

The Dance of Communication

by Susan Duncan (Chicago, IL)

Human beings are able to communicate about many topics with relative ease, despite the fact that natural conversation is full of short and incomplete or otherwise ill-formed spoken utterances. Familiar explanations of this remarkable ability focuses on the speech stream itself, the auditory signal that moves from speaker to listener and that according e.g. to the interactive alignment model, is the sole physical link between two essentially isolated minds. But imagine the following everyday conversational interactions, each of which illustrates how our perspective on communication greatly enlarges upon language processing:

A customer asks a pharmacist about some medication he wants to purchase, the name of which he has forgotten. He says, “it comes in a box like this,” while performing hand gestures that depict the dimensions of the box. He moves his straight, flat palms through space to represent the sides, front and back of the box, in sequence. The pharmacist takes over, gesturally depicting sides, top and bottom, saying, “Ah, in such a box it was?”

A person in the market for a second-hand car attempts to persuade the owner of a repair shop to sell her a vehicle on his lot for less than his asking price. The negotiation plays out as the two walk around the lot, examine the vehicle, and return in the direction of the shop. The customer persistently maneuvers to achieve face-to-face orientation with the shop owner and to keep herself in the focus of his attention. At intervals she claps him on the shoulder or demands eye contact. The coordination of her movements with his seems at times almost choreographed; strikingly so during an interval in which, as the two stride together toward the shop, she suddenly pivots 180 degrees to walk backward, facing the owner, as he continues forward at a brisk pace.

Examples like these make clear that sound is not all conversation partners exhibit to one another! The car buyer was willing even to put herself at risk of tripping and falling over backward in order to conclude her “dance” of negotiation in a preferred interpersonal formation. Also, consider that what happened at the pharmacy was no less than transmission of a mental representation from one person to another via bodily gestures. The seeming simplicity of such an example, however, belies underlying complex issues. Note that the pharmacist’s gestural sequence differed from the customer’s in terms of which surfaces of the box were represented. It was not the movements themselves, rather the thing they represented that the pharmacist internalized.

Current research underway in the context of the ZiF’s research year on “Embodied Communication in Humans and Machines” explores how, in fact, a great deal more is going on in human communication than auditory signaling, organized as arbitrary elements in a linguistic code that travels from a speaker to a listener, who are otherwise isolated from each other. The basis of this research is the sense that no account of human communication will be adequate that confines itself to inferences concerning speaker- and listener-internal symbolic representations and cognitive processes acting on them to encode and decode arbitrary signals. Rather, we held the specification of how language use draws on basic sensorimotor processes, how it incorporates the body as a mapper of meaning and an actor in the world, and how language use is scaffolded by social-interactive and physical environments to be essential to an adequate account of communication.

Inquiries are developing along several intersecting lines that promise to lead to a fuller understanding of how human beings come to understand one another when they converse face-to-face. Our perspective on embodiment thus motivates consideration of theories claiming that human cognitive systems, including the capacity for language, are parasitic on or emerge from a substrate of sensorimotor functioning. Investigation of embodiment obviously—as illustrated in the examples above—must also seek to delineate the roles that different modalities—posture / positioning, gaze, gesture, speech, etc.—play in the conduct of communication and collaborative action and how these are coordinated in real-time interaction. We investigate both *mutual* and *joint* understanding, i.e. the roles that bodily modalities play in interlocutors’ understanding of one another *and* in cooperative situations in which the parties seek to understand some part of the world together.

Defining Communication

A working definition of communication assumed by the research group can most generally be stated as sharing of a content by one or more communicators. When we map this notion to what is seen in communicative behavior, say in insect colonies, reptiles, birds, mammals, and particularly primates including apes and human beings, we can recognize a steady increase in communicative dimensions.

Feedback

In order to be successful, communication requires a feedback system to make sure the contributed information is really shared. Using the cybernetic notion of feedback of Wiener (1948) as a point of departure, we may, as a special case, define a notion of communicative

feedback as follows: unobtrusive (usually short) vocal or bodily expressions whereby a recipient of information informs the contributor about whether he/she is able and willing to (i) communicate (have contact), (ii) perceive the information (perception), and (iii) understand the information (understanding). In addition, feedback information can be given about emotions and attitudes triggered by the information, a special case here being an evaluation of the main evocative function of the current and most recent contributions. A sub-project on feedback envisages several types of feedback occurring with different degrees of awareness and control (indicate, display and signal) and involving bodily coordination, conventional body movements (mostly head movements) and facial gestures as well as conventional vocal signals (words) and prosody.

Rhythm

Interactional rhythm and rhythm in natural language use, observed multimodally, is another domain of inquiry. We consider the capacity to organize behavior according to isochronous and more complex rhythmic patterns as a natural endowment of many animal species. This endowment is exploited by some species, including humans, for communicative ends. In humans, we observe that speech intonational prosody, speech-linked gestures, and larger postural movements may pattern together quasi-rhythmically for intervals of natural interaction. We consider the role this may have in scaffolding cognitive processes underlying speech-gesture production and comprehension, discourse planning, and turn management during natural language exchanges; for example, rhythmic structure as moment-to-moment support for formation of immediate-term expectations concerning the communicative behaviors and intentions of a conversation partner. This inquiry extends to the study of the roles of gesture in other forms of language use; for example, improvised oral poetry such as that of ‘freestyling rappers.’ In such idioms, hand-gestures and other bodily movements seem to facilitate the coordination of impromptu lyrical production with musical *beats*. The latter in ways that then spur our attention back to less-lyrical uses of human language and a more systematic search for multimodal expression of rhythmic patterning there.

Virtual humans

Availability of advanced virtual human technology at the University of Bielefeld in the form of the virtual agent of Max further permits us to attempt to model some aspects of our work, as a number of types of feedback: the words *ja*, *nein*, *ne* and *hm*, modulated by prosody; also, head nods and head shakes, modulated by facial gestures. Rhythm may be the key to

successful multimodal integration of Max's various expressive behaviors which may be systematically manipulated to be more or less rhythmically integrated, multimodally (both in terms of self- and self-other synchrony), in ways that are difficult or impossible to elicit in real human communicative interactions. Such an approach will enable study of the cognitive processing and the social-interactions effects of rhythm.

As an alternative theory of human language processing 'Growth Point' theory of language use as developed by David McNeill posits a central role for visuo-spatio-motoric imagery in language use. The theory has developed on the basis of analysis of speech-co-occurring gestures in natural discourse contexts. We observe how these shape meaning and examine their systematic temporal alignments to constituents of speech and to speech intonational contours. A core claim of the theory is that analog gestures are overt manifestations of a mode of semiosis as integral to language production as categorical linguistic semiosis. This project assesses the possibilities for modeling such a theory of language, the challenge being that it crucially depends on human 'sense of meaning', affective intentionality, and visuo-spatio-motoric representations—*embodiments* not native to a virtual agent.

These inquiries outlined above are steps toward the integrative perspective on embodied communication the research year is aiming at. We will take this goal to be achieved when we can finally explain a seeming simple every-day situation such as those described in the examples above. These, though seemingly simple, nonetheless contain the central elements of our research topic.