

Von Korpus zu Korpus: Herausforderungen und Chancen diachron korpusübergreifenden Arbeitens

Kristin Kopf (Universität Münster/Institut für Deutsche Sprache Mannheim)

Wer eine diachrone Fragestellung über mehrere Jahrhunderte hinweg empirisch verfolgen möchte, ist für das Deutsche meist darauf angewiesen, verschiedene Korpora miteinander zu kombinieren. Dass überhaupt Korpora zugänglich sind, die aneinandergeknüpft vom 8. Jh. bis heute reichen, ist eine Entwicklung der letzten Jahre – lange Zeit war das Bonner Frühneuhochdeutschkorpus das einzige öffentlich zugängliche historische Korpus des Deutschen. Durch die größere Auswahl an diachron angelegten Korpora (s. Abb. 1) ergeben sich aber auch neue Herausforderungen. Eine der wohl wichtigsten Fragen ist die, ob sprachliche Phänomene über Korpusgrenzen hinweg konsistent untersuchbar sind und was dabei beachtet werden muss. Besonders ergiebig sind hier Korpora mit zeitlicher Überschneidung. Daten von Kempf (2016) zur adjektivischen Wortbildung zeigen z.B. bei Bonner Fnhd.-Korpus und GerManC für 1650–1700 gut vergleichbare Anteile von Hapax legomena und Neubelegen. Kopf (2017) betrachtet Anteile von Bindestrichschreibung in Komposita und stellt im klaren Kurvenverlauf keine Brüche zwischen GerManC und dem (noch nicht öffentlichen) Mainzer (Früh-)Neuhochdeutschkorpus fest. Kopf (eingereicht) konstatiert bei der Doppelmajuskelschreibung <Gott> eine konsistente Entwicklung von Mainzer (Früh-)Neuhochdeutschkorpus zu einem DTA-Ausschnitt. Doch welche Aussagekraft haben solche punktuellen Befunde? Für welche sprachlichen Bereiche lässt sich daraus möglicherweise Vergleichbarkeit ableiten? Wie verhält es sich mit Fällen, in denen mit der Korpusgrenze ein deutlicher Bruch sichtbar wird? Worauf lässt sich das zurückführen? Damit verbunden sind letztlich immer auch Fragen nach der Textsorten- und Dialektzusammensetzung der Korpora und der Sensibilität einzelner Phänomene dafür. Im Vortrag wird eine Serie von Einzeluntersuchungen präsentiert, die zeigen, wo und wie korpusübergreifendes Arbeiten mit diachronen Korpora des Deutschen möglich und sinnvoll ist.

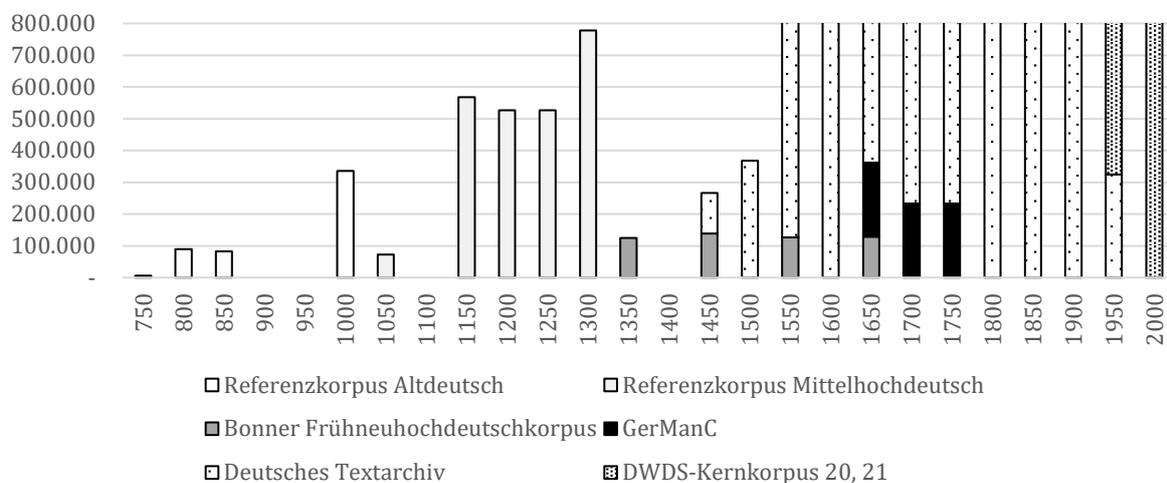


Abb. 1: Diachrone Korpora des Deutschen nach Tokens in Fünfzigjahresschritten (DTA und DWDS aus Skalierungsgründen abgeschnitten).

Quellen

Kempf, Luise. 2016. Adjektivsuffixe in Konkurrenz. Wortbildungswandel vom Frühneuhochdeutschen zum Neuhochdeutschen (Studia linguistica Germanica). Berlin, New York: de Gruyter.

Kopf, Kristin. eingereicht. Was ist so besonders an Gott? Ein grammatischer Abweichler im Frühneuhochdeutschen. In Damaris Nübling, Mirjam Schmuck & Luise Kempf (eds.), Linguistik der Eigennamen.

Kopf, Kristin. 2017. Fugenelement und Bindestrich in der Compositions-Fuge. Zur Herausbildung phonologischer und graphematischer Grenzmarkierungen in (früh)neuhochdeutschen N+N-Komposita.

In Renata Szczepaniak, Nanna Fuhrhop & Karsten Schmidt (eds.), *Sichtbare und hörbare Morphologie*, 177–204. Berlin, New York: de Gruyter.