

Abs-255

Cedric Krummes (Bangor University), Sylvia Jaworska (Queen Mary, University of London), and Astrid Ensslin (Bangor University)

The use of discourse-structuring sequences by advanced learners of German: Corpus-driven investigations

This poster investigates discourse-structuring formulaic sequences produced by advanced learners of German and native speakers of German (cf. Wray 2002). The analysis is based on data collected for the 3-year AHRC/DFG project 'What's Hard in German?' (WHiG), which has been carried out by Bangor University and Humboldt-Universität zu Berlin since July 2009. Both data sets are part of the German error-annotated learner (parent) corpus FALKO (see Lüdeling 2008, Lüdeling et al. 2008, and Zeldes et al. 2008).

Following Stubbs (2002), our approach uses the *data* (and not a theoretical framework) as a starting point ('corpus-driven') in order to investigate the frequency and the use of formulaic sequences (also referred to as lexical bundles, or clusters). We aim to examine (1) the degree to which learners of German produce formulaic sequences compared to their native-speaker counterparts, (2) the types of sequences they over- and underuse compared to native speakers, especially with regard to discourse, reference and stance (cf. Chen and Baker 2010), and (3) the extent to which learners modify canonical structures, for instance, but not exclusively by L1-interference.

An in-depth understanding of which sequences are attested in the L1 corpus, but absent or modified (non-canonical) in the L2 corpus, will provide an impactful contribution to better teaching and learning materials for advanced learners of German.

References

Chen, Y.H. and Baker, P. (2010) 'Lexical Bundles in L1 and L2 Academic Writing', *Language Learning & Technology* 14 (2): 30–49.

Lüdeling, Anke. (2008). "Mehrdeutigkeiten und Kategorisierung. Probleme bei der Annotation von Lernerkorpora". In: Walter, Maik/Grommes, Patrick (eds.), *Fortgeschrittene Lernervarietäten: Korpuslinguistik und Zweitspracherwerbsforschung*. Tübingen: Niemeyer, 119-140.

Lüdeling, Anke/Doolittle, Seanna/Hirschmann, Hagen/Schmidt, Karin/Walter, Maik. (2008). "Das Lernerkorpus Falko". *Deutsch als Fremdsprache* (2), 67-73.

Stubbs, Michael. (2002). "Two quantitative methods of studying phraseology in English". In: *International Journal of Corpus Linguistics*, 7 (2), 215-244.

Zeldes, Amir, Lüdeling, Anke and Hagen Hirschmann. (2008). "What's Hard? Quantitative Evidence for Difficult Constructions in German Learner Data". 3rd Workshop on "Quantitative Investigations in Theoretical Linguistics 3" (QITL-3). Helsinki, Finland, 2-4 June 2008. Available from http://www.ling.helsinki.fi/sky/tapahtumat/qitl/Abstracts/Zeldes_et_al.pdf (accessed 26 November 2010).

Wray, Alison. (2002). *Formulaic Language and the Lexicon*. Cambridge: CUP.