now”) (Figure 19). S responds to T’s pointing gesture starting already 0.3 seconds before his turn. S’s response thus is only early



Figure 19. T points and S turns backward before verbal instruction.

with respect to the verbal part of T's action.³ In addition, the requested action, looking back has already been established as a learnable (see Zemel & Koschman, 2014) and thus is beginning to become routinized and highly expectable as the next step at this point in the action series that is to be performed for reverse parking. The action of looking back starts almost in sync with the onset of the pointing gesture. The gesture and the verbal instruction thus only serve as a cue to indicate when the next step of a series of actions is due. The verbal instruction is produced when S has already turned halfway around to look backward (Figure 19). Yet, when T says *ganz hinten* ("fully back"), T turns her head a little further back, readjusting her gaze to the directional information delivered in the emerging FPP (Figures 20–22). Indeed, the degree particle *ganz* ("fully") in the FPP may respond to the fact that S does not already look far enough into the right direction. The emerging FPP thus could be seen to be a corrective instruction supported by the focal accent on the particle *ganz* ("fully").

In Extract 6, the early response builds merely on embodied conduct (pointing), whereas the verbal cues are delivered only when the complying action is already underway. The verbal part is used by the SPP producer to readjust her early response. The late verbal phase of the implementation of the FPP possibly adapts to an anticipatable inadequacy of the responding SPP. Next, we focus on instances in which the FPP is more fundamentally altered in response to the emerging SPP. We are thus dealing with cases of reciprocal micro-sequential adaptation of FPP and SPP.

Micro-sequential adaptation of FPP to early response

For Extracts 2 through 6, we have shown how the course of the SPP was adapted to the course of the emerging FPP. Yet, there are also cases in which the course of the FPP is adapted to the emerging early response as well. Thus, there may be a not only unilateral but also a reciprocal micro-sequential adaptation of FPP and SPP to each other (see De Stefani, this volume and Mondada this volume, who discuss similar cases). Here, we consider two ways in which the FPP can be adapted to an early SPP: The FPP can be abandoned altogether, when it becomes evident from the early response that its completion is not needed for sequence completion; the FPP is expanded by self-repair, when the early response misfires because of a misunderstanding.

In Extract 7 from a theater rehearsal, the FPP is not completed, because the respondents have already sufficiently complied with the instruction. The actors and the director are discussing trajectories and spatial positions of the actors during the play. AC2 proposes that the white line that runs in parallel to the stage edge will serve as landmark for the basic alignment of the players throughout the play. The three other actors respond to this proposal already before its completion by aligning themselves (more or less) neatly with the white line.

Extract 7: white line/theatre (V_PH-MI_AS_20130212_1a - 03:40-3:48)

- 01 AC2 abber so is glaub ich der GAng hier_rüber;
but so is I guess the trajectory over here
- 02 (0.2)
- 03 AC4 JO;
Yeah
- 04 AC3 (.) JA:.
Yes
- 05 AC2 (.) ^weil wir KÖNnen\$ jetzt_%ja ma \$#grob;
because we can.1PL now PRT just roughly
because we can now y'know just roughly
- ac2-g ^at white line--->
ac2-h \$. \$path gesture--->
Fig #fig.23
ac1/3-g %at AC2's gesture--->
- 06 (0.23)%*\$#
ac2-h ----->
ac1/3-g -----%at white line--->
ac1-b *steps twd white line--->
ac4-b/g \$gazes and steps twd white line--->
Fig #fig.24
- 07 AC2 die_ma- die weiße# LInie,\$ dass da:-\$*\$^
the ?? the white line COMP there
the ma- the white line that there
- ac2-g -----^
ac2-h -----\$, , , , , , , \$
ac1-b -----> *
ac4-b -----> \$
Fig #fig.25
- 08 AC1 #JA;
Yes
Fig #fig.26

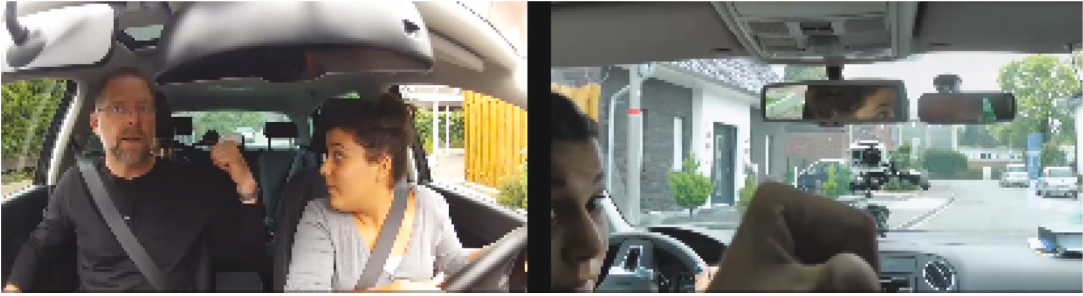


Figure 20. *Ganz* (“fully”).

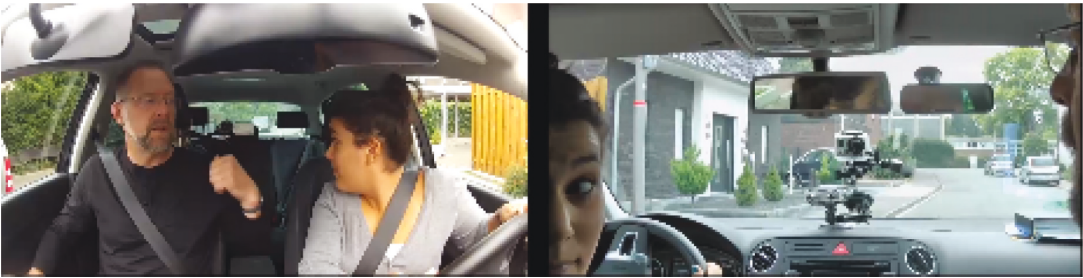


Figure 21. *Hintn* (“back”).

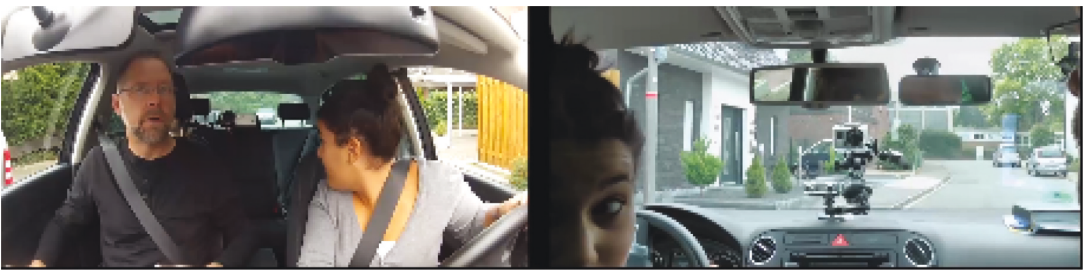


Figure 22. after completion of instruction.

Figure 20-22. S readjusts her looking at *ganz hinten* (“fully back”).

In line 01, AC2 proposes that the actors cross the stage along a certain path. She accounts for this by another proposal (lines 05/07) referring to the white line running in front of them. Having said *wir KÖNnen jetzt ja ma* (“we can now y’know just”), she gazes at the white line and refers to it by a path gesture (Figure 23). AC1 and AC3 follow her gesture by gaze (line 05), and immediately after all actors look at the white line (line 06, Figure 24). Collective attention to the white line is thus established even before AC2 utters the action and the object of her proposal. AC1 and AC4 take a step forward to align themselves with the white line (Figure 25).⁴ While they are in motion, AC2 still traces the white line and now also refers verbally to it (line 07). Before delivering the main verb (which would be required by the auxiliary *können*, “can”), AC2 produces the beginning of a subordinate consecutive clause (“that there”), which, however, she does not complete. Nevertheless, AC1 produces an acknowledgment token, while the ensemble as a whole lines up along the white line (Figure 26). In fact, by this exchange, the white line has been established as the default landmark for the alignment of the ensemble vis-à-vis the audience.

Although AC2’s turn is not formatted as an instruction or a request that is to be followed immediately, all actors align their body with the white line, even before it is mentioned. They do so just by attending to AC2’s gesture and the first person plural verb + auxiliary construction *wir können* (“we can”), which projects a proposal. The actors’ early understanding of and compliance with the

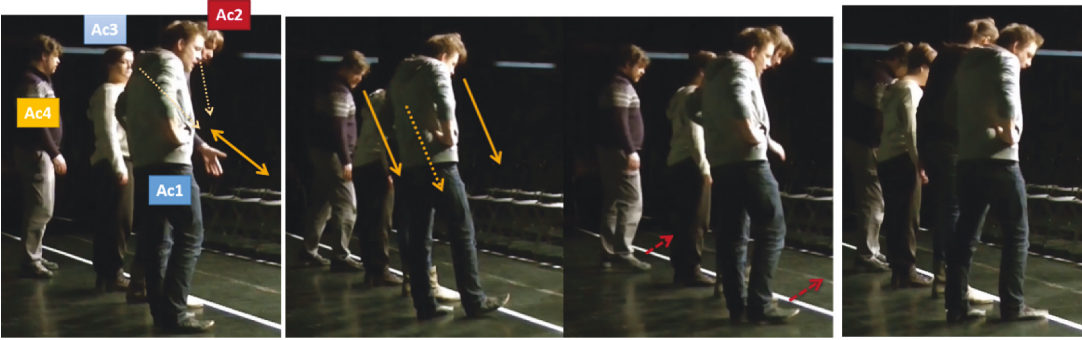


Figure 23–26. **Figure 23.** AC2 gazes at white line and refers to it by path gesture; **Figure 24.** All actors gaze at white line; **Figure 25.** AC1/4 step toward white line; **Figure 26.** Ensemble lined up along white line.

proposal builds on the shared orientation to the current joint project of defining trajectories and positions. The FPP is syntactically not complete, when all addressees bodily display that they understand the reference to the white line as an instruction to line up along the line, and AC2 abandons her turn, which shows that she considers her instruction to be sufficiently complied with (see also Mondada this volume on FPPs being abandoned in reaction to an early response).

In Extract 8, the FPP is adapted to an emerging SPP, not by abandoning the FPP but by self-repair.⁵ The self-repair itself responds to an early response that is moving into the wrong direction. In response, the emerging FPP is self-repaired by embodied means, and the SPP is adapted accordingly. Extract 8 comes from an emergency drill. The paramedic (P) asks the assistant (A) to give him the infusion he wants to apply to the injured person:

Extract 8: infusion/emergency drill (FOLK_E_00142_SE_01_T_01_DF_01 – 09:35-09:41)

- 01 P *§musst einfach gucken wie oft in der miNUTE*
 so you just have to look how often per minute
 a-g *§at patient--->*
sich des hier häng äh hebt und (.) senkt; §ne,
 this here hang uh lifts and lowers
 p-g *§at domain of scrutiny--->*
- 02 (0.33)
 p-g ----->
 a-g ----->
- 03 A *%Joa;#*
 yeah
 p-g --->
 p-h *%...*
 a-g --->
 Fig #fig.27


```

04 P   hãl$=sch#      %grad #emal die## inF$Uision#$ ma% &kurz,$#
      hold.2.SG=you      PRT   the   infusion      PRT   shortly
      can you just hold the infusion quickly

p-g   -----$at A-----$at inf.---->>
p-h   .....%points at inf.-%adjusts point--%--->>
a-g   ---$at hand of P-----$at inf.---->>
a-h                                     &grasps/passes inf.>>
Fig      #fig.28      #fig.29  #fig.30      #fig.31      #fig.32

```

Before P starts his request and in overlap with the end of his previous instruction, he turns his body, gazes to the domain of scrutiny and starts to point in direction of the requested object (Figure 27). A starts gazing at the gesture when P produces the first word of his turn, which makes it recognizable as an upcoming request (*hãlsch*, “hold”) (Figure 28). The apex of the pointing coincides with *grad emal* (“just PRT”) (Figure 29). As a result, A already orients to the domain of scrutiny at the beginning of the request.

However, as there are several possibly relevant objects in the domain of scrutiny, A at first looks at the wrong group of objects. When P realizes this, he adjusts his pointing gesture by slightly changing the position of his hand and his index finger when he starts to produce the article of the object phrase *die* (“the”, Figure 30). P’s adjustment of the pointing is done by a circular movement of his index finger (marked in Figure 30 by a circular arrow) followed by a slight change of the orientation of his index finger (best visible in Figure 32). A gaze-follows P’s adjusted pointing, now looking at the requested object as P names the object (*infusion*) (Figure 31); A starts to grasp the object when P completes his request turn (Figure 32).

In Extract 8, the recipient starts early to orient toward a domain of scrutiny made relevant by a pointing gesture, thus building primarily on embodied cues offered by the FPP producer. However, the domain of scrutiny is not fully correctly recognized and A’s gaze direction does not allow him to find the object. The problem is aggravated, because there are many potentially relevant objects in the



Figure 27–29. **Figure 27.** P gazes, starts to point toward domain of scrutiny; **Figure 28.** A gazes at P’s gesture; **Figure 29.** Apex of P’s pointing.



Figure 30–32. **Figure 30.** Apex of P's pointing; **Figure 31.** P adjusts pointing; **Figure 32.** A looks in right direction.

(possible) domain of scrutiny. Even before completing his turn, the producer of the FPP notices that respondent does not gaze at the right place and readjusts his pointing gesture so that the producer of the SPP manages to identify the correct object.

While in Extract 8 there is a gestural self-repair, Extract 9, from a physiotherapy session, features a verbal self-repair that responds to a failed early response. The therapist (T) proposes/offers the patient (P) to take off his watch before the session starts. Yet, P at first misunderstands T's turn as a request to move back his left arm.

Extract 9: watch/ physiotherapy (FOLK_E_00360_SE_01_T_01_DF_01 - 1:02-1:07)

```

01  T      *%ä:hm      %so*#11      mer die noch-
           uhm          shall.1.PL we DEM
           um shall we that still

           t-b      *walks tw. arm*
           t-h      %.....%points at watch--->
           Fig      #fig.33

02      ^#(0.28)%
           t-h      ,,,,,,,%
           p-p      ^removes arm--->
           Fig      #fig.34

03  T      $soll      ich die #noch-
           shall-1.SG i DEM
           shall I still that

           p-p      ----->
           t-g      $follows patient's arm/watch--->
           Fig      #fig.35

```

```

04      die^ UHR noch abziehe^n?
        the watch take off.inf
        take off the watch

P-P    ---^                ^returns arm to the front--->

05      (0.41)#

P-P    ----->

Fig     #fig.36

06      T      $[ham_ma die ] immer an,# nee.
        do we really have that always on

P      [die is xxx ]^&
        that is

p-g    $at watch--->>

P-P    -----^

p-h    &takes off watch--->>

Fig     #fig.37

```

T starts her turn by using a proposal format with a modal verb in first position (*soll* mer “shall we”). With the onset of her turn, she starts pointing at the watch (which, however, cannot be seen by the patient). She touches it repeatedly when saying *soll* (“shall”) (Figure 33). At this point, instead of taking off the watch, P starts to move his arm backward (Figure 34). Obviously, he misunderstood the demonstrative pronoun (*die*, “this”) and was not able to recognize the right referent (*watch* instead of *hand* or *arm*).⁶ P’s early response builds on the first-person plural modal verb in first position (*soll* “shall” projecting a question or a proposal) and on the touch, which could refer to the watch or the arm. When P realizes T’s wrong response, she self-repairs her FPP twice. First, when P moves his arm backward, she cuts off and changes the subject pronoun from first person plural (*wir*, “we”) to singular (*ich*, “I”), thereby shifting agency from both of them to her alone. This implies that no movement was asked of P. It also changes the action type of the FPP from proposal to offer (Couper-Kuhlen, 2014). However, P continues moving his arm backward (Figure 35).



Figure 33–35. Figure 33. T touches watch/arm; Figure 34. P starts taking arm back; Figure 35. P continues taking arm back.



Figure 36–37. **Figure 36.** P returns arm to the front; **Figure 37.** P takes off the watch.

In response, P cuts her turn off again and self-repairs a second time, now replacing the demonstrative pronoun by an unambiguous lexical NP (*uhr*, “watch”, line 04). Only when T makes the referent (*uhr*, “watch”) explicit and specifies the expected action (*abziehen*, “put off”), P starts to return his arm to the initial position (Figure 36) and takes off the watch himself (Figure 37).⁷

In Extract 9, the FPP is adapted in its course to an early SPP, because the early response exhibits a misunderstanding of the FPP. This is done by a double self-repair, first by changing the agency of the requested action and then by making the reference to the object (and the action to be performed) lexically explicit.

Cases of reciprocal adjustment of the FPP and the emerging SPP give prime evidence for how FPP producers continuously monitor early implementations of the SPP. This allows them to modify their FPP while it is still in progress when problems occur. Still, misfiring early responses that are in need of repair, as in Extracts 8 and 9, suggest a game-theoretic view of early responses. Early responses are potentially risky actions: Their enhanced temporal efficiency (faster than waiting for completion of the request in order to produce the response) builds on anticipation of the not-yet-formulated parts of the SPP, which can lead to wrong inferences. At least theoretically, in consequence, early responses are both more efficient and more risky the earlier they are produced, because earlier responses promote interactive progression more efficiently while resting on less constraining and thus more ambiguous or possibly insufficient projective cues (see also Pekarek-Doehler, this volume, who shows how early responses testify an orientation of respondents to the progressivity of interaction). As “(…) informational and social-affiliative functions of common ground are closely interlinked” (Enfield, 2006, p. 399), early responses can also index close affiliation between participants.

Discussion

This article has studied a specific kind of early bodily responses to requests and instructions, namely, responses whose emerging trajectory is micro-sequentially adapted to the unfolding action to which they respond. Micro-sequential relationships between the time course of the emerging FPP and the emerging SPP are situated at the intersection between simultaneity and sequentiality: The actions are produced in part simultaneously, while individual components or phases of the actions are sequentially coordinated with each other in the sense that the course of an action builds on and adapts to parts of the other action that have been produced already. The mechanism that allows for early responses is projection, operating on several levels. Early responses use bodily and turn-internal projections

concerning the type of action under way and concerning features of the syntactic structure that can be expected—most importantly concerning who will be the subject (agent) and whether the upcoming request or instruction will involve an object and, if so, of which kind. These projections in turn create further projections concerning the kind of next action that is expected from the addressee. Projection thus is the basis for anticipations and inferences on the part of the respondents, which they use for producing an early response.

Earliness means that a response starts before a FPP is complete. In the case of sequences consisting of a verbal instruction and an embodied implementation, there is no “overlap” in the technical sense of Sacks et al. (1974), because talk and embodied actions do not share a floor that is managed according to turn-taking rules. Early embodied responses require a shared multimodal face-to-face context, which they exploit. Their occurrence leads us to rethink actions and their sequential relationships, which are considered as basic for the accomplishment of interaction in conversation analysis (Schegloff, 2007), in at least three ways:

- Multimodal actions are not building blocks but composite actions that include various resources with different temporalities (Deppermann & Streeck, 2018), which are systematically combined and which establish different temporalities of projection—which can be used to initiate and accomplish a response at different points in time.
- A response may be early with respect to a complete turn, but it may be just timely with respect to the projections that have already been made available by certain components of a first action.
- Smooth and maximally efficient sequencing of actions, as is evidenced by early responses, builds on histories of interactional routines, which lay the ground for the projective properties of multimodal resources at a certain sequential position, for a certain action, in the context of a certain joint activity by virtue of prior interactional experience.

Micro-sequential relationships create a timely and tightly matched production of micro-pairs of action components from the requester/instructor and the respondent. Micro-sequential relationships lead to a successive, incremental specification of the response and sometimes of the initiating action as well. We have shown different kinds of micro-sequential relationships between FPP and SPP: Successive specification of the SPP building on the resources that the FPP makes available step by step, repair of the SPP by virtue of later components of the SPP, and reciprocal adaptations of the FPP to the emerging SPP by repairing or abandoning the FPP. Micro-sequential coordination of the production of FPP and SPP thus is not necessarily unilateral but can reciprocally inform both FPP and SPP. Micro-sequential relationships highlight that actions often are not monolithic in their planning and their execution but organized in consecutive temporal phases. These phases amount to a successive specification of the emergent action and sometimes also to self-repair. Altogether these findings support an emergentist or incremental vision of language use and action.

In the analyses of the extracts we have tracked how the resources that are used to construct an ongoing action are used by the partner to build an early response. We have tried to identify the components and their projective potentials to which the responders visibly and by hearing react to on their production. Although our study has not systematically focused on the use of the projective potentials of certain resources, a few observations can be made, which could be explored in more detail in future studies. With respect to requests for immediate bodily responses with present material objects, we could observe that gaze direction and the preparatory phase (Kendon, 2004) of pointing project who is addressed (which is important in multiparty situations) and allow the recipient to identify a domain of scrutiny by gaze-following. Of course, the effect of these factors is supported if requester and addressee already share joint attention. This, in turn, may already allow for object identification, provided there is only one object in (potential) reach that is potentially relevant for the joint action. Object identification is even more facilitated if the object is already used or touched by the respondent. However, if there is more than one potentially relevant object in the domain of scrutiny, the (early) response can only turn to the object later when it is explicitly mentioned (see also Hanna &

Brennan, 2007). Typically, therefore body and gaze orientation characterize the early phases of the response, whereas the precise orientation to the object and the more specific object-related action (such as grasping and handing, removing, attaching, etc.) only comes later. Gaze direction and the preparatory phase of pointing often slightly precede the beginning of the request turn; in addition, the set-up of stationary spatial surroundings and the bodily position and posture of the producer of the FPP are further resources on which the early response can draw, even before the verbal turn begins. The specific projective potentials of different components and phases of the actions and their differential use in the generation of a micro-sequentially coordinated responses give evidence of the decisive role of multimodality for the emergence of early responses. Although we have only studied verbal FPPs, the bodily aspects of these FPPs prove to be decisive for the time course of the early response.

The linguistic structure can also be seen to be used as a temporalized resource (see also Pekarek-Doehler, this volume). In German, the word order of the turn projects early that a request is upcoming. Imperatives and second person (modal) verb-first interrogatives inhabit a clause-initial and therefore mostly also turn-initial position. They create an early projection (Levinson, 2013) of the action type of the turn and thus of the action type of the expected response as well. The imperative and the verb-first interrogative, but not the modal verb-first interrogative, in addition deliver the main verb, which projects a valence frame. This importantly includes the projection of an object, if it is a transitive verb, and possibly even more specifically the semantic class of object (like something edible, readable, tellable, etc.) or an additional specification of the recipient/beneficiary by dative case.

However, the projective properties of components and phases of a turn often are not simply inherent in them. Most early responses in our data are produced as part of routine sequences of everyday life or institutional contexts. The development of the ability to respond early thus seems to be dependent on shared interactional histories and processes of social learning to read the projective potentials of cues (Deppermann, 2018). In addition, the larger sequential context can play an important role: Requests and instructions often are not produced out of the blue. In the cases of early responses in our data, they are mostly next steps within an overarching joint project or, even more predictably, in a series of sequences or actions, which are themselves routine. Future studies thus should also embrace these more extended temporal orders that range beyond the adjacency pair (Deppermann & Streeck, 2018; Deppermann & Schmidt, *in press*; see Clark & Marshall, 1981; Clark & Wilkes-Gibbs, 1986 for experimental evidence): It is a future task to find out in which ways interactional histories and larger joint projects lay the grounds for the projective potentials of components and phases of local actions and how they create preparedness and projections for next actions.

Notes

1. The following abbreviations are used in the transcripts to mark different embodied resources: g = gaze; h = hand; b = whole body/torso; p = body part (e.g., leg).
2. Compare experimental studies that show recipients orient to speakers' embodied conduct as early as possible to start their responses (Clark & Krych, 2004, p. 73).
3. Mondada (this volume) discusses different temporal relationships of earliness of "second" actions with respect to "first" actions.
4. AC3 was already aligned with the white line from the beginning of the extract.
5. The same phenomenon occurred in experiments when instructors self-repaired their instructions when instructed actions seemed to be wrong (Clark & Krych, 2004, 68 et seqq.).
6. Note that the demonstrative pronoun *die* is feminine and thus is congruent with *Uhr* (watch). It could also replace *Hand* (hand) but not *Arm*, which is masculine in German and would have needed to be referred to by the demonstrative *den*.
7. See Kendrick and Drew (2014) on requests that are delivered as offers.

Acknowledgements

Parts of this work were presented at the IPRA conference 2017 in Belfast. We are grateful to Adrian Bangerter, Jon Hindmarsh, Lorenza Monada, Simona Pekarek Doehler, and Giovanni Rossi for their helpful reviews of prior versions of this article. We also thank Aleksandra Gubina for comments and for helping us with data sampling. Mira Rinnebach prepared basic transcriptions of the data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Arnulf Deppermann  <http://orcid.org/0000-0001-7575-8236>

References

- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- Auer, P. (2005). Projection in Interaction and Projection in Grammar. *Text*, 25(1), 7–36.
- Clark, H. H. (1996). Communities, commonalities, and communication. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 324–355). CUP.
- 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., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1), 62–81. <https://doi.org/10.1016/j.jml.2003.08.004>
- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual reference. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). CUP.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39. [https://doi.org/10.1016/0010-0277\(86\)90010-7](https://doi.org/10.1016/0010-0277(86)90010-7)
- Couper-Kuhlen, E. (2014). What does grammar tell us about action? *Pragmatics*, 24(3), 623–647. <https://doi.org/10.1075/prag.24.3.08cou>
- Dausendschön-Gay, U., & Krafft, U. (2009). Preparing next actions in routine activities. *Discourse Processes*, 46(2–3), 247–268. <https://doi.org/10.1080/01638530902728900>
- Deppermann, A. (2006). Deontische Infinitivkonstruktionen: Syntax, Semantik, Pragmatik und interaktionale Verwendung. In S. Günthner & W. I. Wolfgang (Eds.), *Konstruktionen in der Interaktion* (pp. 239–262). de Gruyter.
- Deppermann, A. (2013). Turn-design at turn-beginnings: Multimodal resources to deal with tasks of turn-construction in German. *Journal of Pragmatics*, 46(1), 91–121. <https://doi.org/10.1016/j.pragma.2012.07.010>
- Deppermann, A. (2018). Changes in turn-design over interactional histories – The case of instructions in driving school lessons. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (pp. 293–324). Benjamins.
- Deppermann, A. (2018). Special issue on instructions in driving lessons. *International Journal of Applied Linguistics*, 28(2), 221–225. <https://doi.org/10.1111/ijal.12206>
- Deppermann, A., & Schmidt, A. (in press). On the emergence of shared meanings over an interactional history: *wabi sabi* in a series of theater rehearsals. *Research on Language and Social Interaction*.
- 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). Benjamins.
- Drew, P., & Couper-Kuhlen, E. (Eds.). (2014). *Requesting in social interaction*. Benjamins.
- Enfield, N. J. (2006). Social consequences of common ground. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 399–430). Berg.
- Gergle, D., Kraut, R. E., & Fussell, S. R. (2004). Action as language in a shared visual space. CSCW'04 - Proceedings of the 2004 ACM conference on Computer supported cooperative work (pp. 487–496). <https://doi.org/10.1145/1031607.1031687>
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language* (pp. 97–121). Institute for Ethnomethodology.
- Goodwin, C. (1981). *Conversational organization*. Academic.
- Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 217–241). Lawrence Erlbaum.

- Goodwin, M. H. (2006). Participation, affect, and trajectory in family directive/response sequences. *Text & Talk*, 26 (4–5), 513–541. <https://doi.org/10.1515/TEXT.2006.021>
- Goodwin, M. H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry*, 50(3–4), 303–317.
- Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57(4), 596–615. <https://doi.org/10.1016/j.jml.2007.01.008>
- Hazel, S. (2018). Discovering interactional authenticity: Tracking theatre practitioners across rehearsals. In S. Pekarek Doehler, J. Wagner, & E. Gonzalez-Martinez (Eds.), *Longitudinal studies on the organization of social interaction* (pp. 255–283). Palgrave Macmillan.
- Heath, C., & Luff, P. (2000). *Technology in action*. Cambridge.
- Heath, C., & Luff, P. (2021). Passing touch: Handing and handling tools and implements during surgical procedures. In A. Cekaite & L. Mondada (Eds.), *Touch in social interaction. Touch, language, and body* (pp. 249–268). Routledge. in press
- Hindmarsh, J. (2010). Peripherality, participation and communities of practice: Examining the patient in dental training. In N. Llewellyn & J. Hindmarsh (Eds.), *Organization, interaction and practice: Studies in ethnomethodology and conversation analysis* (pp. 218–240). CUP.
- Hindmarsh, J., Reynolds, P., & Dunne, S. (2011). Exhibiting understanding: The body in apprenticeship. *Journal of Pragmatics*, 43(2), 489–503. <https://doi.org/10.1016/j.pragma.2009.09.008>
- Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica*, 9(1), 47–96. <https://doi.org/10.1515/semi.1973.9.1.47>
- Jefferson, G. (1985). An exercise in the transcription and analysis of laughter. In T. A. van Dijk (Ed.), *Handbook of discourse analysis* (pp. 25–34). Academic.
- Kendon, A. (2004). *Gesture*. CUP.
- Kendrick, K. H., & Drew, P. (2014). The putative preference for offers over requests. In P. Drew & E. Couper-Kuhlen (Eds.), *Requesting in social interaction* (pp. 87–114). Benjamins.
- Koschman, T. (2011). Special issue: Understanding understanding in action. *Journal of Pragmatics*, 43(2), 435–437. <https://doi.org/10.1016/j.pragma.2010.08.016>
- Levinson, S. C. (2013). Action formation and ascription. In T. Stivers & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 103–130). Wiley-Blackwell.
- 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 Blackwell.
- Meyer, C., Streeck, J., & Jordan, S. J. (Eds.). (2017). *Intercorporeality: Emerging socialities in interaction*. OUP.
- Mondada, L. (2013). Coordinating mobile action in real time: The timely organization of directives in video games. In P. Haddington, L. Mondada, & M. Nevile (Eds.), *Interaction and mobility. Language and the body in motion* (pp. 300–342). de Gruyter.
- Mondada, L. (2014). Instructions in the operating room: How the surgeon directs their assistant's hands. *Discourse Studies*, 16(2), 131–161. <https://doi.org/10.1177/1461445613515325>
- Mondada, L. (2015). Introduction de nouveaux référents et gestion de l'attention conjointe: Les apports de l'analyse interactionnelle multimodale. In X. Gradoux, J. Jacquin, & G. Merminoud (Eds.), *Agir dans la Diversité des Langues. Mélanges en l'Honneur d'Anne-Claude Berthoud* (pp. 115–130). De Boeck.
- Mondada, L. (2018a). Multiple temporalities of language and body in interaction. Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (2018b). Driving instruction at high speed on a race circuit: Issues in action formation and sequence organization. *International Journal of Applied Linguistics*, 28(2), 304–325. <https://doi.org/10.1111/ijal.12202>
- 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). Benjamins.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Sauppe, S. (2016). Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences. *Frontiers in Psychology*, 7(95). doi:10.3389/fpsyg.2016.00095
- Schegloff, E. (2007). *Sequence organization in interaction*. CUP.
- Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist*, 70(6), 1075–1095.
- Searle, J. R. (1976). A classification of illocutionary acts. *Language in Society*, 5(1), 1–23. <https://doi.org/10.1017/S0047404500006837>
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P. et al. (2011). A system for transcribing talk-in-interaction: GAT 2. Translated and adapted for English by E. Couper-Kuhlen & D. Barth-Weingarten. *Gesprächsforschung*, 12, 1–51.
- Streeck, J. (1995). On projection. In E. N. Goody (Ed.), *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence* (pp. 87–110). CUP.
- Streeck, J., & Jordan, S. J. (2009). Projection and anticipation: The forward-looking nature of embodied interaction. *Discourse Processes*, 46(2–3), 93–102. <https://doi.org/10.1080/01638530902728777>

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

Stukenbrock, A. (2018). Forward-looking: Where do we go with multimodal projections? In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction* (pp. 31–68). Benjamins.

Zemel, A., & Koschman, T. (2014). 'Put your fingers right in here': Learnability and instructed experience. *Discourse Studies*, 16(2), 163–183. <https://doi.org/10.1177/1461445613515359>