

## Revisiting the German future construction: A corpus-driven approach

The future construction with *werden* + Infinitive, e.g. *Ich werde nach Paris gehen* ‘I will go to Paris’, has triggered a number of controversies in the linguistic literature. The debates that revolved around its genesis, its semantics, and the relative chronology of its meaning variants have been so fierce that Confais (1992: 81) even speaks of a “future war”. In this talk, I discuss how new corpus data can shed light on these long-standing issues. While the *werden* + Infinitive construction has already been the subject of many corpus-based studies – most of them qualitatively-oriented (e.g. Schmid 2000, Diewald & Habermann 2005), but some also quantitative (e.g. Bogner 1989, 1996 and especially Hilpert 2008) –, it is only now that we have access to a reasonable amount of corpus data from the incipient stages of the grammaticalization process to re-address the open questions in a data-driven way. Drawing on data from the Reference Corpus of Middle High German (Klein & Dipper 2016), the Bonn Early New High German Corpus (Solms & Wegera 1998), and the German Text Archive (Geyken & Gloning 2015) as well as a custom dataset of early German bible translations, I will address the following questions:

- a) Which earlier constructions can be considered predecessors of and/or analogical templates for the *werden* + Infinitive constructions? While it is usually claimed that modal verb constructions were used as future constructions before the emergence of *werden* + Infinitive, the corpus data suggest that *werden* + Participle was in the process of being grammaticalized as a future marker but then superseded by *werden* + Infinitive. As such, it might have played a more important role as an analogical template for *werden* + Infinitive than potential competing constructions like the aforementioned modal constructions or inchoative patterns, which have been discussed as potential source constructions in the more recent literature (Krämer 2005, Diewald & Habermann 2005).
- b) How did the *werden* + Infinitive construction as well as its division of labor with other constructions, especially the futurate present (see e.g. Hacke 2009), develop over time? This question will be addressed both from a synchronic and from a diachronic perspective. Random samples of sentences using either the simple present or the *werden* + Infinitive construction from the German Text Archive and from the synchronic DWDS Core Corpus of the 20th century (Geyken 2009) was hand-coded by two annotators for future reference (i.e. does the sentence refer to a future event or state?). An analysis of those attestations that were unambiguously identified as future-related shows that especially in the early modern texts, the strongest predictor for the use of *werden* + Infinitive is actually the text type, which lends support to the hypothesis that *werden* + Infinitive is and has always been strongly tied to written language and especially to more formal contexts.

In sum, re-assessing the German future construction using new corpus data and a combination of explorative and hypothesis-testing methods substantiates some of the hypotheses from earlier research, while others – such as the idea that *werden* + Infinitive emerged from language contact with Czech (Leiss 1992) – seem even less convincing in light of the empirical data.

## References

- Bogner, Istvan. 1989. Zur Entwicklung der periphrastischen Futurformen im Frühneuhochdeutschen. *Zeitschrift für deutsche Philologie* 108. 56–85.
- Bogner, Stephan. 1996. *Periphrastische Futurformen im Frühneuhochdeutschen*. Wien: Praesens.
- Confais, Jean-Paul. 1992. No future ? Les “futurs” du français et de l’allemand. In Gertrud Gréciano & Georges Kleiber (eds.), *Systèmes interactifs : mélanges en l’honneur de Jean David*, 81–93. Metz: Université de Metz, Centre d’Analyse Syntaxique.
- Diewald, Gabriele & Mechthild Habermann. 2005. Die Entwicklung von werden + Infinitiv als Futurgrammem. In Torsten Leuschner, Tanja Mortelmans & Sarah Grootd (eds.), *Grammatikalisierung im Deutschen*, 229–250. (Linguistik - Impulse Und Tendenzen 9). Berlin, New York: De Gruyter.
- Geyken, Alexander. 2009. The DWDS corpus: A reference corpus for the German language of the 20th century. In Christiane Fellbaum (ed.), *Idioms and collocations: corpus-based linguistic and lexicographic studies*, 21–41. London: Continuum.
- Geyken, Alexander & Thomas Gloning. 2015. A living text archive of 15th-19th century German: Corpus strategies, technology, organization. In Jost Gippert & Ralf Gehrke (eds.), *Historical Corpora*, 165–180. Tübingen: Narr.
- Hacke, Marion. 2009. *Funktion und Bedeutung von werden + Infinitiv im Vergleich zum futurischen Präsens*. Heidelberg: Winter.
- Hilpert, Martin. 2008. *Germanic Future Constructions: A Usage-Based Approach to Language Change*. (Constructional Approaches to Language 7). Amsterdam, Philadelphia: John Benjamins.
- Klein, Thomas & Stefanie Dipper. 2016. *Handbuch zum Referenzkorpus Mittelhochdeutsch*. (Bochumer Linguistische Arbeitsberichte 19). Bochum: Ruhr-Universität Bochum. <https://www.linguistics.ruhr-uni-bochum.de/rem/>.
- Krämer, Sabine. 2005. *Synchrone Analyse als Fenster zur Diachronie: Die Grammatikalisierung von werden + Infinitiv*. München: LINCOM.
- Leiss, Elisabeth. 1992. *Die Verbalkategorien des Deutschen. Ein Beitrag zur Theorie der sprachlichen Kategorisierung*. Berlin, New York: De Gruyter.
- Schmid, Hans-Ulrich. 2000. Die Ausbildung des werden-Futurs. Überlegungen auf der Grundlage mittelalterlicher Endzeitprophezeiungen. *Zeitschrift für Dialektologie und Linguistik* 67(1). 6–27.
- Solms, Hans-Joachim & Klaus-Peter Wegera. 1998. Das Bonner Frühneuhochdeutsch-Korpus. Rückblick und Perspektiven. In Rolf Bergmann (ed.), *Probleme der Textauswahl für einen elektronischen Thesaurus. Beiträge zum ersten Göttinger Arbeitsgespräch zur historischen deutschen Wortforschung*, 22–39. Stuttgart: Hirzel.