

Evidence for a non-generic masculine generic in German

Dominic Schmitz¹, Viktoria Schneider¹, Janina Esser²

¹Heinrich-Heine-Universität Düsseldorf

²Association for Diversity in Linguistics

Masculine Generics in German

- in German, role nouns such as *Anwalt* ‘lawyer’ can be used as generic forms

	word	referent gender(s)	grammatical gender	number
target word paradigm	Anwalt	male	masculine	singular
	Anwalt	male or female	masculine	
	Anwältin	female	feminine	
	Anwälte	male	masculine	plural
	Anwälte	male and/or female	masculine	
	Anwältinnen	female	feminine	

- generic forms are not different from explicit masculine forms in their orthographic or phonological form
- they are used to describe individuals of all genders in singular and plural contexts
- generic forms are traditionally assumed to “abstract away” notions of gender; to be “gender-neutral” (Doleschal, 2002)

Previous Research: Findings

- however, previous research has cast doubt on the gender-neutral use of masculine generics
- most (if not all) behavioural studies on the subject find one overall result
 - masculine generics are not gender-neutral but show a clear bias towards the explicit masculine reading (e.g. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001; Trutkowski, 2018)
- even though a masculine generic may be used by a speaker with the intention of considering all genders...
- ...this intention is not fully translated by the receiver's comprehension system
- instead, a reading favouring male individuals is received

Previous Research: Issues

Issue 1: Stereotypicality

Almost no previous research included effects of stereotypicality in their analyses on masculine generics.

Issue 2: Underlying Representations

No previous research investigated the underlying representations of masculine generics in order to account for their masculine bias.

Previous Research: Solutions

Issue 1: Stereotypicality

→ include stereotypicality ratings in analyses

Issue 2: Underlying Representations

→ use linear discriminative learning (e.g. Baayen et al., 2019) to explore semantics

Research Questions

Research Question 1

Is the bias of masculine generics affected by stereotypicality?

Research Question 2

Does linear discriminative learning offer an insight into the underlying nature of the masculine generic's bias?

Method

- we simulate an individual's comprehension by implementing a **linear discriminative learning network** (e.g. Baayen et al., 2019)
- for this, **cues** and **semantics** are required as starting points

Method

Cues

- corpus created based on Leipzig Corpora Collection's (Goldhahn et al., 2012) subcorpus "News"
 - with content lexomes and inflectional lexomes (49,044,960 lexome tokens)
 - 30,000 sentences with target word paradigm members
 - 800,000 sentences with further lexomes
- target words adopted from a study on stereotypicality of role nouns (Gabriel et al., 2008)
- triphones of target word paradigm members and content lexomes

Method

Semantics

- semantic vectors computed based on the 830,000 sentence corpus for content and inflectional lexomes with NDL (e.g. Baayen & Ramscar, 2015)
- originally of 30,887 x 15,023 dimensions, vectors were shortened to 7,511 dimensions based on overall column standard deviations
- semantic vectors of complex words consisted of their individual parts, e.g.

target form	base		number		gram. gender		type
Anwalt	Anwalt	+	singular	+	masculine	+	generic
Anwalt	Anwalt	+	singular	+	masculine	+	explicit
Anwältin	Anwalt	+	singular	+	feminine	+	explicit

Method

- we simulate an individual's comprehension by implementing a **linear discriminative learning network** (e.g. Baayen et al., 2019)

cue
matrix C

semantic
matrix S

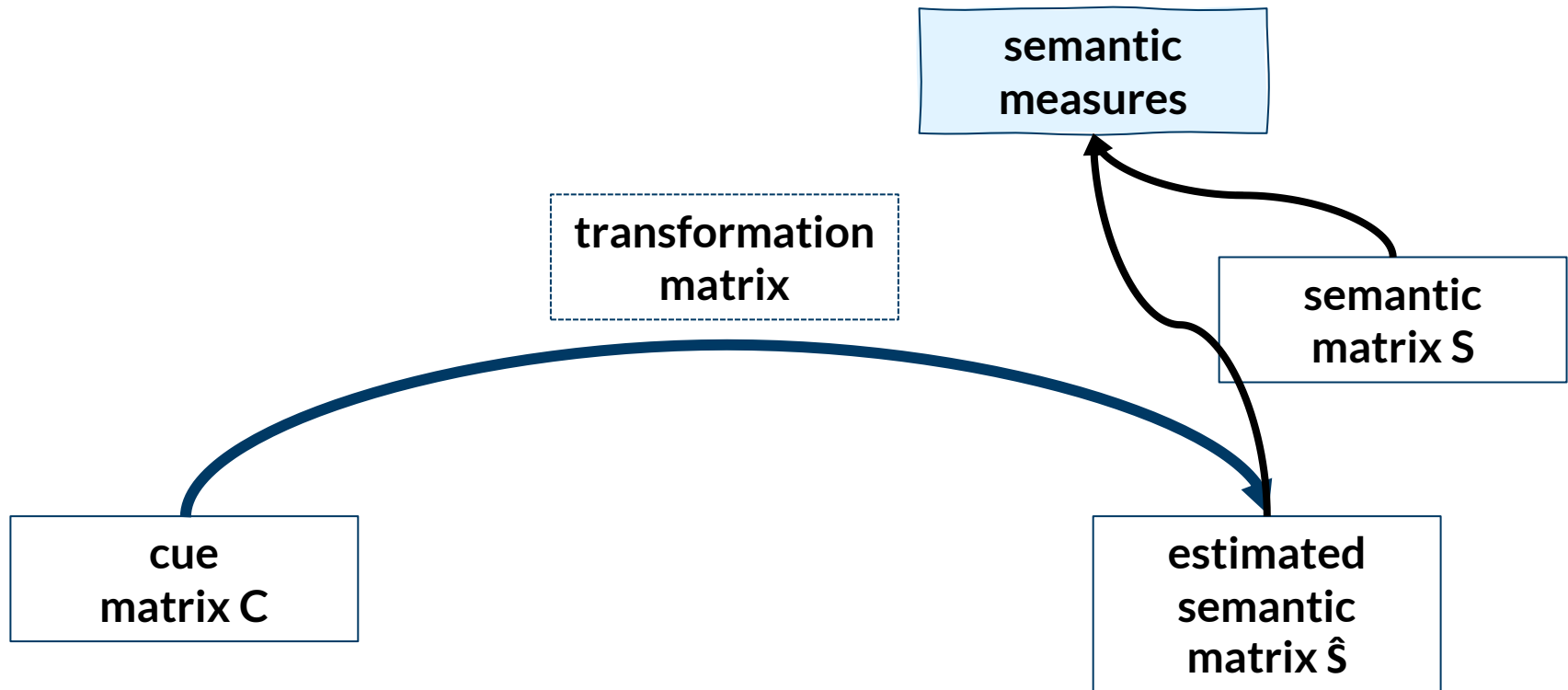
Method

- we simulate an individual's comprehension by implementing a **linear discriminative learning network** (e.g. Baayen et al., 2019)

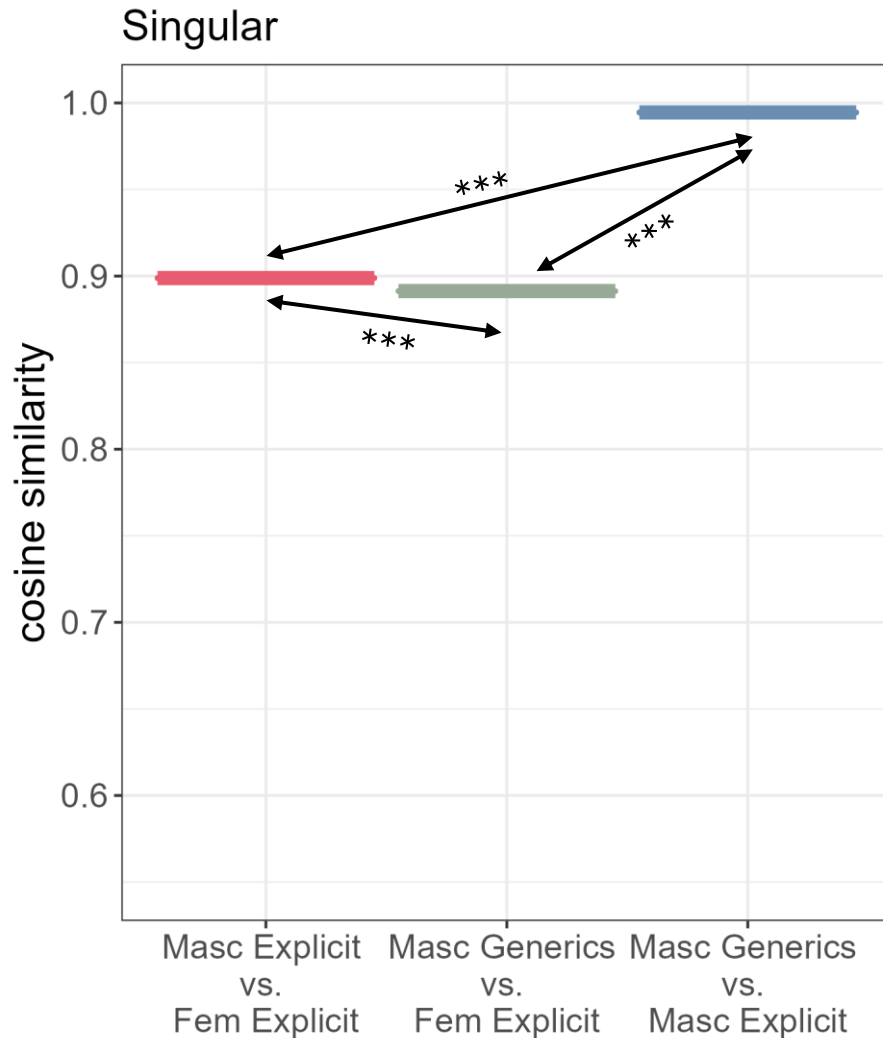


Method

- we simulate an individual's comprehension by implementing a **linear discriminative learning network** (e.g. Baayen et al., 2019)

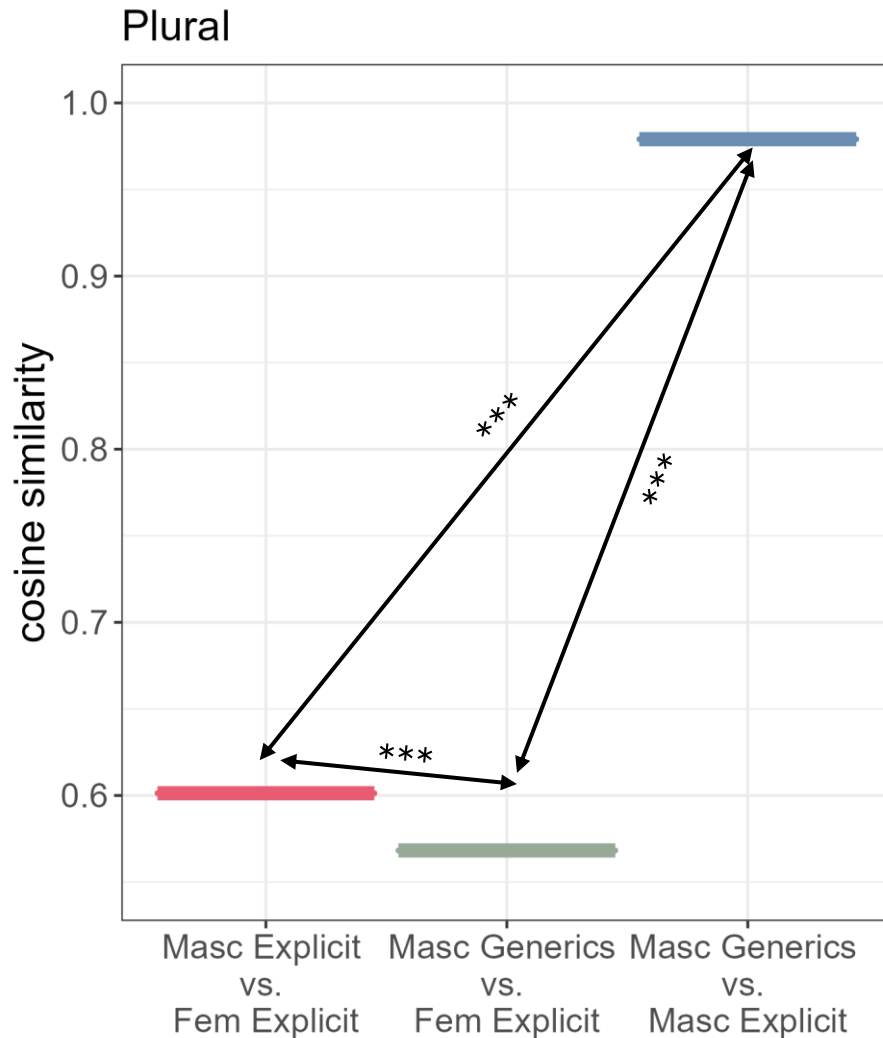


Bias Check



- masculine generics and the explicit masculine are semantically most similar
- the explicit feminine is more similar to the explicit masculine than to masculine generics
- all comparisons are highly significant

Bias Check



- masculine generics and the explicit masculine are semantically most similar
- the explicit feminine is more similar to the explicit masculine than to masculine generics
- all comparisons are highly significant
- differences are more pronounced

Analysis

- stereotypicality ratings are taken from Gabriel et al. (2008)
- measures derived from the LDL implementation are
 - total semantic similarity
correlation of \hat{s} and all other semantic vectors
 - comprehension quality
correlation of \hat{s} and s
 - semantic neighbourhood density
8 nearest neighbours
 - semantic activation diversity 1
Euclidian distance
 - semantic activation diversity 2
Manhattan distance

Analysis

- TYPE predicted by STEREOTYPICALITY ratings of target words and LDL measures via multinomial logistic regression
- TYPE: singular masculine generic; singular masculine explicit; singular feminine explicit
plural masculine generic; plural masculine explicit; plural feminine explicit
- as LDL measures are highly correlated with each other, they are first combined into two principal components
 - **PC1** total semantic similarity, comprehension quality, neighbourhood density
higher = higher similarity/quality/density
 - **PC2** activation diversity 1 & 2
higher = lower activation diversity

$$type \sim stereotypicality + PC1 + PC2$$

Results

singular masculine generics

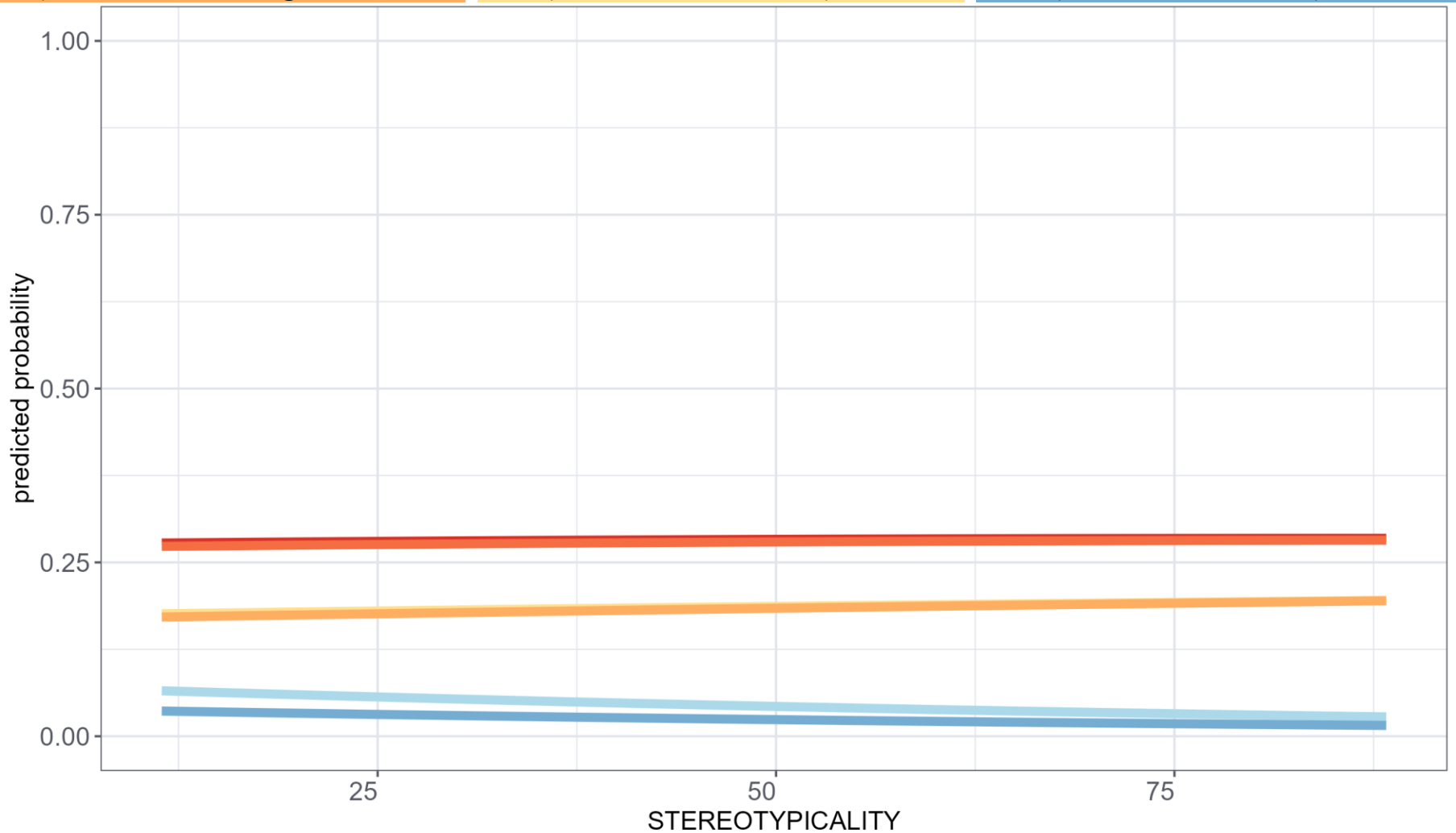
singular masculine explicit

singular feminine explicit

plural masculine generics

plural masculine explicit

plural feminine explicit



Results

singular masculine generics

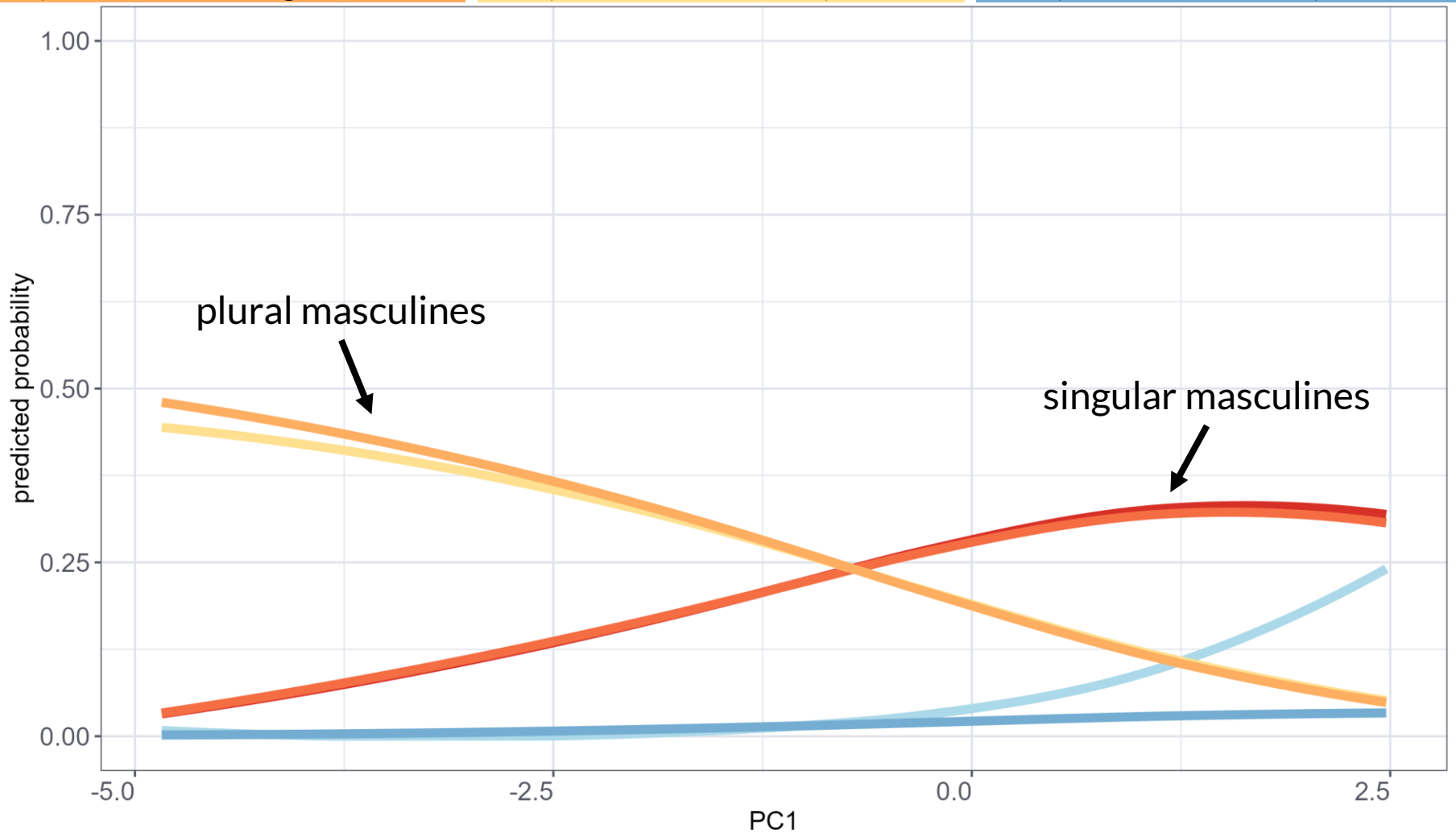
singular masculine explicit

singular feminine explicit

plural masculine generics

plural masculine explicit

plural feminine explicit



Results

singular masculine generics

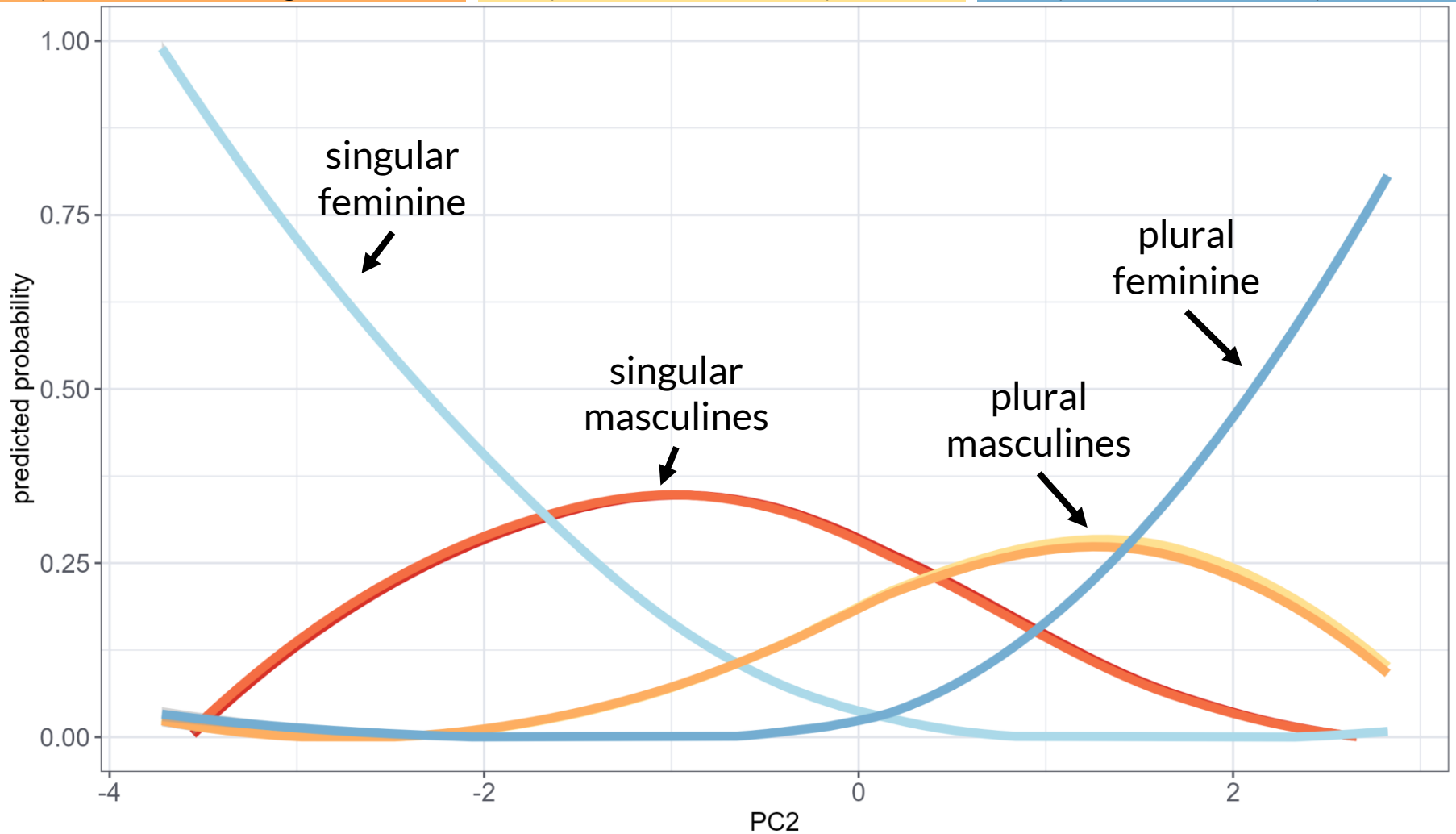
singular masculine explicit

singular feminine explicit

plural masculine generics

plural masculine explicit

plural feminine explicit



Results

- no effect of stereotypicality found
- significant effects found for
 - **PC1** opposite patterns for masculines by number; feminines inconclusive
total semantic similarity, comprehension quality, neighbourhood density
higher = higher similarity/quality/density
 - **PC2** higher = feminine singular; lower = feminine plural; masculines in-between
activation diversity 1 & 2
higher = lower activation diversity

Discussion

Research Question 1

Is the bias of masculine generics affected by stereotypicality?

→ no

Research Question 2

Does linear discriminative learning offer an insight into the underlying nature of the masculine generic's bias?

→ yes

Discussion

- our findings are in line with assumptions found in previous research
 - Stahlberg et al. (2001)
masculine gender of generics has a semantic component of “maleness”
 - Irmen & Linner (2005)
semantic similarity of masculine generics and explicit terms due to their resonance with the lexicon and each other
 - Gygax et al. (2012) and Gygax et al. (2021)
masculine generics activate the underlying representations of masculine explicit terms, leading to a semantic activation of masculine explicit terms, thus a male bias

Conclusion

- the male bias is due to the similar semantic features of the masculine generic and masculine explicit forms
- this leads to a ‘male bias’ in the language system itself
- thus, our findings confirm the bias found in previous behavioural studies (e.g. Demarmels, 2017; Garnham et al., 2012; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Irmen & Linner, 2005; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001; Trutkowski, 2018)
- future research will show
 - whether the LDL measures computed for our data are predictive of behavioural measures
 - how (new) neutral forms perform (e.g. *Anwält*innen*, *AnwältInnen*)

Thank you!

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>
- Demarmels, S. (2017). „Gesucht: Assistentin oder Sekretär der Geschäftsleitung“ – Gendersensitive Formulierungen in Stellenanzeigen aus der Perspektive der Textsorte. In *Stellenanzeigen als Instrument des Employer Branding in Europa*. https://doi.org/10.1007/978-3-658-12719-0_11
- 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>
- 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>
- Garnham, A., Gabriel, U., Sarrasin, O., Gygax, P., & Oakhill, J. (2012). Gender Representation in Different Languages and Grammatical Marking on Pronouns: When Beauticians, Musicians, and Mechanics Remain Men. *Discourse Processes*, 49(6), 481–500. <https://doi.org/10.1080/0163853X.2012.688184>
- 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>
- Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: a challenge for our cognitive system. *Language Sciences*, 83, 101328. <https://doi.org/10.1016/j.langsci.2020.101328>

References

- Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: a challenge for our cognitive system. *Language Sciences*, 83, 101328. <https://doi.org/10.1016/j.langsci.2020.101328>
- 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*. Master's Thesis. 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>
- Schmid, H. (1999). Improvements in part-of-speech tagging with an application to German. In S. Armstrong, K. Church, P. Isabelle, S. Manzi, E. Tzoukermann, & D. Yarowsky (Eds.), *Natural language processing using very large corpora* (pp. 13–25). Springer. https://doi.org/10.1007/978-94-017-2390-9_2
- 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>
- Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name Your Favorite Musician. *Journal of Language and Social Psychology*, 20(4), 464–469. <https://doi.org/10.1177/0261927X01020004004>
- Trutkowski, E. (2018). Wie generisch ist das generische Maskulinum? Über Genus und Sexus im Deutschen. *ZAS Papers in Linguistics*, 59, 83–96. <https://doi.org/10.21248/zaspil.59.2018.437>