

Evidence for a non-generic masculine generic in German

Dominic Schmitz¹, Viktoria Schneider¹, & Janina Esser²

¹Heinrich-Heine-Universität Düsseldorf, ²Universität zu Köln

Linguistic research has repeatedly demonstrated that masculine generics in German show a masculine bias (e.g. Gabriel et al., 2008; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001). That is, grammatically masculine role-nouns such as *Anwalt* ‘lawyer’ can refer to men and women but may favour an interpretation in which mostly or only men are considered as potential referents (e.g. Misersky et al., 2019). However, research on this matter faces two major issues.

First, previous studies have used numerous of such masculine generics and their feminine counterparts to gain insights into their semantics without accounting for language external but potentially confounding influences such as stereotypicality. Second, the majority of studies finds the aforementioned masculine bias; however, very few studies offer a theoretical account of its underlying representations (e.g. Irmen & Linner, 2005). To this date, no investigation has been made to discover whether there is a connection between the male bias and the representation of masculine generics in the mental lexicon.

The present paper offers a solution to both issues. First, the role nouns under investigation are those for which stereotypicality ratings are available (Gabriel et al., 2008). Language external factors are incorporated in the analysis by such ratings. Second, using the general ideas of distributional semantics (Harris, 1954) as well as naïve discriminative learning (e.g. Baayen & Ramscar, 2015) and linear discriminative learning (e.g. Baayen et al., 2019) the underlying nature of masculine generics and counterparts is investigated. The proposed analysis aims at exploring the following question: How semantically similar are masculine generics and their explicitly masculine and feminine counterparts when taking into account stereotypicality?

The following method was employed to tackle this question. An 830,000 sentence (1.7 million words) corpus of contemporary German was created using the Leipzig Corpora Collection (Goldhahn et al., 2012). The corpus included 113 target word pairs which were based on the set of words provided by Gabriel et al. (2008). All target word occurrences were checked for their usage, i.e. whether they were generically or explicitly intended, and annotated accordingly. The corpus was then used to train semantic vectors based on the Rescorla-Wagner equation (Wagner & Rescorla, 1972) as implemented by naïve discriminative learning. Finally, the resulting semantic vectors were then used to train an implementation of linear discriminative learning.

Taking a closer look at the semantic vectors as created by the naïve discriminative learning algorithm, one finds that for the singular generic and explicit masculine role nouns are semantically closest with a mean cosine similarity value of approx. 0.98. Explicit masculine and feminine role nouns are less similar (approx. 0.94), and masculine generic and explicit feminine are least similar (0.93). These differences are highly significant (Mann-Whitney-Wilcoxon, $p < 2.16e-16$) and even more pronounced in the plural. To account for potential influences of stereotypicality, measures derived from the linear discriminative learning implementation were modelled by the stereotypicality measure as provided by Gabriel et al. (2008). As a result one finds that generic and explicit masculines are highly similar in terms of their semantic activation diversity and semantic neighbourhood size, while explicit feminines show higher activation diversities and denser neighbourhoods.

Our results indicate that there is a male bias in masculine generics in German as they exhibit highly similar semantic vectors as well as highly similar levels of semantic activation diversity and semantic neighbourhood size, even when controlled for their stereotypicality. Consequently, generic and explicit masculines show similar underlying representations, while the representations of explicit feminines are less similar. Thus, even though the use of masculine generics might be intended as generic, their resonance with the lexicon, that is more specifically their similarity with explicit masculines, leads to an overall biased association towards male referents.

References

- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative Lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, 2019, 1–39. <https://doi.org/10.1155/2019/4895891>
- Baayen, R. H., & Ramscar, M. (2015). Abstraction, storage and naive discriminative learning. *Handbook of Cognitive Linguistics*, 39, 100–120. <https://doi.org/10.1515/9783110292022-006>
- Gabriel, U., Gygax, P., Sarrasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, 40(1), 206–212. <https://doi.org/10.3758/BRM.40.1.206>
- Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *Proceedings of the 8th International Language Resources and Evaluation (LREC'12)*.
- Gygax, 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. <https://doi.org/10.1080/01690960701702035>
- Harris, Z. S. (1954). Distributional Structure. *WORD*, 10(2–3), 146–162. <https://doi.org/10.1080/00437956.1954.11659520>
- Irmen, L., & Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology*, 57(5), 367–375. <https://doi.org/10.1027/1618-3169/a000044>
- Irmen, L., & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen. *Zeitschrift Für Psychologie / Journal of Psychology*, 213(3), 167–175. <https://doi.org/10.1026/0044-3409.213.3.167>
- Koch, M. (2021). *Kognitive Effekte des generischen Maskulinums und genderneutraler Alternativen im Deutschen – eine empirische Untersuchung*. Technische Universität Braunschweig.
- Misersky, J., Majid, A., & Snijders, T. M. (2019). Grammatical Gender in German Influences How Role-Nouns Are Interpreted: Evidence from ERPs. *Discourse Processes*, 56(8), 643–654. <https://doi.org/10.1080/0163853X.2018.1541382>
- Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen. *Psychologische Rundschau*, 52(3), 131–140. <https://doi.org/10.1026//0033-3042.52.3.131>
- Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–334). Academic Press Inc.