EXPECTATION-BASED ENCODING OF GRAMMATICAL FUNCTIONS IN SPOKEN DISCOURSE CONTEXTS

Thomas Hörberg (Department of Linguistics, Stockholm University)
thomas_h@ling.su.se

Background. In order to facilitate information transfer during communication, language processing is assumed to be expectation-based, drawing on statistical regularities in the input (e.g., Jaeger 2013; MacDonald 2013; Venhuizen et al. 2019). During comprehension, linguistic and extra-linguistic information in the previous discourse set up expectations that facilitate interpretation. During production, information in the discourse is used to balance the upcoming utterance in a way that limits production costs, but also that ensures that the message is informative enough in order to accommodate the comprehenders' expectations. Central to communication is the processing of grammatical functions (GFs) of subject and direct object. GFs express how participants are related to events or states (e.g., *who is doing what to whom*), and involve information structural properties such as topic (e.g. Foley 2011). In many languages, there are multiple ways to encode GFs morphosyntactically (e.g. word order and case), and speakers' encoding preferences depend on an interplay between sentence-level (e.g, NP properties such as animacy), and discourse-level (e.g., information-structural considerations such as topicality) properties (Hörberg 2016). Such *GF information* types therefore work as cues during on-line GF assignment. Here, I present ongoing work that investigates whether Swedish speakers' morphosyntactic choices in the encoding of GFs in transitive sentences is expectation-based, in terms of taking both sentence-level and discourse-level GF information into account in order to limit production costs, but also to provide enough information for the listener.

GF encoding in Swedish and previous findings. Although the preferred word order in Swedish is SVO, OVS word order is also used when the direct object is topical or contrastive (Hörberg 2018). Since lexical NPs lack case marking, OVS sentences with a lexical initial NP (see Example 1) are potentially ambiguous with respect to GFs, and can be costly to comprehend. If the encoding of GFs in Swedish is expectation-based, speakers and writers should therefore balance their use of OVS word order with respect to the availability of GF information in a way that accommodates comprehension. Evidence for this was found in a corpus study of 14500 transitive sentences of written Swedish. This study showed that OVS sentences are more frequently used when sentence-level, morphosyntactic or animacy GF information is available, independent of NP accessibility. In other cases, unambiguous passive constructions are preferred, potential ambiguities thereby being avoided.

GF encoding in spoken discourse contexts. Although these findings indicate that writers' use of OVS is influenced by the availability of GF information, it does not address whether such information causally influences GF encoding in spoken discourse contexts. Since sentence-level ambiguities often can be resolved in context, and even might facilitate communication when resolvable (Piantadosi et al. 2012), the utility of GF information might be limited in discourse contexts (Rahkonen 2006). In an on-going study, I address this issue using a picture-description task (cf. Prat-Sala & Branigan 2000) with syntactic priming. The study investigates whether speakers' propensity for using OVS word order is causally influenced by the availability of animacy information in spoken discourse contexts where discourse-level GF information (i.e., anaphora and contrast) is available. A target sample of 80 participants is to be exposed to spoken short stories accompanied with cartooned images. These stories set up discourse contexts which license OVS word order (see Figure 1). The participants' task is to describe the final scene in the stories, which depict a transitive event in a contrastive context. Crucially, the direct object of the target sentence is either animate or inanimate. If speakers are sensitive to the availability of the animacy information and balance their use of OVS word order accordingly - in order to avoid a potential ambiguity - OVS word order descriptions should be more frequently used when the direct object is inanimate. However, if this information is redundant in these contexts, or if speakers are not inclined to avoid potential sentence-level ambiguities, no effect of animacy on the proportion of OVS sentences should be observed.

1) *Läraren* *gillade* *han* *inte*
teacher.the like he not
"The teacher, he did not like him"

Example 1. Although the sentence word order is OVS, the initial NP '*Läraren*' can initially easily be interpreted as the subject, as it lacks case marking and SVO word order is predominant. The sentence is therefore locally ambiguous with respect to grammatical functions.

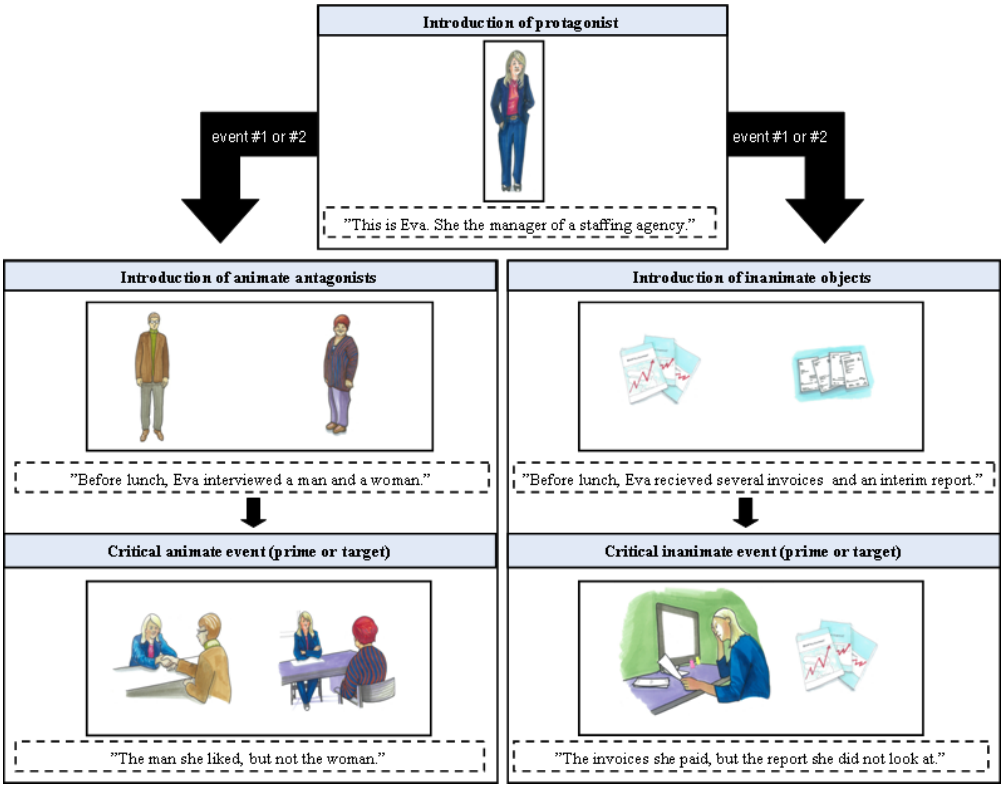

Figure 1. Example of a story discourse context. A protagonist is introduced, followed by Event 1 (prime) and Event 2 (target). The events introduce two story antagonists or story objects, that are involved in a subsequent contrastive event. This sets up a discourse context that motivates the use of OVS word order. Event 1 further functions as a prime for the use of OVS in the picture description of the final target event.

References

Foley, W. A. (2011). A typology of information packaging in the clause. In T. Shopen (Ed.), *Language Typology and Syntactic Description*: *Vol. Volume 1:Clause Structure* (pp. 362–446).

Hörberg, T. (2016). *Probabilistic and Prominence-driven Incremental Argument Interpretation in Swedish* (PhD thesis, Stockholm University).

Hörberg, T. (2018). Functional motivations behind direct object fronting in written Swedish: A corpus-distributional account. *Glossa: A Journal of General Linguistics*, *3*(1), 81.

Jaeger, T. F. (2013). Production preferences cannot be understood without reference to communication. *Frontiers in Psychology*, *4*.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–291.

Prat-Sala, M., & Branigan, H. P. (2000). Discourse Constraints on Syntactic Processing in Language Production: A Cross-Linguistic Study in English and Spanish,. *Journal of Memory and Language*, *42*(2), 168–182.

Rahkonen, M. (2006). Some aspects of topicalization in Swedish declaratives. *Linguistics*, *44*(1), 23–55.

Venhuizen, N. J., Crocker, M. W., & Brouwer, H. (2019). Expectation-based Comprehension: Modeling the Interaction of World Knowledge and Linguistic Experience. *Discourse Processes*, *56*(3), 229–255.