

Background & Motivation

- traditionally, generic masculines in German are assumed to be gender-neutral [cf. 1]
- however, psycholinguistic research has shown that generic masculines are comprehended as biased towards male referents [e.g. 2, 3]
- recently, computational methods resulted in findings in line with psycholinguistic studies [4, 5]
- however, the computational implementations come with two major issues
- issue 1:** the semantic vectors of the MASCULINE and GENERIC were strongly correlated
- issue 2:** genericity, i.e. GENERIC VS. SPECIFIC, was treated as an inflectional feature, which it is not
- aim:** solve these computational issues

Instance vectors as a window to (non-)generic role noun semantics

Methods & Analysis

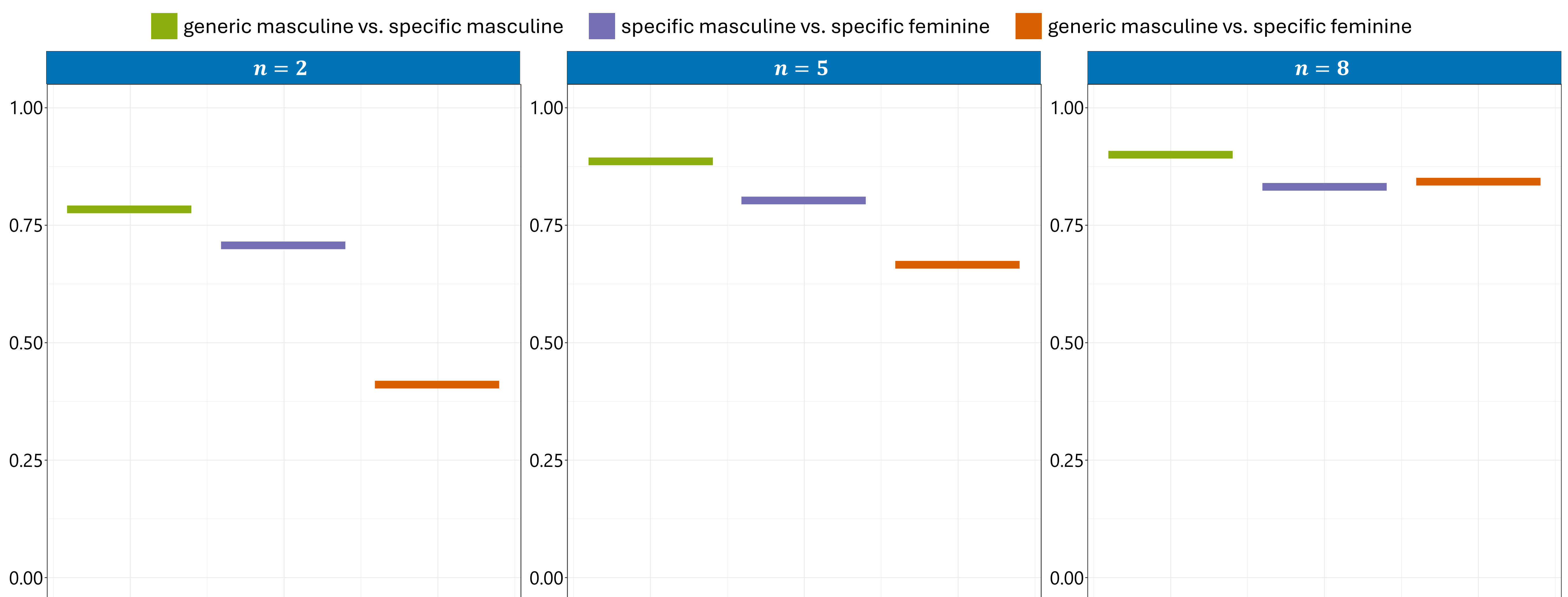
Instance Vectors

- the mean vector of n content word vectors preceding & following a target word token [6] with attestations and contexts taken from [4, 5]
- computed with $n = 2$, $n = 5$, and $n = 8$ to check for influence of context window size
- computed based on semantic vectors generated by fastText [7]
- cosine similarity as measure of semantic similarity between vectors of target words

Analysis

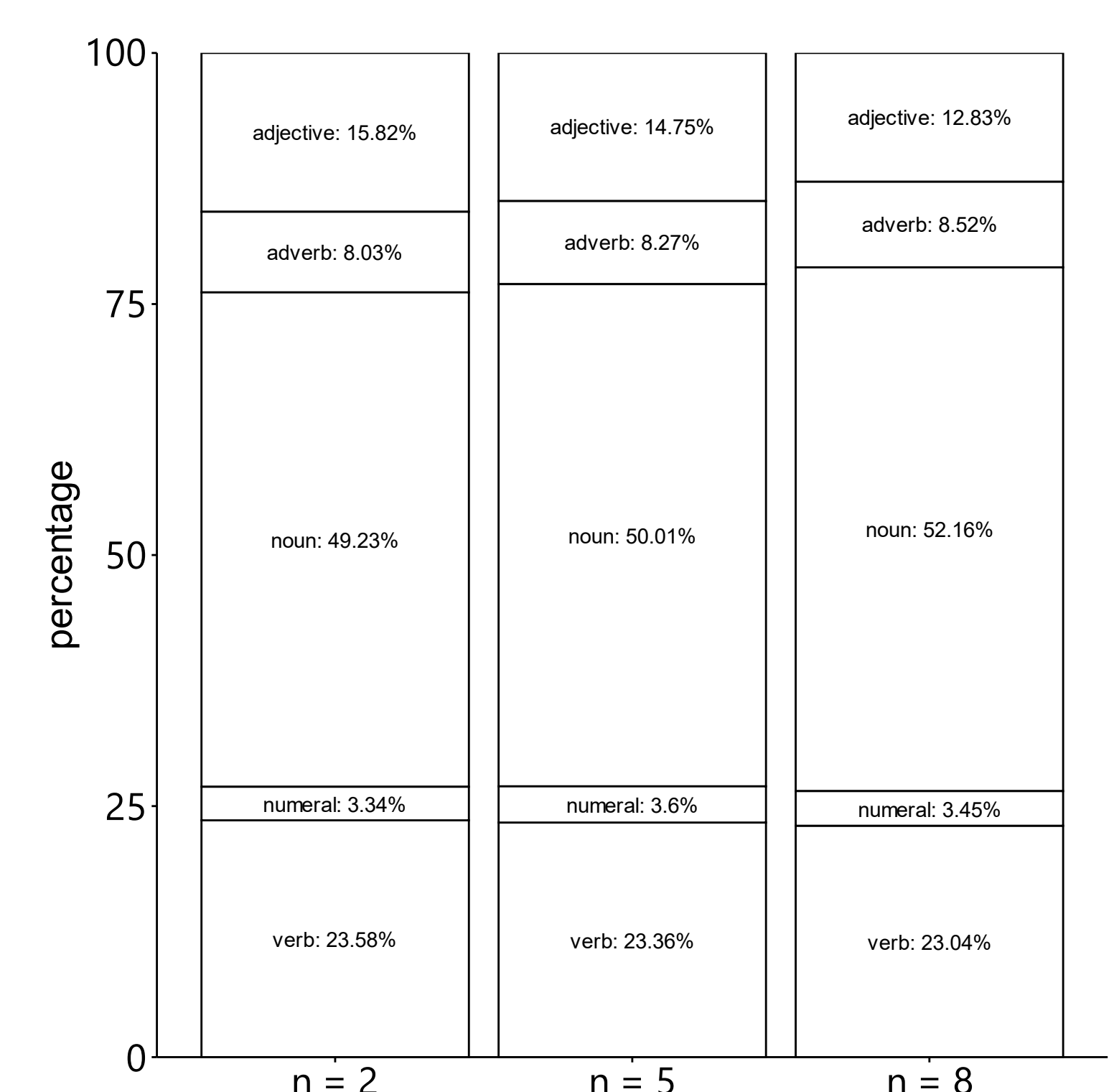
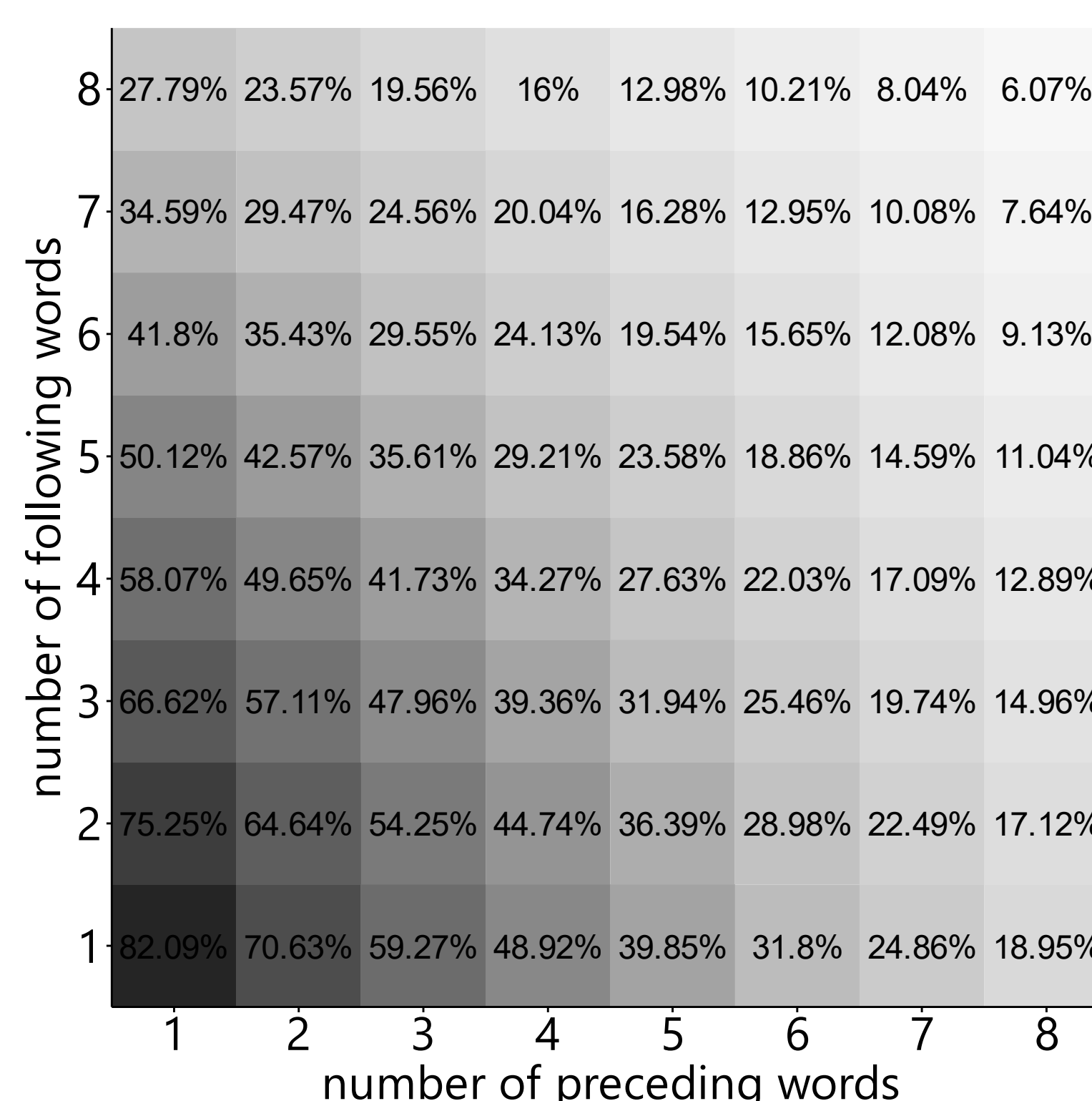
- beta regression in generalised additive mixed models predicting cosine similarity [8]
- predictor of interest is the COMPARISON that belongs to a given cosine similarity value (other variables: number, frequency, stereotypicality)

Results



Discussion

- generic masculines are semantically closer to specific masculines than to specific feminines
 - male bias
- findings in line with previous psycholinguistic [2, 3] and computational research [4, 5]
- the male bias in generic masculines in German is stable across a variety of linguistic methods
- computational methods are a meaningful tool for research on semantic genericity and gender-neutrality



References

[1] Doleschal, U. (2002). Das generische Maskulinum im Deutschen. Ein historischer Spaziergang durch die deutsche Grammatikschreibung von der Renaissance bis zur Postmoderne. *Linguistik Online*, 11(2). [2] Gyga, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23(3), 464–485. [3] Schunack, S., & Binazer, A. (2022). Revisiting gender-fair language and stereotypes - A comparison of word pairs, capital I forms and the asterisk. *Zeitschrift für Sprachwissenschaft*, 41(2), 309–337. [4] Schmitz, D. (2023). In German, all professors are male. In J. Pfeifer, S. Arndt-Lappe, H. Dörgele, G. Kunter, & C. Uffmann (Eds.), *INGO 6.0. The Proceedings. New empirical insights about language, presented on a Great day Out in September*. Preprint. [5] Schmitz, D., Schneider, V., & Esser, J. (2023). No genericity in sight: An exploration of the semantics of masculine generics in German. *Glossa Psycholinguistics*, 2(1). [6] Lapesa, G., Kawalec, L., Plag, I., Andreou, M., Kisselew, M., & Padó, S. (2018). Disambiguation of newly derived nominalizations in context: A Distributional Semantics approach. *Word Structure*, 11(3), 277–312. [7] Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. [8] Wood, S. N. (2017). *Generalized additive models: An introduction with R*. CRC Press.