Gestures Supporting Dialogue Structure and Interaction in the Bielefeld
Speech and Gesture Alignment Corpus (SaGA)

Florian Hahn
CRC 673 Alignment in Communication
Bielefeld University
`fhahn2@uni-bielefeld.de`

Hannes Rieser
CRC 673 Alignment in Communication
Bielefeld University
`hannes.rieser@uni-bielefeld.de`

Abstract

We report about gestures supporting dialogue structure and interaction in the Bielefeld Speech and Gesture Alignment corpus and provide a first classification of them based on Hahn and Rieser (2009-11). Numbers will be given on the poster.

Types of Dialogue Supporting & Interactive Gestures

Our study is based on the Bielefeld Speech and Gesture Alignment corpus (SaGA) containing 25 route description dialogues generated as follows: a Router "drives" through a VR town along a route. His ride is reported to a Follower, who is expected to follow the route by himself. The data contains 5000 indexical and iconic gestures annotated in ELAN and rated (see (Lücking et al., 2010), for results) and approximately 1000 gestures supporting dialogue structure and interaction.

An important trait of the Router-Follower-situation is that it is "layered" (Clark, 1996): We have the route context using the conversational participants' (CPs) gesture spaces detailing the topical or baseline information, the larger embedding context of the experimental situation and the still more encompassing one consisting of the University and Bielefeld City. The discourse-related gestures introduced below can be grouped roughly into gestures used in turn allocation, feed-back gestures in second turn, those indicating assessment of evidence, gestures serving to highlight information, sequences of quick feed-back or monitoring gestures tied to sub-propositional contributions and, finally, truly interactive gestures exclusively social in character. All of these are accompanying speech.

Gestures related to turn allocation: Since the seminal paper of Sacks et al. (1974), valid also for dyads, we assume a regularity for turn allocation in dialogue depending structurally on the larger speech-exchange system: current speaker dominates, he selects next. If not, one of the other speakers can self-select. This option omitted, the first speaker may continue. The SaGA data show that there is more freedom in this schema leaving room for quick interrupts of other. These become acceptable for CPs if interactionally cushioned. Gestures in this class exploit the layeredness property of the situation: current speaker points to other selecting him as next. In contrast, indexing other to take the turn oneself is also a possibility. In the context of a completion current speaker may gesturally invite a contribution from other. Time being a scarce resource, current speaker may indicate a lapse should be tolerated by other and use a finger-to-lip or finger-below-lip gesture to express that. In tightly coordinated discourse there is an interesting "attack-ward-off pair": with a gesture similar to "indexing other" other may indicate that he wants to contribute at a non-turn-transition relevance place. Discouraging that, current speaker may try to fence him off with a posture using palm slanted and ASL-shape B-spread. Under pressure current speaker may give in and offer a "go ahead", palms up, directed against the domineering CP (see (Rieser, 2011)).

Feed-back gestures in second turn. Speaker of a second turn may use an iconic gesture of previ-

ous speaker in order to indicate acknowledgement or accept. As with the indexing, next speaker's gesture imitation uses a topical gesture in a discourse function. Less spectacular means can also be used in second turn, for example pointing without referring. Acknowledging an acknowledgement of second speaker by first may be done in essentially the same way (Bergmann et al., 2011).

Gestures indicating assessment of evidence. Given the Follower's route following task, it is of course vital that he get reliable information about landmarks and directions. This is a pressure on both CPs. We observed two groups of gestures to indicate reliability of information. One is conveying doubt concerning the fit of a description, usually for a landmark or one of its properties. The other one is indicating an agent's epistemic state concerning a situation and characterizing it as weaker than knowing or believing. The first one is aligned with the description in question using ASLs B-spread and a wiggle in handshape or wrist, the other one related to propositional content is a lifting of a hand out of and into a rest position with handshape B-spread accompanied by a head shake.

Gestures to highlight and to downgrade information. We take it for granted that beats are used for emphasis. However, underlining information – often an accented tone group – can also be suggested lifting a G-shaped hand, directing it against the addressee and moving it in a beat-like fashion. On the other hand we have the near-universal "brush-away" gesture indicating that information is considered to be not so relevant.

Sequences of quick feed-back or monitoring gestures tied to sub-propositional contributions. Propositions are of a Fregean design. CPs in near-to-natural task-oriented dialogue often converse quickly and in short thrusts. So we can have a Router's "don't interrupt" followed by the Follower's "let me interrupt" and, finally, the Router's acknowledgement and a "go ahead" gesture. This shows that full-blown dialogue acts do not always matter.

Truly interactive gestures exclusively social in character. To sum up: gestures accompanying turn allocation, feedback gestures in second turn and sequences of quick feed-back or monitoring gestures have to be embedded into suitable adjacency pairs

and reconstructed at the level of dialogue acts. So, in order to explain a gesture's function many parameters have to be considered besides gesture morphology. The embedding speech exchange system in SaGA is a plan-based (memory-based) one on the Router's side and a plan-generating one on the Follower's side providing larger sequential structures. Gestures indicating assessment of evidence and those to highlight and to downgrade information figure at the level of dialogue acts. From these structural features we want to delineate gestures which are truly interactive such as hand and body postures to mollify someone or touching or caressing him. We have calming down and don't bother gestures. Calming down has a B-spread handshape and a slanted palm directed against the torso of other. Don't bother gestures resemble the brush-away gestures in many respects but are also directed against the other's torso.

Acknowledgments

References

Kirsten Bergmann, Hannes Rieser, and Stefan Kopp. 2011. *Regulating Dialogue with Gestures - Towards an Empirically Grounded Simulation with Conversational Agents*, SIGdial 2011.

Herbert H. Clark. 1996. *Using Language*, Cambridge University Press, Cambridge, UK.

Florian Hahn and Hannes Rieser. 2009-11. *Dialogue Structure Gestures and Interactive Gestures. Annotation Manual. CRC 673, Alignment in Communication. Working Paper*, Bielefeld University.

Andy Lücking et al. 2010. The Bielefeld Speech and Gesture Alignment Corpus (SaGA). In: Michael Kipp et al. (Eds.) 2010, *Workshop: Multimodal Corpora*, 92-98.

Hannes Rieser. 2011. *Gestures Indicating Dialogue Structure*. Accepted for SEMdial

Harvey Sacks, Emanuel A. Schegloff, Gail Jefferson 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. In: *Language, 50*: 696-735.