1

Embodied Interaction in the Material World: An Introduction

**Jürgen Streeck, Charles Goodwin, and Curtis LeBaron**

HUMAN ACTION

The chapters in this volume take as their focus the organization of action in human interaction. The question immediately arises as to where and how the structure of human action might be investigated. Different disciplines have taken very different kinds of phenomena, ranging from the mental intentions of individual actors to large, historically shaped social structures, as the proper locus for such a study. Here we take as our point of departure events in which multiple parties are carrying out endogenous courses of action in concert with each other within face-to-face human interaction. A concrete example can make clearer what we mean by this. In Transcript 1.1 Ann, a senior archeologist and director of the field school where the current excavation is taking place, is working with Sue, a new graduate student, as Sue works to outline the shape of an archeological feature faintly visible in the color patterning of the dirt they are examining (this sequence is examined in more detail, from a slightly different perspective in Goodwin (2007a).
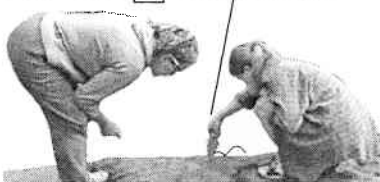
Transcript 1.1. Embodied Interaction.

The actions occurring in Transcript 1.1 are not organized within a single medium, such as talk, but are instead constructed through the simultaneous use of multiple semiotic resources with quite different properties. Thus in line 1, Ann says, "Wha'do you think of:,". In English, of begins a prepositional phrase that requires a noun for its grammatical completion. However, no appropriate noun occurs in Transcript 1.1. A similar argument can be made about "aro:und" in line 3, where the entity being gone around is never specified in the talk. If one focuses only on the talk occurring here, and the linguistic structure emerging within that talk, what is said here does not conform to the requirements of English grammar. However the participants do not in any way treat this talk as defective. Instead the "it" in Sue's line 3 "Does it kinda go aro:und" explicitly ties back to what Ann indicated, and thus not only treats what Ann was talking about as unproblematically understood, but incorporates that recognition into the structure of the utterance responding to Ann's talk.

There is of course no mystery in how Sue was able to appropriately understand what Ann was telling her. As Ann said "of:," in line 1, she used her right arm and index finger to point toward a particular patch of color patterning in the dirt they were working on together. The slot for the noun in the prepositional phrase in the talk was thus filled by the combination of a pointing gesture and the visible structure in the environment it indicated. Ann was showing Sue something in the dirt that should now become the focus of their joint scrutiny and work. Well before she produces "it" in line 3, Sue displays precisely this embodied, work-relevant understanding of the complex structure of Ann's action by moving her own hand and trowel to just the spot in the dirt that Ann indicated. She then uses that positioning as the point of departure for the gesture with the trowel tracing structure in the dirt that accompanies "kinda go around" in line 3.

The interaction between Ann, Sue, and the world that is the focus of their work is organized through the structured exchange of different kinds of signs. These include not only language but also a variety of diverse

signs constituted through the visible organization of the participants' bodies. Ann uses her pointing finger in line 1 to indicate to Sue a specific place in the dirt. Sue's movement of the trowel in line 3 is used to show Ann the path in the dirt that is described in the talk as "kinda go around" and thus constitutes a sign for that path. Each party builds action by producing signs for the other. Thus, to build relevant action in Transcript 1.1, the participants simultaneously make use of a number of quite different kinds of semiotic resources that have different properties and are instantiated in different kinds of semiotic materials (linguistic structure in the stream of speech, signs such as pointing displayed through the visible body, the patterning of phenomena in the environment that is the focus of their work, etc.). The recognizable and consequential actions they are building for each other cannot be found in any single semiotic medium. As noted earlier, by itself the talk is incomplete both grammatically and, more crucially, with respect to the specification of what the addressee of the action is to attend to in order to accomplish a relevant next action. Similarly the embodied pointing movements require the co-occurring talk to explicate the nature and relevance of what is being indicated. Indeed the mutual organization of talk and gesture has long been a central theme in gesture studies (Kendon, 2004; McNeill, 1992). By itself each individual set of semiotic resources is partial and incomplete (Agha, 2007; Goodwin, 2007a). However, when joined together in local contextures of action, diverse semiotic resources mutually elaborate each other to create a whole that is both greater than, and different from, any of its constituent parts (Goodwin, 2000a). Describing how action is built here thus requires an analytic framework that recognizes the diversity of semiotic resources used by participants in interaction, and takes into account how these resources interact with each other to build locally relevant action.

Having the ability to build action by combining resources with diverse properties has clear advantages and greatly expands the repertoire of possible action available to participants. To note one very simple example: In line 3, Sue is tracing with her trowel a complex, irregular shape in the dirt. Describing the precise shape of the phenomena they uncover in the dirt being excavated is crucial to the work of archeology. Suppose the resources available for doing this were restricted to a single semiotic field, such as language. If each different shape encountered in an excavation had to be categorized semantically, the vocabulary of archeology would quickly become unmanageably large – indeed, useless. However if a limited set of semantic categories ("feature," "postmold," "disturbance," etc.) can be supplemented by analogic signs capable of continuous variation (gestures over a shape such as line 3, drawings on maps, etc.), precision and flexibility become not only possible, but quite literally ready at hand as working hands and trowels articulate for others relevant structure in the world they are acting upon together.

To try and demonstrate as clearly as possible how action is built by combining resources with diverse properties that mutually elaborate each other, the discussion has so far been restricted to how talk, gesture, and structure in the world mutually elaborate each other. This might be glossed as the referential domain that the participants are focusing on: what they are talking about and formulating as particular kinds of structure in the dirt they are excavating. However this does not in any way exhaust the different kinds of semiotic resources that are implicated in the organization of their action.

For example, how can Ann unproblematically assume that Sue will take her gesture into account, something an addressee must do in order to properly understand what Ann is telling her and thus build an appropriate next action? Note that Ann places her gesture right in front of Sue's eyes, over the dirt she is already looking at. Ann treats Sue's gaze as a sign for where she is attending and what she is attending to. More generally, the mutual orientation of the participants' bodies creates what Goffman (1964: 64) called an "ecological huddle," which publicly demonstrates through visible embodied practice that the participants are mutually oriented toward each other and frequently toward particular places, objects, and events in the surrounding environment (Heath, Luff, vom Lehn, Hindmarsh and Cloeverly, 2002). Such embodied participation frameworks (Goodwin, 2000a) or F-formations (Kendon, 1990) are central to the organization of action in face-to-face interaction. Like gestures, these displays of mutual orientation are constructed through embodied signs. However, they differ from gesture in a number of important respects. First, they are not "about" the substance of what the participants are talking about (e.g., relevant structure in the dirt these parties are working on), but instead have as their subject matter the orientation of the participants toward each other, and the world that is the focus of their activity. Second, they have a quite different temporal organization. Unlike particular elements of talk, or specific gestures, which disappear and are replaced by other words or gestures almost as soon as they occur, embodied participation frameworks can be sustained over extended stretches of talk and action. Third, even not being about the substance of what is being talked about, they contribute to the organization of that talk in other important ways. For example, the shared orientational frameworks they make publicly visible deictically ground many of the indexical expressions that occur within that talk (including "you," "it," and "there" in Transcript 1.1) while making possible other indexical, context sensitive uses of language, such as the "incomplete" prepositional phrase noted earlier. These embodied orientational frameworks create local environments where participants can treat each other as attending to, and working together within, a shared world of perception and action, something crucial to the way in which Ann and Sue are building action together by attending to

how each other is interpreting and operating on the dirt that is the focus of their work. In essence, the signs used to create and continuously sustain, modify or dismantle participation frameworks (Goodwin, 1981, 2007b; Kendon, 1985) create a public semiotic environment within which other kinds of sign exchange processes, such as talk and gesture, can flourish.

Events of the type found in Transcript 1.1, in which multiple parties are carrying out a course of action together through the use of talk and other embodied action while attending to each other and frequently to the phenomena in the world that are the focus of their scrutiny and activities, provide a perspicuous environment for the systematic investigation of a range of phenomena that are central to the organization of human language, social organization, culture, and cognition. First, insofar as a common course of action is being accomplished through the joint, collaborative work of multiple parties, such events provide pervasive examples of elementary human social organization, a place where one can investigate in detail the actual practices used to build endogenous social order. Simmel (1950: 21–22) argued that "if society is conceived as interaction among individuals, the description of the forms of this interaction is the task of the science of society in its strictest and most essential sense." Such sites, in which action is organized with reference to the properties of embodied co-presence, render clearly visible many of the central features of human interaction noted by Goffman (1963), including mutual monitoring and the reflexivity of embodied interaction. Second, as has long been noted by conversation analysts (Sacks, 1992; Schegloff, 2006), face-to-face interaction is a central place where language emerges in the natural world. Third, if participants are to carry out courses of collaborative action together, they must in some relevant sense understand what each other is doing, and the nature and detailed structure of the events they are engaged in together. Such sites thus permit investigation of the practices of sense making noted by Garfinkel (1967) and of cognition as public practice more generally. They are also central to contemporary work in Europe, such as Linell (2009), which is attempting to rethink language, the mind, and the world dialogically. Fourth, though organized through general practices, the particulars of what participants must see and understand in order to build action together, such as how color patterns in a patch of dirt can be interpreted as archeological features, are lodged within specific communities. Situations such as these are places where the content and organization of culture as practice, as well as the ways in which such knowledge, skills and practices are appropriated by newcomers just entering its distinctive phenomenal world of a community, can be examined in fine detail (Sue is a beginning archeologist at her first excavation). Fourth, in such events, it is possible to investigate both the part played by the individual body in the organization of cognition and action, including how such bodies gain the

skills required for relevant action within specific communities (Ingold, 2000), and how participants see and understand each other's bodies so that they can anticipate what each other is about to do and joint action can be successfully accomplished.

It is not being argued that such events are the only place where human action occurs, or that they are in some sense primordial. Many actions, such as the words now being written, are created by solitary individuals, though ones using culturally structured resources such as language. An individual can come to know the world and its distinctive properties through exploration and work with her own hands (Streeck, 2009), and much phenomenal knowledge is lodged within the experience of an individual embedded within a consequential world. The interactive organization of multi-party action does, however, provide a fruitful arena for investigating from an integrated perspective a host of crucial phenomena that are central to the organization of human action, cognition, and social life.

In brief, by looking at events such as that found in Transcript 1.1, it is possible to systematically examine some of the practices used by human beings to build action in concert with each other. As has long been strongly demonstrated by conversational analysts (Jefferson, 1988; Sacks, Schegloff, and Jefferson, 1974; Schegloff, 2007), sequential organization is central to both the structure of action and the way in which it is understood by the participants themselves (Sue's talk is built in response to what Ann has just said and done, and, as noted earlier, a number of constructional features of her utterance explicitly display this, including the "it," which ties to what Ann has just said and done). One phenomenon that quickly emerges from records that preserve not only the talk but also the bodies of actors, is that action is built through the mutual elaboration of diverse semiotic resources with quite different properties, each of which, including language, can make only a partial, incomplete contribution to the action in progress. The participants themselves attend to both this diversity and to the unique, distinctive contributions made by different kinds of semiotic resources. Thus Sue builds a new action to Ann both with talk and with relevant actions of her body – for example, by moving her own trowel to the place in the dirt indicated by Ann and then using that trowel to outline what she has been asked to see, and on another level, by aligning her body toward both Ann and the patch of dirt they are examining together.

A unifying thread running through all of the papers in this volume, though one developed in very different ways, is the systematic investigation of how multiple participants build action together in the midst of situated interaction, typically by using different kinds of semiotic resources that mutually elaborate each other. One aspect of this process that the current volume is not able to adequately address is prosody. However, this is the focus of rich and important work, much of it in

Europe, by several linked groups of scholars including Elizabeth Couper-Kuhlen, Margaret Selting, Dagmar Beth-Weingarten, Elisabeth Reberm, John Local and his collaborators in York, and many others (see, for example, Couper-Kuhlen and Selting, 1996b). This volume's focus on the organization of action within interaction differentiates it from some other approaches to what is sometimes glossed as multimodality. It is, however, consistent with a growing body of work, in Europe, Japan, and the United States, that has begun to engage in intensive analysis of how action is built through the inter-elaboration of talk, the body, encompassing activities and features of the setting (see, for example, Heath and Luff, 2000; Mondada, 2008a, 2008b; Nishizaka, 2007), and reflexive analysis of the transcription practices that can make such phenomena visible and amenable to analysis (Lindwall and Lymer, 2008; Murphy, 2005; Mondada, 2006).

The approach taken in this volume, with its focus on systematic investigation of the different kinds of semiotic resources and meaning-making practices that participants themselves attend to, and treat as relevant, as they build action within interaction together, seems to us not only fruitful, but straightforward and uncontroversial. The simultaneous use of diverse semiotic resources – currently discussed under the heading *multimodality* (see the fourth section of this chapter) – is pervasive in the organization of endogenous human action. The issue therefore arises as to why the relevance of adopting a perspective that takes this into account must be clearly argued. Briefly, much existing research has avoided the crucial issues posed by the heterogeneous semiosis that sits at the center of actual human action by focusing on the analysis of individual semiotic systems as self-contained wholes. For example, Saussure (1959: 16) envisioned a general science focused on "the life of signs within society." Such a goal is entirely compatible with the work in this volume. However, Saussure then argued that linguistics should confine its study to just one part of this larger field by investigating language as an isolated self-contained whole:

A science that studies the life of signs within society is conceivable; it would be a part of social psychology and consequently of general psychology; I shall call it "semiology" (from Greek *semefon*, "sign"). Semiology would show what constitutes signs, what laws govern them. Since the science does not yet exist, no one can say what it would be; but it has a right to existence, a place staked out in advance. Linguistics is only a part of the general science of semiology; the laws discovered by semiology will be applicable to linguistics, and the latter will circumscribe a well-defined area within the mass of anthropological facts. To determine the exact place of semiology is the task of the psychologist! The task of the linguist is to find out what makes language a special system within the mass of semiological data.

Language is thus demarcated as a "special system" that not only can be, but should be investigated without reference to other semiotic processes with which it characteristically

co-occurs. Delimiting the scope of inquiry in this way, and thus defining the phenomenal and analytic field within which all subsequent inquiry will occur, has had enormous consequences. Such limits defined the scope of formal linguistics, and were carried over as unquestioned assumptions when new fields, such as cognitive science, emerged. Thus it took much creative innovation for cognitive science to reshape itself so that phenomena such as embodiment (Clark, 1997; Gibbs, 2005) and the distribution of cognitive processes beyond the individual brain to encompass the situated practices of communities (Hutchins, 1995; Suchman, 1987) were recognized as essential to the analysis of human cognition.

From a slightly different perspective, human language possesses rich, intricate, and varied structure combined with extraordinarily powerful representational capacities. Moreover, for thousands of years it has been possible to use writing to capture much of this richness in another, more permanent medium. Writing does, however, have the effect of rendering invisible the embodied frameworks within which language in face-to-face interaction is embedded, including the crucial part played by co-present hearers. Rather than simply being constraints, the restrictions and distinctive properties of writing, as a semiotic medium in its own right, make possible new and important ways of using language and preserving some of its detailed structure not only across encounters, but also across generations. In part because of the powerful resources provided by writing, many fields, including some that strongly oppose the formal and monologic assumptions of Saussure and argue persuasively for the crucial importance of dialog (Bakhtin, 1981; Volosinov, 1973), have nonetheless restricted the scope of their inquiry to phenomena that fall within a broad conception of language. While offering a powerful and most important arena for study, such logocentricism – what Linell (2005) calls the written language bias in linguistics – nonetheless renders invisible many of the crucial forms of semiosis that shape human action in actual interaction (for example many of the embodied phenomena found in Transcript 1.1, as well as the crucial role of structure in the world that is a focus of the participants' talk and action).

Not all interaction occurs within the fully embodied frameworks of mutual orientation found in Transcript 1.1. Indeed this is a systematic consequence of the very semiotic structure of such events. Because action is being built through the co-articulation of different semiotic fields, it is possible to remove some of these fields while adapting the structure of others so that the accomplishment of relevant action remains. Throughout human history, from hunter gatherers talking across campfires in the dark to contemporary talk over telephones, human beings have been able to build rich interaction with each other through talk alone. Situations with such restricted semiotic structure do, however, eliminate for participants as well as analysts many of the crucial resources implicated in the organization of action in face-to-face

interaction. Thus, in fully embodied situations, utterances are not constituted exclusively within the stream of speech by the actions of the speaker. Instead the visible actions of hearers, including both orientation toward the speaker and operations on the specifics of the talk as it is being spoken, can systematically lead the speaker to change the structure of a sentence in progress (Goodwin, 1981; M. H. Goodwin, 1980). Many of the consequential actions of the hearer are performed through visible displays of the body rather than with talk. Within such frameworks, both the utterance and the turn-at-talk within which it emerges are not only intrinsically multi-party activities, but also ones built through the interplay of structurally different kinds of semiotic processes, including the talk of the speaker and the visual displays of hearer (the speaker also makes consequential visual displays, for example using gaze to indicate address). Noting this is not to deny the powerful analysis that has been developed from audio recordings of interaction, but it does demonstrate the relevance of analysis that takes into account the distinctive semiotic structure of fully embodied co-presence.

THE INTERACTIONIST PERSPECTIVE

The study of embodied interaction as it is presented in this book takes inspiration from a variety of sources, most of which are familiar names: Mead, Vygotsky, Bakhtin, Bateson, Goffman, Sacks, Schegloff and Jefferson, Kendon. Some would want to include Wittgenstein in the list, others Merleau-Ponty and Heidegger, or Bourdieu, de Certeau, and Marx. Even though there may be minor disagreements about the list, on the whole our field is not lacking in intellectual cohesion. We cannot account for these influences in detail, but want to remind the reader of some especially pertinent intellectual forces that continue to shape the ways in which interactionist researchers think about their subject matter and the proper ways to analyze it.

Of particular importance for work on embodied interaction has been G. H. Mead's critique of methodological individualism (Mead, 1909, 1934), that is, of those accounts of social life and symbolic interaction that posit the self as given and treat meaning, mind, and intersubjectivity as epiphenomena or products of individual minds. Mead (1934: 222–223) maintained that a theory which

... assumes individual selves as the presuppositions, logically and biologically, of the social process or order within which they interact...., cannot explain the existence of minds and selves.... [In contrast, a theory which] assumes a social process or social order as the logical and biological precondition of the appearance of the selves of the individuals involved in that process or belonging to that social order, ... can explain that which it takes as logically prior, namely the existence of the social process of behavior, in terms of such fundamental biological relations and interactions as reproduction.

Mead conceived interaction as a conversation of gestures. Gestures in Mead's conception are not hand gestures as they are studied today, but more broadly early parts of acts, components that can become separated as free-standing units with organic and motivated, yet conventional, relationships to the social acts in which they have emerged. Nevertheless, Mead's conception is quite compatible with interactionist accounts of hand gestures. He observed that

... throughout the entire process of an interaction, we analyze the incipient actions of others by our own instinctive reactions to changes in their postures and other signs of developing social actions (Mead, 1909: 219).

Thus, making gestures that come from and designate acts, we creatively hatch courses of joint action. Through gestures in Mead's sense, we rapidly and incessantly indicate to one another – and thus prepare – what is to come next (McDermott and Roth, 1978). Mead draws our attention to the forward-design of human action. The foreshadowing of imminent actions is made possible not least by the multimodal structure of the human body – its ability to move some of its parts independently from one another and thus create multiple, heterogeneous signs at the same time.

As the self is mediated by interaction, it is also inextricably embedded in a community and draws on this community's historically evolved sense-making tools, in the first place a language and the typified categories of experience that it offers. Vygostky, a near-contemporary of Mead, called such "mediational means" (Wertsch, 1991) psychological tools (Vygotky, 1978). Individual minds are produced through cultural apprenticeship. Bakhtin (1986) proposed an analogous view of language: Speaking means to rent words from a community, to fashion oneself (and one's utterance) by using communal means. In every act of speaking, individual and society are intertwined.

From Gregory Bateson we have learned to think of speech and "nonverbal communication" not as a combination of signs, but as a relation between act and context. Contexts *frame* or *type* behavior. Context can be a meta-message, for example, "[T]his is play" (Bateson, 1956), which instructs us not to take anything that is contextualized by it at face value. But the relation is mutual: The context is also created by the act, a relationship that Gumperz (1992) expresses in his notion of "contextualization cues." The act is "part of the ecological subsystem called context and not ... the product or effect of what remains of the context once the piece which we want to explain has been cut out from it" (Bateson, 1972: 338).

In Goffman's dramatistic view of interaction, characteristic especially of his earlier work (1959, but see 1976), the entire setting insofar as it is under the actor's control can be manipulated to display the committing of acts, to embody the working consensus, or to represent something as something else. He wrote: "[T]he representation

of an activity will vary in some degree from the activity itself and therefore misrepresent it" (Goffman, 1959: 45). He also noted that we cannot separate bodily signs from the settings in which the bodies that make them operate:

The individual gestures with the immediate environment, not only with his body, and so we must introduce this environment in some systematic way ... while the substratum of a gesture derives from the maker's body, the form of the gesture can be intimately determined by the microecological orbit in which the speaker finds himself. To describe the gesture, let alone uncover its meaning, we ... have to introduce the human and material setting in which the gesture is made (Goffman, 1964: 164).

This rarely cited dictum could serve as a motto for this book; it presages the common ground of much contemporary research on embodied and multimodal interaction.

Goffman's term *footing* (Goffman, 1981) also reveals his interest in embodiment, in the question of how aspects of the interaction order are given corporeal form. The term "footing" designates the differing forms of alignment and presence in an utterance that can be taken up by the range of structurally differentiated participants who are implicated in the organization of a strip of talk. For example the current speaker, or animator, may be voicing words authored by either herself or others, and while quoting the words of others can display varying stances toward the talk and action being reported (see also Bakhtin, 1981; Goodwin, 2007b; Hanks, 1996; Levinson, 1988; Volosinov, 1973). Non-speaking participants can have a range of quite different kinds of alignment toward the current utterance, both in terms of typology of different kinds of hearers Goffman offered in footing, and with respect to local operations on the structure of emerging utterances (M. H. Goodwin, 1980). When we observe conversations among people who are standing, we can indeed often read off changes in footing from the reshuffling of the participants' feet, as they reconfigure their spatial arrangement: It was this type of modality-crossing representations of the interaction order that Goffman was especially interested in.

What inspires all contributions to this volume is a view of speakers and listeners as profoundly and inextricably "intervolved" (Dreyfus, 1991) with the material context that they operate in – with the world at hand (Schütz, 1967). When we imagine a speaker, we typically envision her with pen and wrench in hand, or preparing a blood vessel for surgery, or with feet firmly planted in a hopscotch grid. This analytic orientation – to picture speaker and listener at work, doing things *with things* (Streeck, 1996a) – resonates with a certain conception in philosophical anthropology, dating back to the Enlightenment, of humanity as *homo faber*, as makers of artifacts, caught up in the never-ending project of sustaining the world and surviving in it by making and remaking it over and over and over. A phenomenological perspective shapes the work of an increasing number

of linguists, anthropologists, cognitive scientists, and other researchers of communicative practice (Gehlen, 1988; Hanks, 1996). Herder (1772), and Plessner (1965, 1980), among many others, have conceived of human-made material culture and language as an *Ersatz* for a missing biosphere. The human species suffers from its "excentric positionality" (Plessner, 1975) in the world: It is not biologically adapted to a biosphere, but must create its own artifactual work and adapt itself to it, each group to its own, in order to survive. The evolution of the human mind is part of this adaptation. Our ability to adopt a reflective attitude toward our own words and gestures – to regard and scrutinize them as our own objectivations – must have evolved from our primary ability to manufacture – and then behold, probe, and modify – meaningful things. Just like artifacts, words and gestures are external objects brought into existence by human action (Donald, 1991).

Our capacity for manufacture is grounded in specific abilities of hand-eye coordination and certain kinds of precision grip, that is, the ability to closely inspect, rotate, and modify objects while firmly holding on to them (Napier, 1980). The grounding of manufacture and reflexivity in hand-eye coordination, central already to the work of Gehlen (1988) and Plessner (1965) and, much later, Bruner's theory of language acquisition and grammatical relations (Bruner, 1969), is central to any kind of craft (McCullough, 1996; Sennett, 2008; Streeck, 2009). The conception of interaction as multimodal, as it is presented in this book, is consistent with this philosophical-anthropological notion of the excentric positionality of the human species: We have survived by means of our multiple and hetereogeneous objectivations, which include language and artifacts such as tools, skilled practices, rituals, and institutions. These objectivations can only be understood and explained in relation to one another. Such a view contrasts sharply with approaches that seek to abstract language from this nexus and attribute to an innate faculty or claim the centrality of texts to human social life and reproduction. Phenomenological philosophers have given us a notion of the body as a vehicle for being in the world (Merleau-Ponty, 1962) and a primarily haptic – rather than visual – epistemology. Manipulations are our primary understandings of the world (Heidegger, 1962). "Understanding is not in our minds but in our skillful ways of comporting ourselves" (Dreyfus, 1991: 75). It is the body thus conceived – in its concrete, unique, pre-verbal, skilled, and practical coupling with a world – that occupies center stage in the studies of embodied interaction that are collected here.

In another theoretical context, the French social anthropologist Marcel Mauss, nephew and co-worker of Émile Durkheim, proposed the study of *techniques corporelles* (1973), of movement and action skills that people acquire by living in some social milieu. Bourdieu elaborated this focus on the body as practice in the concept *habitus* (Bourdieu, 1977), which designates the

socially contexted bodily dispositions, sensibilities, and skills that permeate our sensory cognition and action skills. Previously, Bateson and Mead had worked from a similar concept when they described the Balinese by focusing on "the way in which they, as living persons, moving, standing, eating, sleeping, dancing, and going into trance, embody that abstraction which (after we have abstracted it) we technically call culture" (Bateson and Mead, 1942: xii).

Anthropologists have produced many textual and visual accounts of embodied culture. As examples for many others, Keller and Keller (1996) have analyzed the sensory cognition of blacksmiths, and Harper (1987) the working knowledge of a car mechanic (see also Csordas, 1994; Ingold, 2000; Jackson, 1989; Strathern, 1996). French anthropologists have developed film-based methods for the praxeological study of cultural transmission (Comolli, 2003; de France, 1983), as exemplified by the *gestes de savoir* (Comolli, 1991) of housewives and violinists.

The phenomenological conception of the body as situated in and "intervolved" (Dreyfus, 1991) with a material-practical world is in many ways a forerunner (sometimes acknowledged, sometimes not) of the currently popular cognitive science program known as *embodied cognition*. Its agenda is neatly summed up in the subtitle of A. Clark's book: "putting brain, body and world together again" (Clark, 1997), whose title, *Being There*, is a direct translation of Heidegger's term *Dasein* (Heidegger, 1962). Cognitive scientists who conceive cognition as embodied widely agree on the following points:

(a) the computational view of the mind, according to which the mind-brain operates by manipulating abstract (amodal) symbols, is rejected;

(b) experience (memory) is modally stored, in the form of "perceptual symbol systems" (Barsalou, 1999); the sensory, perceptual dimensions of experience are retained in the formation of concepts;

(c) the brain is multimodal: it allows us to recode experience, to structure it in terms of schemata from other domains (Deacon, 1997);

(d) the original function of any brain is to control motion – only mobile organisms have brains; other functions of the brain must have evolved from this primary ability (Llinàs, 2001);

(e) cognition and emotion are inseparable; emotion is a form of (embodied and social) cognition (Damasio, 1994, 1999);

(f) perception and motor control are not separate in the brain; perceiving another human being's action means producing an internal (i.e., inhibited, simulated) version of that action (this is known as common coding of motor-control and perception).

Many cognitive scientists interested in embodied cognition, while granting that the body must be conceived as a body in action, even in joint action (Knoblich and

Sebanz, 2006), are reluctant to situate it fully within the material, external, human-made world. Psychologists, keen to maintain the separate integrity of the psychological system(s), have a hard time accepting the idea of distributed cognitive systems as agents of cognitive activity, as proposed, for example, by Hutchins (1995) and contributors to this volume. Thus, Wilson (2002: 126) grants that "cognition is situated, ... takes place in the context of a real-world environment, and ... must be understood in terms of how it functions under the pressures of real-time interaction with the environment, [and] we off-load cognitive work onto the environment." She rejects, however, the notion that, because "the environment is part of the cognitive system, ... the mind alone is not a meaningful unit of analysis" (loc.cit.). For the researchers represented in this volume, an understanding of cognition as socially shared and distributed across mind, communication media, and other artifacts is essential.

EMBODIED INTERACTION

This volume contributes to a stream of research that has gradually emerged and matured during the past four decades. In this section, we seek to account for the convergence of several strands of research and delineate the place of our own attempts in this development.

In the 1970s, scholars from various disciplines began to lament the artificial separation and isolation of so-called "verbal" and "nonverbal" behavior. For instance, Kendon (1972) observed that "it makes no sense to speak of 'verbal communication' and 'nonverbal communication'" (443); he argued that theories of language derived from a study of only speech should be thought of as special language theories, whereas general language theories would show how vocal and visible behaviors function together (Kendon, 1977). In a similar spirit, Margaret Mead (1975) rejected nonverbal research as a "discipline-centric" neglect of vocal phenomena: She argued against Ekman's (1973) theory that facial expressions have universal meanings, suggesting that members of cultures derive meaning from facial expressions by relating them to the context in which they occur, which includes vocal behavior. Such laments in the 1970s were coincident with the mass marketing of a new technology called "videotape," which set the stage for more programmatic explorations of face-to-face interaction.

In the 1980s, a handful of seminal studies clearly and empirically established how talk and embodied behavior co-occur as interdependent phenomena, not separable modes of communication and action. Researchers in the tradition of conversation analysis explored the relationship between talk and eye gaze. Goodwin (1979) examined a videotaped dinner conversation and focused on a single spoken sentence that was shaped and reformed in the process of its utterance as the speaker shifted his gaze among recipients who had different knowledge states – which called into question the linguistic notion of

a sentence as something whose organization was lodged within the mental life of a single individual, the speaker. In an other work, C. Goodwin (1980) analyzed a collection of videotaped instances to show subtle forms of coordination between utterance-initial restarts and shifts in participants' eye gaze (hence attention) toward the speaker. Atkinson (1984) dissected recordings of political speeches to show how politicians elicit applause from audiences, not merely through vocal devices such as "contrastive pairs" and "three-part lists," but also through their rhythmic coordination of talk and gaze shifts toward their audience. Heath (1986) studied the organization of talk and gaze during medical consultations, whereby patients may direct their doctor's attention toward parts of their bodies that need medical attention. Although some prior research had explored the relationship between talk and gaze (e.g., Kendon, 1967), these studies in the 1980s were seminal because they emphasized the sequential unfolding of human activity within specific situations: Rather than code the phenomena and count the frequencies of occurrences, these scholars transcribed and carefully analyzed particular strips of situated interaction.

Researchers who conducted sequence-analytic studies of videotaped interaction also turned their attention to hand gesture (e.g., Kendon, 1983, 1988; Goodwin and Goodwin, 1986), which has become an especially fruitful branch of naturalistic inquiry. When people gesture, they usually talk at the same time, coordinating their behaviors to be understood as an ensemble (e.g., Goodwin, 1986; Goodwin and Goodwin, 1986). Schegloff (1984) considered the connection between gestures and their "lexical affiliates" as evidence for the "projection space" during which an element of talk is in play, without having been uttered, allowing co-interactants anticipatory adaptations. Streeck (1993) showed how gestures may be "exposed" (i.e., made an object of attention during moments of interaction) through their coordination with indexical forms of speech (e.g., words such as "this") and eye gaze (which may perform "pointing" functions). Hands move within three-dimensional spaces that include objects and artifacts, and gestures may be largely recognized and understood through their relationship to the material world within reach (e.g., Goodwin, 1997, 2000b; Heath and Hindmarsh, 2000; LeBaron and Streeck, 2000). Furthermore, gesture may be embedded within extended processes or activities, such that any particular gesture is understood through its relationship to the whole activity (e.g., Koschmann, LeBaron, Goodwin and Feltovich, 2006). During this time, David McNeil (1992) and colleagues at the University of Chicago, including Susan Duncan (2002) and Susan Goldin-Meadow (2003), developed important frameworks for the analysis of gesture that were consistent with their orientation in psychology.

Meanwhile, interaction-focused researchers of gesture demonstrated that communicative acts are always "environmentally coupled" (Goodwin, 2007a), but can also structure the perception of the environment. Working as

an anthropologist in Chiapas, Mexico, Haviland (2000) documented the directional precision of a farmer's pointing gestures, suggesting that his gestures made his "mental map" interactively available, even interactively constructed. Gestures have been explicated as a locus of shared knowledge and emergent understanding (e.g., Enfield, 2008; Koschmann and LeBaron, 2002; LeBaron and Koschmann, 2003), organizing social interaction on the one hand and shaping individual cognition on the other (LeBaron and Streeck, 2000). Such studies of gesture have been more anthropological than psychological (e.g., Sidnell, 2005), emphasizing the public nature of "individual" cognition (Streeck, 2002), treating the human mind as something that extends beyond the skin to include social and material worlds. This research offers an alternative to views that are more psychologically oriented, such as McNeill's, who suggested that "gestures are the person's memories and thoughts rendered visible ... [belonging] not to the outside world, but to the inside one of memory, thought, and mental images" (1985: 12).

All the chapters in Schmitt (2007) focus on the delicate *coordination* of modalities, both intrapersonal and interpersonal, that bring about ordered and intelligible sequences of interaction. Deppermann and Schmitt (2007), who have done much to establish the study of multimodal interaction as a recognized field within European linguistics, conceive the study of multimodality as a study of *coordination*, on the one hand of different strands of bodily action within the single participant (self-organization), and on the other the coordination between co-interactants (interactional organization). The structuring of actions in one modality – for example, gaze – is clearly constrained by, or interacts with, those in another modality – for example, postural configurations or "F-formations" (Tiittula, 2007; cf. Kendon, 1976). As Mondada (2007b) has shown, self-organization is of particular importance where people participate in multiple activities at the same time ("multi-activities" such as conducting a conversation while driving a car or performing surgery while explaining the process to a remote audience). Lindström and Mondada (2009), building on work by Goodwin and Goodwin (1987, 1992), exemplify the multimodal nature of human interaction in a single language game, assessments of which are often performed through careful orchestration of talk, gaze, and facial displays (Ruusuvuori and Peräkylä, 2009), among other modalities. Krafft and Dausendschön-Gay (2007) introduce a useful distinction between "direct coordination" (coordination through the spatial organization of the bodies of the interactants) and coordination via objects, which occurs when participants use gestural and verbal acts of deixis to achieve a shared orientation to the setting of the interaction. Deppermann and Schmitt (2007) point out that research on multimodality complements the analysis of sequencing that is at the core of conversation analysis by an additional focus on *simultaneity*, that is, close attention to which behaviors are produced

at the same time and how such synchronous productions are possible. Simultaneity is a constitutive feature of any interaction, which implies the importance of spatial relations: how participants are positioned in relation to one another or where they look at any point in time is as important as the temporal relations between their talk and movements. This, in turn, points up the relevance of the *materiality* of communication modalities, for example the affordance of gesture to be perceived and processed simultaneously with speech as well as to attract and direct the addressee's visual attention (Heath, 1986; Streeck and Hartge, 1992).

That even speech alone comprises several modalities that must be explicated both in relation to one another and to their shaping, and functions in real-time interaction is the theme of a new paradigm within linguistics known as *interactional linguistics* (Selting and Couper-Kuhlen, 2001). One focus of this conversation-analysis-based field of studies have been the roles of rhythm and prosody in conversational interaction (Auer, Couper-Kuhlen, and Müller, 1999; Couper-Kuhlen and Selting, 1996a, 1996b; Uhmann 1992, 1996); another the emergence and operation of syntactic constructions in interactional contexts (Auer, 2009; Deppermann, Fiehler and Spranz-Fogasy, 2006; Günthner and Imo, 2006; Streeck, 1996b. See also Ford, Fox, and Thompson, 1998; Ochs, Schegloff, and Thompson, 1996.) Although we cannot cover this field here, it is important to note that it is guided by the same view of interaction as multimodal and of structural forms (constructions) as in part interactionally motivated.

Several of the contributors to this volume are linguistic anthropologists. Linguistic anthropology has given us several distinct analytic traditions; it is centrally concerned with the symbolic structuring of behavior. We have learned from linguistic anthropologists to attend to the social-symbolic significance of minimal differences in interactively produced forms (e.g., phonetic choices or prosodic contours; see Gumperz, 1982a, 1982b), but also to investigate such dimensions of embodiment in the context of culturally defined, regulated, and recognized events (Agar, 1975). In *Linguistic Anthropology*, Duranti (1997) presents the study of embodied interaction as one of the standard methodologies in contemporary linguistic anthropology. His own work is a good example of the inevitably "multimodal" nature of anthropological research into linguistic practice: studying the Samoan honorific system (which is expressed in verb morphology), Duranti (1992, 1994) discovered that the system is inextricably bound up with ways in which Samoans position themselves in relation to the place they are in and to one another.

MULTIMODALITY: EMBODIED INTERACTION IN THE MATERIAL WORLD

In a recent review of Tomasello's (2008) *Origins of Human Communication*, Kendon has emphasized, without employing the term, the inherently multimodal nature of human communication:

[T]he transition into referential or language-like expressions involved hands and body, face and voice and mouth, all together, as an integrated ensemble. What so many writers on this topic – "gesture firsters" and "speech firsters" both – pay little attention to is the fact that modern humans, when they speak together in face-to-face situations ... always mobilise face and hands and voice together in complex orchestrations... Every single utterance using speech employs, in a completely integrated fashion, patterns of voicing and intonation, pausings and rhythmicities, which are manifested not only audibly, but kinesically as well, and always, as a part of this, there are movements of the eyes, the eyelids, the eyebrows, the brows, as well as the mouth, ... patterns of action by the head, and ... from time to time variously conspicuous hand and forearm actions or "gestures" (Kendon, 2009: 363).

In the same vein, Stivers and Sidnell write that "face-to-face interaction is, by definition, multimodal interaction in which participants encounter a steady stream of meaningful facial expressions, gestures, body postures, head movements, words, grammatical constructions, and prosodic contours" (Stivers and Sidnell, 2005: 1).

Following Enfield (2005), they distinguish between "vocal/aural" and "visuospatial modalities." In contrast, we regard the abstraction of the interacting body from the material world as an abstraction with problematic consequences and – although we acknowledge the usefulness of terminological distinctions between different kinds or groups of modalities of communication – nevertheless insist that embodied interaction *in the material world*, which includes material objects and environments in the process of meaning making and action formation, is primary. Many of the contributions to this book therefore go beyond the study of the ways in which several bodily "channels" are coordinated in social interaction to show how environmental sources of meaning are drawn into the production of inter-subjective understanding and how interaction, in turn, structures its own semiotic and material environment.

Long before the term "multimodal(ity)" entered the field of interaction studies, it was established as a technical term in two entirely different fields, logistics and therapy. In the logistics industry, "multimodal" refers to the coordinated transportation of goods by air, land, and water; in medicine and psychotherapy, to the combination of multiple therapeutic practices, for example music therapy and talking cure or surgery and radiation. More recently, the term has taken center stage in computer science, where it describes human-computer interfaces that allow for multiple simultaneous input (e.g., by voice and gesture) and heterogeneous representations. Not very different from this usage is the term "multimodal corpora" applied to linguistic research, that is, the production of data representations that combine auditory and visual with textual representations (Kipp, Martin, Paggio, & Heylen, 2002). The term "multimodal communication"

is also used by various groups of researchers who seek to expand the semiotic analysis of texts so as to accommodate text-image combinations, but also other artifacts including films, buildings, and objects of daily use (Kress and van Leeuwen, 2001; Norris, 2004). Some of these researchers draw on Halliday's systemic-functional perspective (O'Halloran, 2004), others develop their own versions of discourse analysis (Levine and Scollon, 2004), but none approach human interaction in the material world with the rigorous microanalytic focus on the formation of action sequences that is characteristic of the contributions to this book.

When exactly the term "multimodal(ity)" entered the microanalysis of interaction is not entirely clear – certainly long before the appearance of Stivers and Sidnell (2005). What is equally certain is that the reconceptualization of embodied interaction as multimodal and the subsequent recognition of the importance of material contexts and artifacts drew a great deal of inspiration from – and partly overlapped with – two new, interdisciplinary research programs: *studies of work* (or *workplace studies*) and *science and technology studies* (see, among many others, Lynch and Woolgar, 1988; see also Heath, Luff, and Knoblauch, 2004). Inspired by these studies, sociologists became interested in the contingent, local production of practical, normatively accountable actions in the context of labor rather than conversational interaction. One of the hallmarks of this research program was recognition of the paramount importance of physical objects – things – in the conduct of work-related activities. Explaining the new research program, Garfinkel wrote that

it was evident from the availability of empirical specifics that there exists a locally produced order of work's things; that they make up as massive domain of organizational phenomena; that classical studies of work, without remedy or alternative, depend upon the existence of these phenomena, make use of the domain, and ignore it (Garfinkel, 1986: vi).

In an early, seminal study, Suchman (1987) demonstrated that normative rules of use are unable to guide (or explain) the operation of technological objects (in her case: copy machines), but that usage of such objects – and the normative accountability of such usage – represents ongoing, situated, contingent, and interpretive accomplishments. Understanding technology-supported action, as well as designing "user-friendly" technologies, thus requires the precise, moment-by-moment study of people's physical actions and the practical reasoning displayed by them, rather than reliance on decontextualized models of cognitive "plans" in the vein of Miller, Gallanter, and Pribram (1960). In another study, Suchman (1996) investigated how competent actors construct shared workspaces and arrange resources and tools to assemble readily interpretable surfaces that facilitate collaborative action. Suchman's work contributed to a growing trend among microanalysts of interaction to investigate talk and embodied communication not apart from, but

within complex material environments that they simultaneously make intelligible and coherent (Button, 1993; Engeström and Middleton, 1996).

A wealth of new research into hitherto unexplored domains of human action and interaction thus emerged. Heath and Luff (2000), in their wide-ranging research in contexts such as control rooms of the London Underground, computer-assisted architectural design, video-conferencing, and software development, focused attention on the difficulties of adapting new technologies to established orders of mundane reasoning and interaction. Rather than simulating face-to-face interaction, communication technologies such as video-conferencing demand that participants reconfigure participation frameworks and practices of turn-taking and speaker-listener coordination. With this widening of scope, compared to the initial focus on conversation over the telephone, ethnomethodological and interactionist researchers began to seriously implement Wittgenstein's vision that the study of a language must encompass the entirety of the community's language games and explicate them as forms of life. As McHoul (2008: 825) writes, "what we are ultimately interested in is taking pretty much any bit of ordinary everyday interaction as a means of understanding forms of life (Lebensformen) as such and not simply for its own sake as a technical object.... Conversation may be our favourite 'game', but it is not the only one in town."

A type of workplace that attracted particular attention were science laboratories, in which the study of work took on the form of studying the practices, instruments, and representations by which scientific findings are assembled and ratified as facts by the relevant community of scientific practice (Knorr-Cetina and Mulkay, 1983; Latour and Woolgar, 1986; Lynch, Potter, and Garfinkel, 1983; Lynch and Woolgar, 1988). Whereas scientific work – especially laboratory work – is inherently multimodal (it is the normatively guided coordination of practices of perceiving, experimenting, measuring, and representing that constitutes legitimate scientific practice), particular attention was paid to the production and interpretation of visual representations. In Latour's (2005) influential conception, dubbed "actor-network theory," agency is seen as being distributed across human actors and material things. Even though this view may not be universally shared by researchers of science, technology, and interaction, Latour's work has undoubtedly contributed to a scientific climate in which it is much easier to find acceptance for the notion that interaction, cognition, and work are inherently multimodal affairs that cannot be studied on the basis of what goes on in a single "channel" alone or by relying on textual representations abstracted from the rich contexts of the phenomena represented by them. The domain of things had rarely entered the picture in studies of conversational interaction, and never in studies of telephone conversation (but see Mondada, 2008b; Whalen, 1995). What is sometimes referred to as the "logocentric" bias in conversation analysis (e.g., Erickson, 2010) certainly has

its basis in the seeming irrelevance of material things to the organization of social interaction conducted over the telephone. In contrast, beyond the evident relevance of material objects, science settings impress on their observers the primordial multimodality of human action and interaction, the skilled, socialized nature of vision, olfaction, and manipulation, and the necessity of coordinating simultaneous, heterogeneous modes of perception and action in order to produce viable, coherent, and normal outcomes, be they manufactured objects, documents, or scientific findings.

The humanities and social sciences are presently witnessing a "spatial turn" (e.g., Warf and Arias, 2009), and the multimodal constitution of interaction spaces has become a very thoroughly studied topic in multimodality research (e.g., Hadington and Keisanen, 2009; McIlvenny, Broth and Haddington, 2009; Mondada, 2007a, 2009a; Müller and Bohle, 2007). Whereas this research takes up a concern that has been central to context analysis (see, for example, Kendon 1973, 1976; Scheflen, 1976), the current focus is on the multiple resources and modalities involved in organizing interaction spaces (e.g., gaze, posture, orientation) rather than on the way space is used to organize interaction. Schmitt and Deppermann write:

Interaction spaces are constituted by the interplay of physical circumstances which, because of their features, have certain implications for the structuring of interaction, and interactive accomplishments in which participants use these features as a resource for their situated ... praxis. Interaction spaces also are connected to certain structures of relevance, which are expressed, for example, through the symbolization of inclusion and exclusion. The concept "interaction space" describes dynamic, constantly changing constellations which partly reveal clear spatial contours (Schmitt and Deppermann, 2007: 96).

Work settings have become a prime site for research on human communication, symbolization, cognition, and interaction. Goodwin studied a variety of professionals at work – archeologists (Goodwin, 1994, 2000a), attorneys (Goodwin, 1994), oceanographers (Goodwin, 1995a), geochemists (Goodwin, 1995a, 1997), and airport technicians (Goodwin, 1996; Goodwin and Goodwin, 1996) – to reveal the nature of what he called "professional vision." He showed how experts interactively direct the attention of others, showing them *what* to see and *how* to see it, using tools to highlight and code their surroundings, while articulating the upshot and import of what is being seen. Such research went beyond the coordination of talk and embodied behavior to consider the entire "contextual configuration," which included "a range of structurally different kinds of sign phenomena in both the stream of speech and the body, graphic and socially sedimented structure in the surround, sequential organization, encompassing activity systems, etc." (Goodwin, 2000b: 1). Scholars have also examined medical consultations (e.g., Beach and LeBaron, 2002; Heath, 2002, 2006; Mirivel, 2007; Modaff, 2003; Robinson and Stivers,

2001), surgeries (e.g., Koschmann, LeBaron, Goodwin, Zemel and Dunnington, 2007; Zemel, Koschmann, LeBaron, Goodwin and Dunnington, 2008;), therapy sessions (e.g., McMartin and LeBaron, 2006), police interrogations (e.g., LeBaron and Streeck, 1997), job interviews (e.g., LeBaron, Glenn, and Thompson, 2009), business negotiations (e.g., Streeck, 1996a), museums (e.g., vom Lehn, 2006; vom Lehn, Heath, and Hindmarsh, 2005), and more. Researchers have closely examined the situated work of anesthesiologists (Hindmarsh and Pilnick, 2007), physicists (Becvar, Holland, and Hutchins, 2005; Ochs, Gonzales, and Jacoby, 1996), architects (e.g., Murphy, 2005), auto mechanics (Streeck, 2002), hairdressers (LeBaron and Jones, 2002), auctioneers (e.g., Heath and Luff, 2007), pilots (Hutchins and Palen, 1997), and so forth. An especially promising stream of workplace studies focuses on the interaction of people with and through technology (e.g., Heath, Luff, and Knoblauch, 2004; Heath, vom Lehn, and Osborne, 2005; Hindmarsh, Heath, and Fraser, 2006; Luff, Hindmarsh, and Heath, 2000). Of particular interest to some of these researchers is a domain virtually non-existent during the early days of context analysis, namely the constitution and configuration of virtual or distributed interaction spaces, for example in online environments or in the context of interactions mediated by visual technologies such as video-conferencing (Heath and Luff, 1992, 1993; Mondada, 2007b). Contributors to McIlvenny, Broth, and Haddington (2009) have investigated interaction in mobile technological environments, including motor-vehicles and mobile-phone networks, in order to establish how talk and interaction are adapted to settings other than the primordial face-to-face formation. These technologies also pose new challenges for research methodology, in particular when interaction mediated by video – that is, interaction whose participants are distributed across different locations – is studied by means of video (e.g. Heath and Luff, 2006; Knoblauch, Schnettler, Raab and Soeffner, 2006; LeBaron and Koschmann, 2003; Mondada, 2006, 2009b; Ochs, Graesch, Mittman, Bradbury and Repetti, 2006; Suchman and Trigg, 1991).

Altogether, this growing body of research gives us a rich understanding of human activity within organizational settings: By carefully examining embodied interaction (including talk), analysts show how people interactively draw on a wide range of social and material resources to negotiate and constitute their institutional lives – which are intellectual and relational, artifactual and technological, cultural and political. Most of the chapters in this book consider multimodal features of human interaction in relation to their organizational or institutional settings. Although workplace studies currently dominate the field, some lines of research have different and important emphases.

Thus, multimodality has become a concern within more traditional fields such as childhood and family communication. For example, Marjorie Goodwin has examined the embodied linguistic resources of children

at play: During games of hopscotch, girls locate and maneuver their bodies relative to each other and their hopscotch grid in order to produce disagreements and arguments (e.g., Goodwin, Goodwin, and Yaeger-Dror, 2002) that achieve a social stance that is simultaneously cooperative and conflictive (M. H. Goodwin, 2006; see also Ochs, Smith, and Taylor, 1989; de León, 1998, 2007.) Even pre-lingual children are capable of rather nuanced and sophisticated forms of social interaction as they produce recognizable courses of action, showing that they expect their actions to be recognized (Lerner and Zimmerman, 2003), relying on subtle patterns of gaze (e.g., Kidwell, 2005) and body orientation (e.g., Kidwell and Zimmerman, 2007) as they engage with other children, objects, and caretakers. Video recordings of naturally occurring interaction sometimes evidence awareness or capability that may escape the notice of more mainstream methods for observation and assessment. For instance, Goodwin's (1995b) study of an aphasic man showed how people with severe disabilities may compensate for their physiological limitations through ways of interacting with other people: Although partially paralyzed and only able to speak three words, the man participated actively in family conversations and decisions as he carefully coordinated his utterances with gestures such that others could articulate his contributions. Analysis of how people with language and other disabilities are able to use the resources provided by the organization of talk and embodiment in interaction has led to important new insights about the nature of such disabilities, and how the lives can be improved by shifting focus from the individual deficits to the collaborative organization of action and meaning within situated human interaction (see, for example, the work of Wilkinson and his colleagues, including Chapter 11 of this volume, and the range of scholars who contributed to Goodwin, 2003). Some of the chapters in this book focus on the interaction of children, including those who are pre-lingual and disabled, which are recent and promising lines of inquiry.

As video-based methods have become more popular, and as empirical studies and findings have accumulated, researchers have been reflective about their approaches and assumptions. Although much has been accomplished, there is nonetheless a "lingering dualism" (Streeck, 2003) in naturalistic research that often maps onto the Cartesian divide between mind and body. Theoretical, conceptual, and empirical achievements (and misfortunes) necessarily emerge from the everyday practices and mundane methods of social science. Too often, analysts regard talk as their starting point, even when talk appears late in the order of things accomplished in face-to-face interaction. Referring to the work of conversation analysts, ten Have (1999) observed that even when videotapes are used, investigators usually start with an audio transcription so that "the verbal production by participants is seen as the base-line for understanding of the interaction, with

selected visual details being added to this understanding [subsequently]" (9). Arguably, analysts should consider visible phenomena from the outset, especially when "body parts are the first mediating elements in our interaction with the people and objects around us" (Duranti, 1997: 322; see also de León, 1998). In a special journal issue of the *Journal of Communication* on the emerging integration of verbal and nonverbal research, Jones and LeBaron (2002) wondered whether audio recordings of face-to-face interaction are sufficient, and suggested that "complete audiovisual records be the basis for future research" (512). When people interact within embodied social frameworks that are structured and changed through their shifting co-presence, analysts should attend to what the participants themselves are treating as important.

OVERVIEW OF THE VOLUME

Founding capacities

The chapters in this initial section demonstrate some of the ways in which the phenomena participants construct within interaction, including units, stances, action, and the forms of understanding they display to each other, are built through the use of diverse linguistic, material, and embodied resources. All of the chapters demonstrate the crucial contributions made—not only by the speaker, but also other usually silent participants—to the organization, comprehension, and integrity of the actions in progress.

Hutchins and Nomura offer a very original analysis of the organization of gesture, talk, and writing within a distributed multi-party, multimodal interactive field. They examine videotapes in which two Japanese airline pilots, who occupy different positions with different duties on the flight deck, are being briefed by an American instructor in preparation for training exercises in a flight simulator. Throughout this process, not only the speaker but also hearers use gesture to enact the embodied actions that will be required to manipulate the aircraft controls in order to accomplish the instructions found in both the instructor's talk and the written documents that specify the procedures to be followed in the cockpit and simulator. The actions being investigated are simultaneously collaborative in that they are constructed through the differentiated actions of multiple parties, and multimodal in that they are built through the mutual elaboration of different kinds of semiotic materials, including talk, gesture, and writing.

Many previous studies of gesture have focused on phenomena such as co-expressive gesture and the relative timing of a gesture and its lexical affiliate. Hutchins and Nomura move beyond this framework to describe how collaboratively constructed conceptual representations that include talk and gesture are situated at the intersection of multiple attributes including temporal placement

(either simultaneous or offset) and different kinds of semantic relationships between talk and gesture. While the semantic relationships between different elements of a multimodal configuration can be congruent, they can also be complementary in two different ways: 1) metonomy, in which relationships between cause and effect are made visible (something that is most relevant to depicting the relationship between an instruction to produce a particular outcome, and the embodied actions with specific controls needed to accomplish that outcome); and 2) synecdoche, where one representation refers to a whole and the other to a part of that whole. Temporal offsets in the placement of gesture provide speakers with resources for working around the constraints imposed by the temporal unfolding of linguistic items in talk by enabling them to highlight a salient action before it is fully specified within the stream of speech. Temporal juxtaposition of gesture by hearers with emerging talk makes it possible for pilots receiving an instruction to demonstrate their embodied understanding of what it means to carry out the requested action, within the specific environment of the cockpit, by displaying manipulation of relevant controls. These are in fact never mentioned in either the printed instructions or the talk.

Two features of this analysis can be briefly noted. First, it provides a strong demonstration of the importance of cognitive ethnography. The analyst's ability to understand the relevance of the movements of the participants' bodies requires thorough knowledge of both the environment that is the focus of their concern, the flight deck with its specific controls and tasks, and of the embodied actions that habitually occur within that environment. This knowledge is also used in very interesting ways by the participants themselves. For example, the instructor's gestures sometimes take into account the specific positions of different pilots within the flight deck (one sitting on the left, one on the right). Second, most prior research has treated gesture as something done by speakers (but see Goodwin, 2000b). In this work, gestures produced by hearers are given an equally prominent role within a process of public, collaborative imagining, as participants discuss together future courses of action.

Lerner, Zimmerman, and Kidwell describe how even children who have not yet acquired the ability to speak can engage in consequential interaction with others by parsing mundane natural activities, such as the process of serving food to another child, into sequential structures where possibilities for participation systematically and visibly emerge. The children thus actively use the visible embodied behavior of others to create interactive contexts that guide the organization of their own action. The sequence the authors examine in the current chapter is drawn from a corpus of videotapes of twelve- to thirty-month-old children in infant-toddler day-care centers. Not only do such young children robustly display the ability to analyze detailed structure in the actions of others, but they also produce their own actions with

an orientation to how their movements (for example, pointing toward something relevant in the local environment) will be treated as forms of action by others. Long before they have mastered the rudiments of language, the children are competent, thoroughly reflexive interactive actors.

The chapter focuses primarily on the actions of Laura, a sixteen-month-old girl who is present as another toddler is fed, but who is not permitted to eat herself. The feeding of the other girl is accomplished through a series of discrete sub-activities, including successively retrieving a bib, washcloth, and food containers from a tray on a railing above the children and placing these in front of the child to be fed. Each of these activities has a projectable sequential structure of its own, including, most crucially, a "task transition space" at the transition between one activity and another. As the feeding is being prepared, Laura systematically places bids to be fed herself (such as pointing gestures toward the food with appropriate vocalizations), and thus be included in the activity, precisely at these transition spaces. She thus demonstrates through her own embodied action that she has parsed the actions she is witnessing into just those sequential structures that would provide for her possible inclusion in the activity, recognizing simultaneously the completion of one task and the relevant projected occurrence of a next.

Eventually it becomes clear that Laura will not be given food. Her recognition of this is displayed through a change in the sequential organization of her action. Rather than placing bids at transition-relevant places, she now produces loud cries of complaint in the midst of the actions being carried out by others. She displays a practical grasp of the emerging organization of a local routine in terms of the possibilities for co-participation each next moment does or does not provide. The authors argue that this ability offers an alternative to models, such as scripts, that posit overall cognitive templates for recognition of larger activity structures. Moreover, such projection of emerging courses of action by others constitutes the basis for what is frequently analyzed as the ability to recognize the goals and intentions of others.

Enfield's chapter proposes a framework for pragmatic analysis in which "the interpreter, not the producer, is the driving force in how utterances come to have meaning." Such a perspective stands in strong contrast to much work in both formal linguistics and fields such as speech act theory, which take as their primary domain of study the mental life and/or sign activity of a speaker constructing an isolated utterance. In shifting analytic focus to the investigation of how what Peirce described as interpretants are constructed, Enfield uses as his point of departure both the classic work of Pierce (1955) and recent research in linguistic anthropology by Kockelman (2007). Enfield notes that whereas all meaning making requires an interpreter, a sender is not necessary for the construction of many relevant signs (consider, for

example, how smoke is seen as a sign for fire, despite the complete absence of anything like a communicative intention on the part of the fire).

Any particular utterance will contain, in addition to lexicon and semantics, a host of other kinds of embodied signs, such as prosody and gesture, and will occur in a particular spatial, material, and cultural environment, all of which will modulate the meaning of what is being said. Enfield seeks to develop a framework for the study of meaning making that encompasses all semiotic modalities, including gesture, gaze, and posture, and which, moreover, takes into account how actors are situated within meaningful, historically structured environments. Peirce provides resources for doing this.

Intentionality is central for prototypical cases of human communication through language. However, rather than focusing on the mental life of the speaker to investigate intentionality, Enfield argues that the task faced by someone producing an utterance is pre-supposing the kinds of interpretative work that will be done by their addressees, and in light of this placing in a public environment appropriate signs that will guide relevant interpretative work: "To communicate is not literally to send a message but to make public the means ... for another person to build an adequate understanding in response."

Enfield argues that the basic unit to be focused on is what Goffman called the move within communicative action sequences. The move is both a building block for larger structures and contains within itself internal complexity drawn from a range of different kinds of semiotic materials. Different signifying materials can be described in fine detail with respect to an array of attributes. Within the process of interpretation, these varied semiotic resources are integrated into a unified sign-vehicle that expresses the move's informative intention. One effect of the analytic perspective developed by Enfield is to shift away from language as an isolated, self-contained system, to focus instead on the rich semiotic and communicative ecology within which language is embedded through human conduct.

Streeck's chapter opens up for investigation interstitial meaning-making practices. In much academic inquiry, the scene within which participants build action together is divided into categorically distinct phenomena, such as language, material objects, gesture, and writing. However a range of crucial hybrid acts come into existence in the spaces created between these canonical distinctions. These activities move across boundaries and intertwine diverse materials into enduring structures that reshape not only the physical and semiotic environment, but also the ways in which participants classify for each other what is happening in their interaction. Two sets of such practices are examined: first, the use of found objects to build interactive action; and second the act of writing, a process that encompasses both the visible activity and the enduring inscriptions that are produced by this process.

Noting that "social interaction is a vociferous process, always hungry for stuff out of which signs, symbols and scenic arrangements can be made," Streeck begins by looking at how two businessmen negotiating a new contract use the physical presence of their product – cookies and the packages they are sold in – to symbolically classify what they are talking about, and the character of their interaction with each other. For example, the physical properties of the aluminum bag used to package the cookies becomes an enduring display of one of the focal issues being discussed in the talk, namely the ability of the cookies to retain their freshness. By putting the bag down with a dismissive gesture, one speaker is able to signal the impossibility of entering the American market. In this process, the instrumental action of placing the bag back on the table and the material structure of the bag are transformed into signs that contribute to the dense organization of the participants' local practices of constructing meaning for each other.

Streeck provides a range of examples of how mundane objects are successively transformed into situated symbols that, because of their enduring physical presence, become external memories of actions accomplished, and topics addressed, within a particular conversation. He then investigates a range of interactive activities accomplished through writing. For example by actively using writing to divide a page into costs and profits allocated among the participants, they are able to rhetorically and dramatically make visible as a staged performance the particulars of the financial relationship they are negotiating, and moreover to leave a permanent trace of that process. Despite its apparent simplicity, the drawing of lines provides participants with powerful interactive, pragmatic, and symbolic resources.

The interstitial position of the practices Streeck examines, and the way in which they establish links between different kinds of domains, creates possibilities for social and cognitive blending as meaningful gestalts are projected from one system to another. The effect is the ongoing transformation of meaning and action through powerful semiotic bricolage across modalities and participants within situated interaction.

Tulbert and Goodwin focus on what they call Choreographies of Attention. They examine how participants use their bodies, local activity systems, the organization of inhabited spaces, and material objects to construct, contest, and negotiate the frameworks of intersubjectivity and co-orientation that ground central forms of human action, such as directives. Their data consists of video recordings of the daily lives and household activities of thirteen families in Los Angeles, California.

To make possible the comparative study of diverse phenomena implicated in the organization of multi-party embodied interaction, Tulbert and Goodwin examine how a single activity – parents getting their children to brush their teeth – is organized in a range of different families. Such directives make relevant mutual alignment

between participants, something that is publicly visible through both the spatial organization of the participants' bodies and the establishment of a joint focus of attention. However, as parents attempted to initiate such directives, the children frequently attended to competing foci of attention such as computers and television sets. Establishing mutual orientation between participants, and alignment to the activity of tooth brushing – the establishment and ratification of a shared phenomenal world implicated in the organization of action – was thus something that required interactive work, and moreover was a project that could fail. Some of the caregivers who succeeded explicitly prefaced their directives by extinguishing competing activities, for example by turning off computers that children were looking at. Alternatively, frameworks of embodied mutual orientation were sometimes choreographed by physically moving a child's head so that the directive occurred within a participation framework where the parties were mutually gazing toward each other. Such phenomena highlight the way in which a successful directive requires cooperative alignment. Because of this, the directives also provided an environment where children could explicitly refuse to align with the caregiver, so that the requested activity never moved forward. Rather than simply existing as a speech act, entirely in the stream of speech, directives – and the shared foci of attention they require – are organized within embodied choreographies of mutual alignment.

Tooth brushing occurs within culturally organized spaces and arrangements of objects within the household. By having their activities guided through the bodies of caregivers, very young children acquire the ability to discern the affordances provided by these spaces, and the tools and artifacts they contain. Even quite young children demonstrate the ability to navigate through the house as an organization of meaningful activity spaces, as they mobilize where they have to go and what they need to carry out a particular task. The importance of the cultural organization of such taken-for-granted spaces is brought into sharp relief in a number of cases in which caregivers tried to initiate the activity of tooth brushing, not by directing children to the bathroom, but instead by carrying toothbrushes to rooms where children were watching television. This not only led to long delays, but sometimes the caregiver's attention shifted to the competing activity as well.

The comparative framework provided by examining how different cohorts of participants accomplish the same activity within the endogenous spaces, and social arrangements that structure their daily activities, makes it possible to see both the diverse alignments of language, bodies, and spaces that must be mobilized to accomplish a particular activity, and the cooperative stances (Garfinkel 1967, Goodwin 2007c) required for a directive to be successful.

Aoki investigates in detail the interactive organization of nods by speakers in Japanese conversation. Her study complements a range of earlier research on recipient head nods. Speaker nods are organized simultaneously with reference to two different kinds of phenomena in the co-occurring interaction: 1) the sequential and linguistic structure of the talk in progress, including positioning within the turn, and the presence of a range of different kinds of particles; and 2) the current, projected, and requested actions of the turn's hearers. Speaker nodding is thus organized with reference to a range of different semiotic fields, including linguistic and sequential structure and the embodied displays of participants who frequently are mutually oriented to each within a participation framework characterized by mutual gaze. In the majority of cases, speaker nods are designed to be seen and responded to. Speakers attend to what their addressees do in response to a nod because that may be consequential for what the speaker will do next. Thus, though produced by a single party, speaker nods are organized as components of multi-party interactive action.

Using a clear and simple system for transcribing nods, Aoki distinguishes two different kinds of speaker nods: a singular regular nod, and a stretched nod that begins with a head raise that is held for variable lengths of time before the head falls to complete the nod. The frequencies of regular and stretched nods are approximately the same. Most occur in the final part of a prosodic unit, with the final fall of the speaker's nodding head occurring over the final mora of a prosodic unit.

In turn-final position, speaker head nods frequently occur with other devices that function to elicit responses from recipients, such as high-rising terminal pitch contour, and a variety of interactive particles. However, speaker head nods can solicit responses from recipients even when produced alone, without other elicitation signals. In the midst of prosodic units, speaker nods can signal to recipients that responses are relevant at specific salient moments within the unfolding structure of the talk. The recipient nods that follow operate on these specific parts of the turn that speaker, through the production of a nod, has indicated as meriting special attention.

Though occurring in a variety of different positions within the organization of the turn, and with respect to the co-occurring talk, speaker head nods function interactively to signal to recipients that a response is immediately relevant. Aoki demonstrates, by examining speaker head nods in a range of different positions, that they provide participants with important resources for calibrating with each other their ongoing understanding of the events they are accomplishing together.

Iwasaki's chapter develops an important new perspective for the analysis of how units are constructed in interaction. It has long been recognized that the units used to build turns at talk (what conversation analysts call Turn Constructional Units, or TCUs) in Japanese have a noticeably segmented or permeable character (Fox, Hayashi, and Jasperson, 1996). Unlike what happens in English, the units that make up individual turns in Japanese

frequently emerge bit by bit, with crucial grammatical and sequential information being provided only at the very end of the unit. This process limits the resources that make it possible for hearers to project the upcoming structure of a unit (Hayashi, 2004b; Tanaka, 2000).

Iwasaki focuses analysis on the subunits found within individual TCUs and demonstrates that these units are themselves constructed through systematic processes of interaction between speaker and hearer(s). These sub-units, which Iwasaki calls Interactive Turn Spaces, are not static, single-party activities, but instead are processes constructed through the mutual monitoring and ongoing interaction of multiple parties – hearers as well as the speaker – that are coordinating their collaborative alignment through not only talk, but also with a range of other embodied modalities, such as facial displays. The locus for projection within Japanese is lodged within these units rather than the turn as a whole.

In order to make her arguments about the multi-modal, multi-party organization of units, Iwasaki found it necessary to develop new methods for transcription to display the intricate, multimodal organization of units she is examining on the printed page.

Some of the phenomena that are central to the analytic arguments in the chapter can be briefly illustrated with one of Iwasaki's examples. In Japanese, grammar nouns have a particle (case or adverbial) attached to them. In excerpt 2 in Iwasaki's paper, a speaker produces a noun, but then pauses without proceeding to the grammatically necessary following particle. Simultaneously his face breaks into a smile, producing a display of stance toward that noun ("drug stuff"), and his gaze moves to his addressee. Only after the addressee produces a reciprocal display of alignment does the speaker at last move to the required particle. What has traditionally been considered an indivisible grammatical structure in Japanese – Noun + Particle – here becomes a site for multi-party, multi-modal interaction (see also Hayashi, 2004a). Processes of interaction between speakers and hearers are thus constitutive of the basic units from which talk in interaction is built.

Transformational ecologies

The chapters in this section investigate how meanings of actions are created or altered and how difficulties in producing coherent and intelligible actions and utterances as a result, for example, of communicative impairments or disorders, are overcome by skilled management of the material, behavioral, and human environment, sometimes creating opportunities for the further development of skills.

Mehus investigates how caregivers in a day-care center use multimodal resources to construct contexts for young children's imminent actions, specifically those that would otherwise threaten to disrupt the social order. Adult-child interaction is characterized by asymmetries

of symbolic competences and capacities to manage one-self and the interaction at hand, and adults often need to provide scaffolds to enable the interaction to proceed. The practices that she describes and which she aptly calls *anticipatory contextualization* alter the significance of actions in which a child is engaged or is about to engage. Instead of directly intervening in or inhibiting an action by a child, a context is provided for it so that it remains within the boundaries of the agreed-on social order. There is not a set repertoire of such recontextualization practices. Rather, caregivers draw on local-cultural resources such as specialized lexicalized items and routines, as well as innovative speaking practices, in an improvisational manner that Mehus characterizes as *multimodal bricolage*. While the methods that she describes are improvisational rather than pre-established, there is always the possibility that routines evolve from them and become sedimented as elements of the practical culture of the local community, a culture that usually interacts with social systems that extend beyond the community at hand, such as historically dominant ideologies of childhood and child-rearing. Mehus' study exemplifies that the significance of an action is not a direct product of the motives and goals that the agent invests in them, but rather issues from its embeddedness in a context, for example the present constellation of co-participant actions and orientations and props for shared and individual activities, so that the meaning and upshot of an ongoing action can be modified by others who may reconfigure that constellation without impacting the physical action itself. Action is *distributed* (see Hutchins and Nomura, Chapter 2, and Goodwin, Chapter 13 of this volume).

Multimodal bricolage does not only describe the logic and semiotic features of a setting as rich in objects and physical activity as a day-care center, but also the practices of meaning making in settings that have traditionally been called "verbal interaction." Gullberg describes how non-native speakers of a language (L2 speakers) deploy and coordinate hand gestures to support and complement their linguistic production and how the joint deployment of speech and gesture may yet require the systematic mobilization of additional resources and dimensions of interactional co-presence such as differential uses of space and gaze direction. Like Mehus and many other contributors to this volume, Gullberg analyzes the perceptual and communicative ecologies of communicative encounters and points out that hand gestures alter the communicative environment; they do not simply add additional forms and significance to what is simultaneously being conveyed by speech, but – in L2 conversations as well as in conversations between native speakers (Goodwin, 1986; Streeck, 1993) – "are carefully deployed to elicit lexical assistance from the interlocutor" (7). Speakers systematically create differential ecological constellations for their gestures – for example, by choosing between a location near their lap or an

exposed position closer to the recipient's line of regard – and thereby constrain and impact the interlocutor's co-participation in different ways: While a gesture can be made in a way to request an overt interpretation (e.g., in the context of a word search), it can also be made in ways that discourage others from contributing. In either case, we see that not only action but also meaning making is distributed between participants and modalities, and that the integration of distributed resources is itself mediated and structured by shifting local ecologies. Ironically, in the context of L2 conversations where one or more participants desire to acquire a language, the short-term benefit of achieving inter-subjectivity and facilitating the progress of the conversation that the systematic deployment of non-linguistic resources such as hand gestures can incur the long-term cost of negatively impacting language learning: The routine use of "compensatory" practices of meaning making can slow down or even stall the acquisition of lexicon and grammar.

A different type of cost-benefit ratio is at issue in Wilkinson, Bloch, and Clarke's study of graphic resources such as writing on paper and hand as well as keyboard-based communication technologies in interactions involving people with communication disorders. Their chapter exemplifies the immense need for microanalytic studies of the deployment of these new technologies in everyday interaction, not only to enable engineers to develop devices that are easy to incorporate in the flow of face-to-face interaction, but also to educate their users and their interaction partners about best ways of adapting interaction practices to the constraints of the technologies. When people who have a difficult time finding or articulating a word temporarily turn to writing to bridge what would otherwise become a gap in the process of understanding, their co-participants usually have no problem shifting their gaze along with them to the surface on which the graphic signs are made, and to return them to the speaker once the interaction is shifted back to the vocal modality. Such shifts in the focus of attention are familiar and commonplace, and the production of multimodal utterances that comprise both vocal and graphic components does not appear difficult at all, despite the differences in the temporal organization of the two modalities. Most participants in interaction are familiar enough with the cognitive and behavioral requirements of writing so that they can easily combine this activity with simultaneous talk. Voice-output communication aides, on the other hand, as they are used by people with cerebral palsy, produce auditory output interrupted by long pauses, and in the situation that Wilkinson, Bloch, and Clarke describe lead to many *anticipatory completions* on the part of the interaction partner. Under this circumstance, the speaker who has the communicative disorder must invent or learn preventive practices to secure authorship of their turns at talk and agency in the interaction. There is also the difficulty of maintaining conversational coherence and to display the type of competence and wit that is usually demonstrated by close timing (e.g., of a response). Rather than just aiding and supplementing verbal communication, every communication technology – just like each bodily modality that is deployed along with speech or in which speech is embedded – comes with its own affordances and constraints and reconfigures the ecology of action and interaction within which the parties operate.

In the case of the blind, these ecologies are still to a large extent *terra incognita*, as Avital and Streeck suggest. In contrast to the tremendeous amount of research on deaf signers and their languages that has been conducted – and notwithstanding the existence of a very active community of educators and psychologists who conduct research on blind children's linguistic and social development and design interventions – there have been virtually no microanalytic studies of social interaction in the everyday lives of the blind. Accordingly, we know little about the ways in which non-visual resources and practices scaffold social interaction, inter-subjective understanding, and joint and distributed activities. In their exploratory ethnographic study in a school for blind and visually impaired students, Avital and Streeck report the use of acoustic signals – which they dub *auditory gestures* – by which these students facilitate each other's navigation and participation in classroom activities, and they describe how voice projection and orientation are used to organize frames of focused interaction. They also explore moments of behavioral coordination that suggest that specific sensory resources (collectively known as *peripersonal perception*) are utilized where sighted individuals rely on visual information to manage shifting frameworks of participation. As is the case in other chapters, their investigation suggests that often more insight can be gained into the processes of meaning making and interactional collaboration by focusing on how the context is being organized than on the individual agent's behavior: They describe instances of stereotyped, repetitive behaviors, known as *blindisms*, for example constant rocking or jumping up and down from a chair, that are common among many blind children and that themselves appear to provide a kind of ecological compensation to children whose ability to move about is otherwise severely constrained by their inability to see; one might expect such behaviors to be highly disruptive, rendering conversation nearly impossible, but the children's interactions with familiar partners such as classmates and teachers who know to disattend these behaviors seem to be completely unaffected by them.

Goodwin shows in his chapter not only that action is multimodal in the sense that it is "constructed from structurally different kinds of sign phenomena that mutually elaborate each other" (6), but also because actors systematically incorporate materials from the actions of others in their own productions – a phenomenon that Bakhtin has explicated under the term *answerability* (Bakhtin, 1990) and that Sacks (1992) has called "tying techniques."

Goodwin analyzes moments of inter-subjectivity in the conversations of an aphasic man, who can utter only three words, and members of his family. He describes the practices that the speaker uses to get his conversation partners to produce actions and talk to which he can then tie his own multimodal utterances – a word, a series of gestures, an intonation contour superimposed into nonsense syllables – so that their significance becomes evident from the context. The speaker, despite lacking "the semantic and syntactic ability to construct sentences that would state [his] proposition by himself, ... is able to vastly expand his repertoire as a speaker by sequentially tying to the particulars of the complex talk and language structure of his interlocutors" (14). Goodwin reveals the enormous complexity of *cooperative semiosis*, of the distributed production of action by interaction participants who draw on vast and shifting arrays of symbolic and environmental resources, notably including their own and each other's bodies. By subtly redirecting their gaze to a sign inscribed in the environment or by borrowing the syntactic and prosodic structure of the other's previous turn at talk, they may profoundly alter the horizon of significances within which bits of behavior are taken to make sense. The processes of interacting with an aphasic man, who is incapable of producing *symbolic speech*, reveal the indexical underpinnings of human action and communication while at the same time showing the difficulties of purely indexical and iconic communication. Symbolic language, in contrast, is a resource that links "the cognitive lives and abilities of different actors ... together in ways that enable the fluent accomplishment of radically new forms of action" (26–7). It allows actors to immediately recognize each other's actions. As Goodwin argues, the example of the aphasic man who, through managing his human environment, is able to produce complex semantic structures and action by tying his limited verbal productions to the complex utterances of others, also suggests that we reconceptualize what a *speaker* does: Instead of being solo productions, action and meaning are more often than not "organized through the cooperative semiosis of multiple actors" (28).

Whereas the other chapters in this section investigate ecologies of embodied co-presence and their refiguration and transformation through multiple modalities and technologies, Keating and Sunakawa study how action and semiosis are restructured in the digital world. Examining both multiple-player online gaming and digital (visual) communication technologies for the deaf, they show how digital technologies transform contexts of social interaction as well as some of the competencies that participants need to develop in order to be able to achieve coherence within and across different interaction spaces and to move back and forth between "real space" and the multiple virtual spaces that the digital realm offers. On the one hand, digital technologies constitute new "extensions" (McLuhan, 1994 [1964]) that give participants new and expanded ways of looking,

moving, and self-transformation. At the same time they also impose new constraints on what can be done "in real space" without disrupting the shared activity. Keating and Sunakawa lay out a research perspective that will enable us to keep track of and understand how human agents develop new skills, new cognitive architectures, and new methods of cooperating and sense making to exploit the possibilities that technology gives them to extend and diversify their abilities to act in the world and shape the actions of others: "Human action and communication in a sociotechnical environment entail both restrictions and enhancements of possibilities compared to face-to-face interaction, and participants manage new challenges through reinterpreting constrastive properties of communicative resources and the environment as well as by distributing meaning across multiple modalities both simultaneously and sequentially" (26).

Professional communities

When researchers examine human activity within organizational and professional settings, they have an opportunity (some might say obligation) to consider how interaction relates to the purposes, policies, practices, and histories of the institutional host. Within the naturalistic research tradition, there is an abiding sense (perhaps consensus) that mundane or everyday forms of interaction are shaped or imprinted by the organizations that appropriate them, according to the organizational purposes that they are used to serve (e.g., Drew and Heritage, 1992). At the same time, of course, forms of interaction also shape and help constitute the appropriating organizations. The chapters in this section explore these issues and contribute to our understanding of embodied interaction within organizational settings that are often complex and subtle.

In the opening chapter of this section, Mondada addresses issues of multimodality through a careful consideration of "multi-activity" – that is, when people are simultaneously engaged in more than one activity, such as talking while driving, driving while eating, or eating while working. Multi-activity settings typically involve complex configurations of spatiality (built spaces that both afford and constrain), materiality (objects, artifacts, tools, representations, etc.), and participation (varying levels of involvement, shifting orientations, and more). Analysts' descriptions of multi-activity, she argues, should be informed by the behavior of the participants themselves, who necessarily organize and coordinate their situated involvements. Eventually, Mondada turns her attention to a complex corpus of videotaped data in which members of a surgical team are simultaneously involved in two distinct activities: While the chief surgeon and his team conduct a surgical procedure, the chief surgeon simultaneously gives demonstrations and instructions to an audience of about one hundred trainees and a few experts, who watch the surgery by video

link and ask the chief surgeon questions via an audio connection. Thus, both activities are accomplished through talk and embodied actions, carefully coordinated so that the multi-activity coheres without running at cross purposes. In great detail, Mondada unpacks the complexities and subtleties of surgical activity that are inherent to the purposes of the host organization – a *teaching hospital.*

Zemel, Koschmann, and LeBaron contribute another study of a surgical team operating within a teaching hospital. In addition to successfully completing a surgical procedure, the senior and most expert surgeon ("Attending") is responsible for the education and training of a less experienced surgeon ("Resident"). The authors focus on a strip of interaction in which Attending asks a question that Resident has difficulty answering. Rather than provide the answer to his own question, Attending pursues a correct response from Resident by deploying various modalities of interaction in order to invoke and navigate the complex referential relevancies that would allow the Resident to produce a proper answer. The work of producing a question that Resident could answer demonstrated the indexical ground shared by Attending and Resident, and indicated the Resident's level of expertise. Talk alone was simply not enough to accomplish all this. Gestures and other embodied referential actions, the patient's own arm, and the talk of Attending and Resident were all used in coordinated and artful ways to make evident the relevant states of the surgical site and the various issues to be addressed in the surgery.

Murphy examines the storytelling practices of professional architects. His chapter begins with a broad overview of how narrative has been conceived and studied, as both a macro- and micro-social form, by scholars of various disciplines, including discourse and conversation analysis. Against this theoretical backdrop, he focuses specifically on the sense-making features and functions of narrative, and how embodied actions help constitute narrative structure and activity in interaction, because stories are not merely *told* in interaction but are also *demonstrated* in rather complex ways. When architects use narrative to do professional work, their stories are less about the empirical past and more about an imagined future. Working with the tools of their trade, such as computer-generated drawings, they perform gestures and embodied maneuvers as the animation of their architectural plans, so that they can "experience" their plans in the present and solve the problems that they "encounter" – before those plans actually take shape as steel, mortar, and glass. Murphy identifies and explicates a recurring narrative form that seems especially useful to architects: He calls it the *embedded skit* – that is, a short and rather undeveloped vignette from an architect's imagination, with enough substance to resonate as relevant to the current conversation, but not set off from the surrounding discourse to constitute a discrete speech activity of typical storytelling. Murphy observes that embedded skits are a discursive strategy, enabling

architects to make arguments about a design at hand. By momentarily stepping into the future, architects acquire "evidence" for the purpose of calibrating their architectural vision and persuading one another to modify their plans. These skits are fundamentally embodied, occurring within the context of hands-on activity, often in relation to a drawing or artifact within reach, performed with the body as a kind of cognitive work, with one person experiencing and thinking aloud in order to clarify and persuade.

Embodied arguments are also featured in the chapter by Mirivel, who notes that argumentation has traditionally belonged to the field of rhetoric, which has favored discourse-centered scholarship and largely ignored (at least until recently) the rhetoric of the visual, artifactual, and embodied. Mirivel locates and studies embodied arguments within a plastic surgery clinic. Drawing on two methodological traditions (discourse analysis and microethnography), he collects a variety of data from documents, interviews, and video recordings. And he quickly puts his finger on an institutional tension – hence an interactional dilemma – that becomes the focus of his analysis: that is, the clinic is both a medical center and a beauty shop; it is about health care and making money. On the one hand, plastic surgeons have a medical responsibility to screen out people in poor health and to educate patients about the risks associated with cosmetic surgery. On the other hand, the doctors need plenty of clients to sign up for elective surgeries so that their business will remain profitable. Mirivel shows how these competing institutional goals are played out during initial consultations between plastic surgeons and new clients. At the beginning of the consultations, surgeons carefully frame their activity as a medical encounter, and then they conduct a medical exam that makes the patient's body an object of scrutiny, visual analysis, and assessment. Through visible and tactile behaviors such as looking, pointing, touching, squeezing, giggling, and gesturing, surgeons move to constitute the patient's body as flawed, at the same time that they formulate what their examination has "discovered." Thus, plastic surgeons make arguments with their whole bodies, such that medical facts become indistinguishable from aesthetic claims, making surgery a logical next step.

In another study of aesthetic assessment, Philabaum takes us into a photography studio where novices learn to use the tools of their trade by watching, talking, and working alongside more experienced photographers. Philabaum focuses on a couple of tools that enable photographers to make decisions through practices of comparison: (1) the contact print and the negative; and (2) the viewing board. Most obviously, the photographic negative is a material artifact – that is, a permanent repository of the original photographic image, providing the raw data for an endless number of final prints. However, within the darkroom, the negative and the contact print function as tools for comparison, enabling photographers to

determine correct exposure times for making new prints. Similarly, the viewing board is more than a surface for affixing prints; it provides a neutral background and basis (white) for assessing whether the color balance of a print is correct. In order to look like a photographer, novices must learn to use the objects of their profession, which are also tools for aesthetic judgment when in the hands of an expert. Philabaum acknowledges that recent advances in digital photography are changing how photographers do their work – for example, there are no negatives. Nevertheless, he observes that new tools have emerged to support practices of comparison, which remain at the center of artistic vision for this profession.

Within organizations, embodied forms of interaction and activity can become unique and highly specialized. Heath and Luff examine a videotaped auction of fine art and antiques to reveal a complex but subtle orchestration of talk, gaze, and gesture, which is critical to the institution's ability to rapidly escalate the price and advance the sale of goods. Rather than focus only on the theatrical performance of the auctioneer, as other researchers have done, Heath and Luff also analyze the behaviors of bidders who are fully and interactively complicit in the high prices that auctioneers attain. The process unfolds incrementally: Auctioneers rhythmically chant to show that next projected values are offered and then received, and at the same time they shift their gaze and gesture around the room, creating specific opportunities for particular individuals to bid, while bidders visibly negotiate their entrance into and exit from a particular "run," all in rapid succession. Heath and Luff emphasize that this entire process is performed for the purview of the participating audience, and that auctioneers may go to some trouble to display significant changes in participation and price, which maintains an orderly and transparent pattern of interaction consistent with the institution's position of "neutrality" and the public's belief or trust in the outcome of auctions. Thus, fundamental purposes of the organization – to make a profit while maintaining trust – are jointly accomplished through embodied forms of interaction between auctioneers and bidders.

In his chapter about musical spaces, Haviland undertakes a rich anthropological exploration of the embodied interaction associated with three different musical groups: a string quartet, a jazz combo, and a trio of musicians (violin, harp, guitar) within a Mayan Indian community of southeastern Mexico. Although music may be fundamentally akin to conversation – that is, highly communicative, jointly accomplished, sequentially organized, and multimodal – studies of musical interaction provide a compelling counterpoint to talk. First, Haviland takes musical space literally by analyzing the physical places where musicians play: Video recordings show the location and orientation of bodies within built spaces that include physical objects and musical instruments. When the participants encounter problems of musical coordination, their ad hoc solutions are shaped by the spatial

arrangements and physical places they occupy. Second, he considers the coordination of action within the three contrasting musical traditions through the notion of musical "dialogues": The musicians "talk" to each other through the instruments they play, negotiating points of entry, exit, and transition; and they respond to the affect of each other's performance through such features as harmony, rhythm, and style. Finally, he observes some of the other semiotic resources that the musicians sometimes employ, including talk, paralinguistic sounds, visible behaviors, and other forms of social music. Haviland's chapter is an appropriate conclusion to this volume about multimodality in human interaction and activity because his exploration foregrounds what the musicians themselves seem to privilege. That is, he first explicates their way of coming and being together, then their coordination of instrumental interaction that has no lyrics, with considerations of talk coming in at the end rather than the beginning. Regarding music both literally and metaphorically, Haviland nudges research on embodied interaction and multimodality in one of the directions that it needs to go.

ACKNOWLEDGMENTS

REFERENCES

Agha, Asif (2007). *Language and social relations*. Cambridge: Cambridge University Press.

Atkinson, J. Maxwell (1984). *Our masters' voices*. London: Methuen.

Agar, Michael (1975). Cognition and events. In M. Sanches & B. Blount (Eds.), *Sociocultural Dimensions of Language Use* (pp. 41–56). New York: Academic Press.

Auer, Peter (2009). On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences, 31*, 1–13.

Auer, Peter, Couper-Kuhlen, Elisabeth, & Müller, Frank (1999). *Language in time. The rhythm and tempo of spoken interaction*. New York: Oxford University Press.

Bakhtin, Mikhail M. (1981). *The dialogic imagination*. Translated by Caryl Emerson and Michael Holquist. Austin: University of Texas Press.

Bakhtin, Mikhail M. (1986/1952-3). The problem of speech genres. In C. Emerson & M. Holquist (Eds.), *Speech Genres and Other Late Essays* (pp. 60–102). Austin: University of Texas Press.

Bakhtin, Mikhail M. (1990 [written 1919–1924; published 1974–1975]). *Art and answerability*. Eds. Vadim Liapunov & Michael Holquist. Trans. Vadim Liapunov. Austin, Texas: University of Texas Press.

Barsalou, Lawrence W. (1999). Perceptual symbol systems. *Behavior and Brain Sciences, 22*(4), 577–609; discussion 610–560.

Bateson, Gregory (1956). The message 'this is play'. In B. Schaffner (Ed.), *Group Processes: Transactions of the Second Conference* (pp. 145–242). New York: Josia Macy, Jr. Foundation.

Bateson, Gregory (1972). *Steps to an ecology of mind.* New York: Ballantine.

Bateson, Gregory, & Mead, Margaret (1942). *Balinese character. A photographic analysis.* New York: New York Academy of Sciences.

Beach, Wayne, & LeBaron, Curtis (2002). Body disclosures: Attending to personal problems and reported sexual abuse during a medical encounter. *Journal of Communication, 52*(3), 617–639.

Becvar, L., Holland, John, & Hutchins, Edwin (2005). Hands as molecules: Representational gestures as cognitive artifacts for developing theory in a scientific laboratory. *Semiotica, 156,* 89–112.

Bourdieu, Pierre (1977). *Outline of a theory of practice.* Cambridge: Cambridge University Press.

Bruner, Jerome S. (1969). Eye, hand, and mind. In D. Elkind & J. H. Flavell (Eds.), *Studies in Cognitive Development* (pp. 223–235). New York: Oxford University Press.

Button, Graham (Ed.) (1993). *Technology in working order: Studies of work, interaction and technology.* London: Routledge.

Clark, Andy (1997). *Being there: Putting brain, body, and the world together again.* Cambridge, MA and London: MIT Press.

Comolli, Annie (1991). *Les Gestes du Savoir* (2eme ed.). Paris: Publidix.

Comolli, Annie (2003). *Eléments de méthode en anthropologie filmique. Travaux en anthropologie filmique* (pp. 5–44). Paris: Publidix.

Couper-Kuhlen, E., & Selting, M. (1996a). Towards an interactional perspective on prosody and a prosodic perspective on interaction. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 11–56). Cambridge: Cambridge University Press.

Couper-Kuhlen, E., & Selting, M. (Eds.). (1996b). *Prosody in conversation. Interactional studies.* Cambridge: Cambridge University Press.

Csordas, Thomas I. (Ed.) (1994). *Embodiment and experience.* Cambridge: Cambridge University Press.

Damasio, Antonio R. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York: Putnam.

Damasio, Antonio R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness.* New York: Harcourt Brace.

de France, Claudine (1983). *L'analyse praxéologique. Composition, ordre et articulation d'un proces.* Paris: Editions de la Maison des Sciences de l'Homme.

de León, Lourdes (1998). The emergent participant: Interactive patterns in the socialization of tzotzil (Mayan) infants. *Journal of Linguistic Anthropology, 8,* 131–161.

de León, Lourdes (2007). Parallelism, metalinguistic play, and the interactive emergence of Zinacantec Mayan siblings' culture. *Research on Language and Scoial Interaction 40*(4), 405–36.

Deacon, Terrence Q. (1997). *The symbolic species: The co-evolution of language and the brain.* New York: W.W. Norton.

Deppermann, Arnulf, Fiehler, Reinhold, & Spranz-Fogasy, Thomas (Eds.) (2006). *Grammatik und Interaktion.* Radolfzell: Verlag für Gesprächsforschung.

Deppermann, Arnulf, & Schmitt, Reinhold (2007). Koordination. Zur Begründung eines neuen Forschungsgegenstandes. In

R. Schmitt (Ed.), *Koordination. Studien zur multimodalen Interaktion* (pp. 15–54). Tübingen: Gunter Narr Verlag.

Donald, Merlin (1991). *Origins of the modern mind.* Cambridge, MA: Harvard University Press.

Drew, Paul, & Heritage, John (Eds.) (1992). *Talk at work: Interaction in institutional settings.* Cambridge, England: Cambridge University Press.

Dreyfus, Hubert L. (1991). *Being-in-the-world. A commentary on Heidegger's "Being and Time."* Cambridge, MA: MIT Press.

Duncan, Susan (2002). Gesture, verb aspect, and the nature of iconic imagery in natural discourse. *Gesture, 2,* 183–206.

Duranti, Alessandro (1992). Language and bodies in social space: Samoan ceremonial greetings. *American Anthropologist, 94*(3), 657–691.

Duranti, Alessandro (1994). *From grammar to politics.* Berkeley: University of California Press.

Duranti, Alessandro (1997). *Linguistic anthropology.* Cambridge: Cambridge University Press.

Ekman, Paul (1973). *Darwin and facial expression; a century of research in review.* New York: Academic Press.

Enfield, N. J. (2005). The body as a cognitive artifact in kinship representations. *Current Anthropology, 46*(1), 51–81.

Enfield, N. J. (2008). *The anatomy of meaning.* Cambridge: Cambridge University Press.

Engeström, Y., & Middleton, D. (1996). *Cognition and communication at work.* Cambridge: Cambridge University Press.

Erickson, Frederick (2010). The neglected listener: Issues of theory and practice in transcription from video in interaction analysis. In J. Streeck (Ed.), *New Adventures in Language and Interaction* (pp. 243–256). Amsterdam: Benjamins.

Fox, Barbara A., Hayashi, Makoto, & Jasperson, R. (1996). Resources and repair: A cross-linguistic study. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 185–237). Cambridge: Cambridge University Press.

Ford, Cecilia E., Fox, Barbara A., & Thompson, Sandra A. (1998). Social interaction and grammar. In M. Tomasello (Ed.), *The New Psychology of Language* (Vol. 2, pp. 119–144). Mahwah, NJ: Erlbaum.

Garfinkel, Harold (1967). *Studies in ethnomethodology.* Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, Harold (Ed.) (1986). *Ethnomethodological studies of work.* London: Routledge.

Gehlen, Arnold (1988 [1958]). *Man. His nature and place in the world.* New York: Columbia University Press.

Gibbs, Raymond W. (2005). *Embodiment and cognitive science.* New York: Cambridge University Press.

Goffman, Erving (1959). *The presentation of self in everyday life.* Harmondsworth: Penguin.

Goffman, Erving (1963). *Behavior in public places: Notes on the social organization of gatherings.* New York: Free Press.

Goffman, Erving (1964). The neglected situation. *American Anthropologist, 66*(6, part 2), 133–136.

Goffman, Erving (1976). *Gender advertisements.* London: Macmillan.

Goffman, Erving (1981). *Forms of talk.* Oxford: Blackwell.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think.* Cambridge, MA: Belknap Press of Harvard University Press.

Goodwin, Charles (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 97–121). New York: Irvington.

Goodwin, Charles (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry, 50*, 277–302.

Goodwin, Charles (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Goodwin, Charles (1986). Gestures as a resource for the organization of mutual orientation. *Semiotica, 62*(1/2), 29–49.

Goodwin, Charles (1994). Professional vision. *American Anthropologist, 96*, 606–633.

Goodwin, Charles (1995a). Seeing in depth. *Social Studies of Science, 25*, 237–274.

Goodwin, Charles (1995b). Co-constructing meaning in conversations with an aphasic man. *Research on Language and Social Interaction, 28*, 233–260.

Goodwin, Charles (1996). Transparent vision. In E. Ochs, E. Schegloff, & S. Thompson (Eds.), *Interaction and Grammar* (pp. 370–404). Cambridge: Cambridge University Press.

Goodwin, Charles (1997). The blackness of black: Color categories as situated practice. In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, Tools, and Reasoning: Essays on Situated Cognition* (pp. 112–140). Berlin: Springer-Verlag.

Goodwin, Charles (2000a). Practices of color classification. *Mind Culture and Activity, 7*(1–2), 19–36.

Goodwin, Charles (2000b). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*, 1489–1522.

Goodwin, Charles (2003). *Conversation and brain damage*. Oxford and New York: Oxford University Press.

Goodwin, Charles (2007a). Environmentally coupled gestures. In S. Duncan, J. Cassell, & E. Levy (Eds.), *Gesture and the Dynamic Dimension of Language* (pp. 195–212). Amsterdam and Philadelphia: John Benjamins.

Goodwin, Charles (2007b). Interactive footing. In E. Holt & R. Clift (Eds.), *Reporting Talk: Reported Speech in Interaction* (pp. 16–46). Cambridge: Cambridge University Press.

Goodwin, Charles (2007c). Participation, stance and affect in the organization of activities. *Discourse and Society, 18*, 53–73.

Goodwin, Charles, & Goodwin, Marjorie Harness (1987). Concurrent operations on talk: Notes on the interactive organization of assessments. *IPRA Papers in Pragmatics, 1*(1), 1–54.

Goodwin, Charles, & Goodwin, Marjorie Harness (1992). Assessments and the construction of context. In C. Goodwin & A. Duranti (Eds.), *Rethinking Context* (pp. 147–190). Cambridge: Cambridge University Press.

Goodwin, Charles, & Goodwin, Marjorie Harness (1996). Formulating planes: Seeing as a situated activity. In David Middleton & Yrjö Engestrom (Eds.), *Cognition and Communication at Work* (pp. 61–95). Cambridge: Cambridge University Press.

Goodwin, Marjorie Harness (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry, 50*, 303–317.

Goodwin, Marjorie Harness (2006). *The hidden life of girls: Games of stance, status, and exclusion*. Malden, MA: Blackwell.

Goodwin, Marjorie Harness, & Goodwin, Charles (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica, 62*(1/2), 51–75.

Goodwin, Marjorie Harness, Goodwin, Charles, & Yaeger-Dror, Malcah (2002). Multi-modality in girls' game disputes. *Journal of Pragmatics, 24*(10–11), 1621–1649.

Gumperz, John J. (1982a). *Discourse strategies*. Cambridge: Cambridge University Press.

Gumperz, John J. (Ed.) (1982b). *Language and social identity*. Cambridge: Cambridge University Press.

Gumperz, John J. (1992). Contextualization and understanding. In Alessandro Duranti & Charles Goodwin (Eds.), *Rethinking Context: Language as an Interactive Phenomenon* (pp. 229–252). Cambridge: Cambridge University Press.

Günthner, Susanne, & Imo, Wolfgang (Eds.) (2006). *Konstruktionen in der Interaktion*. Berlin/New York: de Gruyter.

Haddington, Pentti, & Keisanen, Tiina (2009). Location, mobility and the body as resources in selecting a route. *Journal of Pragmatics, 41*, 1938–1961.

Hanks, William F. (1996). *Language and communicative practices*. Boulder, CO: Westview Press.

Harper, Douglas (1987). *Working knowledge. Skill and community in a small shop*. Chicago: University of Chicago Press.

Haviland, John (2000). Pointing, gesture spaces, and mental maps. In David McNeil (Ed.), *Language and Gesture* (pp. 13–46). Cambridge: Cambridge University Press.

Hayashi, Makoto (2004a). Discourse within a sentence: An exploration of postpositions in Japanese as an interactional resource. *Language in Society, 33*(3), 343–376.

Hayashi, Makoto (2004b). Projection and grammar: Notes on the 'action-projecting' use of the distal demonstrative are in Japanese. *Journal of Pragmatics, 36*, 1337–1374.

Heath, Christian (1986). *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.

Heath, Christian (2002). Demonstrative suffering: The gestural (re)embodiment of symptoms. *Journal of Communication, 52*, 597–616.

Heath, Christian (2006). Body work: The constitution of the body as a clinical object. In J. C. Heritage & D. Maynard (Eds.), *Practicing Medicine: Structure and Process in Primary Care Encounter* (pp. 185–213). Cambridge: Cambridge University Press.

Heath, Christian, & Hindmarsh, Jon (2000). Configuring action in objects: From mutual space to media space. *Mind, Culture & Activity, 7*, 81–104.

Heath, Christian, & Luff, Paul (1992). Media, space and communicative asymmetries. *Human Computer Interaction 7*, 315–46.

Heath, Christian, & Luff, Paul (1993). Disembodied conduct: Interactional asymmetries in video-mediated communication. In G. Button (Ed.), *Technology in working order* (pp. 35–54). London and New York: Routledge.

Heath, Christian, & Luff, Paul (2000). *Technology in action*. Cambridge and New York: Cambridge University Press.

Heath, Christian, & Luff, Paul (2006). Video analysis and organizational practice. In H. Knoblauch, B. Schnettler, J. Raab, & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods. Qualitative Audiovisual Analysis in Sociology* (pp. 35–50). Frankfurt/M.: Peter Lang.

Heath, Christian, & Luff, Paul (2007). Ordering competition: The interactional accomplishment of the sale of fine art and antiques at auction. *British Journal of Sociology, 58*(1), 63–85.

Heath, Christian, Luff, Paul, & Knoblauch, Hubert (2004). Tools, technologies and organizational interaction: The emergence of "workplace studies." In D. Grant, C. Hardy, C. Oswick, &

L. Putman (Eds.), *The Handbook of Organizational Discourse* (pp. 337–359). London: Sage.

Heath, Christian, Luff, Paul, vom Lehn, Dirk, Hindmarsh, Jon, & Cleverly, Jason (2002). Crafting participation: Designing ecologies, configuring experience. *Visual Communication, 1*, 9–33.

Heath, Christian, vom Lehn, Dirk, & Osborne, Jonathan (2005). Interaction and interactives: Collaboration and participation with computer-based exhibits. *Public Understanding of Science, 14*(1), 91–101.

Heidegger, Martin (1962 [1926]). *Being and time.* New York: Harper and Row.

Herder, J. (1772). *Versuch über den Ursprung der Sprache.* Stuttgart: Reclam.

Hindmarsh, Jon, Heath, Christian, & Fraser, Mike (2006). (Im)materiality, virtual reality and interaction: Grounding the "virtual" in studies of technology in action. *The Sociological Review, 54*(4), 794–816.

Hindmarsh, Jon, & Pilnick, Alison (2007). Knowing bodies at work: Embodiment and ephemeral teamwork in anaesthesia. *Organization Studies, 28*(9), 1395–1416.

Hutchins, Edwin (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Hutchins, Edwin, & Palen, Leysia (1997). Constructing meaning from space, gesture, and speech. In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, Tools, and Reasoning: Essays on Situated Cognition* (pp. 24–40). Berlin: Springer-Verlag.

Ingold, Tim (2000). *The perception of the environment: Essays on livelihood, dwelling and skill.* London and New York: Routledge.

Jackson, Michael (1989). *Paths toward a clearing. Radical empiricism and ethnographic inquiry.* Bloomington: Indiana University Press.

Jefferson, Gail (1988). On the sequential organization of troubles-talk in ordinary conversation. *Social Problems, 35*(4), 418–441.

Jones, Stanley, & LeBaron, Curtis (2002). Research on the relationship between verbal and nonverbal communication: Emerging integrations. *Journal of Communication, 52*(3), 499–521.

Keller, Charles M., & Keller, Janet D. (1996). *Cognition and tool use. The blacksmith at work.* Cambridge: Cambridge University Press.

Kendon, Adam (1967). Some functions of gaze direction in two-person conversation. *Acta Psychologica, 26*, 22–63. Reprinted in (1990) *Conducting interaction: Patterns of behavior in focused encounters* (pp. 51–89). Cambridge: Cambridge University Press.

Kendon, Adam (1972). [Review of the book Kinesics and context by Ray Birdwhistell.] *American Journal of Psychology, 85*, 441–455.

Kendon, Adam (1973). The role of visible behaviour in the organization of face-to-face interaction. In M. V. Cranach & I. Vine (Eds.), *Social Communication and Movement: Studies of Interaction and Expression in Man and Chimpanzee* (pp. 29–74). New York: Academic Press.

Kendon, Adam (1976). The F-formation system: Spatial-orientational relations in face-to-face interaction. *Man Environment Systems, 6*, 291–206.

Kendon, Adam (1977). *Studies in the behavior of social interaction.* Bloomington: Indiana University Press.

Kendon, Adam (1983). Gesture and speech: How they interact. In Wiemann & R. Harrison (Eds.), *Nonverbal interaction* (pp. 13–46). Beverly Hills, California: Sage.

Kendon, Adam (1985). Behavioural foundations for the process of frame attunement in face-to-face interaction. In *Discovery Strategies in the Psychology of Action* (pp. 229–253). London: Academic Press.

Kendon, Adam (1988). How gestures can become like words. In F. Poyatos (Ed.), *Crosscultural Perspectives in Nonverbal Communication* (pp. 131–41). Toronto: C. J. Hogrefe.

Kendon, Adam (1990). Spatial organization in social encounters: The F-formation system. In A. Kendon (Ed.), *Conducting Interaction: Patterns of Behavior in Focused Encounters* (pp. 209–238). Cambridge: Cambridge University Press.

Kendon, Adam (2004). *Gesture: Visible action as utterance.* Cambridge: Cambridge University Press.

Kendon, Adam (2009). Language's matrix. *Gesture, 9*(3), 355–372.

Kidwell, Mardi (2005). Gaze as social control: How very young children differentiate "the look" from a "mere look" by their adult caregivers. *Research on Language and Social Interaction, 38*(4), 417–449.

Kidwell, Mardi, & Zimmerman, Don (2007). Joint attention in action. *Journal of Pragmatics, 39*(3): 592–611.

Kipp, M., Martin, J.-C., Paggio, P., & Heylen, D. (Eds.) (2002). *From models of natural interaction to systems and applications.* Berlin and New York: Springer.

Knoblauch, Hubert, Schnettler, Bernt, Raab, Jürgen, & Soeffner, Hans-Georg (Eds.). (2006). *Video analysis: Methodology and methods. Qualitative audiovisual analysis in sociology.* Frankfurt/M.: Peter Lang.

Knoblich, Gunter, & Sebanz, Natalie (2006). The social nature of perception and action. *Current Directions in Psychological Science, 15*(3), 99–13.

Knorr-Cetina, K. D., & Mulkay, M. (Eds.). (1983). *Science observed.* London: Sage.

Kockelman, Peter (2007). Agency. *Current Anthropology, 48*(3), 375–387.

Koschmann, Timothy, & LeBaron, Curtis (2002). Learner articulation as interactional jointing: Studying gesture conversationally. *Cognition and Instruction, 20*(2), 249–282.

Koschmann, Timothy, LeBaron, Curtis, Goodwin, Charles, & Feltovich, Paul (2006). The mystery of the missing referent: Objects, procedures, and the problem of the instruction follower. In S. Greenberg & G. Mark (Eds.), *Computer Supported Cooperative Work* (pp. 373–382). New York: ACM.

Koschmann, Timothy, LeBaron, Curtis, Goodwin, Charles, Zemel, Alan, & Dunnington, Gary (2007). Formulating the "triangle of doom." *Gesture, 7*(1), 97–118.

Krafft, Ulrich, & Dausendschön-Gay, Ulrich (2007). Prozesse interpersoneller Koordination. In R. Schmitt (Ed.), *Koordination. Studien zur multimodalen Interaktion* (pp. 167–194). Tübingen: Gunter Narr Verlag.

Kress, Gunter, & Van Leeuwen, Theo (2001). *Multimodal discourse: The modes and media of contemporary communication.* London: Arnold.

Latour, Bruno (2005). *Reassembling the social: An introduction to actor-network-theory.* Oxford: Oxford University Press.

Latour, Bruno, & Woolgar, Steven (1986). *Laboratory life: The social construction of scientific facts.* Princeton: Princeton University Press.

LeBaron, Curtis, Glenn, Phillip, & Thompson, Michael (2009). Identity work during boundary moments: Managing positive identities through talk and embodied interaction. In J. Dutton & L. Roberts (Eds.), *Exploring Positive Identities and Organizations: Building a Theoretical and Research Foundation* (pp. 187–211). New York: Routledge.

LeBaron, Curtis, & Jones, Stanley (2002). Closing up closings: Showing the relevance of the social and material surround to the completion of an interaction. *Journal of Communication, 52*(3), 542–565.

LeBaron, Curtis, & Koschmann, Timothy (2003). Gesture and the transparency of understanding. In P. Glenn, C. LeBaron, & J. Mandelbaum (Eds.), *Studies in Language and Social Interaction* (pp. 119–132). Mahwah, NJ: Erlbaum.

LeBaron, Curtis, & Streeck, Jürgen (1997). Built space and the interactional framing of experience during a murder interrogation. *Human Studies, 20,* 1–25.

LeBaron, Curtis, & Streeck, Jürgen (2000). Gesture, knowledge, and the world. In McNeill, D. (Ed.), *Language and Gesture* (pp. 118–138). Cambridge: Cambridge University Press.

Lerner, Gene, & Zimmerman, Don (2003). Action and the appearance of action in the conduct of very young children. In P. Glenn, C. LeBaron, & J. Mandelbaum, (Eds.), *Studies in Language and Social Interaction* (pp. 441–457). Mahwah, NJ: Erlbaum.

Levine, Philip, & Scollon, Ron (Eds.) (2004). *Discourse and technology. Multimodal discourse analysis.* Washington, DC: Georgetown University Press.

Levinson, Stephen (1988). Putting linguistics on a proper footing: Explorations in Goffman's concepts of participation. In Anthony Wootton & Paul Drew (Eds.), *Erving Goffman: Exploring the Interaction Order* (pp. 161–227). Boston: Northeastern University Press.

Lindström, Anna, & Mondada, Lorenza (Eds.) (2009). Special issue on Assessments in Social Interaction, *Research on Language and Social Interaction, 42* : 4.

Lindwall, Oskar, & Lymer, Gustav (2008). The dark matter of lab work: Illuminating the negotiation of disciplined perception in mechanics. *The Journal of the Learning Sciences, 17,* 180–224.

Linell, Per (2005). *The written language bias in linguistics: Its nature, origin and transformations.* London: Routledge.

Linell, Per (2009). *Rethinking language, mind, and world dialogically: Interactional and contextual theories of human sensemaking.* Charlotte, NC: Information Age Publishing.

Lllinàs, Rodolfo R. (2001). *I of the vortex. From neurons to self.* Cambridge, MA: MIT Press.

Luff, Paul, Hindmarsh, Jon, & Heath, Christian (Eds.) (2000). *Workplace studies: Recovering work practice and informing systems design.* Cambridge: Cambridge University Press.

Lynch, Michael, Potter, Jonathan, & Garfinkel, Harold (1983). Temporal order in laboratory work. In K. D. Knorr-Cetina & M. Mulkay (Eds.), *Science Observed* (pp. 205–238). London: Sage.

Lynch, Michael, & Woolgar, Steven (Eds.) (1988). *Representation in scientific practice.* Cambridge MA: MIT Press.

MacMartin, Claire, & LeBaron, Curtis (2006). Multiple involvements within group interaction: A video-based study of sex-offender therapy. *Research on Language and Social Interaction, 39*(1), 41–80.

Mauss, Marcel (1973 [1935]). The techniques of the body. *Economy and Society, 2*(1), 70–88.

McCullough, Malcolm (1996). *Abstracting craft. The practiced digital hand.* Cambridge, MA: MIT Press.

McDermott, Ray, & David Roth (1978). Social organization of behavior: Interactional approaches. *Annual Review of Anthropology, 7,* 321–345.

McHoul, A. (2008). Editorial: Questions of context in studies of talk and interaction – ethnomethodology and conversation analysis. *Journal of Pragmatics, 40,* 823–826.

McIlvenny, Paul, Broth, Mathias, & Haddington, Pentti (2009). Editorial: Communicating place, space, and mobility. *Journal of Pragmatics, 41,* 1879–1886.

McLuhan, Marshall (1994 [1964]). *Understanding media: The extensions of man.* Boston: MIT Press.

McNeill, David (1985). So you think gestures are nonverbal? *Psychological Review, 93*(3), 350–371.

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

Mead, George Herbert (1909). Social psychology as a counterpart to physiological psychology. *Psychological Bulletin, 6,* 401–408.

Mead, George Herbert (1934). *Mind, self and society.* Chicago: University of Chicago Press.

Mead, Margaret (1975). [Review of Darwin and facial expression by Paul Ekman.] *Journal of Communication, 25,* 209–213.

Merleau-Ponty, Michel (1962). *Phenomenology of perception.* London: Routledge.

Miller, George A., Galanter, Eugene, & Pribram, Karl H. (1960). *Plans and the structure of behavior.* New York: Holt.

Mirivel, Julien (2007). Managing poor surgical candidacy: Communication problems for plastic surgeons. *Discourse & Communication, 1,* 267–293.

Modaff, Daniel (2003). Body movement in the transition from opening to task in doctor-patient interviews. In P. Glenn, C. LeBaron, & J. Mandelbaum (Eds.), *Studies in Language and Social Interaction* (pp. 411–425). Mahwah, NJ: Erlbaum.

Mondada, Lorenza (2006). Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In H. Knoblauch, B. Schnettler, J. Raab, & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods. Qualitative Audiovisual Analysis in Sociology* (pp. 51–68). Frankfurt/M.: Peter Lang.

Mondada, Lorenza (2007a). Interaktionsraum und Koordinierung. In R. Schmitt (Ed.), *Koordination. Studien zur multimodalen Interaktion* (pp. 55–93). Tübingen: Gunter Narr Verlag.

Mondada, Lorenza (2007b). Operating together through videoconference: Members' procedures for accomplishing a common space of action. In S. Hester & D. Francis (Eds.), *Orders of Ordinary Action. Respecifying Sociological Knowledge* (pp. 51–68). Guildford: Ashgate.

Mondada, Lorenza (2008a). Using video for a sequential and multimodal analysis of social interaction: videotaping institutional telephone calls. *Forum: Qualitative Social Research, 9,* 1–35.

Mondada, Lorenza (2008b). Production du savoir et interactions multimodales. Une étude de la modélisation spatiale comme activité pratique situé et incarné. *Revue d'anthropologie des connaissances,* 219–266.

Mondada, Lorenza (2009a). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics, 41,* 1977–1997.

Mondada, Lorenza (2009b). Video recording practices and the reflexive constitution of the interactional order: Some systematic uses of the split-screen technique. *Human Studies, 32*(1), 67–99.

Müller, Cornelia, & Bohle, Ulrike (2007). Das Fundament fokussierter Interation. Zur Vorbereitung und Herstellung von Interaktionsräumen durch körperliche Koordination. In Reinhold Schmitt (Ed.), *Koordination. Studien zur multimodalen Interaktion* (pp. 129–166). Tübingen: Gunter Narr Verlag.

Murphy, Keith (2005). Collaborative imagining: The interactive use of gestures, talk, and graphic representation in architectural practice. *Semiotica, 156*(1/4), 113–145.

Napier, John (1980). *Hands.* New York: Pantheon.

Nishizaka, Aug (2007). Hand touching hand: Referential practice at a Japanese midwife house. *Human Studies, 30,* 199–217.

Norris, Sigrid (2004). *Analyzing multimodal interaction: A methodological framework.* London: Routledge.

Ochs, Elinor, Gonzales, Patrick, & Jacoby, Sally (1996). When I come down I'm in the domain state: Grammar and graphic representation in the interpretive activity of physicists. In E. Ochs, E. A. Schegloff, & S. Thompson (Eds.), *Interaction and Grammar* (pp. 328–369), Cambridge: Cambridge University Press.

Ochs, Elinor, Graesch, Anthony P., Mittman, Angela, Bradbury, Thomas, & Repetti, Rena (2006). Video ethnography and ethnoarcheological tracking. In M. Pitt-Catsouphes, E. Ernst Kossek & S. Sweet (Eds.), *The Work and Family Handbook* (pp. 387–410). Mahwah, NJ: Lawrence Erlbaum.

Ochs, Elinor, Schegloff, Emanuel, & Thompson, Sandra (Eds.) (1996). *Grammar and interaction.* Cambridge: Cambridge University Press.

Ochs, Elinor, Smith, Ruth, & Taylor, Carolyn (1989). Dinner narratives as detective stories. *Cultural Dynamics, 2:* 238–57.

O'Halloran, Kay L (Ed.) (2004). *Multimodal discourse analysis: systemic-functional perspectives.* London and New York: Continuum.

Peirce, Charles Sanders (1955). *Philosophical writings of Peirce.* New York: Dover Publishers.

Plessner, Helmuth (1965 [1923]). *Die Einheit der Sinne.* Bonn: Bouvier.

Plessner, Helmuth (1975). *Die Stufen des Organischen und der Mensch.* Berlin: de Gruyter.

Plessner, Helmuth (1980). *Anthropologie der Sinne.* Frankfurt/M.: Suhrkamp.

Robinson, Jeffrey, & Stivers, Tanya (2001). Achieving activity transitions in primary-care encounters: From history taking to physical examination. *Human Communication Research, 27,* 253–298.

Ruusuvuori, Johanna, & Peräkylä, Anssi (2009). Facial and verbal expressions in assessing stories and topics. *Research on Language and Social Interaction, 42*(4), 377–394.

Sacks, Harvey (1992). *Lectures on conversation* (2 vols., G. Jefferson, Ed.). Oxford: Blackwell.

Sacks, Harvey, Schegloff, Emanuel, & Jefferson, Gail (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50,* 696–735.

Saussure, Ferdinand de (1959). *Course in general linguistics.* Edited by Charles Bally and Albert Sechehaye, in collaboration with Albert Riedlinger, translated from the French by Wade Baskin. New York: Philosophical Library.

Scheflen, A. (1976). *Human Territories.* Englewood Cliffs, NJ: Prentice-Hall.

Schegloff, Emanuel (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 266–296). Cambridge: Cambridge University Press.

Schegloff, Emanuel (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In N. Enfield (Ed.), *Roots of Human Sociality* (pp. 70–95). London: Berg Press.

Schegloff, Emanuel (2007). *Sequence organization in interaction.* Cambridge, England: Cambridge University Press.

Schmitt, Reinholt (Ed.). (2007). *Koordination. Studien zur multimodalen Interaktion.* Tübingen: Gunter Narr Verlag.

Schmitt, Reinholt, & Deppermann, Arnulf (2007). Monitoring und Koordination als Voraussetzungen der multimodalen Konstitution von Interaktionsräumen. In R. Schmitt (Ed.), *Koordination. Studien zur multimodalen Interaktion* (pp. 95–128). Tübingen: Gunter Narr Verlag.

Schütz, Alfred (1967 [1932]). *The phenomenology of the social world.* Evanston, IL: Northwestern University Press.

Selting, Margret, & Couper-Kuhlen, Elizabeth (Eds.) (2001). *Studies in Interactional Linguistics.* Amsterdam: Benjamins.

Sennett, Richard (2008). *The craftsman.* New Haven, CT: Yale University Press.

Sidnell, Jack (2005). Gesture in the pursuit and display of recognition: A Caribbean case study. *Semiotica, 156*(1/4), 55–87.

Simmel, George (1950). *The sociology of George Simmel.* Translated by Kurt Wolff. Glencoe, IL: Free Press.

Stivers, Tanya, & Sidnell, Jack (2005). Multi-Modal Interaction. *Semiotica, 156*(1/4), 1–20.

Strathern, Andrew J. (1996). *Body thoughts.* Ann Arbor: University of Michigan Press.

Streeck, Jürgen (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs, 60,* 275–299.

Streeck, Jürgen (1996a). How to do things with things. *Human Studies, 19,* 365–384.

Streeck, Jürgen (1996b). A little Ilokano grammar as it appears in interaction. *Journal of Pragmatics, 26,* 189–213.

Streeck, Jürgen (2002). A body and its gestures. *Gesture, 2*(1), 19–44.

Streeck, Jürgen (2003). The body taken for granted: Lingering dualism in research on social interaction. In P. Glenn, C. LeBaron, & J. Mandelbaum (Eds.), *Studies in Language and Social Interaction* (pp. 427–440). Mahwah, NJ: Erlbaum.

Streeck, Jürgen (2009). *Gesturecraft: The manufacturing of meaning.* Amsterdam: John Benjamins.

Streeck, Jürgen, & Hartge, Ulrike (1992). Previews: Gestures at the transition place. In P. Auer & A. di Luzio (Eds.), *The Contextualization of Language* (pp. 138–58). Amsterdam: Benjamins.

Suchman, Lucy A. (1987). *Plans and situated actions: The problem of human machine communication.* Cambridge: Cambridge University Press.

Suchman, Lucy A. (1996). Constituting shared workspaces. In Y. Engeström & D. Middleton (Eds.), *Cognition and Communication at Work* (pp. 35–60). Cambridge: Cambridge University Press.

Suchman, Lucy A., & Trigg, Randall H. (1991). Understanding practice: Video as a medium for reflection and design. In J. Greenbaum & M. Kyng (Eds.), *Design at Work: Cooperative*

Design of Computer Systems (pp. 233–240). San Francisco: Morgan Kaufmann Publishers.

Tanaka, H. (2000). Turn-projection in Japanese talk-in-interaction. *Research on Language and Social Interaction, 33*(1), 1–38.

ten Have, Paul (1999). *Doing conversation analysis: A practical guide*. London: Sage.

Tiittula, Liisa (2007). Blickorganisation in der side-by-side-Positionierung am Beispiel eines Geschäftsgesprächs. In R. Schmitt (Ed.), *Koordination. Studien zur multimodalen Interaktion* (pp. 225–262). Tübingen: Gunter Narr Verlag.

Tomasello, Michael (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Uhmann, Susanne (1992). Contextualizing relevance: On some forms and functions of speech rate changes in everyday conversation. In P. Auer & A. di Luzio (Eds.), *The Contextualization of Language* (pp. 297–336). Amsterdam: Benjamins.

Uhmann, Susanne (1996). On rhythm in everyday German conversation: Beat clashes in assessment utterances. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 303–365). Cambridge: Cambridge University Press.

Volosinov, Valentin Nikolaevic (1973). *Marxism and the philosophy of language*. Cambridge, MA, and London: Harvard University Press.

vom Lehn, Dirk (2006). Embodying experience: A video-based examination of visitors' conduct and interaction in museums. *European Journal of Marketing, 40*: 1340–1359.

vom Lehn, Dirk, Heath, Christian, & Hindmarsh, Jon (2005). Examining exhibits: Interaction in museums and galleries. *Communication & Cognition, 38*(3–4), 229–247.

Vygotsky, Lev (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Warf, Barney, & Arias, Santa (Eds.) (2009). *The spatial turn. Interdisciplinary perspectives*. London: Routledge.

Wertsch, James V. (1991). A sociocultural approach to socially shared cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 85–100). Washington, DC: American Psychological Association.

Whalen, Jack (1995). A technology of order production: Computer-aided dispatch in public safety communications. In Paul ten Have & George Psathas (Eds.), *Situated Order: Studies in the Social Organization of Talk and Embodied Action* (pp. 187–230). Washington DC: University Press of America.

Wilson, Margaret (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9*, 625–636.

Zemel, Alan, Koschmann, Timothy, LeBaron, Curtis, Goodwin, Charles, & Dunnington, Gary (2008). What are we missing? Usability's indexical ground. *Computer Supported Collaborative Work, 17*(1), 63–85.