

Context doesn't matter: The male bias of generic masculines in German remains stable across different context windows

Background

- Traditionally, generic masculines in German are assumed to be gender-neutral [1]
- Evidence for a **male bias** in generic masculines has been brought forward by numerous studies applying a multitude of methodological approaches [2-5]
- ➔ Generically intended masculines are understood not as gender-neutral but as mainly associated with male individuals
- Recently, computational studies entered this field of research [6]
- Only one computational study has considered that context may play a direct role in disambiguating generic from specific masculines [7]

Aim 1: Improve this approach with an LLM-based solution

Aim 2: Explore the role of context in disambiguating specific vs generic masculines

(1) Specific masculine (SM)

*Dort wende ich mich an einen **Mitarbeiter**, bitte ihn um ein Beratungsgespräch.*

'I approach an **employee** there and ask for a consultation.'

(2) generic masculine (GM)

*Jeder **Mitarbeiter** bringt Sachen mit, die er nicht mehr braucht.*

'Every **employee** brings things that they no longer need.'

(3) specific feminine (SF)

*Die nette **Mitarbeiterin** im Reisebüro wusste auch nicht, was man tun könne.*

'The nice **employee** at the travel agency did not know what to do either.'

Method

Language materials

- Corpus of German press texts
 - DPA (Deutsche Presseagentur 'German Press Agency') and the three magazines *Brigitte*, *Zeit Wissen*, and *Psychologie Heute*
- Annotated with a pre-selected set of personal nouns [8]
 - Specific masculine, generic masculine, and specific feminine forms of *Bürgermeister* 'mayor', *Chef* 'boss', *Freund* 'friend', *Gastgeber* 'host', *Geschäftsführer* 'manager', *Künstler* 'artist', *Lehrer* 'teacher', *Leiter* 'leader', *Minister* 'minister', *Mitarbeiter* 'employee', *Pfarrer* 'vicar', *Präsident* 'president', *Richter* 'judge'
- Exclusion of items with too few attestations of at least one form, threshold = 60
 - Singular set: *Bürgermeister*, *Chef*, *Freund*, *Künstler*, *Lehrer*, *Minister*, *Mitarbeiter*, *Präsident*, *Richter*
 - Plural set: *Freund*, *Künstler*, *Mitarbeiter*

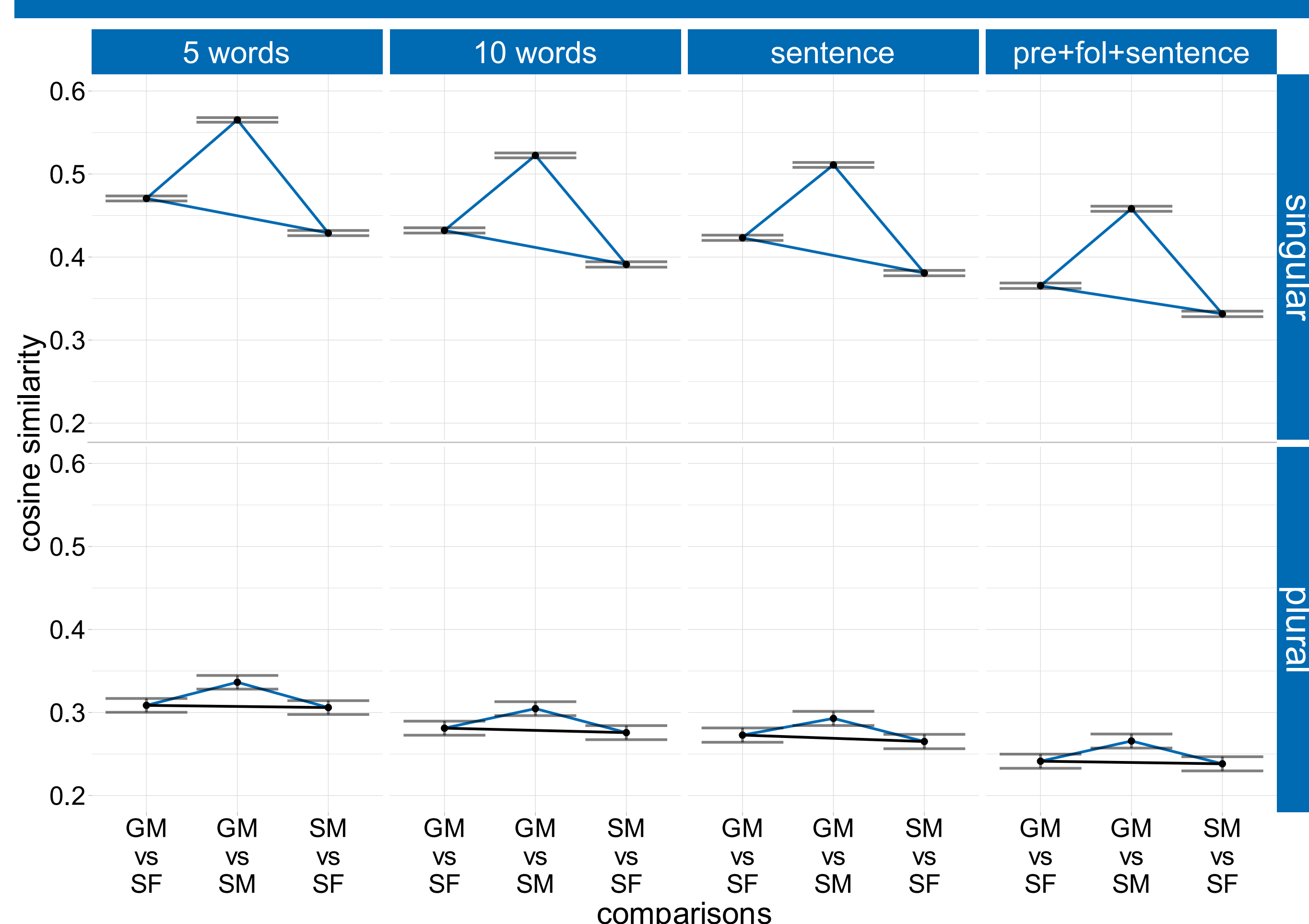
Contextualised embeddings

- Incorporate the context of a target word, producing token-level embeddings that reflect context-specific meanings
- We used the pre-trained bidirectional language model *bert-base-german-cased* (110 million parameters, 768 dimensions) [9]
- Context windows
 - 5 and 10 words before and after the target form
 - The entire sentence
 - The sentence plus the sentence before and after

Analysis of semantic similarity between forms

- Cosine similarity as the measure of choice
- Within target, within number, and within context window
- 10-fold cross-validation approach with n being the number of the comparison (GM vs SF, GM vs SM, SM vs SF) with the fewest available data points

Results



Discussion

- Including more context leads to overall lower semantic similarities, i.e. context does generally disambiguate
- Nonetheless, more context does not alleviate or even remove the male bias of generic masculines
- Singular: three-way pattern
 - Generic masculines and specific masculines are most similar
 - Followed by generic masculines and specific feminines
 - Specific masculines and specific feminines are least similar
- Plural: two-way pattern
 - Generic and specific masculines are most similar
 - Specific feminines are just as similar to generic masculines as they are to specific masculines
- The present findings are in line with previous computational and non-computational studies which found a male bias in generic masculines [2-7]

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). <https://doi.org/10.13092/lo.11.915> [2] Braun, F., Gottburgsen, A., Sczesny, S., & Stahlberg, D. (1998). Können Geophysiker Frauen sein? Generische Personenbezeichnungen im Deutschen. *Zeitschrift für Germanistische Linguistik*, 26(3), 265-283. <https://doi.org/10.1515/zfgl.1998.26.3.265> [3] 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> [4] 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> [5] Glim, S., Körner, A., & Rummer, R. (2024). Generic masculine role nouns interfere with the neural processing of female referents: Evidence from the P600. *Language, Cognition and Neuroscience*, 39(10), 1366-1375. <https://doi.org/10.1080/23273798.2024.2387230> [6] Schmitz, D. (2024). Instances of bias: The gendered semantics of generic masculines in German revealed by instance vectors. *Zeitschrift für Sprachwissenschaft*, 43(2). <https://doi.org/10.1515/zfs-2024-2010> [7] 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). <https://doi.org/10.5070/G6011192> [8] Müller-Spitzer, C., Ochs, S., Kopenig, A., Rüdiger, J. O., & Wolfer, S. (2024). Less than one percent of words would be affected by gender-inclusive language in German press texts. *Humanities and Social Sciences Communications*, 11(1), 1-13. <https://doi.org/10.1057/s41599-024-03769-w> [9] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 1, 4171-4186.