
Die Normalisierung von Komposita in fnhd. Texten am Beispiel des RIDGES-Korpus

Mittwoch
08.03.2017
15:45 – 16:30
B4 1, Foyer

Anke Lüdeling, Carolin Odebrecht, Laura Perlitz, Gohar Schnelle
(Humboldt-Universität zu Berlin, Institut für deutsche Sprache und Linguistik,
Lehrstuhl für Korpuslinguistik und Morphologie)

Anke.Luedeling;odebreca;perlitz;kodshaj@hu-berlin.de

Wie lassen sich Komposita im Frühneuhochdeutschen variationistisch untersuchen? In dieser Sprachstufe existiert keine graphematische Norm, die die Getrennt- und Zusammenschreibung von Komposita regelt. Demzufolge ist ihre zweifelsfreie Identifikation nicht möglich, da konzeptuelle Abgrenzungsprobleme gegenüber Syntagmen wie z.B. Nominalphrasen mit Adjektivattributen (leinen Tu^echlein vs. Leinentüchlein) oder vorangestellten Genitivattributen (Widers Teuffels Biffz vs. Teuffelsbiss) bestehen (vgl. Pavlov 1983, 105). Um Komposita in einem Korpus systematisch suchen und analysieren zu können, ist daher eine Normalisierung notwendig. Dies zeigen wir anhand von RIDGES, einem diachrones Kräuterkundekorpus. Auf folgenden linguistischen Ebenen wird normalisiert (vgl. Belz et al. 2016): graphematisch (Teuffels Biffz → Teuffelsbiss); phonologisch (wieh=rauch → Weihrauch); morphologisch (Kolfewer → Kohlenfeuer). Bei ambigen Fällen greift das Prinzip der Konservativität (auff das Stro das leinen tu^echlein → das leinene Tüchlein). Weitere Annotationen helfen, verschiedene Forschungsfragen über Komposita zu untersuchen, z.B. die Getrennt- und Zusammenschreibung (vgl. Perlitz 2014, 31-32).

References: • Belz et al. (2016): *Annotationsrichtlinien zu Ridges Herbology Version 5.0*, HU Berlin. [https:// hu.berlin/ridges_annotationsrichtlinien_v5](https://hu.berlin/ridges_annotationsrichtlinien_v5). • Lüdeling, A.; Odebrecht, C.; Zeldes, A.: *RIDGES-Herbology* (Version 5.0), Humboldt-Universität zu Berlin. <https://hu.berlin/ridges>. • Odebrecht, C., Belz, M., Zeldes, A., Lüdeling, A., Krause, T. (eingereicht): *RIDGES Herbology - Designing a Diachronic Multi-Layer Corpus*. • Pavlov, V. M. (1983): *Zur Ausbildung der Norm der deutschen Literatursprache (1470- 1730)*. Berlin: Akademie-Verlag. • Perlitz, L. (2014): *Konkurrenz zwischen Wortbildung und Syntax - historische Entwicklung von Benennung* (Bachelorarbeit, Humboldt-Universität zu Berlin). <http://edoc.hu-berlin.de/master/perlitz-laura-2014-08-08/PDF/perlitz.pdf>.

CL Poster