

Corpus-based investigations of the German Vorfeld

(Submission for 25+5 minute talk)

The German Vorfeld is a position that is underspecified with regard to the grammatical function. We assume that the Vorfeld instead has two major (semantic-) pragmatic functions: (i) to relate the current sentence to the prior context, i.e., to support the hearer in integrating the new information into his/her model of the current discourse; (ii) to provide the frame in which the current sentence is to be interpreted.

We have annotated a small corpus of parliament debates with information about the function of the Vorfeld constituent. We distinguish between functions of type (i): coreference relations (including bridging) and discourse relations; and type (ii): “internal” functions, which are not related to the prior context (including frame-setting elements). In addition, there are expletives and situation-deictic functions. We achieved good inter-annotator agreement ($\kappa = 0.66$). This allows us to perform qualitative and quantitative investigations of the German Vorfeld.

The annotations show that in our corpus only 45.1% of the Vorfeld constituents are clearly related to the prior context (i.e., instances of type (i)). This seems to confirm the results by Speyer (2007): in his data, 50.8% of the Vorfelds are occupied by backward-looking centers, i.e., constituents that relate to the prior context (Grosz et al. 1995). However, his data also shows that the type of function of the Vorfeld constituent depends to a large extent on the data’s genre. The genre in Speyer’s investigations that probably comes closest to our parliament debates are scientific radio talks. Here, Speyer finds 71% of Vorfeld constituents that are context-related.

For the presentation, we will focus on the so-called internal functions (i.e., instances of type (ii)). These mainly occur with PPs and adverbials (most of them de-adjectival), which represent common categories for frame-setting elements. We also found a considerable number of subordinated clauses in the Vorfeld with a sentence-internal function. In the presentation, we will discuss a classification of this type of Vorfeld constituents that emerges from our corpus data.

References

- Grosz, Barbara J., Aravind K. Joshi & Scott Weinstein (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 21: 203–225.
- Speyer, Augustin. (2007). Die Bedeutung der Centering Theory für Fragen der Vorfeldbesetzung im Deutschen. *Zeitschrift für Sprachwissenschaft*, 26, 83–115.