Originally published in: Discourse Processes vol. 58 (2021) no. 4, pp. 293-307.
DOI: https://doi.org/10.1080/0163853X.2021.1877516

DISCOURSE PROCESSES
2021, VOL. 58, NO. 4, 293–307
https://doi.org/10.1080/0163853X.2021.1877516

Early Responses: An Introduction

Arnulf Deppermann[a], Lorenza Mondada[b], and Simona Pekarek Doehler[c]

[a]Pragmatics Department, Leibniz-Institut für Deutsche Sprache, Mannheim, Germany; [b]Philosophisch-Historische Fakultät, Departement Sprach- und Literaturwissenschaften, Universität Basel, Basel, Switzerland; [c]Institut des Sciences du Langage et de la Communication, Université de Neuchâtel, Neuchâtel, Switzerland

ABSTRACT

This special issue investigates early responses—responsive actions that (start to) unfold while the production of the responded-to turn and action is still under way. Although timing in human conduct has gained intense interest in research, the early production of responsive actions has so far largely remained unexplored. But what makes early responses possible? What do such responses tell us about the complex interplay between syntax, prosody, and embodied conduct? And what sorts of actions do participants accomplish by means of such early responses? By addressing these questions, the special issue seeks to offer new advances in the systematic analysis of temporal organization in interaction, contributing to broader discussions in the language and cognitive sciences as to the social coordination of human conduct. In this introductory article, we discuss the role of temporality and sequentiality in social interaction, specifically focusing on projective and anticipatory mechanisms and the interplay between multiple semiotic resources, which are crucial for making early responses possible.

Temporality and prospective orientation in language and interaction

Temporality is a core feature of the production and the interpretation of human experience and action and therefore also of talk. Within the phenomenological tradition it has been suggested that the continuity of experience rests on the temporal structure of consciousness, an ever-moving "now-moment" (*Jetztmoment*; Husserl, 1929) that is fringed (see also James, 1909) by retentions (passing present moments that fade into the past) and protentions (more or less vague anticipations of next moments). Consciousness therefore is never self-contained but always ahead of itself (Merleau-Ponty, 1945). A prospective orientation is hence built into the very structure of experience. On a more general level, Heidegger (1927) describes the specific human mode of existence, which he calls *Dasein* ('being-there' as being characterized by *Sich-vorweg-Sein* 'being ahead of oneself'; p. 192). The theory of action considers planning—the future-oriented design of actions in accordance with in-order-to motives (Schütz, 1974, pp. 74–85)—as a key criterion that distinguishes action from mere behavior (Schütz, 1974, p. 78). Schütz (1974) states that action planning includes the anticipation of the completed action in *modo futuri exacti* (p. 81).

Temporality as key to the understanding of language and social interaction has been addressed in Psycholinguistics, as well as in Conversation Analysis and Interactional Linguistics. This insight is captured by the title of this journal, *Discourse Processes*. Linguistic structures and interactional events are time objects: They are produced and understood in time, and at least in the spoken mode, they are evanescent events that fade away upon their very production. The intrinsically temporal nature of language production and comprehension is the topic of Psycholinguistics, with its interest in the time courses of information processing, planning, integration, expectation, and prediction (see, e.g., Levelt,

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

1989; Tanenhaus et al., 1995; Venhuizen et al., 2019). Recent psycholinguistic research has discussed the temporal parameters of speech comprehension, planning, and production that underlie precision-timed turn taking in social interaction (Bögels & Levinson, 2017; Corps et al., 2018), thereby shedding light on the "significant puzzle" that the turn-taking system poses for theories of language processing (Levinson & Torreira, 2015, p. 1).

Conversation Analysis and Interactional Linguistics in turn highlight the intersubjective dimension of this temporality, which is not limited to cognitive processes but centrally pertains to joint engagement in social interaction. For participants in social interaction, the temporal coordination of actions is a ubiquitous task. Coordination largely relies on projection of one's own incipient action trajectories, as well as on anticipation of the trajectories of others' (next) actions (Sacks et al., 1974; Streeck & Jordan, 2009). According to the theoretical framework of Conversation Analysis and Interactional Linguistics, the temporal production of interaction is situated, emergent, incremental, and jointly coordinated (Deppermann & Günthner, 2015), contingent on the local interactional context and not only on what the self does but also on what others do.

In what follows, we discuss the ways in which temporality is managed by participants in social interaction and how it affects its very organization, providing affordances and creating opportunities for successivity and/or simultaneity of actions, including responsive actions. While mainly drawing on insights from Conversation Analysis and Interactional Linguistics, we also address relevant findings from psycholinguistics research. Our purpose is to outline the key features of temporal coordination in social interaction that underlie the phenomenon that is the focus of this special issue—early responses:

- *Sequentiality* refers to the successive, stepwise production of joint action, the way it is responded to, and the way the response reflexively affects the initial ongoing action. It is the fundamental temporal principle that informs both the understanding and the production of actions in social interaction and serves as the general framework for the concatenation of actions in interaction (the second section).
- *Projection* refers to the fact that bits of behavior can foreshadow the continuation of the action as well as future actions. Projection is the key mechanism by which incipient trajectories of interaction on various levels are made more or less expectable; its counterpart is anticipation—the recipient's expectations concerning what comes next, allowing for continuous local adjustments (the third section).
- The temporal organization of social interaction relies on successivity, fundamental for the moment-by-moment emergence and progressivity of sequentiality, and on simultaneity. On the verbal level, simultaneous productions occur in the form of overlap (the fourth section). Even more radically, simultaneity is a crucial feature of multimodal interaction. This begs the question how sequentiality and simultaneity are related to each other—in particular, how multimodal resources are temporally organized to produce intersubjectively coordinated courses of interaction.

This understanding of the nature and the role of temporality in social interaction has developed through five decades of research on the organizational infrastructure of human interaction. Yet one key aspect of the social coordination of actions has largely remained unexplored: the early production of responsive actions. This is so despite earlier work by Jefferson (1983) discussing instances of early overlaps, when "the recipient/next speaker starts up well before anything like a transition place has been reached, but where an understanding of at least the general thrust of the utterance can have been achieved" (p. 20). Despite her fundamental contribution, there has been only scarce research on the workings of early responses (but see the second section).

What has remained unaddressed, in particular, are the conditions that make early responses possible. This is the focus of this special issue: the environments and circumstances that favor the production of early responses—actions (starting to) unfold while the production of the responded-to action is still under way. To better understand these conditions, we relate them to sequentiality, projection, and the complementary relations between successivity and simultaneity.

Sequentiality

Sequentiality is a fundamental principle of social interaction. As such, it has been put on center stage most prominently by research in Conversation Analysis. In *emic* terms, the principle of sequentiality is expressed by the question "Why that now?": Participants scrutinize the sequential position of the details of each other's conduct for what contributes to local meaning making (Schegloff, 2007; Schegloff & Sacks, 1973, p. 299). The question points to the local endogenous orientation of the coparticipants, who can address at any moment (the "now") the relevance of a lexical choice, a prosodic or phonetic feature, the emergent syntax of the turn, a gesture made at some point during an action, or a gaze shift (representing the "that"). These two aspects of the question, the "now" and the "that," constitute the dimensions of position and composition that are fundamental for the production and interpretation of conduct in interaction (Schegloff, 2007). *Composition* refers to the kind of resource and to the configuration of resources mobilized. The composition of linguistic and embodied resources acquires its intelligibility thanks to their position within the turn, the ongoing action, the emergent sequence, and/or the overall organization of the interaction. *Position* thus refers to the specific moment within the temporality of interaction in which some piece of conduct is produced. There is no time-out from the temporal flow of language, embodiment, and action—even though temporality is not uniform but relative and can be accelerated or slowed down. In this sense, an item—for example, a linguistic form, or a practice, or an action—is positioned with respect to what precedes it and what follows it, establishing retrospective as well as prospective relationships with surrounding items (see the third section). Sequentiality informs the temporal unfolding as well as the intelligibility of linguistic and embodied forms composing action. This in turn enables coparticipants to anticipate, adjust, and respond to them.

Schegloff (2007) distinguishes between sequentiality—as a general principle animating all organizational levels of social interaction—and sequence organization. The latter refers to a more specific form of organization, one that holds between two actions where one follows the other, in an adjacent way, the first projecting and making conditionally relevant the second, that is, creating a normative expectation about its nature. Adjacency pairs are the basic form of sequence organization (Schegloff & Sacks, 1973): Given a first, the second is expectable and normatively inspected in this way, generating specific rights and obligations among the participants. The first selects a relevant next, i.e., a responsive action. Its production and its timing hinge on the respondent's recognition of the first action and the point in time at which such recognition is possible (cf. Jefferson, 1983).

Forms of alignment and disalignment, fittedness and unfittedness, agreement and disagreement, affiliation and disaffiliation between the first and the second are all vectors through which dynamic interactional relations are configured between participants (Lee & Tanaka, 2016; Pomerantz, 1984; Raymond, 2003; Stivers, 2008). The informal or institutional nature of the encounter materializes through types of first actions and specific constraints on second actions, as well as through who has the right and obligation to perform these actions; these elements can be considered the building blocks of social order and institutions.

Another fundamental feature of sequentiality and sequence organization is that participants display their understandings of a previous action by responding to it (Moerman & Sacks, 1988). Understanding is not merely a cognitive internal private process, but it is materialized in the next action. Both verbal and embodied responses show (more or less clearly) how a previous action was understood (Deppermann, 2015; Mondada, 2011). Thus, sequentiality is the locus where intersubjectivity is manifested, constituted, and maintained in social interaction.

Projection and anticipation

Projection—and its counterpart, anticipation—is fundamental to the coordination of actions in social interaction. When we interact with others, we format our actions so as to project their incipient course and their completion, as well as to foreshadow possible or expected next actions. At the same time, we

orient to verbal, paraverbal, and embodied cues to anticipate coparticipants' actions and adapt our own actions step-by-step as a function of such anticipation. Already Mead (1934) considered the action-projecting properties of "gestures" (which for him comprise all sorts of acts of meaningful, observable conduct) as the basis of their meaning:

This threefold or triadic relation between gesture, adjustive response, and resultant of the social act that the gesture initiates is the basis of meaning; the existence of meaning depends upon the fact that the adjustive response of the second organism is directed toward the resultant of the given social act as initiated and indicated by the gesture of the first organism (p. 80).

For socio-interactional approaches, anticipation is crucial for participants being able to adapt their own actions to projectable actions and more generally for coordinating and synchronizing actions in a timely and mutually attuned manner (Streeck & Jordan, 2009). Anticipation is equally important for preemptive actions, preventing some projectable course of others' actions from materializing. Knowledge about projectable action sequences is a major source for anticipatory interactive planning (Drew, 1995), as it becomes evident in the initiation of presequences that are produced to check the likelihood of a positive response to some planned initiative that requires the other's collaboration. Already Schütz (1974, p. 20) stresses the fact that communicative action implies anticipation of the other's understanding of one's action, a basic property of turn design, which has come to be termed "recipient-design" (Drew, 2013; Sacks et al., 1974, p. 728; see also work on "audience design" by Clark & Marshall, 1981; Clark & Wilkes-Gibbs, 1986; Isaacs & Clark, 1987). Anticipation in social interaction often uses projections that arise from others' actions. This is a key feature making early responses possible.

In Conversation Analysis, the fundamental role of projection and anticipation in social interaction has prominently been put forward early on the classic paper on turn taking: "Unit-types project, from beginnings, features of their construction, their direction and what it will take to complete them" (Sacks et al., 1974, p. 719). Projections emanate from cues that make expectable the continuation of a course of action (for example, the continuation of a turn) and its completion (for example, the moment at which turn transition becomes relevant), and generate expectations as to the next actions (for example, what type of action should follow; Schegloff, 1968). The anticipation of turn completion makes smooth turn taking possible, minimizing gaps and overlaps as well as making overlaps accountable (Jefferson, 1983). Projection and anticipation can therefore be expected to be among the factors that make early response possible.

The following excerpt, discussed by Sacks (1987), provides a simple yet telling example of how anticipation of others' actions and turns impacts on the fabric of one's own action:

```
Excerpt 1 (Sacks, 1987, p. 64)

01  A:  They have a good cook there?

02      (pause)

03  A:  Nothing special?

04  B:  No everybody takes their turns
```

A utters a question (line 01) that is tilted toward a positive response (also because of its declarative format; Heritage, 2010), such as "yeah, gorgeous"' or "sure." Yet as the recipient fails to respond (line 02), A revises her question by inversing its polarity: *nothing special?* (line 03). She hence appears to construe B's silence as prefiguring a dispreferred response, which the revision of her question's format allows her to avoid, favoring the possibility of a response that aligns with the second question's—now negative—polarity (Boyd & Heritage, 2006): *No everybody takes their turns* (line 04). In short, A's anticipation of the recipient's dispreferred incipient action reflexively shapes her own next action.

Projection and anticipation rest on the fact that at any point in time the occurrence of a stretch of conduct engenders normative expectations as to what happens next, i.e., it prefigures possible trajectories of ongoing turns and actions as well as possible next turns and actions. These expectations are experientially configured through recurrent participation in social interaction, and it is their sharedness that allows participants to recognize projections (Auer, 2005; Goodwin, 2002; Hayashi, 2004; Sacks et al., 1974) as a ground for anticipation. Projections emanate from different linguistic and bodily resources and can concern different orders of interactional organization.

Grammatical projection arises from the fact that at any point in its progression, the grammatical trajectory of an utterance foreshadows possible or even obligatory follow-ups that build on language- and construction-specific principles concerning, for example, linearization (word order), head-complement-ordering, and more specific valency constraints emanating from specific lexical items (Auer, 2005). Auer (2009) states that "grammar is a conventionalized set of formal ways of making projections possible" (p. 180). In many European languages, for instance, the occurrence of a determiner makes expectable the occurrence of a noun (or an adjective + noun combination) as next item (Auer, 2005), and the occurrence of an *if*-clause makes expectable a *then*-clause as a next (Lerner, 1991). Grammatical projection also concerns macrosyntactic continuations, as indexed by projector constructions like "the thing/the problem is" (Auer, 2009; Günthner, 2008), extrapostions ("it's funny"; Couper-Kuhlen & Thompson, 2008), or the initial parts of pseudoclefts of the type "what is interesting is . . . " (Hopper & Thompson, 2008; Pekarek Doehler, 2011, 2015). Importantly, grammatical projections can also operate across turns, as with polar questions projecting a type-conforming answer, which in most languages requires responding by a particle (Enfield et al., 2019; Raymond, 2003, 2013).

Prosodic projection is most important for turn taking and extended sequences/topic developments. Ford and Thompson (1996) and Selting (2000, 2005) have studied the interplay of syntax and prosody to project TCU endings and transition-relevance places; Barth-Weingarten (2009) has shown how prosody is used to project turn continuation at points of syntactic completion; Selting (2007) has identified specific prosodic features projecting that there will be a next item in a list. Couper-Kuhlen (2001, 2004) has shown how high pitch onset at turn beginnings projects that the incipient turn is opening a new sequence or topic. Some of these features may equally play into a recipient's anticipating and timing relevant next actions.

Embodied projection relates to the fact that gesture, gaze, body posture, or the manipulation of material objects may foreshadow possible turn or sequence trajectories as well as possible next actions (Ford et al., 1996; Mondada, 2006, 2015; Streeck, 2009). Gestures and object manipulation can project a claim for the floor (Mondada, 2007), gaze can be used for pursuing a response (Pekarek Doehler, 2019; Stivers & Rossano, 2010), the "hold" of a hand gesture may equally project that a recipient's response is (still) awaited (Streeck, 2007) or that repair is completed (Floyd et al., 2016), and the trajectory of walking can project the initiation of the next action (Mondada, 2014b).

Action projection relates to the fact that specific types of first actions (first pair-parts) make specific types of next actions (second pair-parts) expectable, which is captured by the principle of "conditional relevance" (Schegloff, 1968, 2007). For example, a question makes an answer expectable (Schegloff & Sacks, 1973); an assessment projects an affiliative response (Pomerantz, 1984). Turn beginnings are a most important site for the establishment of projections concerning the upcoming action and its relationship to prior talk (Deppermann, 2013). Turn-initial particles, such as *well* or *so* index aspects of the incipient action (Heritage & Sorjonen, 2018). Bodily conduct, such as leaning forward (Mortensen, 2009; Streeck, 1995), facial expressions (see Kaukomaa et al., 2014 on frowns), walking or stopping to walk (Mondada, 2018a), object use (Mondada, 2006, 2015), and preturn inbreath (Schegloff, 1996) may project more to come.

Sequence projection concerns projections that extend, beyond the next action, to larger sequences of actions. Various types of "pre's" (Schegloff, 2007, pp. 37–47) project a base-sequence, such as a request +complicance, an invitation+acceptance/refusal, or an extended story-telling to follow, produced to check its appropriateness or probable success.

As shown by these different types of projections, its forward-looking nature is a fundamental characteristic of action in interaction that shapes its formatting and the way it is locally interpreted by coparticipants. Projections might be fulfilled on the same level (for example, grammatical formats project the continuation and possible completion of a grammatical structure) but also create expectations on other levels, concerning not only turns but also actions (for example, Hayashi, 2004). For instance, a particle may project the occurrence of a transitional relevance place (Ford & Thompson, 1996); word-order and syntactic dependencies are key factors allowing for early action projection (Auer, 2009; Pekarek Doehler et al., 2015; Pekarek Doehler, this volume).

In a nutshell, projections operating at these various levels share the following properties (cf. Auer, 2005):

- Projections are not deterministic but emerge reflexively from both the format of an action and its interpretation. They have a normative character in the sense that although they do not necessarily have to be fulfilled, deviation from a projected (course of) action is accountable, i.e., needs to be accounted for or may give rise to inferences (Schegloff, 1968).
- The scope of projections can be rather narrow, as with a phonetic feature foreshadowing a next sound or the occurrence of a determiner announcing a noun to follow. It can also be much wider, as in the case of an interrogative format projecting the expected response type or a presequence projecting a specific type of base sequence to follow.
- The projective force can vary as far as its normative obligation is concerned, constraining to various degrees the possible appropriate continuation or next action ranging, for instance, from cases where the next action is made conditionally relevant (for example, in paired greetings; Schegloff, 1968) to cases in which a response is optional (for example, responses to certain kinds of assessments; Stivers & Rossano, 2010).
- Projections can be very constraining, i.e., requiring an action with a determinate format, as in the case of a wedding vow), or they can be rather vague, as in the case of eliciting an autobiographical narrative.
- Projections can be suspended by intervening activities (like parenthetical inserts, see Mazeland, 2007) and may be abandoned or reinterpreted over the course of action (for example, changing the addressee midturn, see Goodwin, 1979; or withdrawing an action in response to recipient's lack of alignment, see Pekarek Doehler, 2019).
- Projection and anticipation are experientially grounded, for example, through repeated and routinized participation in social interaction, and it is their shared character that allows for participants' recognition of projections as a ground for anticipating actions— and configuring their own (responsive) actions.

Conversation Analytic findings converge with other interdisciplinary findings on the role of temporality, projection, and anticipation in human conduct, although other fields have addressed these mostly at the level of individual behavior. For instance, according to Seligman et al. (2013), "navigation into the future is seen as a core organizing principle of animal and human behavior" (p. 119), which includes "evaluative representations of possible future states" and rests on the "prospecting brain" (Seligman et al., 2013, p. 127). Prospection and anticipation operate on different time scales, ranging from the next milliseconds (the Husserlian protentions) over extended action trajectories (with beginning, middle, and end) to "temporal landscapes," which concern biographic phases of a career, chronic illness treatment, or family relationships (Tavory & Eliasoph, 2013). Anticipation of others' actions and planning of own actions are closely intertwined on all of these temporal levels (Tavory & Eliasoph, 2013). Planning itself necessarily involves anticipation because it includes the imagination of goals and the assessment of alternative, possible courses of actions and their possible dependencies and obstacles. Anticipatory action therefore is not simply adaptive in being responsive to expectable events and actions; it involves the adaption of one's own actions to others' actions with an eye to one's own goals and intentions as well (as shown in the previous example cited from Sacks).

More generally, anticipation is forward modeling (Pickering & Garrod, 2013), which builds on prior experiences and often involves one's own behavioral routines. An important mechanism for interactive alignment and quick response, which is distinct from projection, is priming. Priming in interaction involves the mental activation and the behavioral imitation of structures in a prior speaker's talk on various levels—in particular, on the lexical, syntactic, and prosodic level (Foltz et al., 2015; Pickering & Garrod, 2004).

Early responses

Early responses are both conceptually and analytically important for the study of temporality of talk and embodied because they offer unprecedented insights into the workings of sequentiality, projection, and anticipation in social interaction. The timing of an early response shows when and how its producer develops and displays their understanding the previous turn or action; it opens up a window on the precise moment when the point of a turn or action becomes recognized in the course of its very production (see Jefferson, 1973, 1983 on the recognition point of an action-in-progress). This timing also enables us as analysts to reconstruct what resources the participant is orienting to for generating a response. This special issue examines a diversity of contexts, activities, actions, and resources that allow for the emergence of early responses.

One key issue is how to gauge the "earliness" of a response. From the vantage point of turn taking, early responses are overlapping responses, which can be oriented to as either unproblematic or problematic. The picture complexifies further when we consider phenomena of simultaneity in multimodal interaction, where embodied responses can be produced even before the previous action or turn has started (Mondada, 2018b). The simultaneity of embodied resources and actions is not treated by the participants in the same way as the simultaneity of talk (which is why the notion of overlap is not applicable in this case; Deppermann & Streeck, 2018; Schmitt, 2005). In the next sections, we discuss both overlaps in talk and simultaneity in multimodal interaction in an attempt to identify what exactly an early overlapping turn versus an early simultaneous embodied action might reveal about the functioning of social interaction.

Turn taking and overlap

Central for our purpose here is the fact that projection and anticipation are what makes possible the type of precision-timed turn taking that we know from social interaction. Sacks, Schegloff, and Jefferson's observations that "overwhelmingly, one party talks at a time" (Sacks et al., 1974, p. 696), that there is minimization of gaps and overlaps (Sacks et al., 1974, p. 705), and that such minimization basically rests on the "projection"/"projectability" (Sacks et al., 1974, p. 702) of turn ends have been corroborated by a plethora of studies in Conversation Analysis as well as in Psycholinguistics. Research in Interactional Linguistics, itself inspired by Conversation Analysis, has confirmed Sacks et al.'s initial understanding that the anticipation of turn ends involves recipients tracking the lexicosyntactic trajectories of ongoing turns at talk, combined with their pragmatic completion as well as prosodic marking of turn ends (for example, Ford et al., 1996; Selting, 2000, 2005). Phonetic cues to turn ends such as final lengthening and final pitch accents (Local & Walker, 2012) are hypothesized to be mere triggers for the actual production of utterances that have been planned before turn end, and there is evidence that response planning starts as early as possible while the to-be-responded-to turn is in progress (Levinson & Torreira, 2015). Although the complexity of this process has been shown to lead to boundaries of turn ends and transition-relevant points that can be indeterminate or fuzzy (Auer, 2010), findings concur to indicate that at the completion point of a turn at talk, no projection of the turn's (syntactic, prosodic, pragmatic) continuation is in play anymore.

Based on CA insights into the turn-taking mechanisms, psycholinguists have more recently addressed turn taking in terms of the "puzzle" that gaps between turns at talk are in average shorter than the time needed for utterance planning. Gaps average around 200–300 ms (De Ruiter et al., 2006;

Stivers et al., 2009), while planning of a single word requires more than 600 ms and a sentence approximately 1,500 ms (Levinson & Torreira, 2015; Torreira et al., 2015). Research investigating the cognitive and neural processes that allow for effective turn taking has evidenced the anticipatory processes that are involved in it. It has been shown that anticipation of upcoming verbal strings allows for the prediction of turn ends (for example, Bögels et al., 2015; De Ruiter et al., 2006; Magyari & de Ruiter, 2012). Predictions have also been shown to be affected by gesture; Holler et al. (2018) document that gestures the retraction of which begins prior to the end of a question may function as early turn-completion signals, thus facilitating the timing of responses. This work focuses on the processing of syntactic and semantic cues in simple clauses but does not consider complex extended turns. Embodied actions and the larger precontext are only rarely considered, nor is action ascription, which, however, is crucial for the timing of responsive actions: "action ascription by B of A's turn is a prerequisite for the design of B's turn" (Levinson, 2013, p. 103).

In conversation analytic work, Jefferson (1983, 1984) offers a discussion of different types of overlapping turn onsets that is key to our purpose here. She distinguishes between transitional onset and recognitional onset. *Transitional onset* builds on the next speaker's monitoring the turn as to its completion point: It occurs at or just before a possible TRP, for example, in terminal overlap on the last syllable, when—as Schegloff (2000) put it—"one speaker appears to be starting by virtue of a prior speaker's analyzable incipient finishing of a turn" (p. 5). *Recognitional onset*, by contrast, builds on monitoring the point at which the turn's action becomes recognizable: "while a transition place has not been adequately reached, an item has been produced sufficient for recognition and response" (Jefferson, 1983, p. 29; see also Vatanen, 2018). This is one element than can account for early onset of responsive actions.

Early turn onset has mainly been discussed in terms of overlapping talk and the techniques participants may employ to move out of overlap (Drew, 2009; Jefferson, 1984; Schegloff, 2000). Overlaps can be treated as more or less problematic by the parties at talk, i.e., as more or less violating the current speaker's right to the floor. Terminal overlap is typically not treated as problematic and does not need any overlap resolution techniques (Schegloff, 2000, p. 5). There can be conditional access to the turn space, for example, in word searches (Goodwin & Goodwin, 1986) or in collaborative utterance constructions (Lerner, 1991), when a speaker yields the turn in progress to another or invites another to continue the turn. These cases may be treated as noncompetitive overlap (Drew, 2009). Yet other sorts of overlap can index competition and conflict and be treated as interruption (Kotthoff, 1993; Oloff, 2009). Next speakers can display that their turns are competitive (Oloff, 2009). For example, turn-competitive incomings are typically produced with high pitch and loud volume (French & Local, 1983; Wells & Macfarlane, 1998).

It has also been shown that the timing of turn onset is sensitive to participants' embodied conduct (Holler et al., 2018; Mondada, 2007, 2018a) and to its specific temporality (Stivers et al., 2009) with bodily-visual responses tending to be faster than verbal ones. Also, there is variation in the precise timing—though not in the tendency to minimize gap and overlap—depending on language and culture (Stivers et al., 2009; Tannen, 1984).

Embodied early responses

The frequent occurrence of early bodily responses raises the question of the specific temporality and sequentiality characterizing the use of embodied practices and resources, such as gaze, gesture, head tosses, body postures, and body movements. The simultaneity of embodied conduct has a distinctive phenomenology when compared to the simultaneity of talk: Although a persistent simultaneity of talk is often problematic (e.g., in cases of extended overlap, see Schegloff, 2000), simultaneity of embodied behavior is ubiquitous (Deppermann, 2013; Deppermann & Streeck, 2018; Goodwin, 1981; Mondada, 2016). This difference affects the conceptualization of sequentiality: The focus on talk, on the basis of audio recordings, has favored a *one-dimensional successive* vision of sequentiality; the focus on embodied actions, on the basis of video recordings, favors a vision of sequentiality that is constituted by *multiple simultaneous successivities* (Mondada, 2018a).

A participant might mobilize talk, gesture, gaze, and body posture together, in an orderly way, to format an ongoing action; moreover, other participants might mobilize similar resources in a responsive and coordinated way. The first mode of coordination has been referred to as intrapersonal coordination, the latter as interpersonal coordination of activities (Deppermann, 2014; Deppermann & Schmitt, 2007). In intrapersonal coordination, different multimodal resources implementing an action are mobilized together in a form of simultaneity. Nevertheless, they are not properly synchronous—in the sense that the preparation, apex, and retraction of each movement (gesture, gaze, head movement, body posture, etc.) can occur at different moments and unfold at a different pace. There is no strict isochrony between different embodied resources, although they constitute together a single multimodal Gestalt. A pointing gesture, for example, often starts at or even before a turn beginning; the apex of the gesture usually coincides with its lexical affiliate, i.e., the referential term, and it is mostly retracted only after the occurrence of the latter (Kendon, 2004; Schegloff, 1984; Stukenbrock, 2015). Temporal coordination of the various embodied resources depends on their intrinsic temporalities and temporal affordances. For instance, the trajectory of a gesture has not the same temporality as a gaze shift or the construction of a sentence. A participant can respond early to an action exhibiting a complex Gestalt—including, for example, talk, a gesture, a nod, and a gaze shift—by making use of the fact that the different resources are deployed at different moments and have different durations, some resources projecting the referent, the action, the next speaker, etc., earlier than others (see, for example, Amati & Brennan, 2018; Hanna et al., 2020; Neider et al., 2010 on reference and gaze coordination; see Hanna & Brennan, 2007 on the role of gaze vs. talk), although the whole multimodal action is produced in a precisely and timely coordinated manner (Mondada, 2014a, 2018a).

Interpersonal temporal coordination of multimodal resources involves all participants interacting together. Studies of actions accomplishing local reference have shown how the production of pointing gestures strongly relies on a recipient's simultaneous gaze and embodied responsiveness (Goodwin, 2003; Hindmarsh & Heath, 2000; Mondada, 2014b; Stukenbrock, 2018). Goodwin (1981) shows that speakers orient to the presence of the recipient's gaze as a condition for starting a turn and that, if this is not the case, they might mobilize various practices, such as repairs and restarts, to secure their gaze. In the latter case, the ongoing talk is modified, not only orienting to what the recipient does (or does not do) but also prompting or inviting the recipient to do what is expected (to gaze at the speaker). Speakers are constantly monitoring their recipients' behavior (Clark & Krych, 2004; Goodwin, 1979; Heath, 2013), adjusting their ongoing action to them. In particular, they adjust their talk to the kind of continuers (vs. repairs) they are receiving and, more generally, their action to the kind of reception they get, as (dis)aligning and (dis)affiliating with it (Bavelas et al., 2002; Goodwin, 1980; Stivers, 2008).

These distinctions, pinpointing the importance of *when* an embodied action is initiated within a sequence for interpreting it as a response (and a response to a specific action or detail of an action), enable us to distinguish a response from a choral action. A response is always in a relation of successivity to what it responds to and thus rooted in sequentiality. By contrast, choral actions are simultaneous in terms of synchronicity, as in singing or playing music. In talk, this can occur when next turns are strongly projectable, with participants aiming at coproducing them simultaneously (Lerner, 2002; Schwitalla, 1993), as well as turns that are designed in a way that enables audiences to respond in unison, like in churches, classrooms, or assemblies (Atkinson, 1984; Lerner, 2002, p. 343). Choral actions are also done in an embodied way (see Lerner, 2002 for co-enacted same gestures). More generally, for some intercorporeal social activities, simultaneously coordinated engagements are constitutive, for example, for carrying heavy weights together (Mondada, 2016), kissing, dancing, and some sport activities (Meyer et al., 2017). However, these activities are initiated by a specific participant, and at least entry and exit, if not their ongoing unfolding, requires forms of microadjustments (see for kissing Kendon, 1975; for hugging Goodwin, 2017).

Conclusion

Early responses provide a window on participants' orientation to the temporality of social interaction and to the multitude of resources that allow for projection and anticipation of relevant next actions.

They involve (partial) simultaneity, being to a certain extent produced as the action they respond to is still ongoing. They manifest a form of anticipation—of how the first action is about to be completed, of its relevance and meaning, as well as which kind of response it projects. In boundary cases, both actions can start (almost) simultaneously, making it difficult to distinguish which is the first and which is the second, who has initiated, and who has responded. There are even situations in which what would have been expected as second action starts first, before the sequence-initiating action has been produced, making it impossible to speak of first versus second actions (Mondada, this issue).

The circumstances, formats, configurations of resources, and contexts that make these various temporal configurations possible are explored in the contributions to this special issue. They present empirical studies across a variety of settings, ranging from ordinary conversation (Pekarek Doehler; Vatanen et al.) to various types of institutional interactions (Deppermann & Schmidt; De Stefani; Heath & Luff; Mondada). They analyze verbal and embodied responses in terms of the grounds for their production, the sequential environments in which they occur, their orderly coordination with the action that prompts them, and their interactional import: What cues do participants orient to in order to be able to provide an adequate response before the responded-to action is completed? What do such early responsive actions tell us about the complex interplay of syntax, prosody, and bodily-visual conduct? And how do such on-line features interact with participants' orientation to institutional expectancies, shared mutual knowledge and experience, or routine concatenation of actions known to members?

Verbal and embodied resources exhibit different temporal affordances, yet they both presuppose respondents' recognition of the social meaning of the responded-to action. Whereas early verbal responses are constrained by rules of turn taking, there are no such restrictions on the production of early bodily responses, such as nods in response to questions (De Stefani, this volume). The contributions to the special issue scrutinize the properties of first (i.e., initiating) actions (for example, turn-initial particles, syntax of turn beginnings, word order, projections derived from valence patterns, the recognition point; see Deppermann & Schmidt, this volume; Pekarek Doehler, this volume) and their producers' bodily activities (for example, pointing, gaze direction, body movement in space), as well as other contextual sources (such as objects that are possibly relevant for the response; see Deppermann & Schmidt, this volume; Mondada, this volume) that provide for early projection and that are observably used for the production of early responses. They also identify the precise aspects of the expected response that are projected early (for example, who is addressed, the type of responsive action, the object to which the addressee should turn, etc.) and how these provide affordances for producing an early response—and how in turn the very production of such early responses may affect the accomplishment of the responded-to action while it is still in progress (see Deppermann & Schmidt, this volume; Mondada, this volume). In addition, shared interactional routines (Heath & Luff, this volume), the larger ongoing action sequence (Vatanen et al., this volume), and prior interactional histories (Deppermann, 2018) may be crucial for facilitating early responses. In this way, the papers of this special issue contribute to our understanding of how interactional properties of multimodal resources are used and how they differ in terms of their temporal and projective properties.

By addressing these issues, the special issue seeks to offer new advances in the systematic analysis of emporal organization in interaction, contributing to broader discussions in the language and cognitive sciences as to the social coordination of human conduct, the nature and the sources of projection and anticipation, the coordination of joint action, and the interplay between multiple multimodal resources for situated meaning-making in social interaction.

Disclosure statement

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

ORCID

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

References

Amati, F., & Brennan, S. E. (2018). Eye gaze as a cue for recognizing intention and coordinating joint action. In G. Brone & B. Oben (Eds.), *Eye-tracking in interaction* (pp. 21–46). John Benjamins.

Atkinson, J. M. (1984). *Our masters' voices*. Methuen.

Auer, P. (2005). Projection in interaction and projection in grammar. *Text - Interdisciplinary Journal for the Study of Discourse*, *25*(1), 7–36. https://doi.org/10.1515/text.2005.25.1.7

Auer, P. (2009). Projection and minimalistic syntax in interaction. *Discourse Processes*, *46*(2–3), 180–205. https://doi.org/10.1080/01638530902728934

Auer, P. (2010). Zum Segmentierungsproblem in der gesprochenen Sprache [On the segmentation problem of the spoken language]. *InLiSt - Interaction and Linguistic Structures*, *49*. https://www.inlist.uni-bayreuth.de/issues/49/InLiSt49.pdf

Barth-Weingarten, D. (2009). Contrasting and turn transition: Prosodic projection with the parallel-opposition construction. *Journal of Pragmatics*, *41*(11), 2271–2294. https://doi.org/10.1016/j.pragma.2009.03.007

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listeners responses as a collaborative process: The role of gaze. *Journal of Communication*, *52*(3), 566–580. https://doi.org/10.1111/j.1460-2466.2002.tb02562.x

Bögels, S., & Levinson, S. C. (2017). The brain behind the response: Insights into turn-taking in conversation from neuroimaging. *Research on Language and Social Interaction*, *50*(1), 71–89. https://doi.org/10.1080/08351813.2017.1262118

Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific Reports*, *5*(Article), 12881. https://doi.org/10.1038/srep12881

Boyd, E., & Heritage, J. (2006). Taking the history: Questioning during comprehensive history taking. In J. Heritage & D. Maynard (Eds.), *Communication in medical care: Interaction between physicians and patients* (pp. 151–184). Cambridge University Press.

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 knowledge. In A. K. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge University Press.

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

Corps, R. E., Gambi, C., & Pickering, M. J. (2018). Coordinating utterances during turn-taking: The role of prediction, response preparation, and articulation. *Discourse Processes*, *55*(2), 230–240. https://doi.org/10.1080/0163853X.2017.1330031

Couper-Kuhlen, E. (2001). Interactional prosody: High onsets in reason-for-the-call turns. *Language in Society*, *30*(1), 29–53. https://doi.org/10.1017/S0047404501001026

Couper-Kuhlen, E. (2004). Prosody and sequence organization in English conversation: The case of new beginnings. In E. Couper-Kuhlen & C. E. Ford (Eds.), *Sound patterns in interaction: Cross-linguistic studies from conversation* (pp. 335–376). John Benjamins.

Couper-Kuhlen, E., & Thompson, S. A. (2008). On assessing situations and events in conversation: 'Extraposition' and its relatives. *Discourse Studies*, *10*(4), 443–467. https://doi.org/10.1177/1461445608091882

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Predicting the end of a speaker's turn; a cognitive cornerstone of conversation. *Language*, *82*(3), 515–535. https://doi.org/10.1353/lan.2006.0130

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. (2014). Multimodal participation in simultaneous joint projects: Interpersonal and intrapersonal coordination in paramedic emergency drill. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Eds.), *Multiactivity in social interaction: Beyond multitasking* (pp. 247–282). John Benjamins.

Deppermann, A. (2015). Temporality in interaction. In A. Deppermann & S. Günthner (Eds.), *Retrospection and understanding in interaction* (pp. 57–94). John Benjamins.

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). John Benjamins.

Deppermann, A., & Günthner, S. (2015). Introduction: Temporality in interaction. In A. Deppermann & S. Günthner (Eds.), *Temporality in interaction* (pp. 1–23). John Benjamins.

Deppermann, A., & Schmitt, R. (2007). Koordination. Zur Begründung eines neuen Forschungsgegenstandes [Coordination. On the foundations of a new object of research.]. In R. Schmitt Ed., *Koordination. Analysen zur multimodalen interaktion* [Coordination. Analyses of multimodal interaction] (pp. 15–54). Narr.

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. 1–30). John Benjamins.

Drew, P. (1995). Interaction sequences and anticipatory interactive planning. In E. Goody (Ed.), *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence* (pp. 111–138). Cambridge University Press.

Drew, P. (2009). "Quit talking while I'm interrupting": A comparison between positions of overlap onset in conversation. In M. Haakana, M. Laakso, & J. Lindström (Eds.), *Talk in interaction: Comparative dimensions* (pp. 70–93). Finnish Literature Society (SKS).

Drew, P. (2013). Turn design. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 131–149). Wiley-Blackwell.

Enfield, N. J., Stivers, T., Brown, P., Englert, C., Harjunpää, K., Hayashi, M., Heinemann, T., Hoymann, G., Keisanen, T., Rauniomaa, M., Raymond, C. W., Rossano, F., Yoon, K.-E., Zwitserlood, I., & Levinson, S. C. (2019). Polar responses. *Journal of Linguistics*, *55*(2), 277–304. https://doi.org/10.1017/S0022226718000336

Floyd, S., Manrique, E., Rossi, G., & Torreira, F. (2016). Timing of visual bodily behavior in repair sequences: Evidence from three languages. *Discourse Processes*, *53*(3), 175–204. https://doi.org/10.1080/0163853X.2014.992680

Foltz, A., Gaspers, J., Meyer, C., Thiele, K., Cimiano, P., & Stenneken, P. (2015). Temporal effects of alignment in text-based, task-oriented discourse. *Discourse Processes*, *52*(8), 609–641. https://doi.org/10.1080/0163853X.2014.977696

Ford, C. E., Fox, B. A., & Thompson, S. A. (1996). Practices in the construction of turns: The "TCU" revisited. *Pragmatics*, *6*(3), 427–454. https://doi.org/10.1075/prag.6.3.07for

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.

French, P., & Local, J. (1983). Turn-competitive incomings. *Journal of Pragmatics*, *7*(1), 17–38. https://doi.org/10.1016/0378-2166(83)90147-9

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97–121). Irvington.

Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry*, *50*(3–4), 272–302. https://doi.org/10.1111/j.1475-682X.1980.tb00023.x

Goodwin, C. (1981). *Conversational organization*. Academic Press.

Goodwin, C. (2002). Time in action. *Current Anthropology*, *43*(4), 19–35. https://doi.org/10.1086/339566

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. (2017). Haptic sociality: The embodied interactive construction of intimacy through touch. In C. Meyer, J. Streeck, & J. S. Jordan (Eds.), *Intercorporeality: Emerging socialities in interaction* (pp. 73–102). Oxford University Press.

Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, *62*(1–2), 51–75.

Günthner, S. (2008). Projektorkonstruktionen im Gespräch: Pseudoclefts, die Sache ist-Konstruktionen und Extrapositionen mit es [Projector constructions in interaction: Pseudoclefts, die Sache ist-constructions and extrapositions with es]. *Gesprächsforschung*, *9*, 86–114. https://www.gespraechsforschung-online.de/heft2008/ga-guenthner.pdf

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

Hanna, J. E., Brennan, S. E., & Savietta, K. (2020). Eye gaze and head orientation cues in face-to-face referential communication. *Discourse Processes*, *57*(3), 201–223. https://doi.org/10.1080/0163853X.2019.1675467

Hayashi, M. (2004). Projection and grammar: Notes on the "action-projecting" use of the distal demonstrative are in Japanese. *Journal of Pragmatics*, *36*(8), 1337–1374. https://doi.org/10.1016/j.pragma.2004.05.006

Heath, C. (2013). *The dynamics of auction: Social interaction and the sale of fine art and antiques*. Cambridge University Press.

Heidegger, M. (1927). *Sein und Zeit* [Existence and time]. Niemeyer.

Heritage, J. (2010). Questioning in medicine. In A. F. Freed & S. Ehrlich (Eds.), *"Why do you ask?": The function of questions in institutional discourse* (pp. 42–68). Oxford University Press.

Heritage, J., & Sorjonen, M.-L. (Eds.). (2018). *Between turn and sequence: Turn-initial particles across languages*. John Benjamins.

Hindmarsh, J., & Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics*, *32*(12), 1855–1878. https://doi.org/10.1016/S0378-2166(99)00122-8

Holler, J., Kendrick, K. H., & Levinson, S. C. (2018). Processing language in face-to-face conversation: Questions with gestures get faster responses. *Psychonomic Bulletin & Review*, *25*(5), 1900–1908. https://doi.org/10.3758/s13423-017-1363-z

Hopper, P., & Thompson, S. A. (2008). Projectability and clause combining in interaction. In R. Laury (Ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions* (pp. 99–124). John Benjamins.

Husserl, E. (1929). *Vorlesungen zur Phänomenologie des inneren Zeitbewußtseins* [On the phenomenology of the consciousness of internal time]. Niemeyer.
Isaacs, E. A., & Clark, H. H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology. General*, *116*(1), 26–37. https://doi.org/10.1037/0096-3445.116.1.26
James, W. (1909). *The meaning of truth*. Longmans, Green & Co.
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. (1983). Two explorations of the organization of overlapping talk in conversation: Notes on some orderlinesses of overlap onset. *Tilburg Papers in Language and Literature*, *28*, 1–28. https://liso-archives.liso.ucsb.edu/Jefferson/Onset.pdf
Jefferson, G. (1984). Notes on some orderlinesses of overlap onset. In V. d'Urso (Ed.), *Discourse analysis and natural rhetoric* (pp. 11–38). Cleup.
Kaukomaa, T., Peräkylä, A., & Ruusuvuori, J. (2014). Foreshadowing a problem: Turn-opening frowns in conversation. *Journal of Pragmatics*, *71*, 132–147. https://doi.org/10.1016/j.pragma.2014.08.002
Kendon, A. (1975). Some functions of the face in a kissing round. *Semiotica*, *15*(4), 299–334. https://doi.org/10.1515/semi.1975.15.4.299
Kendon, A. (2004). *Gesture*. Cambridge University Press.
Kotthoff, H. (1993). Disagreement and concession in disputes: On the context sensitivity of preference structures. *Language in Society*, *22*(2), 193–216. https://doi.org/10.1017/S0047404500017103
Lee, S.-H., & Tanaka, H. (2016). Affiliation and alignment in responding actions. *Journal of Pragmatics*, *100*, 1–7. https://doi.org/10.1016/j.pragma.2016.05.008
Lerner, G. H. (1991). On the syntax of sentences-in-progress. *Language in Society*, *20*(3), 441–458. https://doi.org/10.1017/S0047404500016572
Lerner, G. H. (2002). Turn-sharing: The choral co-production of talk-in-interaction. In C. Ford, B. Fox, & S. Thompson (Eds.), *The language of turn and sequence* (pp. 225–256). Oxford University Press.
Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. MIT Press.
Levinson, S. C. (2013). Action formation and ascription. In T. Stivers & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 103–130). Wiley-Blackwell.
Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*, Article 731. https://doi.org/10.3389/fpsyg.2015.00731
Local, J., & Walker, G. (2012). How phonetic features project more talk. *Journal of the International Phonetic Association*, *42*(3), 255–280. https://doi.org/10.1017/S0025100312000187
Magyari, L., & de Ruiter, J. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in Psychology*, *3*(Article), 376. https://doi.org/10.3389/fpsyg.2012.00376
Mazeland, H. (2007). Parenthetical sequences. *Journal of Pragmatics*, *39*(10), 1816–1869. https://doi.org/10.1016/j.pragma.2007.05.005
Mead, G. H. (1934). *Mind, self and society from the standpoint of a social behaviorist*. Chicago University Press.
Merleau-Ponty, M. (1945). *Phénoménologie de la perception* [Phenomenology of perception]. Gallimard.
Meyer, C., Streeck, J., & Jordan, J. S. (Eds.). (2017). *Intercorporeality: Emerging socialities in interaction*. Oxford University Press.
Moerman, M., & Sacks, H. (1988). On "understanding" in the analysis of natural conversation. In M. Moerman (Ed.), *Talking culture: Ethnography and conversation analysis* (pp. 180–186). University of Pennsylvania Press.
Mondada, L. (2006). Participants' online analysis and multimodal practices: Projecting the end of the turn and the closing of the sequence. *Discourse Studies*, *8*(1), 117–129.
Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, *9*(2), 195–226. https://doi.org/10.1177/1461445607075346
Mondada, L. (2011). Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, *43*(2), 542–552. https://doi.org/10.1016/j.pragma.2010.08.019
Mondada, L. (2014a). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, *65*, 137–156. https://doi.org/10.1016/j.pragma.2014.04.004
Mondada, L. (2014b). Bodies in action: Multimodal analysis of walking and talking. *Language and Dialogue*, *4*(3), 357–403. https://doi.org/10.1075/ld.4.3.02mon
Mondada, L. (2015). Multimodal completions. In A. Deppermann & S. Günthner (Eds.), *Temporality in Interaction* (pp. 267–307). John Benjamins.
Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, *20*(2), 2–32. https://doi.org/10.1111/josl.1_12177
Mondada, L. (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

Mortensen, K. (2009). Establishing recipiency in pre-beginning position in the second language classroom. *Discourse Processes*, *46*(5), 491–515. https://doi.org/10.1080/01638530902959463

Neider, M. B., Chen, X., Dickinson, C. A., Brennan, S. E., & Zelinsky, G. J. (2010). Coordinating spatial referencing using shared gaze. *Psychonomic Bulletin & Review*, *17*(5), 718–724. https://doi.org/10.3758/PBR.17.5.718

Oloff, F. (2009). *Contribution à l'étude systématique de l'organisation des tours de parole: Les chevauchements en français et en allemand* [A contribution to the systematic study of the organization of turn-taking: Overlaps in French and German] [Doctoral dissertation]. University of Mannheim MADOC. https://ub-madoc.bib.uni-mannheim.de/29617/

Pekarek Doehler, S. (2011). Clause-combining and the sequencing of actions: Projector constructions in French conversation. In R. Laury & R. Suzuki (Eds.), *Subordination in conversation: A crosslinguistic perspective* (pp. 103–148). John Benjamins.

Pekarek Doehler, S. (2015). Grammar, projection and turn-organization: (il) y a NP 'there is NP' as a projector construction in French talk-in-interaction. In A. Deppermann & S. Günthner (Eds.), *Temporality in interaction* (pp. 173–200). John Benjamins.

Pekarek Doehler, S. (2019). At the interface of grammar and the body. *Chais pas* ('dunno') as a resource for dealing with lack of recipient response. *Research on Language and Social Interaction*, *52*(4), 1–23. https://doi.org/10.1080/08351813.2019.1657276

Pekarek Doehler, S., De Stefani, E., & Horlacher, A.-S. (2015). *Time and emergence in grammar: Left-dislocation, right-dislocation, topicalization and hanging topic in French talk-in-interaction*. John Benjamins.

Pickering, M. J., & Garrod, S. C. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190. https://doi.org/10.1017/S0140525X04000056

Pickering, M. J., & Garrod, S. C. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 229–293. https://doi.org/10.1017/S0140525X12001495

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 57–111). Cambridge University Press.

Raymond, G. (2003). Grammar and social organization: Yes/no-type interrogatives and the structure of responding. *American Sociological Review*, *68*(6), 939–967. https://doi.org/10.2307/1519752

Raymond, G. (2013). At the intersection of turn and sequence organization: On the relevance of "slots" in type-conforming responses to polar interrogatives. In B. Szczepek-Reed & G. Raymond (Eds.), *Units of talk–units of action* (pp. 169–206). John Benjamins.

Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J. R. E. Lee (Eds.), *Talk and social organisation* (pp. 54–69). Multilingual Matters.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, *50*(4), 696–735. https://doi.org/10.1353/lan.1974.0010

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, *70*(6), 1075–1095. https://doi.org/10.1525/aa.1968.70.6.02a00030

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. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52–133). Cambridge University Press.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, *29*(1), 1–63. https://doi.org/10.1017/S0047404500001019

Schegloff, E. A. (2007). *Sequence organization in interaction*. Cambridge University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*(4), 289–327. https://doi.org/10.1515/semi.1973.8.4.289

Schmitt, R. (2005). Zur multimodalen Struktur von turn-taking [On the multimodal structure of turn-taking]. *Gesprächsforschung*, *6*, 17–61.

Schütz, A. (1974). *Der sinnhafte Aufbau der sozialen Welt* [The meaningful structure of the social world]. (Original work published 1932). Suhrkamp.

Schwitalla, J. (1993). Über einige Weisen des gemeinsamen Sprechens. Ein Beitrag zur Theorie der Beteiligungsrollen im Gespräch [On some ways of talking together. A contribution to the theory of participation roles in conversation]. *Zeitschrift Für Sprachwissenschaft*, *11*(1), 68–92.

Seligman, M. E., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past? *Perspectives of Psychological Science*, *8*(2), 11–141. https://doi.org/10.1177/1745691612474317

Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, *29*(4), 477–517. https://doi.org/10.1017/S0047404500004012

Selting, M. (2005). Syntax and prosody as methods for the construction and identification of turn-constructional units in conversation. In A. Hakulinen & M. Selting (Eds.), *Syntax and lexis in conversation: Studies on the use of linguistic resources in talk-in-interaction* (pp. 17–44). John Benjamins.

Selting, M. (2007). Lists as embedded structures and the prosody of list construction as an interactional resource. *Journal of Pragmatics*, *39*(3), 483–526. https://doi.org/10.1016/j.pragma.2006.07.008

Stivers, T. (2008). Stance, alignment and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction*, *41*(1), 31–57. https://doi.org/10.1080/08351810701691123

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoonf, K.-E., & Levinson, S. C. (2009). Universality and cultural specificity in turn-taking in conversation. *Proceedings of the National Academy of Science*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stivers, T., & Rossano, F. (2010). Mobilizing response. *Research on Language and Social Interaction*, *43*(1), 3–31. https://doi.org/10.1080/08351810903471258

Streeck, J. (1995). On Projection. In E. Goody (Ed.), *Interaction and social intelligence* (pp. 84–110). Cambridge University Press.

Streeck, J. (2007). Geste und verstreichende Zeit: Stillstand und Bedeutungswandel [Gesture and passing time: Standstill and meaning change]. In H. Hausendorf Ed., *Gespräch als Prozess* [Conversation as process] (pp. 157–177). Narr.

Streeck, J. (2009). Forward-gesturing. *Discourse Processes*, *46*(2–3), 161–179. https://doi.org/10.1080/01638530902728793

Streeck, J., & Jordan, S. (Eds.). (2009). Projection and anticipation in embodied interaction [Special issue]. *Discourse Processes*, *46*(2–3), 93–268. https://www.tandfonline.com/toc/hdsp20/46/2-3.

Stukenbrock, A. (2015). *Deixis in der face-to-face interaktion* [Deixis in face-to-face interaction]. 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). John Benjamins.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Tannen, D. (1984). *Conversational style: Analyzing talk among friends*. Ablex.

Tavory, I., & Eliasoph, N. (2013). Coordinating futures: Toward a theory of anticipation. *American Journal of Sociology*, *118*(4), 908–942. https://doi.org/10.1086/668646

Torreira, F., Bögels, S., & Levinson, S. C. (2015). Breathing for answering: The time course of response planning in conversation. *Frontiers in Psychology*, *6*(Article), 284. https://doi.org/10.3389/fpsyg.2015.00284

Vatanen, A. (2018). Responding in early overlap: Recognitional onsets in assertion sequences. *Research on Language and Social Interaction*, *51*(2), 107–126. https://doi.org/10.1080/08351813.2018.1413894

Venhuizen, N. J., Crocker, M. W., & Brouwer, H. (2019). Expectation-based comprehension: Modeling the interaction of world knowledge and linguistic experience. *Discourse Processes*, *56*(3), 229–255. https://doi.org/10.1080/0163853X.2018.1448677

Wells, B., & Macfarlane, S. (1998). Prosody as an interactional resource: Turn-projection and overlap. *Language and Speech*, *41*(3–4), 265–294. https://doi.org/10.1177/002383099804100403