

# **A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment**

Gerard Kempen<sup>1</sup>, Karin Harbusch<sup>2</sup>

[kempen@fsw.leidenuniv.nl](mailto:kempen@fsw.leidenuniv.nl), [harbusch@informatik.uni-koblenz.de](mailto:harbusch@informatik.uni-koblenz.de)

1. Leiden University & MPI for Psycholinguistics, Nijmegen

2. University of Koblenz-Landau

The grammar of German does not impose hard constraints on the linear order of Subject (SB), Indirect Object (IO) and Direct Object (DO) in the “midfield” of finite complement or adverbial clauses (the region between the subordinating conjunction and the clause-final verb). All six possible orders are acceptable, although with varying degrees of grammaticality (Keller, 2000). Given this flexibility, which factors control the actual linearization preferences of speakers/writers of German?

Many investigations into this question have approached it from a linguistic angle and proposed linear precedence rules in syntactic terms, e.g. SB<<IO/DO; pronominal NPs<<full NPs; IO<<DO (Uszkoreit, 1987; Pechmann et al. 1996). Other studies have explored the impact of conceptual factors, e.g. definiteness (Kurz, 2000) and animacy (Dietrich & van Nice, in press).

In a (semi-)free word order language like German, animacy can influence constituent order in at least two ways: (1) indirectly, via the assignment of grammatical functions; e.g., animate NPs are more likely to become Subject, and Subject NPs tend to precede other NPs (McDonald et al., 1993); and/or (2) directly: NPs with animate referents tend to be conceptualized, assigned a functional role, and attached to the surface structure prior to (and thereby leftward of) inanimate NPs (Feleki & Branigan, 1997; Kempen & Harbusch, 2003).

We present the result of a corpus study aiming at evaluating the second hypothesis. From the German NEGRA-II treebank (Skut et al., 1997), we extracted all adverbial and complement clauses containing (SB,IO) and/or (SB,DO) pairs, possibly with an (IO,DO) pair. Each member NP of a pair was classified as animate/inanimate, definite/indefinite and pronominal/full. Of only two pair types, (SB-full,DO-pronominal) and (SB-full,IO-full), both orderings of the member NPs occur with sufficiently high frequency to allow testing the influence of animacy. In both cases, animacy of an NP significantly increases the likelihood for this NP to precede the other member of its pair. This tendency cannot be attributed to definiteness of the NPs because animacy and definiteness turn out to be uncorrelated.

In conclusion, the corpus data underpin the hypothesis that animacy can affect linearization independently of grammatical function assignment. We will argue that this result supports the view that the assignment of a grammatical function to a constituent is closely interlinked with its position assignment, and does not square well with the standard assumption of successive functional and positional production stages.

## **References**

- Dietrich, R. & van Nice, K.Y. (in press). Belebtheit, Agentivität und inkrementelle Satzproduktion. In: Habel, C. & Pechmann, T. (Eds.), *Sprachproduktion*. Wiesbaden: Westdeutscher Verlag.
- Feleki, E. & Branigan, H. P. (1997). Animacy effects in Greek sentence production: Evidence for single-stage syntactic processing? Poster presented at the Third Conference of Architectures and Mechanisms of Language Processing (AMLaP-97), Edinburgh.
- Keller, F. (2000). *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Unpublished Ph.D. Thesis, University of Edinburgh.
- Kempen, G. & Harbusch, K. (2003). Word order scrambling as a consequence of incremental sentence production. In: Haertl, H. & Tappe, H. (Eds.), *Mediating between concepts and grammar*. Berlin:

Mouton De Gruyter.

- Kurz, D. (2000). A statistical account on word order variation in German. In: Abeille', A., Brants, T., Uszkoreit, H. (Eds.), Proceedings of the COLING Workshop on Linguistically Interpreted Corpora, Luxembourg.
- McDonald, J.L., Bock, K. & Kelley, M.H. (1993). Word and world order: Semantic, phonological and metrical determinants of serial position. *Cognitive Psychology*, 25, 188-230.
- Pechmann, T., Uszkoreit, H., Engelkamp, J. & Zerbst, D. (1996). Wortstellung im deutschen Mittelfeld. Linguistische Theorie und psycholinguistische Evidenz. In: Habel, C., Kanngießer, S. & Rickheit, G. (Eds.), *Perspektiven der kognitiven Linguistik*. Wiesbaden: Westdeutscher Verlag.
- Skut, W., Krenn, B., Brants, T. & Uszkoreit, H. (1997). An annotation scheme for free word order languages. In: Proceedings of the Fifth Conference on Applied Natural Language Processing (ANLP), Washington D.C., 1997.
- Uszkoreit, H. (1987). *Word order and constituent structure in German*. Stanford CA: CSLI Publications.