

Embodied Communication II: An Integrated Perspective

Final Conference

ZiF: Research Group *Embodied Communication in Humans and Machines*

6. – 9. September 2006

Leitung: Ipke Wachsmuth (Bielefeld) und Günther Knoblich (Newark, NJ)

Communication is so much more than just words. It involves turn taking and mirroring, gestures and body postures, rhythm and emotions. But only recently researchers have started to focus on the multidimensional and closely coupled nature of communication. This focus has a potential to lead to a paradigm shift in which the traditional engineering model of signal transmission is replaced, **Ipke Wachsmuth** (Bielefeld) explained in his introduction to the final conference of the ZiF research year *Embodied Communication in Humans and Machines*. The research group launched and explored a new integrated and interdisciplinary perspective, the Embodied Communication Perspective, he summed up. This new framework is a first step towards a multi-modal, multi-level, dynamic, and co-constructive model of communication that will shed new light on recent empirical findings obtained in the cognitive and neurosciences.

A suitable amount of work in the research group was devoted to “learning by doing”, i.e. to create artificial systems to better understand natural ones. Wachsmuth explained the point of this approach by a little story: When he one day realized that his artificial agent Max did not look at him as usual, he found that the emotion component of Max’ architecture was not working. This showed that not only gaze but also emotional feedback is important to look at people in a natural way. Max is also the hero of a new interdisciplinary cooperation between scientists at the University of Bielefeld and scientists at the University of Chicago that aims at implementing the Growth Point Theory of gesture production. One problem in realising gestures in an artificial agent is that there is a lot of variability in the translation from image to gesture. Accordingly, the “GP to the Max”-group, including **Timo Sowa, Susan Duncan, Stefan Kopp, David McNeill, and Ipke Wachsmuth** (Bielefeld and Chigaco, IL), pointed out that gestural image and spoken form are two simultaneously manifesting aspects of the basic interval of language, two faces of one unit. Imagery and spoken form co-express this unit; both unpack from the unitary growth points.

An Embodied Conversational Agent (ECA) should be an agreeable partner. It should show attention and interest, perceive to whom it is talking, and it should be non-obtrusive.

Catherine Pelachaud (Montreuil) presented the gaze system of her ECA, Greta, that is able to detect gaze directions in order to find out whether she is addressed directly. **Justine Cassell** (Evanston, IL) pointed out some problem that people encounter when talking to virtual agents. When people were confronted with either an artificial agent or a map that directed them to a certain location they judged the map to be more efficient and trustworthy than the virtual agent. Different reasons might cause this somewhat disappointing result: the virtual agent might be not useful for providing route information, or language and agent might not fit together well enough. **Christopher Habel** (Hamburg) on the other hand found at least one way in which the body can be useful for route descriptions: Verbal descriptions serve well when giving instructions to act and to communicate the instructor’s experience. However, drawings are more useful than language when solving conflicts about the spatial perspective and alternative reference systems.

In an evening lecture open to the wider public **Karen Emmorey** (San Diego, CA) showed that sign language does not simply consist in speaking with the hands. Sign language use has linguistic, neural, and cognitive consequences. Users of sign language are more willing to map concepts onto the body, their brains are set up to distinguish between sign and

pantomime, and signing can change the perception of action and the perception of the human body.

Sotaro Kita (Birmingham) analyzed cultural variations in pointing. He found that such variations do not occur in a vacuum. Rather, they occur because thought, language and pragmatics vary cross-culturally. For instance, in languages in which the word “swing” does not exist, straight gestures are performed when people are asked to describe a swinging movement. **Matthias Uhl** (Siegen) focused on trans-cultural similarities. Analysing American and Indian blockbuster films he claimed that successful stories reflect universal aspects of human evolution such as the intense struggle for social and economic resources. The plot of successful films is built around issues and is especially interesting for humans because of their relevance in our evolutionary past, he said.

Isabella Poggi (Rome) focused on emotions, and the communication of emotions. She pointed out that normally only one of the communication partners is experiencing a particular emotion. This experience can then be induced in the other communication partner during interaction in different ways that range from explicit communication to unconscious contagion or empathy.

From the perspective of embodied communication one of the most important recent findings is that humans and monkeys understand each other's action using a mirror system. **James Bonaiuto** (Chicago, IL) attempts to model this system in a way that is consistent with what is known about the principles of motor learning. Lesion experiments with cats have shown that cats modify their motor program due to changes in their embodiment. In his attempt to model these processes Bonaiuto starts from the assumption that motor programs are not explicitly represented but emerge through the model's evaluation of priority and via a mirror system that matches observation and execution. Building on Bonaiuto's talk **Michael Arbib** (Chicago, IL) elaborated further on the mirror system hypothesis: The evolutionary basis for language is provided by the evolution of Broca's area atop the mirror system for grasping. Speech is rooted in an earlier system for manual gesture. Thus language is multi-modal by definition. “I don't repeat what you say”, Michael Arbib summed up the role of the mirror system, “but what you say partly shapes what I say.” **Stefan Kopp** (Bielefeld) is trying to make use of the findings on the mirror system for equipping the virtual agent Max with further abilities. Max shall acquire a sense of its own body and be able to employ his motor system for understanding, predicting, adapting to the motor behaviour of others. Imitation via a mirror system comparable to that found in humans provides a form of coupling of perception and action and also between the artificial agent and the demonstrator, he explained. **Cristina Massen** and **Wolfgang Prinz** (Leipzig) analysed the function of such covert imitation, i.e. the activation of motor programs corresponding to a seen motion on the side of the observer. They discussed several of these functions such as the facilitation of overt imitation, action understanding, intention understanding, and predicting the actions of others. Their experiments show that observing others' tool-use actions mainly activates the target-movement-mapping which suggests that action prediction is the main function of covert imitation.

Communication does not depend on perfect understanding. Often, some shared understanding of the content is enough. **Martin Loetzsch** and **Luc Steels** (Paris and Brussels) develop a computational model of embodied agents that perform communicative tasks and are able to align their conceptual and linguistic inventories. In his talk Martin Loetzsch presented recordings of Aibo robots who were able not only to detect balls in their environment but to align their perspective with other aibos, to know that they were jointly looking at the same object, and to communicate about the objects they both were attending at. **Martin Pickering** (Edinburgh) investigates how human interlocutors align their mental models during communication. He found that this effect is largely due to priming, which he described as a

form of unconscious implicit learning. Alignment at one level, he also explained, also enhances the alignment on other levels. The discussion revealed that it is not yet known whether such priming also exists for gestures. **Kristinn Thórisson** (Reykjavik) presented his research on turn taking which can be seen as a further tool for alignment. It ensures that understanding can progress incrementally, he said. It is realized by various information-carrying systems like face, intonation, and tone of voice and should be considered as an emergent phenomenon arising from a complex system.

Jürgen Streeck (Austin, TX) reported on the “multimodal man”. He recorded one day in the life of a car mechanic and determined sequence by sequence how the mechanic used different modalities and practices while interacting with employees and customers. He showed striking examples of how tightly gesture and gaze can become coupled in real-world social interaction.

Gregor Schöner (Bochum) described interaction in terms of the dynamic field theory. This theory focuses on the dynamic and highly interactive nature of the processes in the brain. Given this perspective the question arises how discrete cognitive events arise from more continuous cognitive events. One way to solve this problem, Schöner explained, is preshaping of dynamic fields that could be considered as being equivalent with memory traces. **Scott Jordan** (Normal, IL) proposed a different brand of dynamical systems approach that could potentially replace the discrete information processing accounts of communication. In his view, organisms are self-sustaining systems which embody the constraints in which they had to survive, phylogenetically and ontogenetically. In this sense organisms are necessarily “about” their environment.

Is there any clinical application of the embodied communication perspective? **Elisabeth Ahlsén** (Göteborg) and **Georg Goldenberg** (München) focused on apraxia and aphasia. Topics such as alignment, mirroring, imitation, evolution, and automatic and controlled processing in communication are rarely addressed in clinical work, Ahlsén complained. Taking them into account could lead to new insights in research on aphasia. Goldenberg stressed that imitation and pantomime are not the same thing. For the imitation of meaningless gestures the similarity between the model's and one's own actions can be derived neither from their common meaning nor from their similar impact upon external objects. It must result from direct matching between the own body's configuration and that of the model. Gesture imitation on the other hand evolves in parallel with the development of language skills. **Hans Markowitsch** (Bielefeld) emphasized that a sharp division between psychiatric and psychological or functional and organic is not convincing according to the recent finding in memory research. Memory is a complex bodily process that can be disturbed at different stages.

Conceptual issues were addressed by **John Barresi** (Halifax) and **Jens Allwood** (Göteborg). Focussing on the conceptual foundation of embodied communication research Allwood discussed definitions of embodiment, communication, and understanding with a focus on what is shared and not on what is transmitted. Two communicators form one dynamic system by establishing stabilizing communication links on different levels of control he said, introducing the feedback model as developed by the research year's feedback group. Barresi asked for the conditions which make it possible that people share the same world. Minimal common worlds emerge from the overlap of personal worlds, he explained. With the increasing dependence on social life, personal and social worlds become more intertwined and co-constructed. Barresi looks for a language in which relations between embodied agents of all kinds, from different species of animals, to humans, to robots can be described. With such a language, perhaps we might be able to develop a better knowledge of the laws of social nature that would also lead to a better understanding of our own embodied existence, he said. The context in which communication evolved in birds is a network-like structure and not a dyadic one, **Marc Naguib** (Bielefeld) pointed out. In those networks the signals are

influenced by the presence of an audience that is able to extract information from other's interactions. Males can e.g. obtain information about the strength of intruders from the songs of their neighbours; females react to wins and losses in song contests. Based on the model developed by the research groups' feedback group, **Karl Grammer** and **Lisa Oberzaucher** (Vienna) explained that signals are messages which influence other peoples behaviour based on cost-benefit evaluations. Truth, so their provocative statement, does not matter as long as everyone benefits in the long run. Feedback may have evolved as a probing device to detect cheating or it might serve to manipulate others by demonstrating action tendencies. **Achim Stephan** (Osnabrück) provided an overview of the communicative abilities of animals in order to reflect on how far we can get with artificial communicators in principle. He emphasized that, despite its variety in means, animal communication serves very limited number of purposes such as displaying the attraction of mating partners, warning conspecifics, or pointing to food sources. These are topics, he pointed out, that play no role in the construction of artificial agents. He proposed to keep practical and conceptual questions apart: What is important for us in order to have a good communication with an artificial agent? What is the real nature of an artificial agent? All these results and insights add arguments to the Embodied Communication Perspective and strongly substantiate the multidimensional and closely coupled nature of communication it proposes.