Multimodal Propositions? From Semiotic to Semantic Considerations in the Case of Gestural Deictics

1 Introduction: Gestural Deixis

We call utterances that comprise elements that are perceived by different sense modalities *and* are coded according to a non-linguistic code *multimodal utterances* (cf. (Fricke, 2012)). An example for non-linguistically coded signs are *indexicals* like pointing gestures, which, according to Peirce, bear some *nomological*, *causal* or *attentional* relation to their objects (CP 1.372, 2.248, 2.285 [1]) In situated dialogue, interlocutor's frequently use multimodal utterances like definite descriptions plus pointing gestures in an exophoric way. Accordingly, such deictic acts are a starting point for looking for multimodal propositions: Whereas the linguistic expressions are treated as arbitrary symbols that are interpreted intensionally with respect to a world or a circumstance according to some standard model theory, the nonlinguistic element, if it indeed follows a nonlinguistic code, must, by definition, interpreted in a different way. At least three distinctions have been made at first, however (Levinson, 2008):

- In the most direct way, viz. *gestural deixis*, the pointing identifies a concrete, perceptible object (or event, property, etc., depending on your metaphysical stance).
- The point or region or object in space pointed at in an utterance situation can stand as a proxy for some spatial configuration or referent in the described situation (*transposed deixis*).
- In *symbolic deixis*, the indicated thing is used as a case of *deferred reference* (Nunberg, 1993), that is, standing for something which is somehow related to the indicated entity.

The different uses of deictic gestures are well documented in the literature on co-verbal gestures (see (Fricke, 2007), (Lascarides and Stone, 2009), (Alahverdzhieva and Lascarides, 2011)). They necessitate a distinction between something pointed at in an utterance situation and something referred to in the described situation. We take up the terminology of (Kühnlein, 1999) and call the former *index* and the latter *referent*. Accordingly, the task for a semantics of speech and co-verbal pointing gestures has two aspects:

1. provide an account for the index;
2. provide information for resolving the referent (maybe in pragmatics).

The main focus here is on the first aspect.

2 Significance of Pointing Gestures

Putting theoretical as well as empirical insights together (e.g., (Reimer, 1991), (Bangerter and Oppenheimer, 2006), (Rieser, 2004)), the resulting picture on gestural deictics like *This N* plus pointing is as follows (cf. Figure 1):

1. The demonstrative *This* is an attentional index according to the Peircean distinction motivated above, which shifts the attention of the addressee towards the gesture.
2. The gesture in turn projects a "search space" in terms of a spatial cone extension ((Kranstedt et al., 2006)). The gesture is a causal index, since it is directly affected by the location of the intended referent (cf. the respective remarks in Section 1).
3. The nominal expression *N* finally picks out the referent from the search space by virtue of descriptive conventions.

In order to capture the spatial nature of gestural deictics, we employ a situation semantics-related model with a rich spatial structure, resting on the central notion of oriented vector spaces. An oriented vector space relates to a pointing gesture in the following, twofold way:

[1] We follow the convention to quote the *Collected Papers of Charles Sanders Peirce* (Peirce, 1965) in terms of volume (v) and paragraph (p): 'CP v.p'.

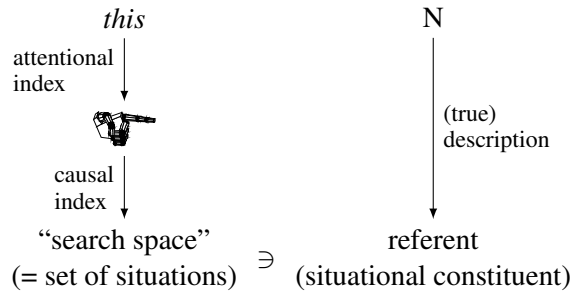

Figure 1: Collaboration in gestural deictics of the form *This N*.

1. The origin of the speaker's vector space **V** in the current utterance situation (i.e., space(speaker(s))) provides the Bühlerean (Bühler, 1999) Origo for pointing gestures;
2. The pointing cone is projected into the direction determined by the orientation of the index finger with reference to the orienting axes.

Having introduced an oriented vector space according to the two steps above, the spatial extension of a pointing gesture can be specified in terms of sets of vectors. Suppose a pointing gestures G ahead, straight away from the speaker's body. Then the region that G encircles is the set of vectors emanating in the origin of **V** in the direction of **FT**. The corresponding cone covers the following subspace $r(G) \in R \subset \mathbf{V}$: $r(G) = \{\mathbf{v} | \text{proj}_{\text{LONG}}\,\mathbf{v} < \text{proj}_{\text{VERT}}\,\mathbf{v} \wedge \text{proj}_{\text{LONG}}\,\mathbf{v} < \text{proj}_{\text{LAT}}\,\mathbf{v}\}$ (where $\text{proj}_y\,\mathbf{u}$ is the orthogonal projection from vector **u** onto line y). The subspace defined this way is quite large so that further constraints for instance in terms of angular specifications should be given. However, angular modification does not affect the account sketched here in principal.

The situational extension of a pointing gesture can then be specified in terms of the set of situations which regions that have relatively maximal intersections with $r(G)$:

(1) $[\![$  $]\!]^{\mathscr{M}} = \{e | \text{region}(e) \cap r(G) \rightsquigarrow \max_i\}$.

'$\rightsquigarrow \max_i$' picks out the i situations that have the largest overlap with the pointing cone. This, the function produces an ordering on situations pointed at, decreasing according to their intersection area with the pointing cone. That is, the spatial extension is assumed to be parameterized. Of course, the best guess at first is to choose the situation with maximal intersection (what corresponds to setting parameter $i = 1$). However, any $i < 2$ does no harm as long the maximal situation provides only on entity that fits the nominal description. This includes plural entities in case of plural nominals. The parameterized treatment leaves a great way of modeling freedom for taking semantic-pragmatic interface issues into account, but is out of scope here. In particular, the spatial model with anchored and oriented vector spaces and the form-based, perspectival interpretation of pointing gestures spells out the cone stipulations verbalized in (Lascarides and Stone, 2009, p. 44). It also gives an account for "referents at certain coordinates" as assumed by (Alahverdzhieva and Lascarides, 2011, p. 17).

Semiotics	*Semantics*
affectedness	form-based interpretation
non-symbolic code	perspectivity
focusing attention	reflexivity

Table 1: Contraposing semiotic and semantic features of gestural deictics.

3 Discussion

In which ways, if any, are propositions for gestural deixis multimodal?

1. Multimodal propositions are *reflexive*: situation s occurs both as the described entity and as part of the description (cf. Figure 1).
2. The interpretation of a gesture G, by means of the determination of $r(G)$, is essentially *affected by the form* of the gesture.
3. The demonstration part of gestural deictics is *perspectival* by depending on the speaker's orientation in space.

The comparison of semiotic theorizing and semantic modeling is summarized in Table 1; it shall not be claimed, however, that pairs of cells are related in a one-to-one manner.

This model not only captures a great deal of semiotic and empirical insights briefly introduced above, it also goes beyond the formal analyses proposed *in this respect* so far.

References

Katya Alahverdzhieva and Alex Lascarides. 2011. An HPSG approach to synchronous speech and deixis. In Stefan Müller, editor, *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar*, pages 6–24.

Adrian Bangerter and Daniel M. Oppenheimer. 2006. Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture*, 6(1):85–102.

Karl Bühler. 1999. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. UTB, Lucius & Lucius, Stuttgart.

Ellen Fricke. 2007. *Origo, Geste und Raum*. Linguistik – Impulse & Tendenzen. Walter de Gruyter, Berlin and New York.

Ellen Fricke. 2012. *Grammatik multimodal. Wie Wörter und Gesten zusammenwirken*, volume 40 of *Linguistik – Impulse und Tendenzen*. De Gruyter, Berlin and Boston.

Alfred Kranstedt, Andy Lücking, Thies Pfeiffer, Hannes Rieser, and Ipke Wachsmuth. 2006. Deictic object reference in task-oriented dialogue. In Gert Rickheit and Ipke Wachsmuth, editors, *Situated Communication*, pages 155–207. Mouton de Gruyter, Berlin.

Peter Kühnlein. 1999. Dynamics of complex information. Proc. of ESSLLI. Utrecht, Netherlands.

Alex Lascarides and Matthew Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics*, 26(4):393–449.

Stephen C. Levinson. 2008. Deixis. In *The Handbook of Pragmatics*, chapter 5, pages 97–121. Blackwell.

Geoffrey Nunberg. 1993. Indexicality and deixis. *Linguistics and Philosophy*, 16(1):1–43.

Charles Sanders Peirce. 1965. Collected papers of Charles Sanders Peirce. repr. from 1932.

Marga Reimer. 1991. Do demonstrations have semantic significance? *Analysis*, 51(4):177–183.

Hannes Rieser. 2004. Pointing in dialogue. In *Proceedings of the Eighth Workshop on the Semantics and Pragmatics of Dialogue*, Catalog '04, pages 93–100, Barcelona.