

1 Introduction to embodied communication: why communication needs the body

Ipke Wachsmuth, Manuela Lenzen, and Günther Knoblich

1.1 The embodied communication perspective

Over the last decade, embodiment has become a key concept in language, speech, and communication research. Converging insights have been accumulated in the cognitive and neurosciences indicating that communication among social partners cannot be reduced to the transfer of abstract information. They have revealed shortcomings of “classic” communication models that emphasize symbolic information transfer. Such models neglect the decisive role of non-symbolic information transmitted by the body and especially in face-to-face communication. At the same time, researchers all around the world have started to explore the cognitive and brain mechanisms supporting interpersonal action coordination. Major discoveries are being made which have impact on, and are fostered by, research in embodied artificial intelligence, humanoid robotics, and embodied human–machine communication. While the empirical evidence is rapidly growing, an integrative view bridging the gap between low-level, sensorimotor models and their role in the “social loop” and the higher-level, functional models of communicative mechanisms is lacking.

The aim of this book is to launch and explore a new integrated and interdisciplinary perspective, the *Embodied Communication Perspective*. The embodied communication perspective creates a new framework to (re-)interpret empirical findings in the cognitive and neurosciences, and to integrate findings from different research fields that have explored similar topics without much crosstalk between them. At the same time the embodied communication perspective can serve as a guide for engineers who construct artificial agents and robots who should be able to interact with humans.

The book reflects the progress of a research year on embodied communication¹ that took place at the Center for Interdisciplinary Research of Bielefeld University (Wachsmuth and Knoblich 2005a, b).

Why is this new perspective needed? It starts from the observation that cognition arose in living organisms, is inseparable from a body, and only makes sense in a body. Likewise, natural communication and human language developed in intimate connection with body. When a person speaks, not only symbols (words, sentences, conventionalized gestures) are transmitted. One can indicate the size and shape of an object by a few hand strokes, direct attention to a referenced object by pointing or gaze, and modify what is being said with emotional facial expressions. Practical actions create affordances inviting other actors to participate in joint action, for example when trying to lift an object too heavy to be moved by one person (Richardson *et al.* 2007). The meanings transmitted in this way are multimodally encoded, strongly situated in the present context, and to a large extent expressed in bodily movements. Thus bodily communication is a topic of central interest for the biological, psychological, and social sciences because it may well be the most basic form of communication. It is likely that bodily communication preceded verbal communication in phylogenesis (Rizzolatti and Arbib 1998) and it may be the first communicative ability developing during ontogenesis (Tomassello and Camaioni 1997). In modern communication technology bodily communication has increasingly come into focus as a central aspect of intelligent behavior that artificial agents should be able to perform.

Of course, the communicative function of bodily movements has long been addressed, for instance, in connection with rhetoric and drama. However, the last decades have seen rapid developments in the study of bodily communication, partly related to improved facilities for recording and analyzing human movements (cf. Allwood 2002). Pioneering work in the modern study of bodily communication was performed in the 1930s when Gregory Bateson filmed communication on Bali (cf. Lipset 1980) and in the 1950s when Carl Herman Hjortsjö (1969) started his investigations of the anatomical muscular background of facial muscles, later to be completed by Paul Ekman and Wallace Friesen (1969, 1975). Another breakthrough was Gunnar Johansson's (1973) point light technique. Filming moving people dressed in black with

¹ <http://www.uni-bielefeld.de/ZIF/FG/2005Communication/>

white reflective spots on their main joints in front of a black background, he succeeded in isolating “pure” movement information. Further important steps using filmed data were taken by Michael Argyle (1975), Desmond Morris (1977), Adam Kendon (1981), William Condon (1986), and David McNeill (1979, 1992). Finally, in the late 1990s, another barrier was crossed when it became possible to study gestures using computer simulations in a virtual reality environment (cf. Cassell *et al.* 2000). For an overview of the whole field and its development see, for example Knapp 1978; Key 1982; Armstrong *et al.* 1995; Cassell *et al.* 2000.

In previous research, bodily communication has often been considered as being less flexible and abstract than verbal communication. However, it seems that this is not necessarily the case. If one considers the descriptive framework for communication introduced by Charles Sanders Peirce (1902/1965) it becomes immediately clear that the three basic types of signs, namely, *iconic*, *indexical*, or *symbolic* signs are all present in bodily communication. An icon is a characterizing sign that carries meaning in itself (by being related through similarity to the information that is being shared). Showing the size of a ball with both hands is one example of how iconic signs are used in bodily communication. Indexicals point to a contextual content and, of course, have their origin in manual pointing gestures. Symbols (e.g. words) that require a shared social background, a convention, and symbolic signs in bodily communication are abundant in dance, sports, and everyday conversations (e.g. thumbs up, victory sign, etc.). In human (multimodal) communication, we normally use a combination of these three types of signs.

A further important aspect of communication highlighted by the embodied communication approach is the purpose or function of communication. This is best understood in the light of competition and cooperation among members in a social group. One prevailing use of communication is social manipulation, that is to influence the behavior of conspecifics to one’s own advantage. However, communication also serves to establish social cohesion, and joint action coordination, that is to cooperate with conspecifics in achieving joint goals. A focus on the function of communication can create new links between the rapidly expanding research on social cognition and communication research.

The embodied communication approach also stresses that reception and sharing

of information is not always conscious but involves a dynamic process at diverse levels of awareness of what is being transmitted. As mentioned above, bodily movements can be used to convey symbolic information, as in “OK” gestures or by signers/viewers of deaf sign language. However, on the most basic level bodily movements also can convey meaning without the use of a conventionalized code leading to a reciprocal understanding that is based on inhabiting similar bodies and shared action repertoires (Rizzolatti and Craighero 2004). We may commonly assume a variation in the extent to which communicators are aware of what they are doing and variation regarding how intentional their actions are. Hence we propose a very broad definition of embodied communication to entail *any exchange of information among members in a social group that depends on the presence of an expressive body and its relation to objects and other expressive bodies.*

Accordingly, the core claim of the Embodied Communication perspective is that human communication involves parallel and highly interactive couplings between communication partners. These couplings range from low-level systems for performing and understanding instrumental actions, like the mirror system, to higher systems that interpret symbols in a cultural context. For instance, emotions can be communicated through instrumental actions such as smashing a dish, words can be replaced by gestures and looks, and the same action can be meaningless in one culture or an offensive communicative act in another (e.g. spitting at the floor while engaged in a conversation). The challenge for the embodied communication perspective is to identify interpersonal couplings, to identify individual cognitive mechanisms that enable such couplings, and to determine how these different mechanisms get aligned to create shared perceptions, shared references, shared beliefs, and shared intentions. We believe that our attempt to face these challenges should be interesting to a wide interdisciplinary audience ranging from cognitive neuroscientists who are interested in identifying basic mechanisms of social interaction to cognitive scientists and engineers who are interested in modeling the human mind or constructing intelligent machines.

In the following sections we describe the type of research contributions from the different fields and disciplines that set the context for the embodied communication perspective. Such an integrated perspective will, on the one hand, decisively advance our understanding of how primates (especially humans) produce, perceive, and

understand bodily gestures and how they utilize such gestures in order to coordinate their actions and exchange symbolic and non-symbolic information (Section 1.2). On the other hand, embodied communication is seen as a research metaphor to foster technology advancement in areas like anthropomorphic human–machine interfaces and artificial humanoid agents, such as virtual humans and humanoid robots. The cognitive modeling challenge is to devise theoretically grounded and empirically guided models that specify how mental processes and embodiment work together in communication (Section 1.3).

Further important input comes from brain research in general, and social neuroscience, in particular. For instance, a large number of empirical findings indicate the crucial role of the motor system during action observation, imitation, and social interaction. Computational neuroscience has started to examine the parallels between the processes involved in controlling bodily actions and understanding observed actions. Moreover, it has been proposed that communicative signals might provide a specific context for the motor commands controlling the body (e.g. forward models predicting the consequences of actions in the context of social interaction; Section 1.4).

Together, the contributions of this book reflect the embodied communication perspective in that communication should no longer be understood simply as an exchange of a series of abstract signals. Rather, it should be seen as a dynamic system of cross-modal attunement, decisively depending on embodiment, and constrained by cultural practices that structure the ways in which people interact, be it verbally or non-verbally. An outline of the chapters is given in Section 1.5.

1.2 Embodied communication in humans and other primates

Language has long been conceived of as an isolable natural object with formal properties that can be investigated independently of communicative events and their participants. Speech has often been looked at merely as “spoken language”. However, a more complete and correct picture of human communication may require researchers to include non-verbal communication and its intimate connection to speech in social interaction. A good starting point to achieve this is the embodied cognition perspective that has advanced our understanding of individual cognition by pointing out that it is spread across the mind, the body, and the various artifacts located in the environment

(Wilson 2002; Núñez 2000; Cruse 2003). The fundamental difference between embodied and cognitivist perspectives lies in the role ascribed to the body, its characteristics, and its interactions with the environment. This emerging view is well articulated in a statement by A. Clark (1999, p. 506): “Biological brains are first and foremost the control systems for biological bodies. Biological bodies move and act in rich real-world surroundings.” An important implication of this view is that communication calls systematically on physical and biological resources beyond those of natural language. Thus a new understanding of communication should explain how living beings (and primates, in particular) produce, perceive, and understand bodily gestures and how they utilize such gestures in order to understand, represent, and coordinate their actions and how they exchange symbolic and non-symbolic information.

Understanding and representing actions is closely connected with issues of communication and language (cf. Meggle 1997; Glenberg and Kaschak 2002). While traditional linguistics has tended to embrace very idealized assumptions about language, more recent approaches have brought the importance of deviations from this clean picture to the forefront. When partners in a social group cooperate, natural language is used face to face and it is situated in a non-verbal context. Research on situated communication has shed new light on the highly flexible use of language in such settings, its interaction with non-verbal means of communication, such as facial and hand gestures, and its rich grounding in visual context (Rickheit and Wachsmuth 2006). This has led to new insights on fundamental processes of communication, such as the reference to objects or their spatial relations, the coordination of speakers, the linking of dialog with ongoing actions, emotion and attitude, and the grounding of language in bodily states (Goodwin 2000; Brennan 2002, 2005; Streeck 2002; Glenberg and Kaschak 2003; Glenberg *et al.* 2005). The importance of bodily communication is illustrated by estimates that more than 65 percent of the information exchanged during a face-to-face interaction is expressed through non-verbal information in human–human communication (Argyle 1988) and that as much as 90 percent of speech in natural discourse is accompanied by gestures (Nobe 2000).

It should be mentioned that *cultural variation* is considerable for most types of body movements. This is especially well studied with regard to facial gestures, head

movements, gaze, arm and hand movements, distance, spatial orientation, as well as touch (e.g. Heeschen *et al.* 1980; Grammer *et al.* 1988). Cross-linguistic studies have led to further insights about how gestures support speech (e.g. Kita and Özyürek 2003), and attempts are being made to set up dictionaries of the communicative gestures most frequently used in everyday life (e.g. Müller and Posner 2004).

The *ontogeny of gestures and intentionality* are closely connected. Children begin to use gestures between 9 and 12 months of age. Many of these gestures originate from actions performed on objects and become intentional actions about objects (Bates *et al.* 1975). As Adamson (1995) notes, behaviors that accomplish other functions are progressively transformed into ritualized gestures. For instance, the gesture with which infants ask to be lifted up starts out with the infant grasping and trying to climb up the adult's legs. After repeated instances—and because the adult understands what the child wants—the grasping and climbing behaviors are substituted by the outstretched arms display. Communicative gestures precede first words; when gestures and speech first co-occur, they are sequential, with synchronous word and gesture combinations emerging between 16 to 18 months of age (Iverson and Thelen 2000). Later, children also use gaze to infer word meanings (Baldwin 1991), and there are a number of developmental changes in pointing gestures that go hand in hand with the development of joint attention (Moore and D'Entremont 2001).

Gesture has also been extensively studied in non-human primates (e.g. Tomasello *et al.* 1994, 1997). For instance, chimpanzees extend their arm to beg for food, clap their hands to raise others' attention, and young chimpanzees touch their mother's side to request transport to a different location. Gestures with tactile or auditory components are used independently of where the addressee is looking. In contrast, visual gestures, like “hand-beg”, are only used when the recipient is facing the actor. Some apes have learned to use pointing gestures that are not part of their natural behavioral repertoire to request food from humans (e.g. Call and Tomasello 1994; Leavens *et al.* 1996). Human-reared apes have also been observed to use pointing gestures to request things other than food (Call and Tomasello 1996). Furthermore, there seem to be some similarities between apes and human infants in the development of gestural communication (e.g. Tomasello and Camaioni 1997). It has also been argued (Stephan 1999) that non-human animals can intentionally use symbols to communicate,

at least to some extent.

What seems to differentiate humans from all other species is the large-scale use of symbolic communication. But as soon as we look at spoken verbal communication and include intonation and bodily movements, we notice that even this type of interaction is not purely symbolic. Instead, there are many iconic and indexical elements. Therefore, traditional approaches focusing on language perception and production (e.g. syntactic structures, word patterns, lexical cues, phonology) appear to be insufficient for a complete understanding of what senders intend to communicate and what listeners are capable of comprehending (Clark 1996; Allwood 2002). The same is true for the linguistic system of sign languages (Liddell 2003; Kita *et al.* 1998; Duncan 2003). Conversations are organized not by speech alone, but rather through a dynamic process of interaction. Both speakers and listeners are mutually involved through different forms of embodiment (eye gaze, gesture, posture, facial expression, etc.) in the organization of talk and action.

The distribution of meaning across speech and gesture is sometimes redundant and sometimes complementary (Kendon 1987). Careful analyses of speech and gesture reveal that language is inseparable from imagery as illustrated by speech-synchronized, coexpressive gestures (Nobe 2000; McNeill and Duncan 2000; McNeill *et al.* 2002; Duncan 2002a). Iconic gestures appear to play a vital role in organizing imagistic information about complex scenes into packages that can be verbalized within single speech-production cycles (Kita 2000). Furthermore, prosodic cues are essential for turn-taking and conceptual grounding, as demonstrated in computational models of turn-taking that enable real-time predictions in dyadic interactions (Cahn and Brennan 1999; Brennan 2000). Additional insight into the structure of a conversation comes from analyzing postural mirroring between conversants (Rotondo and Boker 2003).

Other findings have revealed forms of *rhythmic organization* for both the production and the perception of utterances. Just as the coordination of rhythmic limb movement (Schöner and Kelso 1988), speech production and gesturing requires the coordination of a huge number of disparate biological components. When a person speaks, her arms, fingers, and head move in a structured temporal organization (self-synchrony) (Condon 1986). The gesture stroke is often marked by a sudden stop that is closely coupled to speech, with temporal regularities observed between stressed syl-

lables and accompanying gesture. Moreover, hearers readily pick up the rhythm behind a speaker's utterances (interactional synchrony). The body of a listener, after a short latency following sound onset, entrains to the articulatory structure of a speaker. It has been claimed that there are interpersonal gestural rhythms (McClave 1994) and body movement may be important in interactive communication management (Davis 1982; Jaffe *et al.* 2001). Rhythm phenomena have been reported both for speech production (Fant and Kruckenberg 1996; Cummins and Port 1998) and perception (Martin 1972, 1979; Pö ppe 1997). Wachsmuth (2002) has suggested that rhythmic patterns provide an important mechanism in intraindividual and interindividual coordination of multimodal utterances and that the analysis of communicative rhythm could help to improve human-machine interfaces.

Pertaining to the association between *body language and affective states*, it has been suggested that attitudes such as openness and shyness are expressed through body movement (e.g. Argyle 1988). Darwin (1872/1965) observed long ago, across a far wider range of mammalian species than just the primates, that the facial expressions of conspecifics provide valuable cues to their likely reaction to certain courses of behavior, a rich complex summarized as “emotional state”. This work has had enormous impact and continues to do so (Ekman *et al.* 2003). Recent studies have suggested that *motion* carries far more information than the semantic content and that communication can work without involving direct cognitive processing (e.g. Grammer *et al.* 2002, 2003). In contrast, research on *body posture* is almost non-existent in non-verbal behavior analysis (see Shockley *et al.* 2003, for an exception), partially due to methodological problems (Grammer *et al.* 1997).

However, the observations about the crucial role of bodily communication will ultimately have to be put in context with representation and content. For instance, in Glenberg's (1997a, b) approach, a representation is embodied if it is constrained by how one's body can move and how it can manipulate objects. This view seems to be in accordance with the prevailing concept of embodiment in current cognitive science (Feldman 1997; Ballard *et al.* 1997), but the assumption of an *analogical structure* of cognitive representations does not follow from the fact that cognition is somehow constrained by bodily features. A distinction must be made between (1) the idea that cognitive representations are constrained by possible bodily interactions and (2) the

hypothesis that these representations are analogically related to properties of the world (Kurthen *et al.* 2003). Without assuming the existence of representations that are not directly embodied, the use of knowledge abstracted from direct experience cannot be accounted for (Habel *et al.* 1997).

In conclusion, body movements are an essential part of interactive face-to-face communication, where gestures normally are integrated with speech to form a complex whole (Streeck 2003). However, the integration of communicative body movements into a perspective that also includes speech and language requires a new understanding of the complex relations that exist between content and expression. This kind of integration is needed as a counterbalance to the traditional view that has emphasized writing over speech, speech over body, and symbolic over iconic and indexical communication (Allwood 2002).

1.3 Embodied communication in machines

A growing body of work in artificial intelligence, robotics, and agent research takes up questions that can be related to embodied communication in a technical way. From a basic research perspective, these areas can advance our understanding of key aspects of cognition, embodiment, and cognitive processes in communication. From an application perspective, they are positioned to provide well-grounded support to enable “anthropomorphic” interfaces for assistance systems in many application areas. The view that human language crucially depends on embodiment and that this would be a major challenge among many other ones for creating “Intelligent Machinery” was already envisioned by Alan Turing (1948), in stating: “Of all the above fields the learning of languages would be the most impressive, since it is the most human of these activities. This field seems however to depend rather too much on sense organs and locomotion to be feasible.”

Artificial Intelligence (AI), originally a field of the study of intelligence by computational theories of symbol use (overview see Wachsmuth 2000), has over the past decade undergone a paradigmatic shift toward the scientific study of embodied artificial agents in artificial life, humanoid robots, and virtual humans. In applied research this shift resulted in new topics of study such as perceptive or anthropomorphic human-machine interfaces and interface agents (e.g. Terada and Nishida 2002). These

efforts are complemented by the novel interface technologies for display and sensing becoming broadly available. These include force and position sensors, miniaturized cameras, touch sensitive or immersive visual displays. The first hardware platforms of humanoid robots have reached the edge of commercial availability, offering a basis for physical assistance systems in home or public environments. Interfaces are about to become less rigid and more integrated and are expected to revolutionize the human–technology interface that we know today.

The paradigmatic shift in AI also led to new research directions referred to as “Behavior-based AI”, “Situated AI”, or “Embodied AI”. In all of these new directions, agent–environment interaction, rather than disembodied and purely mental problem solving is considered to be the core of cognition and intelligent behavior (e.g. Agre and Chapman 1990; Brooks 1991a, b; Maes 1994; Agre and Rosenschein 1995; Arkin 1998; Pfeifer and Scheier 1999; Pfeifer and Bongard 2006). The aim is to build artificial agents, which interact with and adapt to new environments, previously unknown to them. Through their embodiment, such agents are continuously coupled to the current real-world situation (i.e. situated). Researchers in embodied AI and behavior-based robotics believe that *embodiment and situatedness* are also main features of natural intelligent agents and that they could be decisive in solving the problem of how symbols are grounded in sensory, non-symbolic representations (Harnad 1990).

This new AI paradigm has also led to new types of models, as in biorobotics, which uses robots to model specific behavioral phenomena observed in animals (Webb 2001). Models in the field of biorobotics generally work at a neuroethological (or in some cases neurophysiological) level of explanation. Notably, they are empirical, in that artificial neural networks are embodied in robot models that are tested under the same conditions that animals encounter in the real world, for example in the study of gait patterns in locomotion (Dean *et al.* 1999) or in sensorimotor control (Möller 1999). Another modeling approach is to construct robots that illustrate how a behavior observed in natural intelligent agents (e.g. to “learn” or to “imitate”) can be implemented. In such models, the aim is not to reproduce data that has been collected in a controlled environment, but rather to get a detailed understanding of a cognitive ability in a situated and embodied context (e.g. Pfeifer and Scheier 1997; Brooks *et al.* 1998; Ritter *et al.* 2003; Rickheit and Wachsmuth 2006). Demonstrable by robotic

appearances of expressive faces, limbs and hands, efforts include the simulation of human-like abilities, such as attention and emotional expression (e.g. Breazeal and Scassellati 1999; Kleinjohann *et al.* 2003), imitation of grasping (e.g. Steil *et al.* 2004), and the development of protolanguage (Billard 2002; Billard *et al.* 2004).

A further important issue in embodied AI is the empirical study of language evolution by way of synthetic modeling approaches with both robotic and simulated agents (Steels 1997a, 2000). As Steels and Vogt (1997) argue, robots need to be equipped with at least basic communication abilities in order to move on from agents that can solve basic spatial tasks, such as object avoidance and navigation, towards agents that could be said to exhibit “cognition”. These abilities must be developed bottom-up by the agents themselves, and the communicated concepts as well as the means of communication must be grounded in the sensorimotor experiences of the robot (Steels 1997b). This way, robots can be used to study the origins of language and meaning in self-organization and coevolution (Steels 1998a). A number of experiments were carried out with robotic and software agents to study the emergence of reference and meaning (Steels 1996a), lexicon (Steels 1996b, 1997c), and syntax (Steels 1998b).

An attempt to study communication in (predesigned, largely controlled) simulated environments is undertaken in *virtual humans research*. Researchers across a wide range of disciplines have begun to work together toward the goal of building virtual humans (Gratch *et al.* 2002)—also known as “embodied conversational agents” (Cassell *et al.* 2000) or “perceptive animated interfaces” (Cole *et al.* 2003). These are software entities that look and act like people and can engage in conversation and collaborative tasks in virtual reality. Clearly such an agent does not have a body in the physical sense (cf. Becker 2003), but it can be equipped with a synthetic voice, verbal conversational abilities, visual and touch sensors, etc., and employ its virtual body to express non-linguistic qualities such as gesture and mimicked emotions. The focus of virtual human research is on capturing the richness and dynamics of human communication behavior, and its potential applications are considerable. A variety of applications are already in progress in the domains of education and training, therapy, marketing, and virtual reality construction (e.g. Johnson *et al.* 2000; Marsella *et al.* 2000; André *et al.* 2000; Kopp *et al.* 2003).

By engaging in face-to-face conversation, conveying emotion and personality,

and otherwise interacting with the synthetic environment, virtual humans impose fairly severe behavioral requirements on the underlying animation system that must render their virtual bodies. Animation techniques must span a variety of body systems: locomotion, gestures, hand movements, body pose, faces, eyes, gaze, and speech. Research in human figure animation has addressed all of these issues (e.g. Badler *et al.* 1993; Terzopoulos and Waters 1993; Tolani *et al.* 2000). But at a more fine-grained level, it is necessary to determine the specific spatial and temporal relations among modalities, with timing emerging as a central concern. For instance, speech-related gestures must closely follow the voice cadence (Cassell *et al.* 2001; Wachsmuth and Kopp 2002). First attempts have been made to integrate these multimodal behaviors in computer-animated human models with sufficient articulation and motion generators to effect both gross and subtle movements with visual acceptability and real-time responsiveness (Kopp and Wachsmuth 2004). A related technical effort is to assemble software tools and to reach interface standards that will allow researchers to build on each other's work (Gratch *et al.* 2002).

A research challenge at the heart of the study of embodied communication is *imitation* of non-verbal behaviors such as gestures demonstrated by a human interlocutor (Kopp *et al.* 2004a). For instance, gestural movements derived from imagistic representations in working memory must be transformed into patterns of control signals executed by motor systems (Kopp *et al.* 2004b). Another research challenge is *emotion*, that is can a virtual human express emotions related to internal parameters that are driven by external and internal events. In communication-driven approaches, a facial expression is deliberately chosen on the basis of its desired impact on the user (e.g. Poggi and Pelachaud 2000). In contrast, simulation-based approaches view emotions as arising from an agent's valenced reaction to events and objects in the light of goals (e.g. Becker *et al.* 2004), where the current emotional states of the agent are communicated by consistent facial expression, intonation, and further behavioral parameters.

The realization of synthetic agents engaging in natural dialog has drawn attention to questions on how to model social aspects of conversational behavior in mixed-initiative dialog, in particular, feedback signals and turn-taking, a basic interactive mechanism for scheduling the speaker role in conversation. Whereas conversation analysis has emphasized the context-free and rule-based character of this mechanism

(Sacks *et al.* 1974), empirical studies by Duncan (1974) and successors have documented the role of interactive signals for the negotiation of the speaker role. Both aspects are reflected in theories that emphasize the interactive character of dialog (e.g. Goodwin 1981; Clark 1996). In this line, the work by Thórisson (1997, 1999, 2002) and Cassell (Cassell *et al.* 1998) has paved the way for computational models of turn-taking in human–machine communication.

In summary, the design of human–machine interactions with robotic agents and virtual humans is of great heuristic value in the study of communication because it allows researchers to isolate, implement, and test essential properties of interagent communications. Creating artificial systems that reproduce certain aspects of a natural system can help to understand the internal mechanisms that have led to the particular results. Such modeling should draw both on cognitive and brain research. It should include approaches to simulate behaviors and processes in neuroinformatics as well as artificial intelligence approaches that address a wide range of functions supporting communication, ranging from bodily action to language.

1.4 The role of basic social interaction in embodied communication

In the past, there has hardly been any crosstalk between action research and communication research. However, new findings in the domain of social cognition suggest that many primates (including humans) are equipped with basic functions for social interaction that reside in the perception action system. This raises the question of whether more sophisticated forms of verbal communication are grounded in basic sensorimotor loops for social interaction that serve to understand and predict conspecifics' behavior and support basic action coordination.

Ideomotor theories (e.g. Jeannerod 1999; Prinz 1997) claim that the specific actions of others can selectively affect one's own actions, as observed in mimicry (Chartrand and Bargh 1999), priming (Wegner and Bargh 1998), and imitation (Brass *et al.* 1999; Iacoboni *et al.* 1999; Prinz and Meltzoff 2002). According to these theories, actions are coded in terms of the perceptual events resulting from them. Observing an event that regularly resulted from one's own actions induces a tendency to carry out this action. Thus it is assumed that perceiving events produced by others' actions activate the same representational structures that govern one's own planning and control of these

actions. Further findings provide evidence that actions at the disposal of another agent are represented and have an impact on one's own actions, even when the task at hand does not require taking the actions of another person into account (Sebanz *et al.* 2006, 2003). These and other results (e.g. Barresi and Moore 1996; Shiffrar and Pinto 2002) suggest that social interactions depend on a close link between perception and action.

Ideomotor theories have gained strong empirical support from neuroscience through the finding of "mirror neurons". These neurons do not only fire when a monkey performs a particular goal-directed action but also when a monkey observe another monkey or the experimenter perform a similar action (Gallese *et al.* 1996; Rizzolatti and Arbib 1998; Gallese 2003). The idea of a direct perception action match is further supported by functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies in humans. Several areas, such as premotor cortex (Iacoboni *et al.* 1999; Rizzolatti *et al.* 2001; Rizzolatti and Craighero 2004), posterior parietal cortex (Ruby and Decety 2001), and the cerebellum (Grossman *et al.* 2000), are activated when an action is imagined or carried out as well as when the same action is observed in others (Blakemore and Decety 2001; Grèzes and Decety 2001). Further neuroimaging and magnetic stimulation studies have shown that areas associated with action are also active during imitation (Fadiga *et al.* 1995, 2002; Iacoboni *et al.* 1999; Grèzes *et al.* 2001). Premotor systems are also activated when subjects view manipulable tools (e.g. Grafton *et al.* 1997; Weisberg *et al.* 2007) or action verbs (e.g. Hauk *et al.* 2004). The finding of mirror systems suggests that we don't necessarily need conventional sign systems in order to get aligned with others. Mirroring seems to provide a mechanism that allows us to understand others' actions by matching them to our own action repertoire.

Another important mechanism for motor control that could have implications for embodied communication is the real time simulation of action (e.g. Kawato *et al.* 1987; Miall and Wolpert 1996; Jeannerod 2001; Wolpert and Flanagan 2001). It is now well established that forward models predict the sensory and perceptual consequences of one's own actions in order to compensate for the time that it takes for the reafferences to arrive in the central nervous systems. More recent is the proposal that others' actions can be predicted using the same forward models that are used to predict the consequences of own actions once the mirror system has established a match between

the observed action and one's own action repertoire (Wilson and Knoblich 2005). Such predictions could ensure that one stays aligned with the actions others will perform during joint action, particularly when precise timing is important (Knoblich and Jordan 2002, 2003). It has also been speculated that similar processes support alignment during verbal discourse (Pickering and Garrod 2007).

Learning by imitation is another essential part of human motor behavior that could be crucial for embodied communication and seems very limited in other primates, even chimpanzees (Tomasello *et al.* 2005). Although seemingly a trivial “copying” task, learning by imitation poses a series of computational challenges including: (i) how to map the perceptual variables (e.g. visual and auditory input) into corresponding motor variables; (ii) how to compensate for the difference in the physical properties and control capability of the demonstrator and imitator; and (iii) how to understand the intention of action from observation of the resulting movements (Schaal *et al.* 2003). This illustrates that, although imitation may use mirroring mechanisms, mirroring is not sufficient to explain imitation. Arbib (2005) emphasizes that the evolution of communication may have crucially hinged on an extension of the mirror system that supported the complex imitation abilities found in humans. Such an extension could also have provided a basis for the development of gestural pantomime and the gradual development of a combinatorially open repertoire of manual gestures that ultimately led to the evolution of a language-ready brain.

Wolpert, Doya, and Kawato (2003) have explored the parallels between the computations that occur in motor control and in action observation, imitation, and social interaction. In particular, they have examined the extent to which motor commands acting on the body can be equated with communicative signals acting on other people and suggest that computational solutions evolved for motor control in natural organisms may have been extended to the domain of social interaction.

According to Wolpert and colleagues (2003) social interaction involves that an actor generating motor commands causes communicative signals which, when perceived by another person, can cause changes in their internal states that in turn can lead to actions which are perceived by the actor. The authors suggest that their approach to action understanding provides an efficient mechanism for performing the computations needed in social interaction that may contribute to a theory of mind that is

based on difference modeling between one's own and others' internal states. From a philosophical perspective, it has been speculated that observed action, together with the simulation component of action memory, forms a major building block for an understanding of other minds (Proust 2000). Under a representationalist analysis, this process can be conceived of as an internal, dynamic representation of the intentionality-relation itself and, once in place, could later function as a building block for social cognition and for a more complex, consciously experienced representation of a first-person perspective (Gallese and Metzinger 2003; Metzinger and Gallese 2003).

Further insights come from clinical studies on communication disorders on patients with aphasia (i.e. the loss of power of expressing or of comprehending language, e.g. Ahlsén 1991) and apraxia (i.e. the loss of the ability to carry out purposeful movements, e.g. Rose and Douglas 2003; Goldenberg 2001; Goldenberg *et al.* 2003). Parkinson's disease also causes decrements in motor outputs—including speech and general motor systems—that also lead to a reduction in spontaneous gesturing during interactive communication (Duncan 2002b). Thus studying verbal and non-verbal communication in different patient groups may help to illuminate the architecture of the human communication device.

To conclude, including the contributions of perceptual and motor systems in the study of embodied communication is likely to help us establish the urgently needed links between research on social cognition in primates and cognitive and traditional language research. This should also allow us to better understand to what extent basic sensorimotor functions are reused and reshaped to enable a wide variety of communicative behaviors.

1.5 Outline of contents

Bringing together a selection of articles from the cognitive and neurosciences as well as the computer sciences, this book aims to develop the new perspective of embodied communication. The 18 chapters to follow focus on several aspects of embodied communication to elaborate a comprehensive understanding of the processes that give rise to the exchange of information by verbal and, in particular, non-verbal means.

The first eight chapters address basic sensorimotor, cognitive, and brain mechanisms that enable the social couplings between humans that are crucial for any

form of social interaction and discuss the evolutionary forces behind these mechanisms.

In Chapter 2 “Some boundary conditions on embodied agents sharing a common world”, *John Barresi* defines some general constraints that any embodied agent, human or machine, must meet in order to effectively work together with other agents of the same kind. He starts with the observation that such agents will have personal worlds that are characterized through relations with the environment that embody the agent’s purposes (intentional relations). “Common worlds” between agents emerge when their personal worlds overlap or interact. Barresi applies his framework to a number of findings from research on evolution and child development. He also proposes a thought experiment involving a robot community (the “Cyberiad”) to illustrate his framework. He points out that this framework should be understood as an attempt to develop a common language that captures basic principles of social life.

The “wild systems” approach *Jerome Scott Jordan* proposes in Chapter 3 “Toward a theory of embodied communication” is similarly ambitious. The fundamental assumption here is that organisms need to be understood as systems that survive through energy transformations. In this perspective cognition and communication are functions that are enabled by a dynamical control system. Each layer of this hierarchically organized system embodies aspects of the contexts organisms need to survive in, at different scales. Meaning, in this approach, is conceptualized as “embodied aboutness” and thus tightly linked to function. Communication is conceptualized as a special case of control where organisms jointly gain control over the environment. This is a provocative proposal because it treats intentionality as primary and knowledge as secondary, the reverse of what traditional cognitive science theories suggest.

In Chapter 4 “Synchrony and swing in conversation: Coordination, temporal dynamics and communication”, *Daniel Richardson, Rick Dale, and Kevin Shockley* provide an overview of their empirical research on interpersonal synchrony during conversation. This research is guided by the assumption that there is a continuum between thinking and action and that higher-level cognition is tightly linked to perception and action. One way to test this claim is to look at the temporal alignment of people’s body movements and eye movements while they converse. The authors introduce a new method (recurrence analysis) that they have used to study such

temporal alignments. The results of their studies make a very strong point for the notion of embodied communication. Hearing each other speak is sufficient to make conversation partners move in a similar rhythm and mutual understanding is improved when their eyes are temporally aligned in scanning the same objects in a scene.

Chapter 5 “The visual perception of dynamic body language”, by *Maggie Shiffrar*, addresses the perceptual processes enabling us to derive cues from movements that support basic forms of emotional and intentional understanding. The human brain is without doubt an organ of a social organism. Maggie Shiffrar shows that visual social information derived from others’ movements is indeed processed in a different way as non-social information derived from movements. She further shows that visual processing is affected by the similarity of motion representations in the observer and the observed actor. Thus the human visual system seems not to be a general-purpose processor but an inherently social organ that allows people to read the bodily expression of others with ease in their daily lives.

A look at “mirrors for embodied communication” is taken by *Wolfgang Prinz* in Chapter 6. He starts with a discussion of the manifold cultural uses for mirrors: they provide means for people to perceive themselves in new ways and in different perspectives. He then shows how the mirror metaphor can be used to describe mental functions and representations (“mirrors inside”) as well as social functions that constrain people’s actions (“mirrors outside”), and applies these metaphors to a wide range of phenomena that are of central interest to cognitive scientists and neuroscientists alike. In his view, the mirror metaphor will not only help us to understand how people mimic each other, imitate each other, and engage in joint action. It also provides a way to explain how people create a sense of self for themselves that “is tantamount to creating a homunculus” within their own body.

In Chapter 7 “The role of the mirror system in embodied communication”, *Natalie Sebanz and Günther Knoblich* discuss which aspects of embodied communication mirroring mechanisms can explain and which aspects they cannot explain. They start with an overview of the recent empirical evidence from cognitive neuroscience that leaves few doubts that while observing others we “recreate” their actions, emotions, and sensations in our own minds. Mirroring creates a basic social link that helps us to understand others, to predict what they will do next, and to create

emotional bonds with them. However, Sebanz and Knoblich also point out that it is important to recognize the limits of mirroring. More sophisticated social interactions that involve imitation, joint attention, joint action, mind reading, or verbal communication require additional cognitive mechanisms. However, it seems likely that these additional mechanisms interact and make use of the powerful mirroring machinery that is already in place in monkeys.

Like the human body, the human mind was shaped by evolutionary constraints and requirements. In Chapter 8 “Everything is movement: on the nature of embodied communication”, *Elisabeth Oberzaucher and Karl Grammer* interpret the ability of humans to analyze other people’s body language as a tool to identify honest signaling and to detect cheaters. They present empirical studies on motion quality and the expressiveness of body motions demonstrating that body language is not easily disguised. The difficulty to suppress expressive motion signals makes them enormously valuable as veridical cues to what others feel and intend and is indeed intensively analyzed by human communication partners. These observations lead the authors to a multilayered dynamic model of communication going beyond the traditional “ping-pong” theories of signaling.

Nature is a great toolbox for engineers and so is the communicative behavior of living beings. In Chapter 9 “Communication and cooperation in living beings and artificial agents”, *Achim Stephan, Manuela Lenzen, Josep Call, and Matthias Uhl* compare the communicative and cooperative behaviors of living and artificial beings. In their view, highlighting similarities and differences between these behaviors will help us to better understand the phenomenon of communication and embodiment in communication in general. They present a fine-grained typology of the very diverse and complex ways in which living beings communicate and cooperate and then apply these distinctions to artificial agents. A large amount of cooperation, as it turns out, is possible without intentional communication. Complex forms of cooperation needing communication involve a social dimension that is mostly absent in artificial beings. Finally, they discuss whether artificial beings will ever develop genuine understanding. Six further chapters discuss how thoughts, intentions, and bodily gestures are integrated during embodied communication to form a close, multilayered coupling between communication partners. To begin, Chapter 10 “Laborious intersubjectivity: Attentional

struggle and embodied communication in an auto-shop” by *Jürgen Streeck* shows how fine-grained speech and bodily signaling interact in an every-day discourse. Using the methodology of microethnography Streeck analyzes a tiny dialogue in an auto-shop. In his view, there is neither a single mechanism nor an automatic procedure responsible for achieving intersubjectivity. Rather, intersubjectivity emerges out of a heterogeneity of bodily mechanisms, practices, and resources. The communication partners use them in a flexible way that develops during their conversation. Achieving intersubjectivity works not only from “the inside out”, that is by using oneself as a model for the other but also from “the outside in”, by visually attending to one’s own gestures and how they are registered by the other.

In Chapter 11 “The emergence of embodied communication in artificial agents and humans”, *Bruno Galantucci and Luc Steels* propose a genuinely interdisciplinary approach for studying the emergence of sign systems. This is one of the relatively rare cases where research in cognitive psychology and computer science converged, although the researchers did not even know of each other’s work. Inspired by Wittgenstein’s notion of language games, Galantucci and Steels assume that the evolution of communication was tightly linked to solving practical problems in particular environments and in real time. Steels studies, in experiments involving multiple robots, how the need for coordination in such practical social interactions can attach meaning to arbitrary symbols and how it can generate abstract syntactical structures. Galantucci studies the same question in humans in a controlled laboratory setting where participants have to invent new ways of communicating because all conventional channels are cut. Both lines of research provide exciting new evidence that abstract communication can emerge from concrete, practical interactions.

Chapter 12 “Dimensions of embodied communication—Towards a typology of embodied communication”, by *Jens Allwood*, discusses how various types of content, function, and organizational features of communication are embodied. He stresses that even though new research areas are characterized by a certain fluidity of researchers’ concepts, it is important to strive for definitional clarity. Then he analyzes the concepts “embodiment”, “body”, and “communication”. Based on this analysis he develops an extensive agenda of what could and should be included in embodied communication research, concluding that there is no overwhelming risk that embodied communication

research will run out of work in the near future.

Turning to application, Chapter 13 “Neurological disorders of embodied communication”, by *Elisabeth Ahlsén*, analyzes whether findings and hypotheses on embodied communication may be useful for clinical diagnosis and the treatment of communication disorders like aphasia. After reviewing relevant theories and findings from embodied cognition research, she discusses the shortcomings of classical clinical frameworks on communication disorders. Then she shows in the light of concrete examples what it would mean to take embodiment issues into consideration when dealing with patients with communication disorders. Finally Ahlsén discusses a new model of “embodied communication disorders”.

Chapter 14 “Gestural imagery and cohesion in normal and impaired discourse”, by *Susan Duncan*, focuses on errors that are not predicted by formalist models of language production and that support the assumption that language production is an embodied cognitive process. The analyses of speech and coverbal gestures presented in this chapter draw on videotaped stories told by healthy individuals and by individuals with Parkinson’s disease. Unrehearsed storytelling performances of both speaker groups are examined and compared for evidence that coverbal gestures may function as embodied representations of meaning that help build and maintain cohesive storylines. Duncan concludes that this line of research could contribute to reconsider the modularist, amodal symbol manipulation models of human language use that have dominated psycholinguistic research for decades.

In Chapter 15 “Conversational metacognition”, *Joëlle Proust* sets out to create a link between embodied communication and psychological and philosophical theories of metacognition. To establish this link she provides a general definition of metacognition that covers not only assessing and monitoring the cognitive adequacy of one’s own information processing performance (the classical definition), but also assessing and monitoring one’s “conversational adequacy”. She then proceeds to describe a number of metacognitive gestures that can be understood as being distributed over the conversation partners and as ensuring joint control over the interactions that take place during a conversation. This allows her to define metacognitive functions in conversation and to demonstrate that the functions of conversational metacognition can neither be reduced to mirroring mechanisms nor to theory of mind mechanisms. The chapter ends

with discussing the implications of the proposal for conceptualizing cooperation and defection.

The last four chapters explicitly turn to the computational modeling of communicative behavior. In Chapter 16 “Imitation in embodied communication—from monkey mirror neurons to artificial humans”, *Stefan Kopp, Ipke Wachsmuth, James Bonaiuto, and Michael Arbib* approach the roles imitation plays in embodied communication from two different directions. The “mirror system” of the macaque brain is looked at in the first approach, assessing models of neurons, which are active both when the monkey performs a particular instrumental action, and when the monkey sees another monkey or a human executing a similar action. In the second approach, a “virtual human” is studied to make computationally explicit the ways in which enabling an artificial agent to imitate can help it attain better capabilities of communicating with humans. Both these efforts then serve to discuss the role of imitation, its underlying functions and mechanisms in communicative behavior as well as in building a general theory of embodiment, which could both advance our understanding of human communication and patterns of communication between humans and future robots.

Gesturing is an essential feature of lively communication that is often admired in humans and not often seen in artificial agents. But what exactly is the role of gestures? In Chapter 17 “Persuasion and the expressivity of gestures in humans and machines”, *Isabella Poggi and Catherine Pelachaud* analyze how gestures can make a discourse more persuasive. After an overview of the history of gesture research and studies on the expressivity of gestures from antiquity onwards they present a model of persuasive discourse in terms of goals and beliefs. They illustrate their model using case studies on the gestural behavior of famous politicians. Finally, they discuss how such a model can be used to implement persuasive gesturing in an embodied conversational agent.

Computer simulations of multimodal behavior are an increasingly popular method to test and to refine cognitive models of language production. Chapter 18 “Implementing a non-modular theory of language production in an embodied conversational agent”, by *Timo Sowa, Stefan Kopp, Susan Duncan, David McNeill, and Ipke Wachsmuth*, assesses which aspects of McNeill’s Growth Point theory of language production can be implemented in an artificial agent. So far such agents have been largely based on assumptions borrowed from modularist views of speech production.

Focusing on the model architectures of two communicative agents, the authors contrast these views with the assumptions and implications of Growth Point theory and outline how some of these could be modeled computationally. They discuss which communicative advances can be expected for conversational agents that conform to Growth Point theory and, more generally, how predictive computational models of language and gesture production can further the cognitive modeling of multimodal behavior.

Finally, Chapter 19 “Towards a neurocognitive model of turn-taking in multimodal dialogue”, by *James Bonaiuto and Kristinn Thórisson*, seeks to investigate hierarchically organized actions in communication. One essential, but often overlooked, feature of natural dialogue is turn-taking. The seemingly simple human ability to smoothly take turns while communicating becomes obvious in its complexity when one tries to teach turn-taking to artificial agents. Bonaiuto and Thórisson assume that turn-taking during conversation exists primarily for the purpose of helping participants to reduce cognitive load during conversation. They develop a hybrid cognitive model of turn-taking enhanced with a detailed, neural model of action selection. Then they present experiments demonstrating how turn taking emerges in this model. It turns out that their hybrid model, with little or no overlap in speech, is able to learn turn-taking and to process “social” turn-taking cues.

The authors and the editors hope that this volume will stimulate further discussion and that it will inspire research that further enriches the embodied communication perspective: to identify individual cognitive mechanisms that enable interpersonal couplings and to determine how these different mechanisms get aligned to create shared perceptions, shared references, shared beliefs, and shared intentions. They also hope that the detailed study of modeling issues will lead to novel ideas advancing work on anthropomorphic human–machine interfaces and artificial humanoid agents.¹ Finally, they hope that the embodied communication perspective will help to boost joint research and improved communication between the various disciplines involved.

¹ A related book is published as: *Modeling Communication with Robots and Virtual Humans* (I. Wachsmuth, G. Knoblich, eds.), Berlin, Springer, April 2008.

Acknowledgements

The editors would like to thank the Center for Interdisciplinary Research at Bielefeld University (ZiF) for hosting the research group on “Embodied Communication in Humans and Machines”, the ZiF staff for their professional support, our reviewers for valuable comments, and all fellows of the research group for taking up the manifold challenges associated with interdisciplinary research and for an exciting year of debate and cooperation.

This chapter has appeared in Lenzen M, Wachsmuth I, and Knoblich G, eds. (2008). *Embodied Communication in Humans and Machines*, pp. 1–28. Oxford: Oxford University Press.

References

- Adamson LB (1995). *Communication Development During Infancy*. Boulder CO, Westview Press.
- Agre PE and Chapman DR (1990). What are plans for? In P Maes, ed. *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*, pp. 17–34. Cambridge, MA; London, UK, MIT Press.
- Agre PE and Rosenschein SJ, eds. (1995). *Computational Theories of Interaction and Agency*. Cambridge MA; London UK, MIT Press.
- Ahlsén E (1991). Body communication and speech in a Wernicke’s aphasic—a longitudinal study. *Journal of Communication Disorders*, 24, 1–12.
- Allwood J (2002). Bodily communication dimensions of expression and content. In B Granström, D House, and I Karlsson, eds. *Multimodality in Language and Speech Systems*, pp. 7–26. Dordrecht NL, Kluwer.
- André E, Rist T, van Mulken S, Klesen M, and Baldes S (2000). The automated design of believable dialogues for animated presentation teams. In J Cassell *et al.*, eds. *Embodied Conversational Agents*, pp. 220–55. Cambridge MA, MIT Press.
- Arbib MA (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105–67.
- Argyle M (1975). *Bodily Communication*. London, Methuen.
- Argyle M (1988). *Bodily Communication*, 2nd edn. New York, Methuen & Co.
- Arkin RC (1998). *Behavior-Based Robotics*. Cambridge MA; London UK, MIT Press.
- Armstrong DF, Stokoe W, and Wilcox S (1995). *Gesture and the Nature of Language*. Cambridge MA, Cambridge University Press.
- Badler N, Phillips C, and Webber B (1993). *Simulating Humans: Computer Graphics, Animation, and Control*. New York, Oxford University Press.
- Baldwin DA (1991). Infants’ contribution to the achievement of joint reference. *Child Development*, 62, 875–90.
- Ballard DH, Hayhoe MM, Pook PK, and Rao RPN (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20, 723–67.

- Barresi J and Moore C (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, 19, 107–54.
- Bates E, Camaioni L, and Volterra, V (1975). The acquisition of performatives prior to speech. *Merril-Palmer Quarterly*, 21, 205–26.
- Becker B (2003). Marking and crossing borders: Bodies, touch and contact in cyberspace. *Body Space and Technology Journal*, 3. Available at <<http://people.brunel.ac.uk/bst/vol0302/>>, accessed 27 Feb 2008.
- Becker C, Kopp S, and Wachsmuth I (2004). Simulating the emotion dynamics of a multimodal conversational agent. In E. André *et al.*, eds. *Affective Dialogue Systems*, pp. 154–65. Berlin, Springer.
- Billard A (2002). Imitation: a means to enhance learning of a synthetic proto-language in an autonomous robot. In K Dautenhahn and CL Nehaniv, eds. *Imitation in Animals and Artifacts*, pp. 281–311. MIT Press.
- Billard A, Epars Y, Calinon S, Cheng G, and Schaal, S (2004). Discovering optimal imitation strategies. *Robotics and Autonomous Systems, Special Issue: Robot Learning from Demonstration*, 47, 65–7.
- Blakemore S-J and Decety J (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2, 561–67.
- Brass M, Bekkering H, and Prinz W (1999). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, 106, 3–22.
- Breazeal C and Scassellati B (1999). A context-dependent attention system for a social robot. *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI99)*. Stockholm, Sweden, pp. 1146–51. Denver: Morgan Kaufmann.
- Brennan SE (2000). Processes that shape conversation and their implications for computational linguistics. *Proceedings 38th Annual Meeting of the ACL. Hong Kong*, Association of Computational Linguistics.
- Brennan SE (2002). Visual co-presence, coordination signals, and partner effects in spontaneous spoken discourse. *Journal of the Japanese Cognitive Science Society*, 9, 7–25.
- Brennan SE (2005). How conversation is shaped by visual and spoken evidence. In J Trueswell and M Tanenhaus, eds. *Approaches to Studying World-Situated Language Use: bridging the language-as-product and language-action traditions*, pp. 95–129. Cambridge MA, MIT Press.
- Brooks RA (1991a). Intelligence without representation. *Artificial Intelligence*, 47, 139–60.
- Brooks RA (1991b). New approaches to robotics. *Science*, 253, 1227–32.
- Brooks R, Breazeal C, Marjanovic M, Scassellati B, and Williamson M (1998). The Cog project: Building a humanoid robot. In C Nehaniv, ed. *Computation for Metaphors, Analogy, and Agents*, pp. 52–87. New York, Springer (LNCS 1562).
- Cahn JE and Brennan SE (1999). A psychological model of grounding and repair in dialog. In SE Brennan, A Giboin, and D Traum, eds. *Proceedings, AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*, pp. 25–33. North Falmouth, MA: American Association for Artificial Intelligence.
- Call J and Tomasello M (1994). Production and comprehension of referential pointing by orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, 108, 307–17.
- Call J and Tomasello M (1996). The effect of humans on the cognitive development of apes. In AE Russon, KA Bard, and ST Parker, eds. *Reaching into Thought. The minds of the great apes*, pp. 371–403. New York, Cambridge University Press.

- Cassell J, Bickmore T, Billingham M, Campbell L, Chang K, Vilhjálmsón H, and Yan H (1998). An architecture for embodied conversational characters. In *Proceedings of the first Workshop on Embodied Conversational Characters*, October 12–15 1998, Tahoe City, California.
- Cassell J, Sullivan J, Prevost S, and Churchill E, eds. (2000). *Embodied Conversational Agents*. Cambridge, MA, MIT Press.
- Cassell J, Vilhjálmsón H, and Bickmore T (2001). BEAT: The behavior expression animation toolkit. *Proceedings of SIGGRAPH 01*, Los Angeles, CA. Association for Computing Machinery.
- Chartrand TL and Bargh JA (1999). The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893–910.
- Clark A (1999). An embodied cognitive science? *Trends in Cognitive Science*, 3, 345–51. Clark HH (1996). *Using Language*. Cambridge UK, Cambridge University Press.
- Cole R, van Vuuren S, Pellom B, Hacıoglu K, Ma J, Movellan J, Schwartz S, Wade-Stein D, Ward W, and Yan J (2003). Perceptive animated interfaces: first steps toward a new paradigm for human-computer interaction. *Proceedings of the IEEE*, 91, 1391–405.
- Condon WS (1986). Communication: Rhythm and structure. In J Evans and M Clynes, eds. *Rhythm in Psychological, Linguistic and Musical Processes*, pp. 55–77. Springfield IL, Thomas.
- Cruse H (2003). The evolution of cognition—a hypothesis. *Cognitive Science*, 27, 135–55.
- Cummins F and Port RF (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, 26, 145–71.
- Darwin C (1872/1965). *The Expression of the Emotions in Man and Animals*. Chicago, University of Chicago Press.
- Davis M (1982). *Interaction Rhythms*. New York, Human Sciences Press.
- Dean J, Kindermann T, Schmitz J, Schumm M, and Cruse H (1999). Control of walking in the stick insect: from behavior and physiology to modeling. *Autonomous Robots*, 7, 271–88.
- Duncan S, Jr (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society*, 3, 161–80.
- Duncan S (2002a). Gesture, verb aspect, and the nature of iconic imagery in natural discourse. *Gesture*, 2, 183–206.
- Duncan S (2002b). Preliminary data on effects of behavioral and levodopa therapies on speech-accompanying gesture in Parkinson’s disease. In JHL Hansen and B Pellom, eds. *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP2002)*, pp. 2481–2484, Denver, Colorado, USA, September 16–20, 2002, ISCA. Available at <http://www.isca-speech.org/archive/icslp_2002>, accessed 27 Feb 2008. (cf <http://www.isca-speech.org/archive/icslp_2002/i02_2481.html>)
- Duncan S (2003). Gesture in language: Issues for sign language research. In K Emmorey, ed. *Perspectives on Classifier Constructions in Signed Languages*. Mahwah NJ, Lawrence Erlbaum Associates.
- Ekman P, Campos JJ, Davidson RJ, and de Waal FBM (2003). *Emotions inside out—130 years after Darwin’s The Expression of the Emotions in Man and Animals* (Annals of the New York Academy of Sciences, Vol. 1000). New York, New York Academy of Sciences.
- Ekman P and Friesen W (1969). The repertoire of nonverbal behavior: categories, origins, usage and coding. *Semiotica*, I, 49–98.
- Ekman P and Friesen WV (1975). *Unmasking the Face*. Prentice-Hall.

- Fadiga L, Craighero L, Buccino G and Rizzolatti G (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, *15*, 399–402.
- Fadiga L, Fogassi L, Pavesi G and Rizzolatti G (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–11.
- Fant G and Kruckenberg A (1996). On the quantal nature of speech timing. In *Proceedings of the 4th International Conference on Spoken Language Processing ICSLP 1996*, pp. 2044–47, Philadelphia, PA, USA, Oct 3–6, 1996. Available at <http://www.isca-speech.org/archive/icslp_1996>. Accessed 27 Feb 2008. (cf. <http://www.isca-speech.org/archive/icslp_1996/i96_2044.html>)
- Feldman JA (1997). Embodiment is the foundation, not a level. *Behavioral and Brain Sciences*, *20*, 746–7.
- Gallese V (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Philosophical Transactions of the Royal Society, London B*, *358*, 517–28.
- Gallese V, Fadiga L, Fogassi L, and Rizzolatti G (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.
- Gallese V and Metzinger T (2003). Motor ontology: the representational reality of goals, actions and selves. *Philosophical Psychology*, *16*, 365–88.
- Glenberg AM (1997a). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–19.
- Glenberg AM (1997b). What memory is for: creating meaning in the service of action. *Behavioral and Brain Sciences*, *20*, 41–55.
- Glenberg AM and Kaschak MP (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9*, 558–65.
- Glenberg AM and Kaschak MP (2003). The body's contribution to language. In B Ross, ed. *The Psychology of Learning and Motivation*, V43, pp. 93–126. New York, Academic Press.
- Glenberg AM, Havas D, Becker R, and Rinck M (2005). Grounding language in bodily states: The case for emotion. In R Zwaan and D Pecher, eds. *The Grounding of Cognition: the role of perception and action in memory, language, and thinking*, pp. 115–28. Cambridge, Cambridge University Press.
- Grafton ST, Fadiga L, Arbib MA, and Rizzolatti G (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, *6*, 231–36.
- Grammer K, Fieder M, and Filova V (1997). The communication paradox and possible solutions. In A Schmitt, K Atzwanger, K Grammer, and K Schäfer, eds. *New Aspects of Human Ethology*, pp. 91–120. New York, Plenum Press.
- Grammer K, Fink B, and Renninger, LA (2002). Dynamic systems and inferential information processing in human communication. *Neuroendocrinology Letters*, *23*, 15–22.
- Grammer K, Keki V, Striebel B, Atzmueller M, and Fink B (2003). Bodies in motion: a window to the soul? In E Voland and K Grammer, eds. *Evolutionary Aesthetics*. Heidelberg, Springer.
- Grammer K, Schiefenhövel W, Schleidt M, Lorenz B, and Eibl-Eibesfeldt I (1988). Patterns on the face: The eyebrow flash in crosscultural comparison. *Ethology*, *77*, 279–99.
- Gratch J, Rickel J, André E, Cassell J, Petajan E, and Badler N (2002). Creating interactive virtual humans: some assembly required. *IEEE Intelligent Systems*, *17*, 54–63.
- Grèzes J and Decety J (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, *12*, 1–19.
- Grèzes J, Fonlupt P, Bertenthal B, Delon-Martin C, Segebarth C and Decety J (2001). Does perception of biological motion rely on specific brain regions? *NeuroImage*, *13*, 775–85.

- Goldenberg G (2001). Imitation and matching of hand and finger postures. *NeuroImage*, 14, 132–6.
- Goldenberg G, Hartmann K, Schlott I (2003). Defective pantomime of object use in left brain damage: apraxia or asymbolia? *Neuropsychologia*, 41, 1565–73.
- Goodwin C (1981). *Conversational Organization: Interaction between speakers and hearers*. New York NY, Academic Press.
- Goodwin C (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32, 1489–522.
- Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, and Blake R (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, 12, 711–20.
- Habel C, Kelter S, and Kaup B (1997). Embodied representations are part of a grouping of representations. (Commentary on Glenberg’s Article “What memory is for”). *Behavioral and Brain Sciences*, 20, 26.
- Harnad S (1990). The symbol grounding problem. *Physica D*, 42, 335–46.
- Hauk O, Johnsrude I, and Pulvermüller F (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41, 301–7.
- Heeschen V, Schiefenhövel W, and Eibl-Eibesfeldt I (1980). Requesting, giving and taking. The relationship between verbal and nonverbal behavior in the speech community of the Eipo, Irian Jaya (WestNew Guinea). In MR Key, ed. *The Relationship of Verbal and Nonverbal Communication—Contributions to the Sociology of Language*, pp. 139–66. Den Haag, Mouton.
- Hjortsjö CH (1969). Människans Ansikte och Mimiska Språket. Studentlitteratur, Malmö. (quoted after Allwood 2002).
- Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, and Rizzolatti G (1999). Cortical mechanisms of human imitation. *Science*, 286, 2526–8.
- Iverson JM and Thelen E (2000). Hand, mouth, and brain: The dynamic emergence of speech and gesture. In R Núñez and WJ Freeman, eds. *Reclaiming Cognition: The Primacy of Action, Intention, and Emotion*. Thorverton UK, Imprint Academic.
- Jaffe J, Beebe B, Feldstein S, Crown CL, and Jasnow MD (2001). Rhythms of dialogue in infancy. *Monographs of the Society for Research in Child Development*, 66, No. 2. Boston, Blackwell.
- Jeannerod M (1999). The 25th Bartlett Lecture: to act or not to act: perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology, Human Experimental Psychology*, 52A, 1–29.
- Jeannerod M (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, 14, 103–9.
- Johansson G (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201–11.
- Johnson WL, Rickel JW, and Lester JC (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11, 47–78.
- Kawato M, Furawaka K, and Suzuki R (1987). A hierarchical neural network model for the control and learning of voluntary movements. *Biological Cybernetics*, 56, 1–17.
- Kendon A (1981). *Nonverbal Communication, Interaction, and Gesture*. The Hague, Mouton Publishers.

- Kendon A (1987). On gesture: its complementary relationship with speech. In A Siegman and S Feldstein, eds. *Nonverbal Behavior and Communication*, pp. 65–97. Hillsdale, Lawrence Erlbaum.
- Key MR (1982). *Nonverbal Communication Today*. Berlin, Mouton.
- Kita S (2000). How representational gestures help speaking. In D McNeill, ed. *Language and Gesture*, pp. 162–85. Cambridge University Press.
- Kita S and Özyürek A (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16–32.
- Kita S, van Gijn I, and van der Hulst H (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I Wachsmuth and M Fröhlich, eds. *Gesture and Sign Language in Human-Computer Interaction*, pp. 23–35. Berlin: Springer (LNCS 1371).
- Kleinjohann B, Kleinjohann L, Stichling D, and Esau N (2003). MEXI—Machine with Emotionally eXtended Intelligence. *Proceedings of the 4th International Scientific and Technical Conference on Intellectual and Multiprocessor Systems (IMS 2003)*. Gelendzhik, Russia, Sept. 2003.
- Knapp M (1978). *Nonverbal Communication in Human Interaction*. New York, Holt, Rinehart and Winston.
- Knoblich G and Jordan S (2002). The mirror system and joint action. In M Stamenov and V Gallese, eds. *Mirror Neurons and the Evolution of Brain and Language*, pp. 115–24. Amsterdam, John Benjamins.
- Knoblich G and Jordan JS (2003). Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1006–16.
- Kopp S, Jung B, Leßmann N, and Wachsmuth I (2003). Max—a multimodal assistant in virtual reality construction. In *KI—Künstliche Intelligenz*, 4, 11–17. Available at <<http://www.kuenstliche-intelligenz.de/index.php?id=%3ANO-5813>>. Accessed 27 Feb 2008.
- Kopp S, Sowa T, and Wachsmuth I (2004a). Imitation games with an artificial agent: from mimicking to understanding shape-related iconic gestures. In A Camurri and G Volpe, eds. *Gesture-based Communication in Human-Computer Interaction*, pp. 436–47. Berlin, Springer (LNCS 2915).
- Kopp S, Tepper P and Cassell J (2004b). Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proceedings of the International Conference on Multimodal Interfaces (ICMI'04)*, pp. 97–104. New York, ACM Press.
- Kopp S and Wachsmuth I (2004). Synthesizing multimodal utterances for conversational agents. *Journal of Computer Animation and Virtual Worlds*, 15, 39–52.
- Kurthen M, Grunwald T, Helmstaedter C, and Elger CE (2003). The problem of content in embodied memory. *Behavioral and Brain Sciences*, 26, 641–50.
- Leavens DA, Hopkins, WD, and Bard KA (1996). Indexical and referential pointing in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 110, 346–53.
- Liddell SK (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge MA, Cambridge University Press.
- Lipset D (1980). Gregory Bateson—the legacy of a scientist. Englewood Cliffs NJ, Prentice Hall. Maes P, ed. (1994). *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*. Cambridge MA; London UK, MIT Press.

- Marsella SC, Johnson WL, and LaBore C (2000). Interactive Pedagogical Drama. In C Sierra, M Gini, and JS Rosenschein, eds. *Proceedings of the Fourth International Conference on Autonomous Agents*, pp. 301–8. ACM Press.
- Martin JG (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79, 487–509.
- Martin JG (1979). Rhythmic and segmental perception. *Journal of the Acoustical Society of America*, 65, 1286–97.
- McClave E (1994). Gestural beats: The rhythm hypothesis. *Journal of Psycholinguistic Research*, 23, 45–66.
- McNeill D (1979). *The Conceptual Basis of Language*. Hillsdale, Lawrence Erlbaum.
- McNeill D (1992). *Hand and Mind: what gestures reveal about thought*. Chicago, University of Chicago Press.
- McNeill D, Quek F, McCullough K-E, Duncan S, Bryll R, Ma X-F, and Ansari R (2002). Dynamic imagery in speech and gesture. In B Granström, D House, and I Karlsson, eds. *Multimodality in Language and Speech Systems*, pp. 27–44. Dordrecht, Kluwer.
- McNeill D and Duncan S (2000). Growth points in thinking-for-speaking. In D McNeill, ed. *Language and Gesture*, pp. 141–61. Cambridge UK, Cambridge University Press.
- Meggle G (1997). Communicative actions. In G Holmström-Hintikka and R Tuomela, eds. *Contemporary Action Theory*, Vol. 2, pp. 251–72. Dordrecht, Kluwer.
- Metzinger T and Gallese V (2003). The emergence of a shared action ontology: building blocks for a theory. *Consciousness and Cognition*, 12, 549–71.
- Miall RC and Wolpert DM (1996). Forward models for physiological motor control. *Neural Networks*, 9, 1265–79.
- Möller R (1999). Perception through anticipation—A behavior-based approach to visual perception. In A Riegler, A von Stein, and M Peschl, eds. *Understanding Representation in the Cognitive Sciences*, pp. 169–75. New York, Plenum Press.
- Moore C and D’Entremont B (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development*, 2, 109–29.
- Müller C and Posner R, eds. (2004). *The Semantics and Pragmatics of Everyday Gestures*. The Berlin conference. Berlin, Weidler Buchverlag.
- Morris D (1977). *Manwatching*. Oxford, Elsevier.
- Nobe S (2000). Where do most spontaneous representational gestures actually occur with respect to speech? In D McNeill, ed. *Language and Gesture*, pp. 186–98. Cambridge UK, Cambridge University Press.
- Núñez R (2000). Could the future taste purple? Reclaiming mind, body and cognition. In R Núñez and WJ Freeman, eds. *Reclaiming Cognition: The Primacy of Action, Intention, and Emotion*. Thorverton UK: Imprint Academic.
- Peirce CS (1902/1965). *Collected Papers of Charles Sanders Peirce*. Cambridge MA, The Belknap Press of Harvard University Press.
- Pfeifer R and Bongard J (2006). *How the body shapes the way we think: a new view of intelligence*. Cambridge MA, MIT Press.
- Pfeifer R and Scheier C (1997). Sensory-motor coordination: The metaphor and beyond. *Robotics and Autonomous Systems*, 20, 157–78.
- Pfeifer R and Scheier C (1999). *Understanding Intelligence*. Cambridge MA; London UK, MIT Press.
- Pickering MJ and Garrod S (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11, 105–10.

- Poggi I and Pelachaud C (2000). Performative facial expressions in animated faces. In J Cassell, J Sullivan, S Prevost, and E Churchill, eds. *Embodied Conversational Agents*, pp. 155–88. Cambridge MA, MIT Press.
- Pöppel E (1997). A hierarchical model of temporal perception. *Trends in Cognitive Science*, 1, 56–61.
- Prinz W (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–54.
- Prinz W and Meltzoff AN (2002). An introduction to the imitative mind and brain. In W Prinz and AN Meltzoff, eds. *The Imitative Mind: Development, Evolution, and Brain Bases*, pp. 1–15. Cambridge MA, Cambridge University Press.
- Proust J (2000). Awareness of agency: three levels of analysis. In T Metzinger, ed. *The Neural Correlates of Consciousness*. Cambridge, MIT Press.
- Richardson MJ, Marsh KL, and Baron RM (2007). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 845–59.
- Rickheit G and Wachsmuth I, eds. (2006). *Situated Communication*. Berlin, Mouton de Gruyter.
- Ritter R, Steil J, Nölker C, Röthling F, and McGuire P (2003). Neural architectures for robotic intelligence. *Reviews in the Neurosciences*, 14, 121–43.
- Rizzolatti G and Arbib MA (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–94.
- Rizzolatti G and Craighero L (2004). The mirror neuron system. *Annual Review of Neuroscience*, 27, 169–92.
- Rizzolatti G, Luppino G, and Matelli M (1998). The organization of the cortical motor system: new concepts. *Electroencephalography and Clinical Neurophysiology*, 106, 283–96.
- Rizzolatti G, Fogassi L, and Gallese V (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2, 661–70.
- Rose M and Douglas J (2003). Limb apraxia, pantomime, and lexical gesture in aphasic speakers: Preliminary findings. *Aphasiology*, 17, 453–64.
- Rotondo JL and Boker SM (2003). Behavioral synchronization in human conversational interaction. In M Stamenov and V Gallese, eds. *Mirror Neurons and the Evolution of Brain and Language*, pp. 151–62. John Benjamins.
- Ruby P and Decety J (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature Neuroscience*, 4, 546–50.
- Sacks H, Schegloff EA, and Jefferson G (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Schaal S, Ijspeert AJ, and Billard A (2003). Computational approaches to motor learning by imitation. *Philosophical Transactions of the Royal Society: Biological Sciences*, 358:1431, 537–47.
- Schöner G and Kelso JAS (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, 239, 1513–20.
- Sebanz N, Bekkering H, and Knoblich G (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70–6.
- Sebanz N, Knoblich G, and Prinz W (2003). Representing others' actions: Just like one's own? *Cognition*, 88, B11–B21.
- Shiffrar M and Pinto J (2002). The visual analysis of bodily motion. In W Prinz and B Hommel, eds. *Common Mechanisms in Perception and Action: attention and performance, Vol. XIX*, pp. 381–99. Oxford, Oxford University Press.

- Shockley K, Santana MV, and Fowler CA (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 326–32.
- Steels L (1996a). Perceptually grounded meaning creation. In M Tokoro, ed. *Proceedings of the International Conference on Multi-Agent Systems*, pp. 338–44. Cambridge MA: MIT Press.
- Steels L (1996b). Emergent adaptive lexicons. In P Maes, MJ Mataric, J-A Meyer, J Pollack, and SW Wilson, eds. *From Animals to Animats 4, Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*. Cambridge MA: MIT Press.
- Steels L (1997a). The synthetic modeling of language origins. *Evolution of Communication*, 1, 1–35.
- Steels L (1997b). Constructing and sharing perceptual distinctions. In M van Someren and G Widmer, eds. *Proceedings of the European Conference on Machine Learning, ECML'97*, pp. 4–13. Berlin, Springer.
- Steels L (1997c). Self-organizing vocabularies. In C Langton and T Shimohara, eds. *Artificial Life V: Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems*, pp 179–184. Cambridge MA: MIT Press.
- Steels L (1998a). Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. In J Hurford, C Knight, and M Studdert-Kennedy, eds. *Evolution of Human Language: Social and Cognitive Bases*, pp. 384–404. Edinburgh: Edinburgh University Press.
- Steels L (1998b). The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103, 1–24.
- Steels L (2000). The puzzle of language evolution. *Kognitionswissenschaft*, 8, 143–50.
- Steels L and Vogt P (1997). Grounding adaptive language games in robotic agents. In C Husbands and I Harvey, eds. *Proceedings of the Fourth European Conference on Artificial Life (ECAL' 97)*. London, MIT Press.
- Steil J, Röthling F, Haschke R, and Ritter H (2004). Situated robot learning for multi-modal instruction and imitation of grasping. *Robotics and Autonomous Systems* (Special Issue on Imitation Learning), 47, 129–41.
- Stephan A (1999). Are animals capable of concepts? *Erkenntnis* (Special Issue on Animal Mind), 51, 79–92.
- Streeck J (2002). A body and its gestures. *Gesture*, 2, 19–44.
- Streeck J (2003). The body taken for granted: Lingering dualism in research on social interaction. In P Glenn, CD LeBaron and J Mandelbaum, eds. *Studies in Language and Social Interaction. In Honor of Robert Hopper*, pp. 427–40. Lawrence Erlbaum.
- Terada K and Nishida T (2002). Active artifacts: for new embodiment relation between human and artifacts. In *Proceedings of the 7th International Conference on Intelligent Autonomous Systems (IAS-7)*, Marina del Rey, California.
- Terzopoulos D and Waters K (1993). Analysis and synthesis of facial images using physical and anatomical models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, 569–79.
- Thórisson KR (1997). Gandalf: An embodied humanoid capable of real-time multi-modal dialog with people. In *Proceedings First International Conference On Autonomous Agents*, pp. 536–7.
- Thórisson KR (1999). A mind model for multimodal communicative creatures and humanoids. *International Journal of Applied Artificial Intelligence*, 13, 449–86.

- Thórisson KR (2002). Natural turn-taking needs no manual: computational theory and model, from perception to action. In B Granström, D House and I Karlsson, eds. *Multimodality in Language and Speech Systems*, pp. 173–207. Dordrecht, Kluwer.
- Tolani D, Goswami A, and Badler N (2000). Real-time inverse kinematics techniques for anthropomorphic limbs. *Graphical Models*, 62, 353–88.
- Tomasello M, Call J, Nagell K, Olguin R, and Carpenter M (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 35, 137–54.
- Tomasello M, Call J, Warren J, Frost, GT, Carpenter M, and Nagell K (1997). The ontogeny of chimpanzee gestural signals: A comparison across groups and generations. *Evolution of Communication*, 1, 223–59.
- Tomasello M and Camaioni L (1997). A comparison of the gestural communication of apes and human infants. *Human Development*, 40, 7–24.
- Tomasello M, Carpenter M, Call J, Behne T, and Moll H (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–91.
- Turing AM (1948). Intelligent machinery. Published in B Meltzer and D Michie, eds. (1969): *Machine Intelligence*, Vol. 5, pp. 3–23. Edinburgh, Edinburgh University Press.
- Wachsmuth I (2000). The concept of intelligence in AI. In H Cruse, J Dean, and H Ritter, eds. *Prerational Intelligence—Adaptive Behavior and Intelligent Systems without Symbols and Logic*, Vol. 1, pp. 43–55. Dordrecht NL, Kluwer Academic Publishers.
- Wachsmuth I (2002). Communicative rhythm in gesture and speech. In P Mc Kevitt, S O’Nuáallain, and C Mulvihill, eds. *Language, Vision and Music*, pp. 117–32. Amsterdam, Benjamins.
- Wachsmuth I and Kopp S (2002). Lifelike gesture synthesis and timing for conversational agents. In I Wachsmuth and T Sowa, eds. *Gesture and Sign Language in Human-Computer Interaction*, pp. 120–33. Berlin, Springer (LNCS 2298).
- Wachsmuth I and Knoblich G (2005a). Embodied communication in humans and machines. *AI Magazine*, 26, 85–6.
- Wachsmuth I and Knoblich G (2005b). Embodied communication in humans and machines—a research agenda. *Artificial Intelligence Review*, 24, 517–22.
- Webb B (2001). Can robots make good models of biological behaviour? *Behavioral and Brain Sciences*, 24, 1033–50.
- Wegner DM and Bargh JA (1998). Control and automaticity in social life. In DT Gilbert, ST Fiske and G Lindzey, eds. *The Handbook of Social Psychology*, pp. 446–96. Boston MA, McGraw-Hill.
- Weisberg J, Turenout M, and Martin A (2007). A neural system for learning about object function. *Cerebral Cortex*, 17, 513–21.
- Wilson M (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9, 625–36.
- Wilson M and Knoblich G (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–73.
- Wolpert DM and Flanagan JR (2001). Motor prediction. *Current Biology*, 11, R729–732.
- Wolpert DM, Doya K and Kawato M (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society: Biological Sciences*, 358, 593–602.