The Interplay between the Speaker's and the Hearer's Perspective

Petra Hendriks, Helen de Hoop & Henriëtte de Swart

Suppose you witnessed the killing of Harry by Frank and wish to report on that. To express this unfortunate event, you may consider using a sentence such as *Frank killed Harry*, or *A man killed Harry*, or *He killed Harry*, and perhaps even briefly consider the reverse word order *Harry killed Frank*. When choosing the form to be used, however, it is important that a hearer will understand the chosen form correctly. This means that the hearer should be able to identify the intended referents on the basis of the referring expressions used. If a pronoun is used, it should be clear to the hearer to whom this pronoun refers. In addition, the hearer should be able to distinguish between an interpretation according to which Frank is the killer and Harry the victim and an interpretation such that Harry is the killer and Frank the victim. In English, this distinction is made with the help of word order. That is, for an English hearer, the sentence *Frank killed Harry* can only mean that Frank is the agent and Harry the patient in the killing event, as Frank is mentioned first. If the speaker had wanted to express the meaning that Harry killed Frank, the speaker would have mentioned Harry first and thus would have said *Harry killed Frank*. In Dutch, like in English, agent-patient argument structure is generally reflected in word order, although in particular contexts and with special intonation the patient may sometimes appear first.

Now suppose you did not witness the killing event but only saw Harry's body and would like to know the identity of his killer. In English, you would ask *Who killed Harry?* and in Dutch *Wie heeft Harry vermoord?* For someone hearing the English question, it would be perfectly clear what you are after. If you knew that Harry was the killer instead of the victim and were interested in the identity of Harry's victim, you would have asked *Who did Harry kill?* In Dutch, however, you would have asked the exact same question *Wie heeft Harry vermoord?* in the latter situation, too. The fact that the same form is used by Dutch speakers to inquire about the identity of the agent and about the identity of the patient makes this form ambiguous for Dutch hearers. Although differences in form generally signal differences in meaning, wh-questions in Dutch illustrate that in particular situations a difference in form (in this case, word order) is neutralized

by a stronger force in the language, such as the obligatory fronting of wh-words. This then results in the availability of only one form for two fundamentally different meanings.

Recent years have seen a surge of interest in frameworks of language and communication that seek to explain the patterns of forms and meanings in a language as resulting from the interaction between the speaker's and the hearer's perspective. Rather than viewing language as mainly determined by individual language users' linguistic competence and performance, such frameworks consider language to be crucially dependent on the interaction between speakers and hearers. Formal accounts of such interactions have been proposed in terms of, e.g., bidirectional Optimality Theory (Blutner, 2000; Blutner, de Hoop, & Hendriks, 2006; Hendriks, de Hoop, Krämer, de Swart, & Zwarts, 2010) and Game Theory (Parikh, 2001; van Rooij, 2004). As these frameworks generally predict a one-to-one pairing of forms and meanings, the ambiguity of wh-questions in Dutch discussed above and other asymmetries in the pairing of forms and meanings provide a challenge for these linguistic frameworks.

The neutralization of contrasts in form or meaning that is sometimes observed in language is at odds with the classical view that language is a systematic one-to-one pairing of forms and meanings. This classical view is reflected in Williams' (1997:578) observation of a "hatred of synonymy": "if two forms exist (in syntax or morphology), they must have different meanings; if two forms cannot be assigned different meanings, then one of them cannot exist". Williams' Blocking Principle is in accordance with general principles of economy in language (it is not economical to have two forms that express the exact same meaning), as well as with the cross-linguistic generalization that marked forms typically have marked meanings (Levinson, 2000). The generalization that marked forms go with marked meanings does not need to be posited as an independent principle of language. In the framework of bidirectional Optimality Theory (Blutner, 2000), which arose out of the theory of grammar Optimality Theory (Prince & Smolensky, 1993/2004), this generalization turns out to be epiphenomenal and emerges as a result of the evaluation procedure.

In bidirectional Optimality Theory, the speaker's perspective and the hearer's perspective are integrated in a formal model of communication. Generation involves taking the speaker's perspective by selecting an optimal form for an input meaning *and* taking the hearer's perspective by recovering that meaning from the optimal form. Likewise, interpretation involves taking the hearer's perspective by selecting an optimal interpretation for an input form *and*

taking the speaker's perspective by reproducing that form for the optimal interpretation. Thus, bidirectional optimization is more than the sum of two unidirectional processes of optimization. In contrast to unidirectional Optimality Theory, in bidirectional Optimality Theory evaluation takes place over form-meaning pairs rather than over forms or meanings separately. As a result, the non-recursive or strong version of bidirectional Optimality Theory captures cases of total blocking (the non-existence of a form because an alternative form already exists for the same meaning), and the recursive or weak version is able to capture partial blocking or iconicity (because a marked form is used for a marked meaning, the other – unmarked – form takes up the other – unmarked – meaning). Furthermore, bidirectional optimization can account for patterns of cross-linguistic variation as in the expression and interpretation of negation (de Swart, 2010) and its difficulty for children may explain why production sometimes precedes comprehension in language acquisition (Hendriks & Spenader, 2005/6).

A drawback of the standard versions of bidirectional Optimality Theory as proposed by Blutner (2000) is that they are not flexible enough to deal with exceptions to the one-to-one mapping between form and meaning in adult language. We already discussed the pattern of one form for two meanings with Dutch wh-questions. In general, bidirectional Optimality Theory has problems accounting for ambiguity and optionality in natural language (Beaver & Lee, 2004). Linguists have noted that, whereas marked forms tend to be used for marked interpretations, unmarked forms can often be used to express unmarked as well as marked meanings, even when there is a marked form available in the language (de Hoop, Haverkort & van den Noort, 2004). The standard versions of bidirectional Optimality Theory cannot deal with these patterns of variation in form and variation in meaning encountered in natural language.

As communication is more than the mere combination of the two unidirectional processes of generation and interpretation, we need to fine-tune our modeling of communication. Several studies have pointed out the connection between bidirectional Optimality Theory and Game Theory (e.g., Benz & Mattausch, 2011). Bidirectional Optimality Theory is rooted in grammar and as such is concerned with the constraints and their interaction; it has no direct means to represent speaker and hearer strategies. Game Theory, on the other hand, is a theory of strategic interaction; it can provide us with clear criteria for distinguishing between strategies in communication but is not concerned with the representation of grammar. Thus Bidirectional Optimality Theory and Game Theory complement each other. The papers in this special issue

arose from a series of workshops that were organized to explore variants of bidirectional Optimality Theory and Game Theory as models of the interplay between the speaker's and the hearer's perspective. The papers in this special issue provide additional motivation for making a distinction between the speaker's perspective and the hearer's perspective and address the issue of how the two perspectives must interact to account for patterns of variation in form and meaning in language and communication.

Legendre and Smolensky show that there is a production/comprehension asymmetry in children's acquisition of third-person subject pronouns in French. This asymmetry is explained by these children's non-adult-like constraint ranking. Due to the direction-sensitivity of constraints in Optimality Theory, this non-adult-like constraint ranking results in errors of reference assignment in comprehension but correct forms in production.

Klimek-Jankowska provides an account of aspect in bare habitual statements in Polish in the framework of bidirectional Optimality Theory. To capture the observed variation in production and ambiguity in interpretation, this paper employs stratified bidirectional Optimality Theory, which allows for partial constraint rankings.

Bouma and Hendriks explain partial freezing of word order in Dutch in terms of stratified bidirectional Optimality Theory. Their bidirectional model predicts that speakers disprefer object fronting if hearers have no other clues than word order to determine the argument structure. This prediction is supported by empirical evidence from a quantitative corpus study.

Grønn and Sæbø provide a Game-Theoretic account of the interpretation of indefinites. They claim that indefinites are unmarked for novelty and non-uniqueness, and argue that these aspects of the interpretation of indefinites are derived via competition with definite determiners on the one hand and 'another' indefinites on the other.

Benz argues that the intentional and unintentional errors produced by speakers, and the hearers' strategies to cope with these errors, can be modeled in a version of Game-Theoretic pragmatics. This provides a uniform method for explaining ambiguities, relevance implicatures, quantity implicatures, and presupposition accommodation.

Franke and Jäger explore the relation between bidirectional Optimality Theory and Game Theory. They argue that bidirectional optimization is to be grounded in signaling games rather than strategic games, and compare two implementations of such games: a model of rational step-by-step reasoning and a model of reinforcement learning.

Acknowledgments

As guest editors of this special issue we wish to thank the reviewers for their thorough and helpful reviews of the submitted manuscripts and the participants of the workshops for interesting and inspiring discussions. Furthermore, we gratefully acknowledge the Netherlands Organisation of Scientific Research (NWO) for financial support.

References

Beaver, D. & H. Lee (2004). Input-output mismatches in OT. In: R. Blutner & H. Zeevat (Eds.), *Optimality Theory and Pragmatics* (pp. 112-153). Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

Benz, A. & J. Mattausch (2011). Bidirectional Optimality Theory: An introduction. In: A. Benz & J. Mattausch (Eds.), *Bidirectional Optimality Theory* (pp. 1-31). Amsterdam: John Benjamins.

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17, 189-216.

Blutner, R., H. de Hoop, & P. Hendriks (2006). *Optimal Communication*. Stanford, CA: CSLI Publications.

Hendriks, P. & J. Spenader (2005/6). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition* 13, 319-348.

Hendriks, P., H. de Hoop, I. Krämer, H. de Swart, & J. Zwarts (2010). *Conflicts in Interpretation*. London: Equinox Publishing.

de Hoop, H., M. Haverkort, & M. van den Noort (2004). Variation in form versus variation in meaning. *Lingua* 114, 1071-1089.

Levinson, S.C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Parikh, P. (2001). *The Use of Language*. Stanford, CA: CSLI Publications.

Prince, A. & P. Smolensky (2004). *Optimality Theory. Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell. (Also appeared as Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, April 1993).

van Rooij, R. (2004). Signaling games select Horn strategies. *Linguistics and Philosophy* 27, 393-450.

de Swart, H. (2010). *Expression and Interpretation of Negation*. Dordrecht: Springer.

Williams, E. (1997). Blocking and anaphora. *Linguistic Inquiry* 28, 577-628.