

## POSTPRINT

# The body in interaction

## Its multiple modalities and temporalities

Arnulf Deppermann and Jürgen Streeck

### 1. The rising interest in temporalities of bodily interaction

In linguistics, but also in sociology, philosophy, psychology and education, human communication was for a long time considered to be an exchange of signs (Saussure 1959). Neither the materiality of the signs, their sounds, visual shapes, haptic qualities, etc. nor the bodily, psychophysical nature of the participants in a communicative episode mattered to the concept of communication and social interaction. Participants have been treated as sign-printing and -decoding machines, as symbol systems (Newell 1982), which can be reduced to abstract processing modules. The last years, however, have seen an increasing interest in the humanities to move away from abstract understandings of communication to a recognition of the irreducible role of the body in social interaction (Norris 2004; Stivers and Sidnell 2005; Streeck et al. 2011; Deppermann 2013a; Hazel et al. 2014; Nevile 2015). This perspective links up with an increasing recognition of the importance of temporality as a fundamental property of all meaningful conduct (Auer 2009; Streeck and Jordan 2009; Auer and Pfänder 2011; Deppermann and Günthner 2015). The current volume brings together these two lines of research. It asks how social interaction is organized as a multi-modal, multi-sensory process of bodily activities. Each chapter in this volume reveals ways in which participants in social interaction coordinate linguistic and physical action and the expressive modalities of the body (facial display, gaze, gesture, etc.) and incorporate the environment in their activities as they structure and make sense with each other from moment to moment.

Although the topic of this volume articulates a current trend in research on social interaction, there have been important predecessors on whose work this enterprise can build. Using video-data, studies from the school of “context analysis” (Birdwhistell 1970; Schefflen 1972) charted the bodily matrix of kinesic-visible actions and their temporal trajectories in precise detail (for a historical account see Leeds-Hurwitz 2010). Following Schefflen’s lead, researchers such as Kendon (1970, 1990, 2004), and Erickson and Shultz (1979) investigated body motion and changing spatial relations among the interactants’ bodies as “contextualization cues”

(Gumperz 1982, 100). The study of what came to be known as “nonverbal communication” (Argyle 1975), begun during the 1950s (Ruesch and Kees 1972), led to a typology of nonverbal behaviors, provided by Ekman and Friesen (1969), which focused on individual behavior, in line with the psychological orientation of the researchers. Subsequently, much research on embodied behavior has been conducted with a focus on what it reveals about the individual or individual cognitive processes, considered ‘internal’, of which body motion is an ‘external’ expression. This is true also of much of the large amount of research on hand gestures that has been conducted in recent years, especially of the influential work of McNeill (1992) and Goldin-Meadow (2006), who regard gesture as a ‘window onto thought’. Kendon (1990, 2004), however, and others (e.g. Schegloff 1984; Goodwin 2003; Heath and Luff 2007; Streeck 2009a, 2009b) have insisted that gesture must be explained by what it contributes to the interactional process and intersubjective understanding, not what it may or may not directly reveal about ‘mental’ processes. These lines of study have drawn our attention to a rich and diverse range of bodily behaviors contributing to face-to-face interaction and dialogue that had previously been ignored. Psychological research largely relied on codings of the perceived gestural form, without considering that a given gestural form can function quite differently in different interactional contexts: a gesture’s significance cannot be established on the basis of its form alone.

In contrast, interactionist approaches to embodied communication adopted the emic (cf. Pike 1956) perspective of linguistic anthropology and, later, ethnomethodology (Garfinkel 1967) and conversation analysis (Schegloff and Sacks 1973). Identifiable patterns of behavior (or action formats) were investigated with a view to the significance assigned to them by co-participants. Combining an emic, situational perspective with close attention to the fine-grained details of the sequential organization of talk in interaction (Schegloff 1968; Sacks 1992), in the late 1970s and 1980s C. Goodwin (1979, 1980, 1981), M. Goodwin (1980), Heath (1986) and Streeck (1988) put the study of embodied interaction on the conversation analytic agenda. Rather than abstracting a single bodily modality (e.g., gesture) and examining its ‘use’ by individual participants, they aimed to study the entire bodily conduct of participants within activities such as dinner-table-talk, doctor-patient-interaction, peer-instruction, and various other activities in and out of the workplace. While conversation analysis focused almost exclusively on talk (but see Sacks and Schegloff 2002[1975]; Schegloff 1984 for an early interest in gesture) and context analysis on kinesic behaviors and postural relations, the Goodwins, Heath and Streeck began to analyze the full range of multi-modal resources and their coordination in interaction holistically, aiming to understand their mutual coordination and elaboration. No modality was given priority a priori. Instead, all communicative modalities and all other sense-making resources that the participants draw

from are treated by the methodology as potentially being equally relevant for social conduct: how each component contributes to the moment's understanding is an analytic question that must be worked out case by case. Like other modalities, talk is seen as embedded in and produced by bodily action and interaction, but not necessarily as the main, let alone the only, relevant resource for sense-making and interactional organization.

Temporality – the temporal dynamics of body movements – has been a main concern for both context analysts and psychologists studying nonverbal communication from early on. Birdwhistell (1970) and Scheflen (1972) documented changes in the various modalities during the course of an interactional episode in microscopic detail. Kendon developed a system for representing the temporal trajectory of compound 'gesture units' and of the gesture phrases of which they are comprised (preparation, apex/stroke, hold/freeze, retraction; see Kendon 2004). Ekman and Friesen (1978) created the "facial action coding system" to capture facial 'micro-expressions', which may subconsciously inform viewers' attributions of meaning, emotion, and intent. These studies yielded an enormous wealth of knowledge about processes of nonverbal communication, their temporal parameters, and the patterning of co-occurring movements. The turn to situated action in conversation analysis and multimodal analysis led to growing scepticism towards the 'units-and-implementation' view of action that is characteristic, for example, of speech act theory (Searle 1969) and construction grammar (Langacker 1987), but which also seems to inform some understandings of conversation analysis (for a critique see e.g. Ford 2004). The 'units-and-implementation'-view regards types of actions as abstract units, belonging to an inventory of cultural (pragmalinguistic) knowledge, which are implemented in talk as unanalyzed wholes. The detailed study of naturally occurring action, in contrast, reveals the flexible, reflexive and situated nature of practices of actions (cf. Heritage 2010; Deppermann et al. 2016). The practices approach highlights that actions and linguistic structures are produced 'online' and incrementally (Auer 2009). They are sensitive to the ever-changing situated contingencies of their production, using self-regulatory processes of monitoring of self and others, and resting on knowledge which is open-ended and only schematically defined.

## 2. Temporalities of multimodal conduct

The range of temporal orders that matter to interactional conduct extends from centiseconds, as when a face is reconfigured (Ekman and Friesen 1968), to the phylogenetic evolution of meaningful gestures (cf. Mead 1934). Of course, different time scales impinge on multimodal conduct in quite different ways. We can

distinguish different roles temporality plays in the emergence of communicative phenomena (see also Bergmann 1983):

- a. Temporally constituted phenomena (time-objects), like a pause, a stretched syllable, the trajectory of a gesture, or an action sequence;
- b. the temporal coordination and ordering of various phenomena, like the coordination between gesture and speech or the temporal order of various activities produced over the course of the opening of an encounter;
- c. the ways in which activities are affected by various orders of time;
- d. the temporal perspectives of participants on (interactional) events (retrospective vs. prospective, at different temporal distances).

The most intensely studied temporalities concern the temporal trajectories of moment-by-moment coordination by which distinct modal resources – actions, signs, artifacts, etc. – are aligned in situated conduct (a. and b. above). Whereas these temporal orders are interaction-intrinsic, brought about in and by interaction, there are also interaction-extrinsic aspects of temporality which leave their mark on activities. Streeck (this volume) reminds us that an exclusive focus on the micro-temporalities of emergent coordinated actions neglects that larger socio-cultural, individual, and biological time-scales are also in play (c.). Among them are age, biography, the time of the day, times which are normatively allocated to certain activities in a society (Zerubavel 1981), the ‘arc of work’ of extended organizational processes of the division of labor (Strauss 1985), the standardization of time by schedules, calendars and other measurement systems (Zerubavel 1985), the ritualized and rhythmic nature of social time, economies of time in a given society and temporal effects on the body (like fatigue, see also Keevallik this volume; for an overview over concepts of ‘time’ in sociology see Bergmann 1983). While these categories are commonly invoked by observers as explanations for peculiarities of behavior; their relevance may also be indexed by the participants themselves. A case in point are interactional histories of enduring relationships between participants which extend over a number of encounters. They leave their fingerprint on the very design of actions as being performed for the first time in this relationship vs. building on shared experience (Deppermann this volume). They account for the intelligibility of implicit and economical ways of interacting which are opaque for outsiders who do not share this interactional history, i.e. who have not been “growing old together” (Schütz 1974[1932]).

Interactional phenomena are temporally constituted and organized on different levels of granularity: The encounter, the (action) sequence (or a stretch of topical talk), the turn, the turn-constructional unit (TCU, the action). Below, there are further meaningful linguistic units: the syntactic, intonation and gesture phrase,

prosodic and phonetic cesurae (Barth-Weingarten 2016), words, syllables and phonemes. Already Simmel (1916) argued that the meaningfulness of events is tied to a certain degree of granularity, below which the meaning of the phenomenon gets lost. Simmel noted also that together with this mode of holistic perception, temporally constituted units acquire their identity and their meaning by segmentation of the continuous flow of time and by their place in a temporal succession of units. We have to distinguish between physical time and lived time (*la durée*, Bergson 1889; see also Husserl 1928; Schütz 1974[1932]): While physical time is without qualities, but measurable, lived time changes in quality and is structured by the type of event and its experienced tone (Heidegger 1927). ‘Atmospheres’ like urgency (Mondada 2013a; Deppermann in press) or boredom (Keevallik this volume), inform, and are indexed in, interaction by the design of actions. The same applies to normative temporal economies. The normativity of action, often subject to negotiation, does not only concern what to do, how to do it, and who must or can do it, but also when and how long an activity should take. Oshima (this volume) shows this temporal normativity at work in customers’ inspections of the beautician’s work at the end of a hair-cut: The inspection is due at a defined moment and should neither be too short, so that sufficient information needed for a valid assessment can be gained, nor too long, so as not to delay the efficient progression and completion of the activity and making it possible for another client to be served.

From both cognitive-phenomenological and interactional points of view, lived time is an ever-moving present (Husserl 1928). Lived present and immediate future are incessantly transformed into past. The continuity of experience rests on the structure of intentionality. Intentionality in its basic phenomenological sense means ‘aboutness’, ‘the property of being directed towards some object or content’ (Brentano 1874). Intentionality implies self-transcendence of the lived present, which is always surrounded by (constantly changing) protentions of the immediately impending future and retentions of each passing moment (Husserl 1928; Merleau-Ponty 1945). Intentional consciousness is always ahead of itself (cf. Heidegger 1927). This temporal orientation is also constitutive of actions in social interaction (Deppermann and Günthner 2015). Even seeing distant objects involves anticipation, as we can see what we will feel once we come closer and touch it (Gehlen 1988).

The temporal transcendence of interaction units is illustrated by the fact that an important – and sometimes the only – aspect of the meaning and the organizational import of a social action is its projective potential: What it foreshadows, what it makes expectable as a next action (Streeck 1995; Auer 2005). Mead (1934) had already defined the social meaning of an action by its relationship to the response it engenders and the joint project of which it is (an initiating) part.

Projection concerns both the relationship between activities of one participant, e.g., sound-stretches, hesitation phenomena, and filled pauses that project imminent self-repair (Goodwin 1980; Schegloff 2013), and the relationship between the action of one participant and the next action of a responding participant (Sacks et al. 1974; Auer 2005). It is most visibly at work in the mechanism of conditional relevance that accounts for the contingency of adjacency pairs such as summons-answer (Schegloff 1968; cf. Stukenbrock this volume). But projection also concerns other orders than action, for example the anticipated completion of a syntactic structure (and potentially of the turn), the choice of code and style, or the continuation of a topic (cf. Auer 2005). Projections rest on repeated, routine patterns of turn- and action-construction and of sequential trajectories; as expectations, they build on knowledge and memory of past interactional experience. Projection is the active counterpart of the inferential notion of ‘anticipation’: Whereas one behavioral phenomenon is produced to project the occurrence of a next, the next itself can be anticipated by virtue of what has just occurred. Projection and anticipation are fundamental to both understanding and action coordination: In order to be able to coordinate one’s own action with those of others, it is necessary to anticipate both the kind of next action and the moment of its occurrence. E.g., the ability to shift body positions in unison can only be explained by anticipation, which, as Sunakawa (this volume) shows in her study of how a conducting instructor’s motions are shadowed by his students, rests on shared interactional experience and learning. Understanding and action coordination thus build on sedimented patterns of action accumulated through histories of mutual interaction. The fundamental role that projection and anticipation play in interactional collaboration is likely the main reason for the existence of stable, recurrent projection-practices across language communities. As adaptations to the needs for projection and anticipation, they favor re-use of prior solutions, that is, trust in the iterability of solutions for coordination problems that have previously worked under similar circumstances (cf. Feilke 1996). Obviously, underlying bodily and cognitive capacities (memory) are needed for anticipation. Jeannerod (2006) claims that ‘forward models’ of motor cognition, which predict certain patterns of bodily action as next steps in a behavioral sequence, provide for intersubjectivity (see also Morganti et al. 2008). Forward models make use of mirror neurons. They allow for the simulation of others’ current and projected behavior and its imitation (Rizzolati and Craighero 2004). They can be considered the neurophysiological basis of culturally shared action types, resting themselves on associative learning from prior experiences of social interaction (Heyes 2009; Cook et al. 2014) – only what is familiar can be “mirrored”.

Verbal actions are transient vocal phenomena that vanish during production. They are therefore only able to project next actions, but they do not linger on during these next actions. Schmidt (this volume) shows that other modal resources like

placing objects, adopting a posture, or certain gestures allow for the creation of time-enduring, “frozen” structures (see also below). The latter not only project next actions, but prepare material arrangements upon which next actions can be built (e.g. adopting a bodily configuration necessary for some collective action, placing an object at a place in which it is needed for a certain action). The term ‘preparation’ captures this distinct way of how an earlier action enables the occurrence of a later action by virtue of the temporal affordances of the modality in which is carried out (e.g., by posture, which can be “frozen”). The different temporal characteristics of various modalities (i.e. their duration, the possibility of sustaining modal configurations, the production of objectivations that survive the actions that created them) are consequential for the ways in which they can partake in the production of social action and social reality.

As we have said, the identity and meaning of a social action is in part its future-oriented, prospective import, that is, what it foreshadows about what can come next. At the same time, each activity is responsive to its context and has retrospective meaning (Deppermann and Günthner 2015). This Janus-faced nature of actions in interaction is captured by Heritage’s (1984, 242) dictum that every turn-at-talk is both context-shaped (i.e. retrospective) and context-renewing (i.e. prospective). Each turn can be inspected for how it relates to and displays an understanding of prior context (retrospective meaning), the type of action it is (present) and what it projects about the action that can come next (prospective meaning; Deppermann 2015).

Human actions, including “mental actions”, are bodily actions, even when the overt bodily component seems of little significance to the way in which it makes meaning – or is absent altogether (as during the activity we call ‘thinking’, which, after all, has evolved from patterns of motor control; Llinas 2001). As bodily acts, human actions are constrained and shaped by features of the component “systems” involved, each with its own temporal order: vocalizing a syllable takes less time than extending an arm, rendering a linguistic description may take longer than producing a gestural depiction, and so on. As a consequence, coordination is a ubiquitous requirement in all complex forms of human action, including face-to-face communication. The temporality of human actions is in part a function of the materiality of the media and resources involved: while vocal calls and cries and blinks of the eye are well adapted to rapid deployment, a linguistic expression by which an assertion is made or a question put forth takes considerably more time to run its course, which is not a trivial matter because it needs sustained attention by others, who may have to suspend ongoing projects of their own to give it. A clear example of the nexus of materiality and temporality are physical objects, which endure. Maintaining their shape unless subjected to outside forces (such as human action) and thus being immutable for all practical purposes (Latour 1993),

an object can keep meaning available until further notice, that is, until the object and its situated meaning is modified by action.

The term ‘modality’, which has largely replaced the prior, broadcast-inspired term ‘channel’, is not very well defined. Most commonly, it refers to the communicative or expressive modalities of the body: speech, gaze, facial display, hand gestures, posture, manipulation of objects, etc. At the same time, the term ‘modality’ invokes sensory modalities, that is, the senses that are implicated in each communicative modality: gaze relates to vision, posture and gesture are kinesic phenomena, touch is haptic. Each sense connects us to the world in a particular way, and each partakes in a distinct way in which humans make sense when they interact with one another. However, some communicative modalities are actually ‘multimodal’ in themselves: gesture “translates” haptic, tactile, and kinaesthetic experience into visible forms, vision is informed by touch (so that we know how a familiar distant object feels when we see it), and facial display is entirely visual for the perceiver and not visual at all for owner of the face.

‘Multimodality’ is a term originally adopted by the logistics industry to refer to shipments transported via a variety of modes or vessels, by truck, train, ship, and so on (United Nations Conference on Trade and Development 1980). In our field, we similarly tend to think that the modalities that we distinguish each carry a bit of the message, each in its own “modal” way, not only in succession as in the transportation industry, but importantly also simultaneously. By implication, the term ‘modality’ has then been transferred to types of communicative action connected with these senses and the organs associated with them, so that we speak of the ‘modalities’ of gaze, gesture, posture, speech, and so on. Moreover, varieties of human action that involve things, especially artifacts, have also come to be subsumed under a single modality-category (‘uses of objects’). Altogether, these modalities, as well as the varieties of artifact and technology involved in an ongoing activity, are then classified together by the term ‘multimodality’. Nowadays, this term refers to a wide range of different approaches in linguistics, psychology and the social sciences. Their common denominator is to abstain from treating communication as a mono-modal, e.g. only verbal, textual process, but to make the interplay of the different bodily, semiotic, mediated modalities the subject matter of research (for an overview see Bateman et al. 2017). In most of this research, (computer-)mediated communication and the relationship between text and visual modalities (figures, images and moving pictures) is in focus, the approach being primarily inspired by semiotics, psychology and systemic-functional linguistics (Baldry and Thibault 2006; Kress 2010). The term ‘multimodality’ is used to refer to rather heterogeneous phenomena and open to quite different understandings. In the context of this volume, we will take a different approach. We focus on the divergent temporalities of



the orders of phenomena involved in cooperation and symbolic communication in embodied interaction, as well as their coordination and entrainment.

We will continue to use the term ‘modality’ to refer to communication modalities, and we will treat practical action (i.e. the lifting of a pitchfork) as a modality in its own right when we focus on how it contributes to meaning-making, acknowledging that these actions are usually ‘multimodal’ in and of themselves. We use the term ‘multimodal(ity)’ to refer to concurrent actions in different modalities, and ‘modal’ (or ‘modal resource’) whenever we refer to a single modality.

### 3. Indigenous temporal orders of multimodal resources

The different modal resources out of which multimodal interaction is fashioned have different temporal properties. Depending on their materiality, different modes of communicative behavior, as much as the environmental resources that they incorporate, show differences in duration and durability, require different sorts of preparation, and vary in terms of their potential to be coordinated with other behaviors.

We have already noted that talk is fleeting, its transience being the very condition for its efficacy. Talk therefore is a purely presentist medium, needing its own decay for its progression, lest there be echoes of all prior talk (a constraint), but this constraint also accounts for the flexibility of its affordances. But we store the utterances of others in short-term memory and thus are able to perceive and recognize gestalts in time, such as sentence formats and prosodic contours (Couper-Kuhlen 2002). Along with pitch and volume, tempo is one of the three basic parameters involved in prosody. It is the relevant parameter in phenomena such as acceleration (which can index parenthetical and background information, Bergmann 2012; Déhé and Kavalova 2007), deceleration (which can index heightened importance, Gumperz 1982) and rhythm. The study of metrics and rhythm has a long-standing tradition in poetics since antiquity. More recently, interaction researchers including Auer et al. (1999), Condon (1980), and Sczcepek Reed (2006), among others, have shown the importance of shared rhythm for sharing a conversational key and mutual involvement and as an index of mutual understanding and harmony, as well as for establishing a connection with the audience during oral performances (Hymes 1975; Bauman 1975; Tannen 1989). Cognitive researchers and neuroscientists investigate the neural underpinnings of the human body’s ability to keep rhythm and to share it with others (Meltzoff and Prinz 2002). It is well known that temporal phenomena are also most important for turn-construction and turn-taking. ‘Rush-throughs’ (Schegloff 1998, 241) and ‘abrupt-joins’ (Local and

Walker 2004) are ways of turn-keeping that work by systematic deviation from usual pace-patterns of unit-production, -completion and -continuation (speeding up, avoiding micro-pauses which could be understood as transition-spaces, cf. Levinson 2015). Pauses and hitches can index dispreferred actions (Pomerantz 1984), but may also be related to requirements of bodily coordination with talk, for example, the need to establish the proper spatial framework for successful reference (Mondada 2009), and to lapses due to other concurrent bodily activities (Hoey 2015, 2017; Keevallik this volume).

‘Durability’, that is, the possibility to “freeze” some bodily action for some time, is one temporal property with important interactional consequences. Gaze, facial expression, gesture and posture are bodily modalities that allow for the freezing or continuation of an action longer than a speech sound can be extended. The post-stroke hold of a gesture may be used to bracket off a larger semantic unit of speech (Kita et al. 1998), or the persistent claim to the floor during overlap, when the turn is only temporarily yielded to a competitor (Schmitt 2003; Oloff 2013). More generally, freezing a gesture shows that a turn or an action is only suspended, but not abandoned, projecting its resumption at the next possible occasion (Deppermann 2014). When the gesture is made turn-finally, it is used to elicit a response from the interlocutor (Streeck 2007). Streeck (this volume) shows that postures can act as framing devices which index the lingering relevance of an activity type, a topic, or a participation framework over an extended sequence of actions. Especially if the body adopts a rest position, which requires only minimal physical effort to be sustained, bodily configurations can endure for longer stretches of interaction. Extreme durability is achieved by object manipulations that result in altered states of objects (spatial positions, compositions of objects, transformations due to cleaning, cooking, cutting, etc.). These states and the meaning embodied in them can linger on independently of continued human action, and they can impinge on (possibilities for) future actions (cf. Schmidt this volume on the interactional relevance of arrangements of objects in theater rehearsals).

Another temporal property is changeability (mutability): How fast can an actor change from one behavior to another? While facial expressions can change fast, as can the action performed in the talk, spatial positions of the body often take longer to reconfigure in a new spatial arrangement and participation framework needed for some next action, especially if corresponding adaptive alignments by other participants are needed as well (cf. Mondada this volume on questioning, answering and walking in multi-party settings). The different temporal properties of different modal resources thus provide participants with potentials to operate on different orders of interactional conduct (a phoneme, a word, a TCU, a turn, a sequence, an activity type, the whole encounter), which in turn may be used to highlight, elaborate, frame, key, or otherwise contextualize the talk.

#### 4. The temporal coordination of multimodal resources

In each moment, bodily action involves the temporal coordination of all body parts in space. To a considerable degree, bodily practice is organized as holistic multimodal *gestalts* (Mondada 2014a, 2016a). This organization requires intra- and interpersonal coordination (Deppermann and Schmitt 2007; Deppermann 2014). Intrapersonal coordination concerns the synchronous and sequential coordination of the various multimodal resources of one participant, e.g. the coordination of gaze, a pointing gesture and its lexical affiliate when referring to an object (Schegloff 1984; Kendon 2004). Interpersonal coordination concerns the synchronous and sequential coordination of the bodies of several participants *vis-à-vis* each other, e.g. the coordination of speaker's pointing and the gaze of speaker and addressee when referring to an object (see Stukenbrock this volume). Intra- and interpersonal coordination are interdependent, because the temporal relationships between one participant's multiple concurrent "streams" of action are adapted to the ongoing actions of other participants, to what these actions project as next actions, and to the others' receptive capacities at the moment. Interpersonal coordination in multi-party settings can involve the bodily alignment and synchronization of multiple participants beyond those who are individually addressed (see Schmidt this volume for theater rehearsals and Mondada this volume for guided tours).

Human actions emerge as the product of coordination of several component activities, such as gesturing, verbal deixis, and reaching for an object, but the modalities involved can also be devoted to several concurrent activities in which the participant is involved. However, there are clear restrictions on how many modal resources can be combined in a single action project. Deppermann (2014) found combinations of 'looking (+ gesturing) and talking (+ gesturing)', 'listening and looking + pointing', and 'manual action and talk' to be routinely used by paramedics in first aid interactions. A major restriction on multi-activity patterns is how much attention (Deppermann 2014) and physical effort (Keevallik this volume) the main activity requires. Multi-activity can engender uncoordinated responses, because a modality needed may already be occupied with carrying out a different line of action within the multi-activity (see e.g. Oloff this volume on misplaced other-completions because of lack of visual monitoring devoted to manual action of others). Still, multi-activity more often comprises rapid switches back and forth between different activities ("embedded multi-activity", Mondada 2014b, 37–40) than sustained involvement in two activities at the same time (see the papers in Haddington et al. 2014). A particularly striking way of performing embedded multi-activity is body torque (Schegloff 1998), that is, the torquing of the torso towards another participant (object or event) while maintaining the trunk's orientation. Body-torque is a way to coordinate different lines of action, indexing a focal

short-time involvement (by the torso) while confirming the ongoing relevance of a suspended, long-term involvement, to be resumed soon, by the orientation of the trunk (see also Mondada this volume for cases of walking and also Schmitt and Deppermann 2007).

Multimodal coordination can mean synchrony, as in the case of the precision-timed simultaneity of the production of focal accents, lexical items encoding new information, and beat gestures (Schönherr 1997), or in the instruction of rhythmic patterns entraining sound production and movement in dance classes (Keevallik 2015). Most patterns of coordination, however, involve asynchronous relationships between component actions. These can project, prepare, complement, elaborate, modulate, complete, etc. other component actions. Actions often are accomplished by an ordered succession of component acts, each of which can bring additional, environmental sense-making resources into play. Verbal actions can be disambiguated, commented on, mitigated, etc. by post-completion stance-markers (like shrugs or nods, Ford et al. 2012). Bodily actions often precede and foreshadow upcoming verbal actions, which in turn allows recipients to adapt to the projected verbal action. Examples are the use of bodily resources to claim the turn (Mondada 2007: by pointing; Mortensen 2009; Li 2014: by posture shifts), to project various features of the upcoming turn like word-searches (Streeck 1995: by gestures), self-repairs (Rasmussen 2014: by posture shifts), reference (see above), and modal qualities (Streeck 1995, 2009a; Stukenbrock 2010: by iconic gestures; Keevallik 2013: by movements of the whole body), to project an upcoming question (Mondada this volume: by walking towards and catching up with the addressee), the private vs. public character of an answer (Mondada this volume: by walking vs. stopping and bodily orientation), or disaffiliative responses such as disagreement and refusal (Kaukomaa et al. 2014: by frowning; Ursi 2016: by gaze-aversion). The accomplishment of actions by the concerted, precision-timed interplay of various resources takes their different temporalities into account. This becomes evident when some activities are delayed and a slot is created for the contextually fitted execution of others. Verbal production can be delayed to allow the unfolding of complex iconic gestures (Kendon 2004), writing (Mondada and Svinhufvud 2016), or the repositioning of the body and the reconfiguration of interactional space in direction-giving (Mondada 2009). Inversely, pointing can be sustained over longer stretches of referential talk (Kendon 2004). Intrapersonal coordination is closely linked to inter-personal concerns, for example when a fledgling action is delayed to enable joint attention (Goodwin 1980; Heath 1986; Kendon 2004; Mondada 2009). Just like verbal actions – and often in combination with them –, gestures and gaze project next actions that are expected from recipients, such as taking the turn (Streeck 2007; Bohle 2007), gazing at the speaker's gesture (Streeck 1995), or identifying an object (Stukenbrock this

volume). Handling an object may not make a material contribution to preparing the next action (Schmidt this volume), but nevertheless carry the activity forward by virtue of the projection that it makes as a symbolic act (Deppermann et al. 2010: a folder in a meeting; Mondada 2015: a map).

Interpersonal coordination, however, does not only start when a next participant responds to a complete prior turn or action by their partner. Beginning in the late 1970s, the Goodwins have shown how the accomplishment of interactional structure rests on quasi-simultaneous, microsequential actions of speakers and hearers within the scope of the production of a turn (C. Goodwin 1979, 1980, 1981; M.H. Goodwin 1980). Both the meaning and the emergent construction of a turn can often be shown to be joint, intersubjectively organized achievements. Once we are sensitive to the concerted, composite nature of actions, we will notice that a segmental, building-block understanding of actions as bounded units is rather inadequate (cf. also Ford 2004 for a critique). Beginnings, transitions, and closings are cases in point: they are fuzzy and often brought about by extended, tentative negotiations. This is true both on a macroscopic level, where we might want to know when, precisely, an encounter begins (see Mondada and Schmitt 2010), and on a microscopic level where we examine the step-by-step, multimodal beginnings of turns at talk (Schegloff 1996; Deppermann 2013b). The same applies to closings of turns, sequences, and whole encounters, even though the latter are typically bounded by verbal closings, withdrawals of objects, turning away, walking away, and other overt maneuvers (Schmitt and Deppermann 2010; Broth and Mondada 2013; Mondada 2015).

## 5. Methodological requirements of studying the temporalities of multimodal interaction

Studying bodily interaction among participants in co-presence requires video-recordings (Heath et al. 2010; Mondada 2013b). This is a corollary of the constraint to use data which allow the analyst to have access to the same actions and events as the participants in the encounter had. This is necessary in order to have a comprehensive documentation of their actions and the resources they are using; in addition, we also have to be able to analyze what they can perceive and what they are responding to. If we use only field notes from participant observation or similar methods (e.g. autoethnographic documents), a detailed account of the precise temporal unfolding of bodily action is impossible and action analyses will risk to be prematurely constrained by the interpretive schemata of the researcher while other relevant events are ignored. In contrast to the set-up of recordings for psychological research or for broadcasting, it is necessary not to focus on single

actors (i.e. patients, speakers, protagonists), but to document the whole interactive configuration, including the interactional space to which participants orient in their encounter. This often requires the use of several cameras in order to capture all relevant perspectives in sufficient detail.

None of the data-analyses in this volume, however, rest on video-recordings alone. The authors also work with transcripts produced according to the common standards for vocal transcription used in Conversation Analysis (Hepburn and Bolden 2017). Transcripts are amended by more or less detailed annotations of visual conduct, mostly using the conventions established by Mondada (2018). Conventions for multimodal transcription capture visible kinesic activities in a descriptive language that is neither physicalistic nor overly interpretive. It should allow for the observational intelligibility of participants' actions, without attributing meanings, motives and intentions that can be accounted for only by detailed data analysis. Most importantly, multimodal transcripts also represent the trajectories of activities, i.e. their phases of preparation, apex, possibly freezing, retraction and transition (cf. Kendon 2004), and the precise temporal coordination with onsets and offsets of other activities (Mondada 2018). While these formal properties make for a set of parameters that are fairly general and apply to virtually every activity, transcripts are necessarily selective. It is impossible, and would not make sense, to try to transcribe „everything“ that is visible on a video-screen. What to transcribe depends on which aspects of visual behaviors the participants make relevant in their interaction and the research question. The production of a multimodal transcripts therefore already contains much more analysis of what the participants were doing than a purely auditory transcript, whose production is constrained to a much higher degree solely by the transcription conventions.

Reading and producing multimodal transcription requires training and forces the analyst to represent relevant activities in a concise way within the confines of the transcript line. A more detailed description can be given in the text of the analysis. The advantage of transcribing (and not only describing) multimodal conduct lies in the fact that only transcription allows for a continuous representation of the precise temporal conduct of multimodal activities. For the analyst, this is a precondition for tracking the trajectories of, and the relationships between multiple modal resources that the parties to an interaction draw upon at certain ordered points. For the reader, the transcript allows to check the transcription in a more rigorous way than is possible if only descriptions of selected interactional moments are provided.

While transcripts and stills from video-recordings are used in all contributions to this volume, other more invasive methods of data-collection which are common in experimental set-ups like camera glasses or motion capturing are avoided in studies with a conversation analytic background. This is because a prime objective

here is to preserve the naturalness of social setting being studied. The only exception is the contribution by Stukenbrock (this volume), in which a part of the study builds on data from mobile eye-tracking.

## **6. Consequences of a temporal and multimodal perspective on social interaction for Conversation Analysis (CA)**

From its very beginning, temporality has played a major role in conversation analysis. The notion of ‘sequentiality’ lies at the heart of most of its main concepts, such as turn-taking, sequence, conditional relevance, and adjacency pair (cf. Schegloff 2007). The two leading analytical questions that conversation analysts seek to answer, “why that now?” (Sacks 1984) and “what’s next?” (Schegloff 2007), point to the irreducibly temporal nature of conversational phenomena. Importantly, conversation analysis has revoked the conception of linguistic units as timeless abstractions (types) whose real-life temporal properties are irrelevant to its structural features. Their position, composition, and meaning are all temporally organized phenomena (see also Cliff 2016). By insisting to focus on single turns only within their sequential contexts, CA revealed turns to be both retrospectively oriented and context-shaped on the one hand, and prospectively oriented and context-renewing (Heritage 1984, 238–242) on the other. The threefold temporal indexicality of turns (interpreting the interactional past, performing a social act in the present, and thereby projecting an interactional future) accounts for why the meaning of a single turn in interaction is already multifold (Deppermann 2015). The sequential organization of interaction has been revealed as the basic mechanism on which the accomplishment of intersubjectivity in conversational time rests (Schegloff 1992; Sidnell 2014; Deppermann 2015). While attention to sequentiality is the hallmark of conversation analysis, simultaneity plays only a minor role in its investigations. It has been studied mainly in the context of overlap, which initially appeared to be a deviant case of turn-organization, but upon analysis can be explained largely by the working of the sequential principles of the turn-taking-system itself (cf. Sacks et al. 1974; Jefferson 2004; see also Egbert 1997).

Even though conversation analysis is known for its focus on verbal interaction, already in the 1970s its founders showed an interest in visual phenomena as well (see Sacks and Schegloff 2002[1975] on gestures’ home position; Schegloff 1984 on gestures and their role in the projection of talk to come). When interactional linguistics developed as a distinctive strand of CA-inspired linguistic research, prosody and conversational phonetics became major topics (Couper-Kuhlen and Selting 1996, 2001, 2004; but see also Schegloff 1998), which added vocal phenomena to the analysis of talk. Early research on multimodal interaction was strongly

influenced by conversation analysis (e.g. Goodwin 1981; Heath 1986), and much groundbreaking research on multimodal interaction carried out today is informed by concepts, findings, and methodological strategies of conversation analysis. Much research on multimodal interaction is conceived as a multimodal enlargement of the conversation analytic approach (Stivers and Sidnell 2005; Heath et al. 2010; Deppermann 2013a; Heath and Luff 2013; Mondada 2014a, 2016a; Hazel et al. 2014; Neville 2015). The papers in this volume also build, albeit to different degrees, on CA-concepts and -methodology.

To capture the full phenomenological richness and constitutional grounds of bodily interaction, however, it is not sufficient to simply enlarge or elaborate the analytic framework of conversation analysis. Instead, several fundamental assumptions need to be revised. We will discuss three issues:

- a. Speakers and hearers,
- b. Sequential organization,
- c. Turns and actions.

It has sometimes been argued that CA has a verbal bias (e.g. Hazel et al. 2014; Schmitt 2015). The tendency to neglect the relevance of other modalities than talk for interactional conduct becomes most evident in the use of the terminology of ‘speaker and hearers’. Already Goffman (1979) criticized this terminology for being insensitive to the different ways in which both speakers and recipients can be aligned to some stretch of talk. In particular, Goodwin’s insights into the active role of the recipient shows how the notion of ‘hearer’ is misleading in underrating the importance of their ongoing behavior for the co-construction of turns-at-talk (Goodwin 1981). Instead of assigning talk a priori a privileged position, only data-analysis can reveal which modalities matter (most) in accomplishing an interactional episode (Mondada 2016a, 2018). Therefore ‘participant’ is the more fundamental notion than ‘speaker’ and ‘hearer’ (cf. Goodwin and Goodwin 2004). By referring to participants, one is able to inspect both verbal and non-verbal activities for the ways in which they contribute to the interactional construction of structure and meaning. The notion of ‘participants’, moreover, does justice to the fact of the simultaneity of interactional participation (cf. Goodwin 1981, 2017): Non-speaking participants can simultaneously express stances, be involved in silent activities, display their (non-)alignment with talk in various ways by gaze, facial expression, posture, etc. Their simultaneous bodily behavior thus is co-constitutive of the interaction, which therefore emerges as a joint accomplishment (see Deppermann and Schmitt 2007). The insufficiency of the ‘speaker-hearer’-dichotomy and the indispensability of a notion of ‘participation’ has been shown to be particularly important



- when participants are involved in multiple activities in the same interactional episode (Haddington et al. 2014). Multiple involvements are usually dealt with by using different modal resources (see above);
- when talk is not the only focus of interaction, but talk only asks for, instructs, comments on, etc. bodily actions or if the main activity even does not involve barely any talk (like in athletic sports, Meyer et al. 2017);
- in multi-party settings in which different participants contribute to and are aligned with some line of action in different ways at the same time (see Schmidt this volume; Oloff this volume);
- when physical objects (tools or objects of shared concern; Nevile et al. 2014), spatial arrangements (Mondada 2009; Hausendorf et al. 2016), and movement between places (Haddington et al. 2018, 2013; Deppermann in press; Mondada this volume) matter to interactional organization.

Focus on sequential organization has always been the main focus of CA (cf. Schegloff 2007), and the most distinctive and ground-breaking findings of classic CA-studies concern this order. The phenomena which the study of bodily interaction discovers to be central, namely simultaneity, multi-party settings, the materiality of multi-modal resources, multi-activity, objects, space and movement, however, are bound to affect and alter our view of sequentiality. The importance of the simultaneity of activities of various participants (Goodwin 1981; Iwasaki 2011, this volume), the relevance of continuously monitoring others' actions (M.H. Goodwin 1980; Heath 1986; Heath and Luff 1996; Heath et al. 2002) while co-constructing and adapting actions on the fly, speak against a building block model of interactional structure as solely consisting of sequentially enchainned, paired actions. Rather, a dynamic, more fluid picture of interactional structures as emerging over time coalesces, which shows actions as joint products requiring continuous response, monitoring, and mutually affecting, moment-by-moment micro-sequentiality. Stukenbrock (this volume) proposes that microsequential relationships within an ongoing turn be considered as multimodal adjacency pairs. The example she discusses are deictic elements that serve as summons by which the recipient's gaze – in other words: a response – is solicited. This systematically projected action-response pair does not involve speaker-change. In addition to the collaborative construction of turns and turn-components (see also Goodwin 2017), early responses, i.e. the onset of the production of a second pair part (e.g. an instructed action) already during the production of the first (e.g. Broth and Keevallik 2014), in unison or even before the first are cases in point. In her study of bodily shadowing of conductors' instructions, Sunakawa (this volume) shows how instructed simultaneous action is mediated by the shared rhythm of music. Like Deppermann (this volume) shows for turn-construction, Sunakawa shows for sequential organization that the emergence

of new interactional patterns (here: simultaneous bodily shadowing instead of sequentially following a lead instruction) build on shared interactional histories and the development of shared routines. Two orders of time, the online organization of interaction and the larger interactional histories beyond the single encounter, thus intersect here. Interactional history and the type of interaction in play both contribute contexts that affect how the first pair-part of a sequence is constructed (Deppermann this volume, 2018). First pair-parts do not appear out of nowhere, but are themselves tailored to, and occasioned by, prior context. A larger temporal perspective thus allows us to widen our notion of responsivity and retrospectivity of interactional conduct beyond the confines of the organization of positions within sequences. The complexity of different temporal orders that can simultaneously be in play in interactional conduct also invite us to reconsider the prospective orientation of the participants. Oshima (this volume) shows this with respect to the preference for progressivity. What counts as an aptly progressive action is assessed differently in relation to different temporal orders (sequential expectations, interaction-type related expectations, topical relevancies, time-constraints for an interactional episode), which can lead to pragmatic dilemmas and conflicting stances of participants.

What applies to quasi-simultaneous actions is also valid more generally for action sequences. The mechanisms of projection and conditional relevance operate in embodied action in much the same way as they do in verbal action. Neither the first nor the second pair-part within a sequence needs to be verbal, both can be completed by a non-linguistic bodily act (e.g. Clark 2012; Keisanen and Rauniomaa 2012; Mondada 2013a, 2014c; Rossi 2015; Deppermann 2018). The complexity of multi-party interaction and multi-activity also impinges on sequential organization: for example, second pair-parts may be complex when multiple recipients are expected to respond in different, but coordinated ways to a first pair-part (Schmidt this volume). In multi-activity contexts, when talk is coordinated with manual labor (Keevallik this volume) or with driving (Mondada 2012), the autonomy and integrity of the sequential organization of each component activity is in question. The lack and latency of responses to talk during bodily work is taken by Keevallik (this volume) as proof of the non-autonomy of talk in certain kinds of bodily practice. Multi-activity can drastically alter talk-related sequential expectations, for example by creating tolerance for silences and lapses, and the bodily activity itself accounts for – explains and justifies – the absence of a response (Hoey 2015).

For CA, the turn-taking system is the most basic organizational order of interactional organization – the organization of sequences can be regarded as a specialization of and restriction of turn-taking by defining which kind of action is due after turn-transition (Schegloff 1968). The representation of patterns of sequential organization sometimes seems to convey the idea that producing a turn just

means to implement a local instance of an action type. In contrast, the temporal perspective of online-syntax (Auer 2009) posits that turn-construction is a process of online production. It argues against an understanding of TCUs and turns as segmental units (Auer 2010). Instead, they are shown to be emergent, continuously unfolding, not fully planned units – they are produced in increments (see also Auer 1996). Whereas in classical CA, interactivity and intersubjectivity are handled by the turn-taking system (cf. the next turn proof procedure, Sacks et al. 1974, 728) and sequential organization (Schegloff 1992), studies of the relevance of parallel recipient activities provide evidence that even the TCU itself is a locus of interactive production and intersubjectivity. Iwasaki (this volume) shows how delays uttered after the production of an assessable, but before the overt production of the assessment, routinely elicit recipients' stances, which can then be incorporated into the further construction of the ongoing turn. Again, this collaborative production is possible only by virtue of the multimodal participation of the addressee, who can rely on posture, gaze, head gestures, and facial expression, in shaping the response. Oloff's study of delayed completions (this volume) shows that reaching completion of a turn can itself be the result of a negotiation comprising multiple communication modalities. Again, resources beyond syntax are crucial for signaling what counts as a legitimate part of a turn.

It is undisputed by now that various communicative resources, as well as their combinations, can play roles in turn-taking in face-to-face interaction (e.g. Goodwin 1980; Schmitt 2005; Mondada 2007; Rossano 2013; Deppermann 2013b); turns often have a composite structure, crucially involving gaze, gesture and facial expression (Enfield 2009). Still, the notion of 'multimodal turns' (see e.g. Ford et al. 2012; Keevallik 2014) seems to be problematic, at least if it is taken to mean that non-vocal actions (like granting a request by handing an object or making a bid for a turn by raising a hand) are to be regarded as turns as well. When the notion of 'turn' is abstracted from the verbal-vocal domain, fundamental principles of turn-organization do not apply anymore: it makes no sense to conceive of non-vocal bodily actions as 'inhabiting the floor', or as 'overlapping' each other. Nor can they in many cases be neatly segmented into discrete TCUs. Instead of assimilating bodily actions to turns, the challenge rather consists in relating the organization of (bodily) action to the organization of sequences of turns. Both orders do seem to fall into one wherever interaction consists only of verbal turns (as in phone-calls), but they may be configured according to quite different principles whenever participants are visually available to each other, especially when more than two parties are involved. The relationship between turns and (bodily) actions has many facets that are still underexplored. Unresolved issues concern the constitution and identity criteria of the temporal gestalts of multimodal actions, the role of the body in the formation and interpretation of verbal acts (see Mondada

this volume), the internal composite structure of turns, the mapping of boundaries and phases of verbal TCUs onto boundaries of bodily acts, and the action-status of bodily micro-activities that often do not seem to be accountable and meaningful as such (like some kinds of gesticulation, swallowing or postural shifts), but merely serve organizational purposes or physiological needs or are otherwise, preparatory or supportive of ongoing accountable actions. While an action may well cover a larger stretch of talk than a single TCU (Levinson 2013), a TCU also may include micro-actions such as referring or calling for attention (Stukenbrock this volume) or affiliation (Iwasaki this volume) that could not constitute full turns. That the relationship between turns and actions is highly complex has become apparent in studies of openings (Mondada and Schmitt 2010) and closings (Broth and Mondada 2013; Mondada 2015). These studies have proved the important role of multimodal preparations of openings and closing, of embodied availability-displays and monitoring (see also Heath and Luff 2000; Schmitt and Deppermann 2007), and of mutual perception and joint attention – all factors that cannot be accommodated by a turn-taking view of social interaction.

In which ways and to what extent there is a need for a fundamental revision of CA-concepts will finally also depend on what conversation analysis considers its domain of inquiry: Is it talk-in-interaction or social action more generally? Initially the choice of audio recordings was opportunistic and not motivated by a particular interest in conversation (cf. Sacks 1984). However, the strength of CA-concepts lies in their close correspondence to the data at hand, and these have for a long time primarily been audio-data with talk as the participants' prime concern (see also Keevallik this volume). Now that video-data have become available that allow for a rigorous study of a heterogeneous array of social actions, including kinesic-visual actions performed without talk, the choice will be to either stay with the original domain of study (talk-in-interaction), but including a richer and more comprehensive data basis of video-recordings with an enhanced understanding of bodily interaction. The other option is to revise the definition of our domain of inquiry precisely because of the availability of video-data. Depending on this choice, the implications for the turn to multimodal interaction will be quite different. Whereas in the first case, a refinement and some revisions of CA-concepts will be necessary, in the latter case, a more thorough search for renewing foundational concepts and for rethinking of the subject matter will be in order.

## References

- Argyle, Michael. 1975. *Bodily Communication*. London: Methuen.
- Auer, Peter. 1996. "On the Prosody and Syntax of Turn-continuations." In *Prosody in Conversation*, ed. by Elizabeth Couper-Kuhlen, and Margret Selting, 57–100. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511597862.004
- Auer, Peter. 2005. "Projection in Interaction and Projection in Grammar." *Text* 25 (1): 7–36. doi:10.1515/text.2005.25.1.7
- Auer, Peter. 2009. "On-line Syntax: Thoughts on the Temporality of Spoken Language." *Language Sciences* 31: 1–13. doi:10.1016/j.langsci.2007.10.004
- Auer, Peter. 2010. *Zum Segmentierungsproblem in der gesprochenen Sprache*. Bayreuth: Universität, InLiSt 49: 1–19.
- Auer, Peter, Couper-Kuhlen, Elizabeth, and Frank E. Müller. 1999. *Language in Time – The Rhythm and Tempo of Verbal Interaction*. Oxford: Oxford University Press.
- Auer, Peter, and Stefan Pfänder (eds). 2011. *Constructions: Emerging and Emergent*. Berlin: de Gruyter. doi:10.1515/9783110229080
- Baldry, Anthony, and Paul J. Thibault. 2006. *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook*. London: Equinox.
- Barth-Weingarten, Dagmar. 2016. *Intonation Units Revisited – Cesuras in Talk-in-Interaction*. Amsterdam: Benjamins.
- Bateman, John, Wildfeuer, Janina, and Tuomo Hiippala. 2017. *Multimodality. Foundations, Research and Analysis – A Problem-oriented Introduction*. Berlin: de Gruyter.
- Bauman, Richard. 1975. "Verbal Art as Performance." *American Anthropologist* 77 (2): 290–311. doi:10.1525/aa.1975.77.2.02a00030
- Bergmann, Werner. 1983. "Das Problem der Zeit in der Soziologie : ein Literaturüberblick zum Stand der "zeitsoziologischen" Theorie und Forschung." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 35 (3): 462–504.
- Bergmann, Pia. 2012. "The Prosodic Design of Parentheses in Spontaneous Speech." In *Prosody and Embodiment in Interactional Grammar*, ed. by Bergmann, Pia, Brenning, Jana, Pfeiffer, Martin, and Elisabeth Reber, 103–141. Berlin: de Gruyter.
- Bergson, Henri. 1889. *Essai sur les données immédiates de la conscience*. Paris: Alcan.
- Birdwhistell, Ray L. 1970. *Kinesics and Context: Essays on Body Motion Communication*. Philadelphia: University of Pennsylvania Press.
- Bohle, Ulrike. 2007. *Das Wort ergreifen – Das Wort übergeben. Eine explorative Studie zur Rolle redegleitender Gesten in der Organisation des Sprecherwechsels*. Berlin: Weidler.
- Brentano, Franz. 1874. *Psychologie vom empirischen Standpunkt*. Leipzig: Duncker and Humblot.
- Broth, Mathias, and Leelo Keevallik. 2014. "Getting Ready to Move as a Couple: Accomplishing Mobile Formations in a Dance Class." *Space and Culture* 17 (2): 107–121. doi:10.1177/1206331213508483
- Broth, Mathias, and Lorenza Mondada. 2013. "Walking Away. The Embodied Achievement of Activity Closings in Mobile Interactions." *Journal of Pragmatics* 47: 41–58. doi:10.1016/j.pragma.2012.11.016
- Clark, Herbert H. 2012. "Unspoken Questions, Unspoken Answers." In *Questions: Formal, Functional and Interactional Perspectives*, ed. by de Ruiter, Jan-P., 81–102. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139045414.006
- Clift, Rebecca. 2016. *Conversation Analysis*. Cambridge: Cambridge University Press.

- Condon, William S. 1980. "The Relation of Interactional Synchrony to Cognitive and Emotional Processes." In *The Relationship of Verbal and Nonverbal Communication*, ed. by Mary R. Key, 49–66. The Hague: Mouton.
- Cook, Richard, Bird, Geoffrey, Catmur, Caroline, Press, Clare, and Cecilia Heyes. 2014. "Mirror Neurons: From Origin to Function." *Behavioral and Brain Sciences* 37 (2): 177–192. doi:10.1017/S0140525X13000903
- Couper-Kuhlen, Elizabeth. 2002. "Prosody." In *Handbook of Pragmatics*, ed. by Jef Verschueren, Jan-Ola Östman, Jan Blommaert, and Chris Bulcaen, 1–19. Amsterdam: Benjamins.
- Couper-Kuhlen, Elizabeth, and Margret Selting (eds.). 1996. *Prosody in Conversation. Interactional Studies*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511597862
- Couper-Kuhlen, Elizabeth, and Margret, Selting (eds.). 2001. *Studies in Interactional Linguistics*. Amsterdam: Benjamins.
- Couper-Kuhlen, Elizabeth, and Cecilia E. Ford (eds). 2004. *Sound Patterns in Interaction: Cross-linguistic Studies from Conversation*. Amsterdam: Benjamins. doi:10.1075/tsl.62
- Dehé, Nicole, and Yordanka Kavalova (eds). 2007. *Parentheticals*. Amsterdam: Benjamins. doi:10.1075/la.106
- De Saussure, Ferdinand. 1959. *Course in General Linguistics*. New York: Philosophical Library.
- Deppermann, Arnulf (ed). 2013a. "Conversation Analytic Studies of Multimodal Interaction." *Journal of Pragmatics* 46: 1–7. doi:10.1016/j.pragma.2012.11.014
- Deppermann, Arnulf. 2013b. "Turn-Design at Turn-Beginnings: Multimodal Resources to Deal with Tasks of Turn-Construction in German." *Journal of Pragmatics* 46 (1): 91–121. doi:10.1016/j.pragma.2012.07.010
- Deppermann, Arnulf. 2014. "Multi-modal Participation in Simultaneous Joint Projects: Interpersonal and Intra-personal Coordination of Paramedics in Emergency Drills." In *Multitasking in Social Interaction: Beyond Multitasking*, ed. by Haddington, Pentti, Keisanen, Tiina, Mondada, Lorenza, and Maurice Nevile, 247–281. Amsterdam: Benjamins.
- Deppermann, Arnulf. 2015. "Retrospection and Understanding in Interaction." In *Temporality in Interaction*, ed. by Deppermann, Arnulf, and Susanne Günthner, 57–94. Amsterdam: John Benjamins.
- Deppermann, Arnulf. (2018). "Instruction Practices in German Driving Lessons: Differential Uses of Declaratives and Imperatives." *International Journal of Applied Linguistics*. doi:org/10.1111/ijal.12198
- Deppermann, Arnulf. this volume. Changes in Turn-design over Interactional Histories – the Case of Instructions in Driving School Lessons.
- Deppermann, Arnulf, and Reinhold Schmitt. 2007. "Koordination. Zur Begründung eines neuen Forschungsgegenstandes." In *Koordination. Analysen zur multimodalen Interaktion*, ed. by Schmitt, Reinhold, 15–54. Tübingen: Narr.
- Deppermann, Arnulf, and Susanne Günthner. 2015. "Introduction: Temporality in Interaction." In *Temporality in Interaction*, ed. by Deppermann, Arnulf and Susanne Günthner, 1–24. Amsterdam: Benjamins.
- Deppermann, Arnulf, Schmitt, Reinhold, and Lorenza Mondada. 2010. "Agenda and Emergence: Contingent and Planned Activities in a Meeting." *Journal of Pragmatics* 42 (6): 1700–1718. doi:10.1016/j.pragma.2009.10.006
- Deppermann, Arnulf, Feilke, Helmuth, and Angelika Linke. 2016. "Sprachliche und kommunikative Praktiken: Eine Annäherung aus linguistischer Sicht." In *Sprachliche und kommunikative Praktiken. Jahrbuch des Instituts für Deutsche Sprache 2015*, ed. by Deppermann, Arnulf, Feilke, Helmuth, and Angelika Linke, 1–23. Berlin: de Gruyter.

- Egbert, Maria. 1997. "Schisming: The Collaborative Transformation from a Single Conversation to Multiple Conversations." *Research on Language and Social Interaction* 30: 1–51. doi:10.1207/s15327973rlsi3001\_1
- Ekman, Paul, and Wallace V. Friesen. 1968. "Nonverbal Behavior in Psychotherapy Research." *Research in Psychotherapy* 3: 179–216. doi:10.1037/10546-011
- Ekman, Paul, and Wallace V. Friesen. 1969. "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding." *Semiotica* 1 (1): 49–98. doi:10.1515/semi.1969.1.1.49
- Ekman, Paul, and Wallace V. Friesen. 1978. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto CA: Consulting Psychologists Press.
- Enfield, Nick J. 2009. *The Anatomy of Meaning*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511576737
- Erickson, Frederick, and Jeffrey Shultz. 1982. *The Counselor as Gatekeeper: Social Interaction in Interviews*. New York: Academic.
- Feilke, Helmuth. 1996. *Sprache als soziale Gestalt*. Frankfurt/Main: Suhrkamp.
- Ford, Cecilia E. 2004. "Contingency and Units in Interaction." *Discourse Studies* 6 (1): 27–52. doi:10.1177/1461445604039438
- Ford, Cecilia E., Thompson, Sandra A., and Veronika Drake. 2012. "Bodily-Visual Practices and Turn Continuation." *Discourse Processes* 49 (3–4): 192–212. doi:10.1080/0163853X.2012.654761
- Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. New York: Prentice-Hall.
- Gehlen, A. 1988 (1958). *Man. His Nature and Place in the World*. New York: Columbia University Press.
- Goffman, Erving. 1979. "Footing." *Semiotica* 25 (1–2): 1–29. doi:10.1515/semi.1979.25.1-2.1
- Goldin-Meadow, Susan. 2006. "Talking and Thinking with our Hands." *Current Directions in Psychological Science* 15: 34–39. doi:10.1111/j.0963-7214.2006.00402.x
- Goodwin, Charles. 1979. "The Interactive Construction of a Sentence in Natural Conversation." In *Everyday Language: Studies in Ethnomethodology*, ed. by Psathas, George, 97–121. New York: Irvington.
- Goodwin, Charles. 1980. "Restarts, Pauses, and the Achievement of Mutual Gaze at Turn-Beginning." *Sociological Inquiry* 50 (3–4): 272–302. doi:10.1111/j.1475-682X.1980.tb00023.x
- Goodwin, Charles, 1981. *Conversational Organization. Interaction between Speakers and Hearers*. New York: Academic.
- Goodwin, Charles. 2003. "Pointing as Situated Practice." In *Pointing: Where Language, Culture and Cognition Meet*, ed. by Kita, Sotaro, 217–241. Mahwah, NJ: Lawrence Erlbaum.
- Goodwin, Charles. 2017. *Co-operative Action*. Cambridge: Cambridge University Press.
- Goodwin, Charles, and Marjorie H. Goodwin. 2004. "Participation." In *A Companion to Linguistic Anthropology*, ed. by Duranti, Alessandro, 222–244. Malden, MA: Blackwell.
- Goodwin, Marjorie H. 1980. "Processes of Mutual Monitoring Implicated in the Production of Description Sequences." *Sociological Inquiry* 50 (3–4): 303–17. doi:10.1111/j.1475-682X.1980.tb00024.x
- Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511611834
- Haddington, Pentti, Keisanen, Tiina, and Maurice Nevile (eds). 2012. "Meaning in Motion: Interaction in Cars." *Semiotica* 191 (1–4): 171–198.
- Haddington, Pentti, Mondada, Lorenza, and Maurice Nevile (eds). 2013. *Interaction and Mobility. Language and the Body in Motion*. Berlin: De Gruyter. doi:10.1515/9783110291278
- Haddington, Pentti, Keisanen, Tiina, Mondada, Lorenza, and Maurice Nevile (eds). 2014. *Multiactivity in Social Interaction: Beyond Multitasking*. Amsterdam: Benjamins.

- Hausendorf, Heiko, Schmitt, Reinhold, and Wolfgang Kesselheim (eds.). 2016. *Interaktionsarchitektur, Sozialtopographie und Interaktionsraum*. Tübingen: Narr.
- Hazel, Spencer, Mortensen, Kristian, and Gitte Rasmussen (eds.). 2014. "A Body of Resources – CA Studies of Social Conduct." *Journal of Pragmatics* 65: 1–9. doi:10.1016/j.pragma.2013.10.007
- Heath, Christian, and Paul Luff. 1996. "Convergent Activities: Line Control and Passenger Information on the London Underground." In *Cognition and Communication at Work*, ed. by Engeström, Yngwe, and David Middleton, 96–129. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139174077.005
- Heath, Christian, and Paul Luff. 2000. *Technology in Action*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511489839
- Heath, Christian, and Paul Luff. 2007. "Gesture and Institutional Interaction: Figuring Bids in Auctions of Fine Art and Antiques." *Gesture* 7 (2): 215–240. doi:10.1075/gest.7.2.05hea
- Heath, Christian, and Paul Luff. 2013. "Embodied Action and Organizational Activity." In *The Handbook of Conversation Analysis*, ed. by Sidnell, Jack, and Tanya Stivers, 283–307. Oxford: Wiley-Blackwell.
- Heath, Christian, Sanchez Svensson, Marcus, Hindmarsh, Jon, Luff, Paul, and Dirk vom Lehn. 2002. "Configuring Awareness." *Computer-Supported Cooperative Work* 11 (3–4): 317–347. doi:10.1023/A:1021247413718
- Heath, Christian. 1986. *Body Movement and Speech in Medical Interaction*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511628221
- Heath, Christian, Hindmarsh, Jon, and Paul Luff. 2010. *Video in Qualitative Research*. London: Sage.
- Heidegger, Martin. 1927. *Sein und Zeit*. Tübingen: Niemeyer.
- Hepburn, Alexa, and Galina B. Bolden 2017. *Transcribing for Social Research*. London: Sage.
- Heritage, John. 1984. *Garfinkel and Ethnomethodology*. Oxford: Polity.
- Heritage, John. 2010. "Conversation Analysis: Practices and Methods." In *Qualitative Sociology*, ed. by Silverman, David, 3rd ed., 208–230. London: Sage.
- Heyes, Cecilia. 2009. "Where do Mirror Neurons Come from." *Neuroscience and Biobehavioral Reviews* 343 (4): 575–583.
- Hoey, Elliott. 2015. "Lapses: How People Arrive at, and Deal with, Discontinuities in Talk." *Research on Language and Social Interaction* 48 (4): 430–453. doi:10.1080/08351813.2015.1090116
- Hoey, Elliott. 2017. "Sequence Recompletion: A Practice for Managing Lapses in Conversation." *Journal of Pragmatics* 109: 47–63.
- Husserl, Edmund. 1928. *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins*. Halle/Saale: Niemeyer.
- Hymes, Dell. 1975. "Breakthrough into Performance." In *Folklore: Performance and Communication*, ed. by Ben-Amos, Dan, and Kenneth S. Goldstein, 11–74. Den Haag: Mouton. doi:10.1515/9783110880229.11
- Iwasaki, Shimako. this volume. *Suspending Talk: Multimodal Organization of Participation and Stance in Japanese*.
- Iwasaki, Shimako. 2011. "The Multimodal Mechanics of Collaborative Unit Construction in Japanese Conversation." In *Embodied Interaction. Language and Body in the Material World*, ed. by Streeck, Jürgen, Goodwin, Charles, and Curtis LeBaron, 106–120. Cambridge: Cambridge University Press.
- Jeannerod, Marc. 2006. *Motor Cognition*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198569657.001.0001



- Jefferson, Gail 2004. "A Sketch of Some Orderly Aspects of Overlap in Conversation." In *Conversation Analysis. Studies from the First Generation*, ed. by Lerner, Gene H., 43–59. Amsterdam: Benjamins. doi:10.1075/pbns.125.05jef
- Kaukoma, Tiimo, Peräkylä, Anssi, and Johanna Ruusuvauro. 2014. "Foresadowing a Problem: Turn-Opening Frowns in Conversation." *Journal of Pragmatics* 71: 132–147. doi:10.1016/j.pragma.2014.08.002
- Keevallik, Leelo. 2013. "Decomposing Movement: Spatial Deixis in Dance Instruction." In *Interaction and Mobility. Language and the Body in Motion*, ed. by Haddington, Pentti, Mondada, Lorenza, and Maurice Nevile, 345–370. Berlin: De Gruyter. doi:10.1515/9783110291278.345
- Keevallik, Leelo. 2014. "Turn Organization and Bodily-Vocal Demonstrations." *Journal of Pragmatics* 65: 103–120. doi:10.1016/j.pragma.2014.01.008
- Keevallik, Leelo. 2015. "Coordinating the Temporalities of Talk and Dance." In *Temporality in Interaction*, ed. by Deppermann, Arnulf, and Susanne Günthner, 309–336. Amsterdam: Benjamins.
- Keevallik, Leelo. this volume. The Temporal Organization of Conversation while Mucking Out a Sheep Stable.
- Kendon, Adam. 1970. "Movement Coordination in Social Interaction." *Acta Psychologica* 32: 1–25. doi:10.1016/0001-6918(70)90094-6
- Kendon, Adam. 1990. *Conducting Interaction. Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.
- Kendon, Adam. 2004. *Gesture. Visual Action as Utterance*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511807572
- Kita, Sotaro, van Gijn, Ingeborg, and Harry van der Hulst. 1998. "Movement Phases in Signs and Co-Speech Gestures, and Their Transcription by Human Coders." In *Gesture and Sign Language in Human-Computer Interaction: International Gesture Workshop Bielefeld, Germany, September 17–19, 1997*, ed. by Wachsmuth, Ipke, and Martin Fröhlich, 23–35. Berlin: Springer. doi:10.1007/BFb0052986
- Kress, Gunther. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- Langacker, Ronald. 1987. *Foundations of Cognitive Grammar*. Vol. 1. Stanford, CA: Stanford University Press.
- Latour, Bruno. 1993. *We have Never Been Modern*. Harvard University Press.
- Leeds-Hurwitz, Wendy (ed). 2010. *The Social History of Language and Social Interaction Research: People, Places, Ideas*. Cresskill, NJ: Hampton Press.
- Levinson, Steven C. 2013. "Action Formation and Ascription." In *The Handbook of Conversation Analysis*, ed. by Tanya Stivers, and Jack Sidnell, 103–130. Malden, MA: Wiley-Blackwell.
- Levinson, Steven C., and Francisco Torreira. 2015. "Timing in Turn-Taking and its Implications for Processing Models of Language." *Frontiers in Psychology* 6: 731. doi:10.3389/fpsyg.2015.00731
- Li, Xiaoting. 2014. "Leaning and Recipient Intervening Questions in Mandarin Conversation." *Journal of Pragmatics* 67: 34–60. doi:10.1016/j.pragma.2014.03.011
- Llinas, Rodolfo R. 2001. *I of the Vortex. From Neurons to Self*. Cambridge, MA: MIT Press.
- Local, John, and Gareth Walker. 2004. "Abrupt-Joins as a Resource for the Production of Multi-Unit, Multi-Action Turns." *Journal of Pragmatics* 36: 1375–1403. doi:10.1016/j.pragma.2004.04.006

- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago IL: University of Chicago Press.
- Mead, George. H. 1934. *Mind, Self, and Society: From the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.
- Meltzoff, Andrew N., and Wolfgang Prinz. 2002. *The Imitative Mind: Development, Evolution, and Brain Bases*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511489969
- Merleau-Ponty, Maurice. 1945. *Phénoménologie de la perception*. Paris: Gallimard.
- Meyer, Christian, Streeck, Jürgen, and Scott Jordan (eds). 2017. *Intercorporeality. Emerging Socialities in Interaction*. Oxford: Oxford University Press.
- Mondada, Lorenza. 2007. "Multimodal Resources for Turn-Taking: Pointing and the Emergence of Possible Next Speakers." *Discourse Studies* 9 (2): 195–226. doi:10.1177/1461445607075346
- Mondada, Lorenza. 2009. "Emergent Focused Interactions in Public Places: A Systematic Analysis of the Multimodal Achievement of a Common Interactional Space." *Journal of Pragmatics* 41: 1977–1997. doi:10.1016/j.pragma.2008.09.019
- Mondada, Lorenza. 2012. "Talking and Driving: Multiactivity in the Car." *Semiotica* 191: 223–256.
- Mondada, Lorenza. 2013a. "Coordinating Mobile Action in Real Time: The Timely Organization of Directives in Video Games." In *Interaction and Mobility. Language and the Body in Motion*, ed. by Haddington, Pentti, Mondada, Lorenza, and Maurice Neville, 300–341. Berlin: De Gruyter. doi:10.1515/9783110291278.300
- Mondada, Lorenza. 2013b. "The Conversation Analytic Approach to Data Collection." In *Handbook of Conversation Analysis*, ed. by Sidnell, Jack, and Tanya Stivers, 32–56. New York: Blackwell-Wiley.
- Mondada, Lorenza. 2014a. "The Local Constitution of Multimodal Resources for Social Interaction." *Journal of Pragmatics* 65: 137–156. doi:10.1016/j.pragma.2014.04.004
- Mondada, Lorenza. 2014b. "The Temporal Orders of Multiactivity: Operating and Demonstrating in the Surgical Theatre". In *Multiactivity in Social Interaction: Beyond Multitasking*, ed. by Haddington, Pentti, Keisanen, Tiina, Mondada, Lorenza, and Maurice Neville, 33–75. Amsterdam: Benjamins.
- Mondada, Lorenza. 2014c. "Instructions in the Operating Room: How Surgeons Direct Their Assistant's Hands." *Discourse Studies* 16 (2): 131–161. doi:10.1177/1461445613515325
- Mondada, Lorenza. 2015. "Multimodal Completions." In *Temporality in Interaction*, ed. by Deppermann, Arnulf, and Susanne Günthner, 267–307. Amsterdam: Benjamins.
- Mondada, Lorenza. 2016. "Challenges of Multimodality: Language and the Body in Social Interaction." *Journal of Sociolinguistics* 20 (2): 2–32.
- Mondada, Lorenza. 2018. "Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality." *Research on Language and Social Interaction* 51 (1), 85–106.
- Mondada, Lorenza. this volume. Questions on the Move: The Ecology of Question/Answer Sequences in Mobile Settings.
- Mondada, Lorenza, and Reinhold Schmitt (eds). 2010. *Situationseröffnungen. Zur multimodalen Herstellung fokussierter Interaktion*. Tübingen: Narr.
- Mondada, Lorenza, and Kimmo Svinhufvud. 2016. "Writing-in-Interaction: Studying Writing as a Multimodal Phenomenon in Social Interaction." *Language and Dialogue* 6 (1): 1–53.
- Morganti, Francesca, Carassa, Antonella, and Giuseppe Riva (eds). 2008. *Enacting Intersubjectivity*. Amsterdam: IOS.
- Mortensen, Kristian. 2009. "Establishing Reciprocity in Pre-Beginning Position in the Second Language Classroom." *Discourse Processes* 46 (5): 491–515. doi:10.1080/01638530902959463

- Nevile, Maurice. 2015. "The Embodied Turn in Research on Language and Social Interaction." *Research on Language and Social Interaction* 48 (2): 121–151.  
doi:10.1080/08351813.2015.1025499
- Nevile, Maurice, Haddington, Pentti, Heinemann, Trine, and Mirka Rauniomaa (eds). 2014. *Interacting with Objects: Language, Materiality, and Social Activity*. Amsterdam: John Benjamins.
- Newell, Allen. 1982. "Physical Symbol Systems." *Cognitive Science* 4 (2): 135–183.  
doi:10.1207/s15516709cog0402\_2
- Norris, Sigrid. 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. London: Routledge.
- Oloff, Florence. 2013. "Embodied Withdrawal after Overlap Resolution." *Journal of Pragmatics* 46 (1): 139–156. doi:10.1016/j.pragma.2012.07.005
- Oloff, Florence. this volume. Revisiting Delayed Completions: The Retrospective Management of Co-participant Action.
- Oshima, Sae. this volume. Embodiment of Activity Progress: The Temporalities of Service Evaluation.
- Pike, Kenneth Lee. 1956. *Language in Relation to a Unified Theory of Structure of Human Behavior*. The Hague: Mouton.
- Pomerantz, Anita. 1984. "Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes." In *Structures of Social Action: Studies in Conversation Analysis*, ed. by Atkinson, J. Maxwell, and John Heritage, 57–101. Cambridge: Cambridge University Press.
- Rasmussen, Gitte. 2014. "Inclined to Better Understanding: The Coordination of Talk and 'Leaning Forward' in Doing Repair." *Journal of Pragmatics* 65: 30–45.  
doi:10.1016/j.pragma.2013.10.001
- Rauniomaa, Mirka, and Tiina Keisanen. 2012. "Two Multimodal Formats for Responding to Requests." *Journal of Pragmatics* 44 (6–7): 829–842. doi:10.1016/j.pragma.2012.03.003
- Rizzolatti, Giacomo, and Laila Craighero. 2004. "The Mirror-Neuron System." *Annual Review of Neuroscience* 27: 169–192. doi:10.1146/annurev.neuro.27.070203.144230
- Rossano, Federico. 2013. "Gaze in Conversation." In *The Handbook of Conversation Analysis*, ed. by Jack Sidnell, and Tanya Stivers, 308–329. Malden, MA: Wiley-Blackwell.
- Rossi, Giovanni. 2015. *The Request System in Italian Interaction*. Nijmegen: MPI Psycholinguistics.
- Ruesch, Jürgen, and Kees, Weldon. 1972. *Nonverbal Communication*. Berkeley: University of California Press.
- Sacks, Harvey. 1984. "Notes on Methodology." In *Structures of Social Action*, ed. by Atkinson, John Maxwell, and John Heritage, 21–27. Cambridge: Cambridge University Press.
- Sacks, Harvey. 1992. *Lectures on Conversation*. Vol. 2. Oxford: Blackwell.
- Sacks, Harvey, Schegloff, Emanuel A. and Gail Jefferson. 1974. "A Simplest Systematics for the Organization of Turn-Taking in Conversation." *Language* 50 (4): 696–735.  
doi:10.1353/lan.1974.0010
- Sacks, Harvey, and Emanuel A. Schegloff. 2002[1975]. "Home Position." *Gesture* 2: 133–146.  
doi:10.1075/gest.2.2.025ac
- Schefflen, Albert Edward. 1972. *Body Language and Social Order: Communication as Behavioral Control*. Englewood Cliffs, NJ: Prentice-Hall.
- Schegloff, Emanuel A., and Harvey Sacks. 1973. "Opening Up Closings." *Semiotica* 8 (4): 289–327.  
doi:10.1515/semi.1973.8.4.289

- Schegloff, Emanuel A. 1968. "Sequencing in Conversational Openings." *American Anthropologist* 70: 1075–1095. doi:10.1525/aa.1968.70.6.02a00030
- Schegloff, Emanuel A. 1984. "On Some Gestures' Relation to Talk." In *Structures of Social Action*, ed. by Atkinson, John Maxwell, and John Heritage, 266–298. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. 1998. "Body Torque." *Social Research* 65 (3): 535–596.
- Schegloff, Emanuel A. 1992. "Repair after Next Turn. The Last Structurally Provided Defence of Intersubjectivity in Conversation." *American Journal of Sociology* 97 (5): 1295–1345. doi:10.1086/229903
- Schegloff, Emmanuel A. 1996. "Turn Organization: One Intersection of Grammar and Interaction." In *Interaction and Grammar*, ed. by Ochs, Elinor, Schegloff, Emanuel A. and Sandra A. Thompson, 52–133. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620874.002
- Schegloff, Emanuel A. 2007. *Sequence Organization*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511791208
- Schegloff, Emanuel A. 2013. "Ten Operations in Self-initiated, Same-turn Repair." In *Conversational Repair and Human Understanding*, ed. by Hayashi, Makoto, Raymond, Geoffrey and Jack Sidnell, 41–70. Cambridge: Cambridge University Press.
- Schmidt, Axel. this volume. Prefiguring the Future: Projections and Preparations within Theatrical Rehearsals.
- Schmitt, Reinhold. 2003. "Inszenieren: Struktur und Funktion eines gesprächsrhetorischen Verfahrens." *Gesprächsforschung* 4: 186–250, <http://www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2003/ga-schmitt.pdf>, last access June 8, 2016.
- Schmitt, Reinhold. 2005. „Zur multimodalen Struktur von turn-taking." *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 6: 17–61. <http://www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2005/ga-schmitt.pdf>, last access June 8, 2016.
- Schmitt, Reinhold. 2015. "Positionspapier: Multimodale Interaktionsanalyse." In *Ko-Konstruktionen in der Interaktion. Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen*, ed. by Dausendschön-Gay, Ulrich, Gülich, Elisabeth, and Ulrich Krafft, 43–51. Bielefeld: transcript. doi:10.14361/9783839432952-004
- Schmitt, Reinhold, and Arnulf Deppermann. 2007. "Monitoring und Koordination als Voraussetzungen der multimodalen Konstitution von Interaktionsräumen." In *Koordination. Analysen zur multimodalen Interaktion*, ed. by Schmitt, Reinhold, 95–128. Tübingen: Narr.
- Schmitt, Reinhold, and Arnulf Deppermann. 2010. "Die Transition von Interaktionsräumen als Eröffnung einer neuen Situation." In *Situationseröffnungen. Zur multimodalen Herstellung fokussierter Interaktion.*, ed. by Mondada, Lorenza, and Reinhold Schmitt, 335–386. Tübingen: Narr.
- Schönherr, Beatrix. 1997. *Syntax – Prosodie – Nonverbale Kommunikation. Empirische Untersuchungen zur Interaktion sprachlicher und parasprachlicher Ausdrucksmittel im Gespräch*. Tübingen: Niemeyer. doi:10.1515/9783110926156
- Schütz, Alfred. 1974 (1932). *Der sinnhafte Aufbau der sozialen Welt*. Frankfurt am Main: Suhrkamp.
- Szczepek Reed, Beatrice B. 2006. *Prosodic Orientation in English Conversation*. Basingstoke: Palgrave Macmillan. doi:10.1057/9780230625273
- Searle, John R. 1969. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139173438

- Sidnell, Jack. 2014. "The Architecture of Intersubjectivity Revisited." In *The Cambridge Handbook of Linguistic Anthropology*, ed. by Enfield, Nick J., Kockelman, Paul, and Jack Sidnell, 423–446. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139342872.020
- Simmel, Georg. 1916. *Das Problem der historischen Zeit. Philosophische Vorträge. Veröffentlicht von der Kantgesellschaft. Nr. 12*, ed. by Liebert, Arthur. Berlin: von Reuther und Reichard.
- Stivers, Tanya, and Jack Sidnell (eds). 2005. "Multimodal Interaction." *Semiotica* 156 (1–4): 1–10. doi:10.1515/semi.2005.2005.156.1
- Strauss, Anselm. 1985. "Work and the Division of Labor." *The Sociological Quarterly* 26 (1): 1–19.
- Streeck, Jürgen. 1988. "The Significance of Gestures: How It is Established." *Papers in Pragmatics* 2 (1–2): 60–84. doi:10.1075/1prapip.2.1-2.03str
- Streeck, Jürgen. 1995. "On Projection." In *Interaction and Social Intelligence*, ed. by Goody, Edward, 84–110. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511621710.007
- Streeck, Jürgen. 2007. "Geste und verstreichende Zeit: Stillstand und Bedeutungswandel." In *Gespräch als Prozess*, ed. by Hausendorf, Heiko, 157–177. Tübingen: Narr.
- Streeck, Jürgen. 2009a. *Gesturecraft. The Manu-Facture of Meaning*. Amsterdam: Benjamins. doi:10.1075/g5.2
- Streeck, Jürgen. 2009b. "Forward-Gesturing." *Discourse Processes* 46 (2–3): 161–179. doi:10.1080/01638530902728793
- Streeck, Jürgen. this volume. Times of the Rest. Temporalities of Some Communicative Postures. Streeck, Jürgen, Goodwin, Charles and Curtis LeBaron (eds). 2011. *Embodied Interaction. Language and Body in the Material World*. Cambridge: Cambridge University Press.
- Streeck, Jürgen, and Jordan, J. Scott. 2009. "Projection and Anticipation: The Forward-Looking Nature of Embodied Communication." *Discourse Processes* 46 (2–3): 93–102. doi:10.1080/01638530902728777
- Stukenbrock, Anja. 2010. "Überlegungen zu einem multimodalen Verständnis der gesprochenen Sprache am Beispiel deiktischer Verwendungsweisen des Ausdrucks „so“." In *Beschreibungen für gesprochenes Deutsch auf dem Prüfstand*, ed. by Dittmar, Norbert, and Nils Bahlo, 165–193. Frankfurt/Main: Lang.
- Stukenbrock, Anja. this volume. Forward-Looking: Where do We Go with Multimodal Projections?
- Sunakawa, Chiho. this volume. Learning to Be an Orchestral Conductor.
- Tannen, Deborah. 1989. *Talking Voices*. Cambridge: Cambridge University Press.
- Ursi, Biagio. 2016. *Le refus en interaction. Une approche syntaxique et séquentielle de la negation*. Lyon: Unpublished PhD-thesis at University Lyon 2.
- Zerubavel, Eviatar. 1981. *Hidden Rhythms: Schedules and Calendars in Social Life*. Chicago: University of Chicago Press.
- Zerubavel, Eviatar. 1985. *The Seven-Day Circle: The History and Meaning of the Week*. New York: Free Press.