
Linear distribution of register features in the context of Early New High German register and genre research

Jürg Fleischer, Lena Haden, Gohar Schnelle & Lars Erik Zeige

Humboldt-Universität zu Berlin

[juerg.fleischer, lena.haden.1, gohar.schnelle, lars.zeige]@hu-berlin.de

The topic of the workshop is pursued under the specific conditions of historical research, where data are limited and knowledge of cultural contexts and practices is not based on experience or intuition (Schnelle 2025). We will build on results of previous studies in which text-based classification of linguistic and functional features led to unexpected groupings, and explore whether a dynamic view of unfolding text structures can lead to a better understanding of observed differences.

Based on the RegiL corpus of Early New High German texts by Martin Luther (v0.1: <https://doi.org/10.5281/zenodo.14702151>), Schnelle (2025) identified a set of seven register categories using common register-related features (see Table 1, 2nd column). The text units of RegiL are also annotated for genre. A comparison of registers and genres revealed both communalities and differences between both categories (see Table 1, 3rd column).

Reg.	Main characteristics	Genres
#1	tense symmetry in indicative mode	NT, fables, hymns
#2	attributive adjectives, subordination, auxiliaries, low lexical diversity	missives, sermons, treatises, prefaces
#3	dominant reference to the present	fables, hymns
#4	dominant reference to the present, modals	hymns, prefaces
#5	present, no past tense, high lexical diversity	hymns, prefaces
#6	subjunctive, few indicatives, subordination	hymns
#7	subjunctive, few modals, subordination, lexical verbs	hymns

Table 1. Register categories established for RegiL v. 0.1

Focusing on these differences and on atypical texts, we will examine if the characteristics of our registers benefit from an integration of the linear distribution of features. In doing so, we will take into account both the sequence and the relative distance or even the clustering of features. Possible patterns in linear distributions will be tied back to the concept of genre. For instance, linear distributions can help to improve cross-genre mapping by, for example, disentangling treatises and sermons (register #2) or by understanding the differences between various prefaces or hymns.

References: • Schnelle, G. (2025). Register als variationsgrammatische Kategorie. Korpusbasierte Registermodellierung in den Werken Otrfrids von Weißenburg und Martin Luthers. In L. Schnee, H. Huber & S. Hartmann (eds.), *Historische Grammatik(en)*. Berlin: De Gruyter.