Areal variation in Middle High German: Methodological and quantitative aspects

Using a large, electronically available corpus of Middle High German deeds (originally compiled by Friedrich Wilhelm; cf. Wilhelm 1932–2004), we demonstrate that this text genre offers fine-scaled information on areal variation in the language period in question. Unlike most other historical documents, deeds can usually be located exactly on the basis of textual evidence. We discuss and evaluate two methods for retrieving such information, i.e. (1) regular expressions (on the basis of spelling variants extracted from the *Wörterbuch der mittelhochdeutschen Urkundensprache* (WMU) 'dictionary of Middle High German language of deeds' (Kirschstein et al. 1994–2010) and (2) queries based on semi-automated annotations (with the aid of Schmid's 2019 *RNNTagger*). Patterns of areal variation are indispensable for gaining a deeper understanding of fine-scaled grammatical variation (Weinreich 1954; Bresnan et al. 2007) or the diffusion of linguistic innovations (cf. Nerbonne 2010), in particular the layering nature of grammaticalization processes that manifest in competing constructional variants (Hopper 1991; Girnth 2000). What is more, they can be useful for localizing manuscripts whose provenience is unknown and/or controversial (cf. Fleischer 2019 who demonstrates that already the 100 most frequent word forms per manuscript allow for a general assessment).

We focus on the following two phenomena:

1. So-called "contraction forms", i.e. cases of shortening that occur with several verbs in this language period, e.g. *lâzen/lân* 'let', *haben/hân* 'have', etc. and are assumed to show a clear areal profile (Paul et al. 2007: 280–284; Klein et al. 2018: 948–949), which is also confirmed by our data (cf. map 1 on *haben*[1]). Such forms cannot be solely explained by phonological deletion processes but are also sensitive to morphological factors like e.g. paradigmatic position and even syntactic effects (Nübling 1995). As cases of (weak) suppletion, they are also of a more general interest for morphological theory and typology (Nübling 2000; Corbett 2007).
2. Causative diathesis: In Middle High German, a contrast can be observed between intransitive *hangen* 'hang' (strong verb) – occasionally also in its contracted form, i.e. *hâhen* – and transitive/causative *hengen/henken* 'hang' (weak verb). The latter forms stem from *jan*-derivations (with *i*-umlaut) (Paul 2007: 157, §L103, Anm. 2). Our data show a clear areal contrast between these verbs, with *hengen/henken* being mostly confined to the west, while *hangen* can take up both transitive and intransitive functions in the eastern and west central dialects (cf. map 2 on hangen/hengen[2]). These findings are very informative with regard to the diffusion of umlaut in German and its functionalization as a diathesis marker. In the further course of language history, *hängen* develops into a highly heteroclitic verb (Stump 2006; Rißleben 1931).

Both phenomena are attested in modern dialects of German (which are usually well-documented in regional atlases, grammatical descriptions, or dictionaries) so that it is possible to compare and relate our findings to these data.

References

Bresnan, Joan, Ashwini Deo und Devyani Sharma (2007): Typology in variation: a probabilistic approach to *be* and *n't* in the Survey of English Dialects. *English Language and Linguistics* 11(2), 301–346.

---

1 Link: https://www.dropbox.com/s/jgqnwdzs6el9pcq/haben_h%C3%A2n_contracted_south.pdf?dl=0
2 Link: https://www.dropbox.com/s/ttrszymvvb9xk7i/2020-03-06_hahen_Formen.png?dl=0

Corbett, Greville (2007): Canonical Typology, Suppletion, and Possible Words. *Language* 83(1), 8–42.

Fleischer, Jürg (2019): Zur sprachlichen Einordnung mittelhochdeutscher Zeugnisse. Grundsätzliche und quantitative Explorationen anhand der „Kaiserchronik"-Überlieferung. *Zeitschrift für deutsches Altertum und deutsche Literatur* 148(2), 258–284.

Girnth, Heiko (2000): Untersuchungen zur Theorie der Grammatikalisierung am Beispiel des Westmitteldeutschen. (Reihe Germanistische Linguistik; 223). Tübingen: Niemeyer.

Hopper, Paul J. (1991): On some principles of grammaticalization. In: Elisabeth C. Traugott and Bernd Heine (eds.): Approaches to grammaticalization. Vol. 1: Focus on theoretical and methodological issues. (Typological Studies in Language; 19.1). Amsterdam, Philadelphia: Benjamins, 17–35.

Kirschstein, Bettina [et al.] (eds.) (1994–2010): Wörterbuch der mittelhochdeutschen Urkundensprache; auf Grundlage des Corpus der altdeutschen Originalurkunden bis zum Jahr 1300. Berlin: Erich Schmidt Verlag [3 Vols.].

Klein, Thomas, Hans-Joachim Solms und Klaus-Peter Wegera (2018): Mittelhochdeutsche Grammatik. Vol. 2: Flexionsmorphologie. Berlin, Boston: De Gruyter.

Nerbonne, John (2010): Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 3821–3828.

Nübling, Damaris (1995): Kurzverben in germanischen Sprachen: Unterschiedliche Wege – gleiche Ziele. *Zeitschrift für Dialektologie und Linguistik* 62(2), 127–154.

— (2000): Prinzipien der Irregularisierung: eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen. (Linguistische Arbeiten; 415). Tübingen: Niemeyer.

Paul, Hermann [et al.] (2007): Mittelhochdeutsche Grammatik. (Sammlung kurzer Grammatiken germanischer Dialekte. A, Hauptreihe; 2). Tübingen: Niemeyer. 25th ed.

Stump, Gregory T. (2006): Heteroclisis and Paradigm Linkage. *Language* 82(2), 279–322.

Rißleben, Fritz (1931): Die Geschichte der Verbgruppe hâhan-hangên-hengen-henken. Dissertation, University of Greifswald.

Schmid, Helmut (2019): Deep Learning-Based Morphological Taggers and Lemmatizers for Annotating Historical Texts. *DATeCH2019 – Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage,* 133–137.

Weinreich, Uriel (1954): Is a structural dialectology possible? *Word* 10, 388–400.

Wilhelm, Friedrich [et al.] (eds.) (1932–2004): Corpus der altdeutschen Originalurkunden bis zum Jahr 1300. Lahr, Berlin: Erich Schmidt Verlag [5 Vols.].