Hannes Rieser & Florian Hahn
Universität Bielefeld, Germany

Corpus-based gesture typology

Questions of gesture typology have always been a topic in the literature on gestures co-occurring with speech, see for example Kendon (2004). However, the situation for typology work in this field has drastically changed due to annotated MM corpora based on sophisticated data acquisition techniques and the application of statistical methods. We report about joint work on gesture typology in the Bielefeld CRC 673, Alignment in Communication, starting with our corpus. Then we go into the technology developed for gesture investigation, called grid-technology. In the sequel, the statistical underpinning for a typology of the whole corpus is given. Finally, we show the use of statistics-based typology and partial ontology in the description of a MM dialogue.

Our typology work is based on the Bielefeld Speech-And-Gesture-Alignment corpus, SAGA. It comprises 25 MM dialogues, route and landmark descriptions of a sight-seeing tour through a virtual town given by a Router to a Follower. The dialogues are video-filmed and recorded using speech recordings, body-tracking techniques and eye-tracking. Of the corpus containing some 6000 gestures, roughly 5000 have been annotated in ELAN. Annotation is carried out inter alia with respect to agents, hands and features for e.g. hand-shape. The annotation is manual-based and has been rated, (agreement score 75%).

The annotation features provide the basis for the typology set up in two steps:

In the first step, a complete typological grid is developed for one video film (see Rieser 2010) as follows: The features originating from the annotation are grouped into single features, feature complexes, and objects of various dimensions. The grouping is motivated by gesture identity conditions and considerations of compositionality, i.e. by observations concerning which feature bundles are recurring in the grid-datum. Obviously, the feature configurations have to have a specifiable distribution. They are assembled in a multiple inheritance hierarchy using a special graph-theoretical tool. Independently, every feature configuration is mapped onto a partial ontology description giving its semantics: A gesture using specific hand-shape, wrist-movement, palm-direction, and back-of-hand-direction can depict a cylinder partially. Depictional content can in turn interface with the lexical semantics of the noun “tower”.

The second step takes the typological grid as a hypothesis for the whole SAGA corpus. Using up-to-date statistical methods it checks, how the categories set up for the grid are distributed in SAGA (henceforth “SAGA-groupings”). The statistics uses only gesture morphology, it does not incorporate partial ontology. As a consequence, we get an independent estimate as to whether some partial ontology description newly stipulated for a SAGA-grouping matches the partial ontology attributed to a grid item, for example, whether a configuration of gesture features in the grid partially depicts a particular entity in the WHOLE SAGA corpus.

Finally, gesture typology is put to work: using a two-turn exchange of a MM dialogue, two things are shown: the interface of gesture meaning and verbal meaning yielding MM propositions and the function of gesture meaning for dialogue management.

References

Kendon, A.: *Gesture: Visible Action as Utterance*. CUP, (2004)

Rieser, H.: On Factoring out a Gesture Typology from the Bielefeld Speech-And-Gesture-Alignment Corpus (SAGA). In: Proceedings of GW 2009, (to appear 2010).

Rieser, H., & Poesio, M. (2009). Interactive Gesture in Dialogue: a PTT Model. In P. Healey et al. (Eds.), *Proceedings of the SIGDIAL 2009 Conference* (pp. 87–96). London, UK: ACL.

Isabel Rodrigues[1] & Carla Fernandes[2]
(1) Universidade do Porto, Portugal; (2) Universidade Nova de Lisboa, Portugal;

Dance and spaces: Multimodal metaphors for emotions, interpersonal relations, and society

This paper is about the creation of spaces and mutimodal metaphors in a contemporary dance performance from Rui Horta.

The way people position themselves in relation to one another in space and the different spaces in interaction they create convey relevant cues for the interpretation of both interpersonal relations and the engagement and disengagement in interaction/conversation. These aspects were considered in the turn-taking rules, in the descriptions of the rights and obligations attached to the roles of speaker and hearer in a conversation and in other regulating principles for conversation.

The positioning of interacting people in space has been the object of research of some scholars, such as Hall (1969) and Kendon (1990), more recently Haviland (2000), Özyürek (2002), Sweetser/Sizemore (2006). Kendon (1990) observed the occupation of space in multiparty conversations in time and uses the expression F-formation system to describe the way participants adjust their positions in relation to each other during face-to-face interactions. This kind of studies is now being pursued in the I.A., for the development of embodied agents, such as Tapus/Mataric (2007).

Categories like Kendon’s F-formation system, o-space and face-address system, or different classifications by others for personal space, gesture space and interactional space represent the theoretical background for the description of the relation between body movements, speech and position in space.

The analysis of dancers’ body movements and of the spaces they create (or not) is a way to find out how the choreographer uses proxemics to express abstractions like freedom, standardized society, as well as emotions (Camurri et al., 2003) in relations between the single and the group.

These concepts correspond to the theme of the performance, according to Horta, the idea of space for freedom and for singularity in the standardized society. The oscillations between distance and body contact function as concrete representations for the a.m. abstract dimensions, i.e., as metaphors. Furthermore, movements taken from the video material recorded during the creation processes of Horta will be interpreted with the background of conventional vs. idiosyncratic metaphor and multimodal metaphor (Müller/Cienki, 2009). Understanding that metaphor occurs in other modes than language alone, the analysis of metaphors on dance videos will be framed within the Multimodal Metaphor theory as presented on Forceville/ Urios-Aparisi (2009). We suggest that conceptual metaphors find expression in the dance movements themselves, and indeed in some visual sign used by the choreographer, in ways that are not always translatable into language. They may be taken, as “direct” manifestations of the conceptual metaphors unmediated by language.

References

Camurri, A./Lagerlöf, I., Volpe, G. (2003)”Recognizing emotion from dance movement”. In: *Int. J. Human Computer Studies*, 59, 213–225.

Forceville, C./ Urios-Aparisi, E. (2009) *Multimodal Metaphor*. Berlin, de Gruyter.

Hall, E. (1966) *The Hidden Dimension*. New York, Doubleday.

Haviland, J. (2000) “Pointing, gesture spaces and mental maps”. In: McNeill, D. (ed.) *Language and Gesture*. Cambridge, CUP.

Kendon, Adam (1990) *Conducting Interaction*. Cambridge, CUP.

Müller, C./Cienki, A. (2009) “Words, Gestures and beyond”. In: Forceville, C./Urios-Aparisi (eds.) *Multimodal Metaphor,* Berlin, de Gruyter, 297–328.

Özyürek, A. (2002) “Do speakers design their co-speech gestures for their addresses?”. In: *J.of Memory and*