

Generic masculines in German – Why all teachers are masculine

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 *Lehrer* ‘teacher’ can be used as generic forms

	word	referent gender(s)	grammatical gender	number
target word paradigm	<i>Lehrer</i>	male	masculine	singular
	<i>Lehrer</i>	male or female	masculine	
	<i>Lehrerin</i>	female	feminine	
	<i>Lehrer</i>	male	masculine	plural
	<i>Lehrer</i>	male and/or female	masculine	
	<i>Lehrerinnen</i>	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

- 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)
- 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

Issues of previous research

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.

Issues of previous research

Issue 1: Stereotypicality

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

→ include stereotypicality ratings in analyses

Issue 2: Underlying Representations

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

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

Research questions

Research Question 1

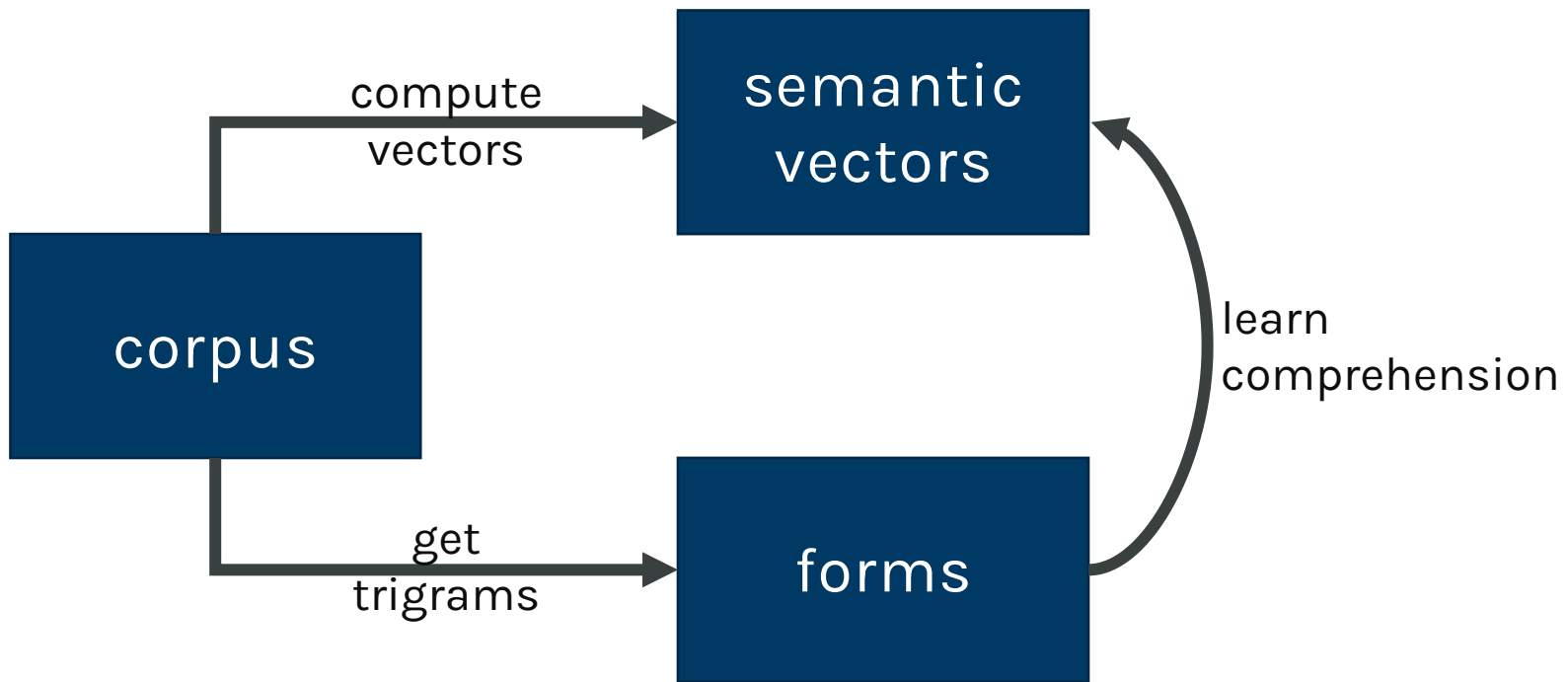
Is the bias of masculine generics affected by stereotypicality?

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 mental lexicon by implementing a linear discriminative learning network (e.g. Baayen et al., 2019)
- for this, semantics and word forms are required as starting points



Method

Corpus

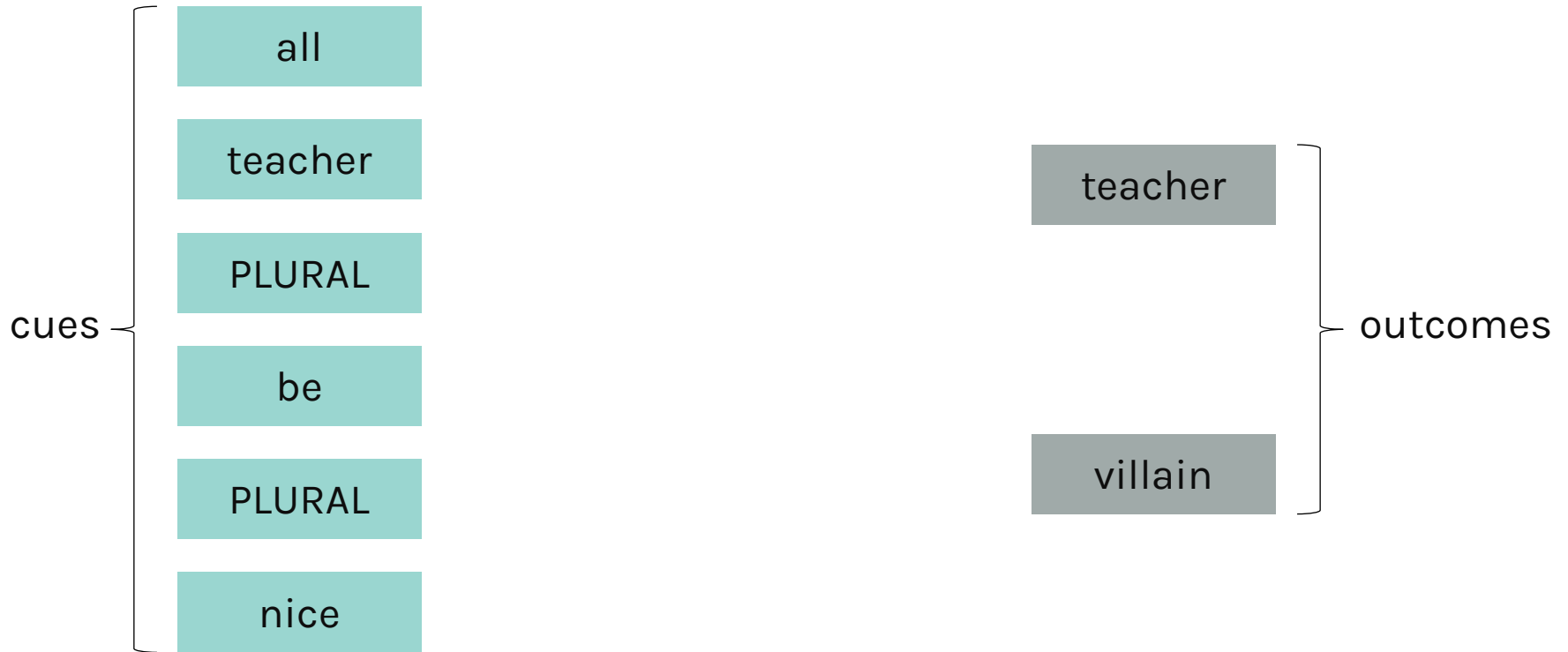
- corpus created based on Leipzig Corpora Collection's (Goldhahn et al., 2012) subcorpus "News"
 - 49,044,960 word form tokens
 - 30,000 sentences with target word paradigm members
 - 800,000 sentences with further word forms
- target words adopted from a study on stereotypicality of role nouns (Gabriel et al., 2008)
- 113 target word paradigms in total

Method

Semantic vectors

- semantic vectors computed based on the 830,000 sentence corpus for words and inflectional functions with Naive Discriminative Learning (NDL; e.g. Baayen & Ramscar, 2015)
 - semantic vectors for bases, function words, and inflection
- NDL follows the Rescorla-Wagner rules (Rescorla & Wagner, 1972)
- most importantly, these rules state that
 - outcomes (word forms) are predicted by cues (words/inflection)
 - the associative strength between an outcome and a cue is represented by a single number
- we used each sentence to predict each individual word within the sentence by the other words in that sentence

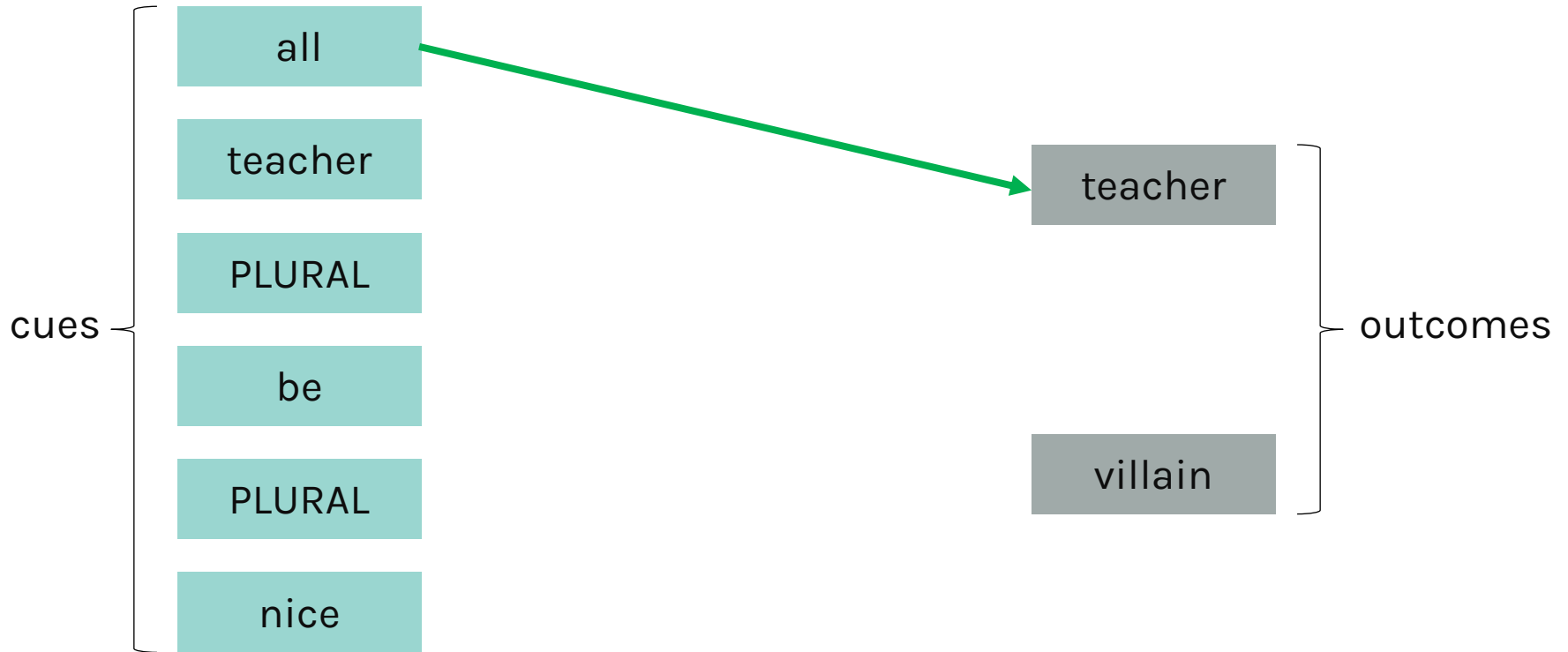
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher							
villain							

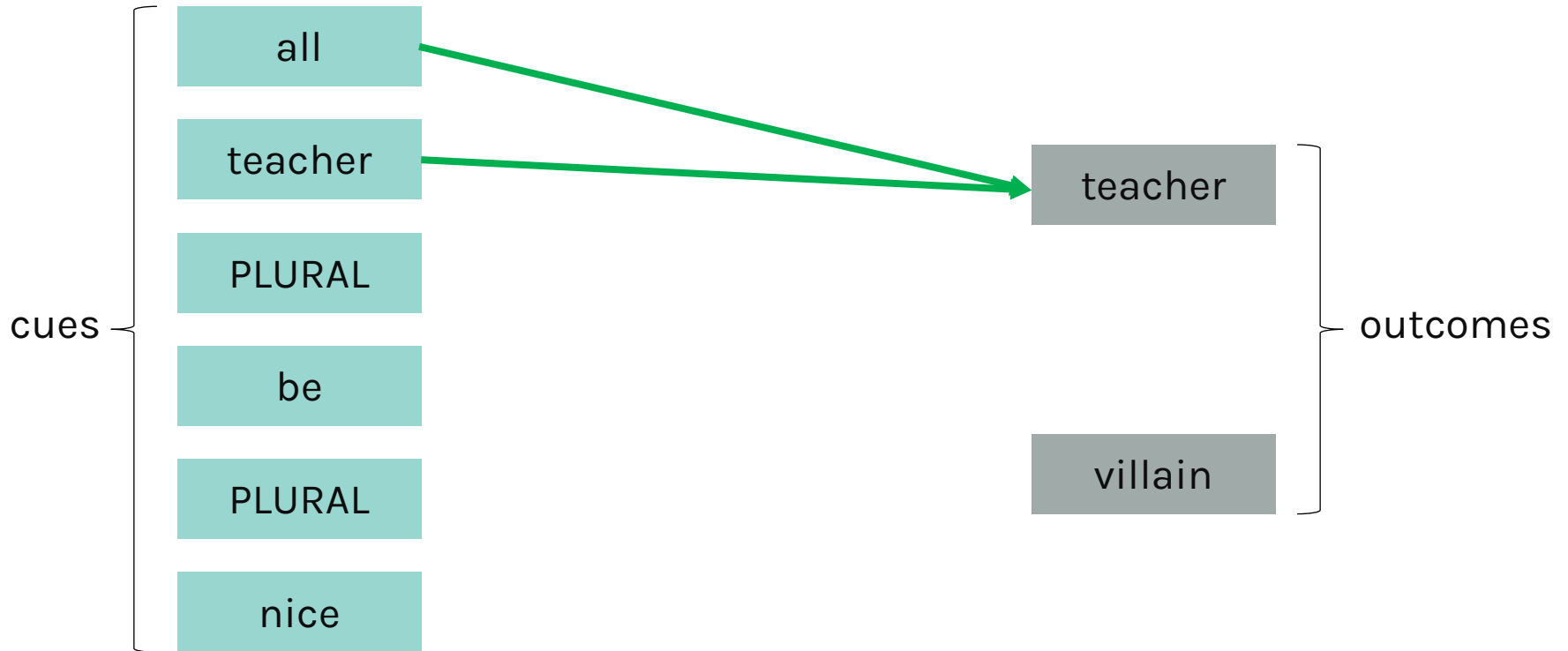
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+						
villain							

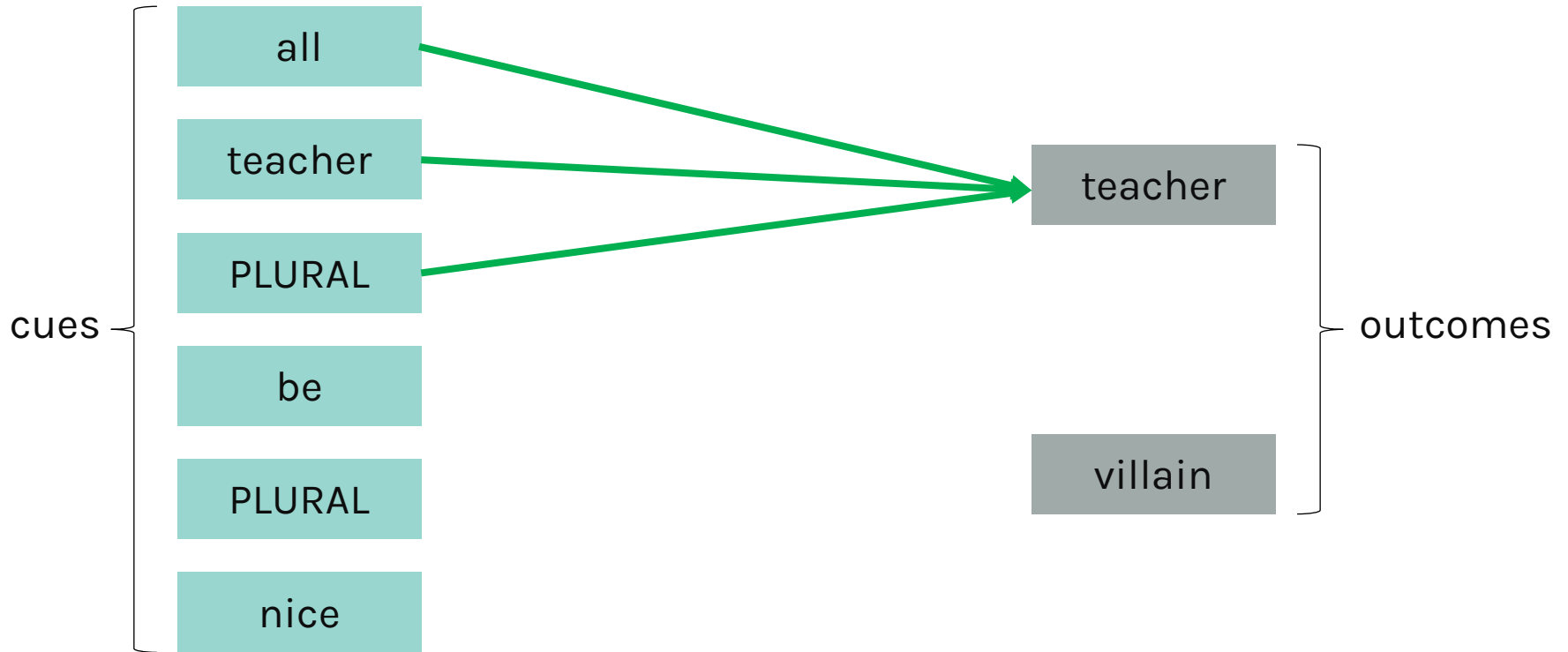
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+	+					
villain							

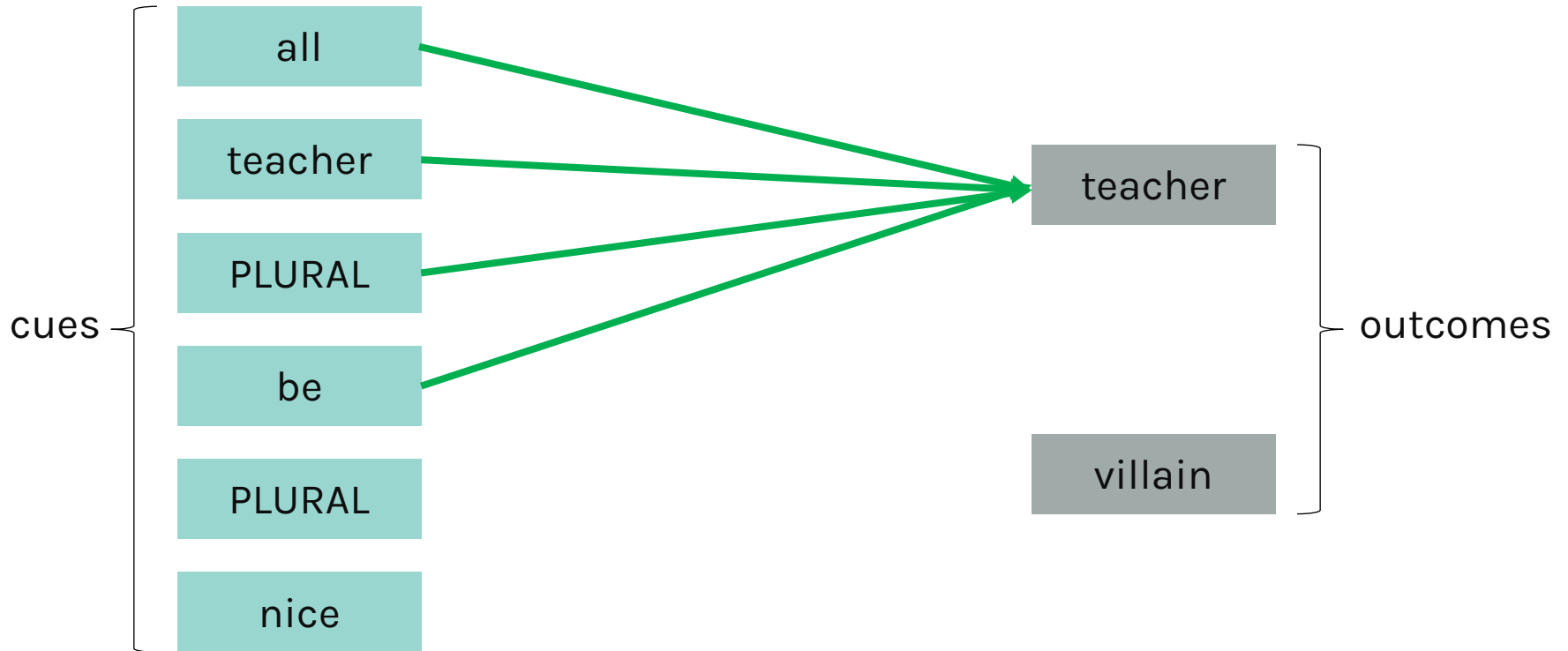
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+	+	+				
villain							

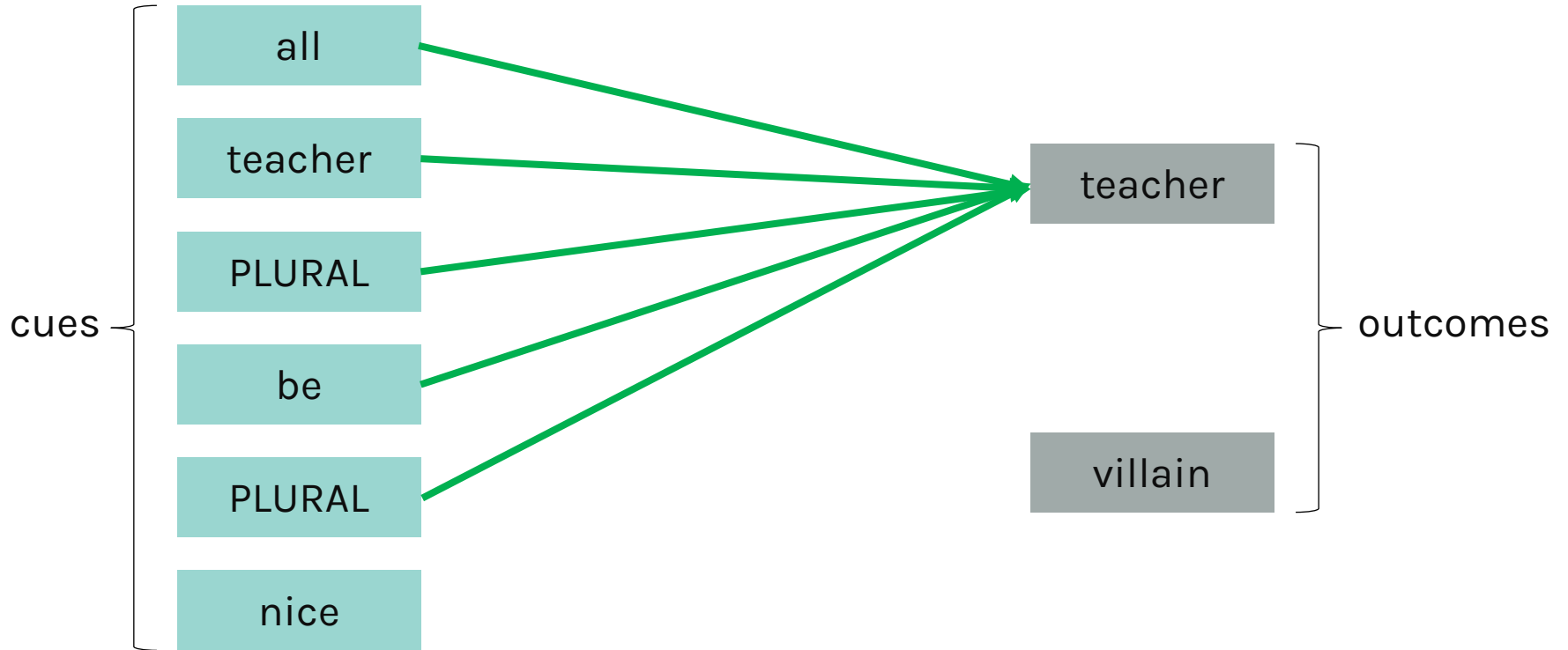
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+	+	+	+			
villain							

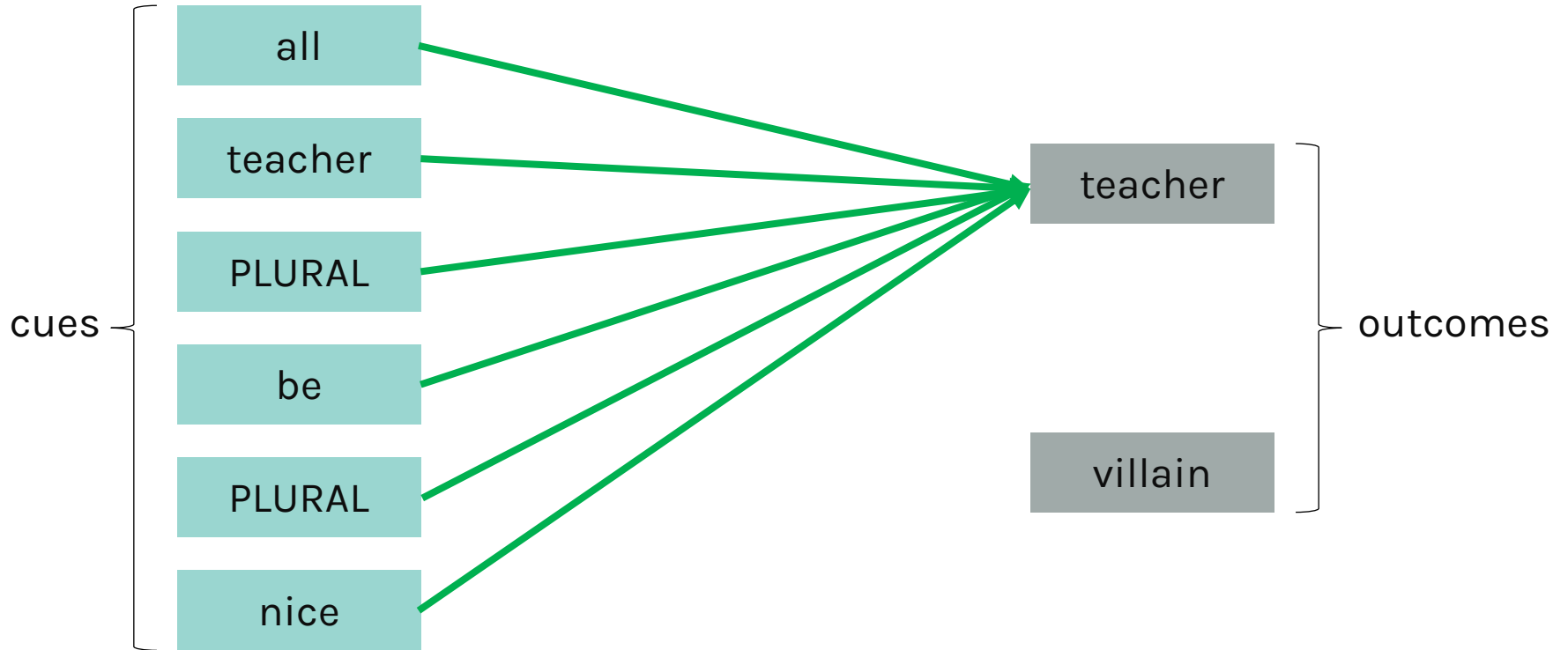
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+	+	++	+			
villain							

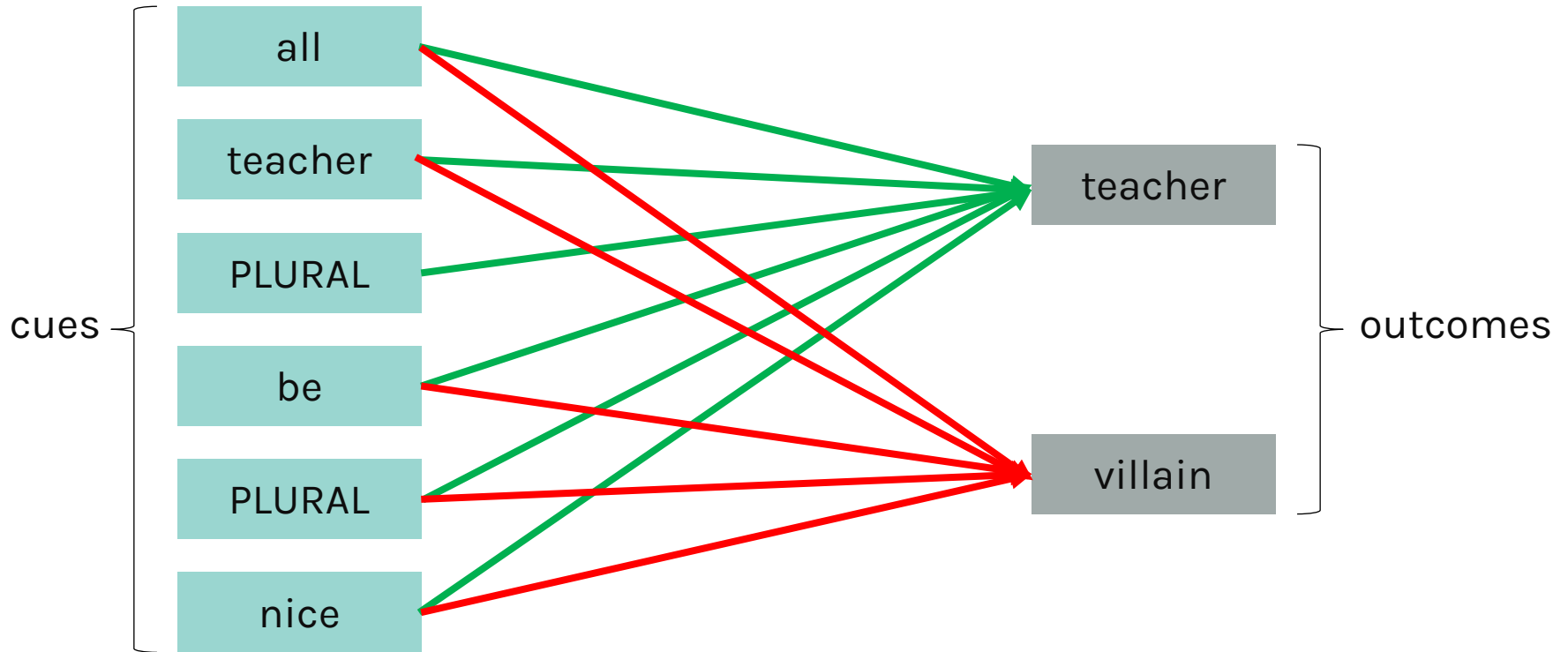
Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+	+	++	+	+		
villain							

Method



Example: *All teachers are nice.*

	all	teacher	PLURAL	be	nice	villain	evil
teacher	+	+	++	+	+	-	-
villain	-	-	-	-	-		

Method

Semantic vectors

- repeating this procedure for 830,000 sentences, we obtained association weights for all target words, inflectional functions, and a huge number of other words
- taking these rows of association weights, we obtain semantic vectors of individual words and inflectional functions of length 7,500
- for example:

<i>teacher</i>	all	teacher	PLURAL	be	nice	villain	evil
teacher	0.31	1.0	0.57	0.43	0.15	0.00071	0.0007
villain	0.0003	0.001	0.0005	0.0004	0.0091	1.0	0.96

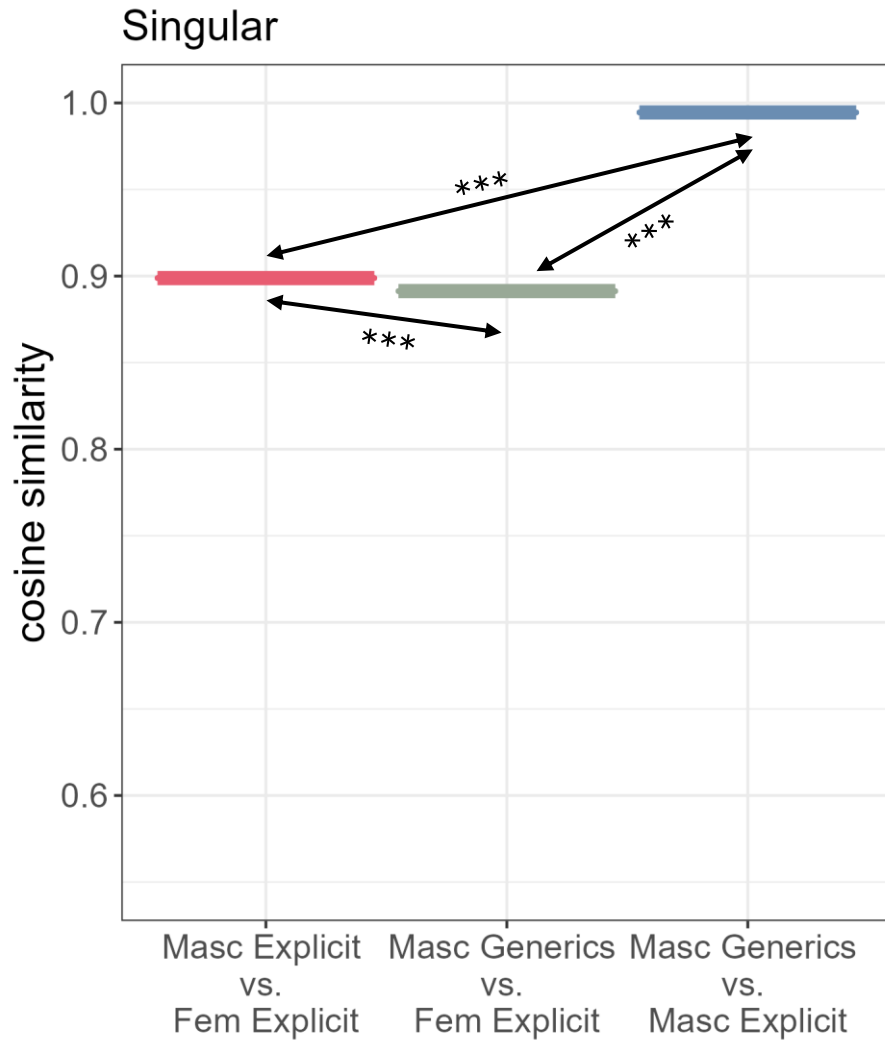
Method

Semantic vectors

- the members of our target word paradigms are complex words
- thus, their semantics need to be assembled

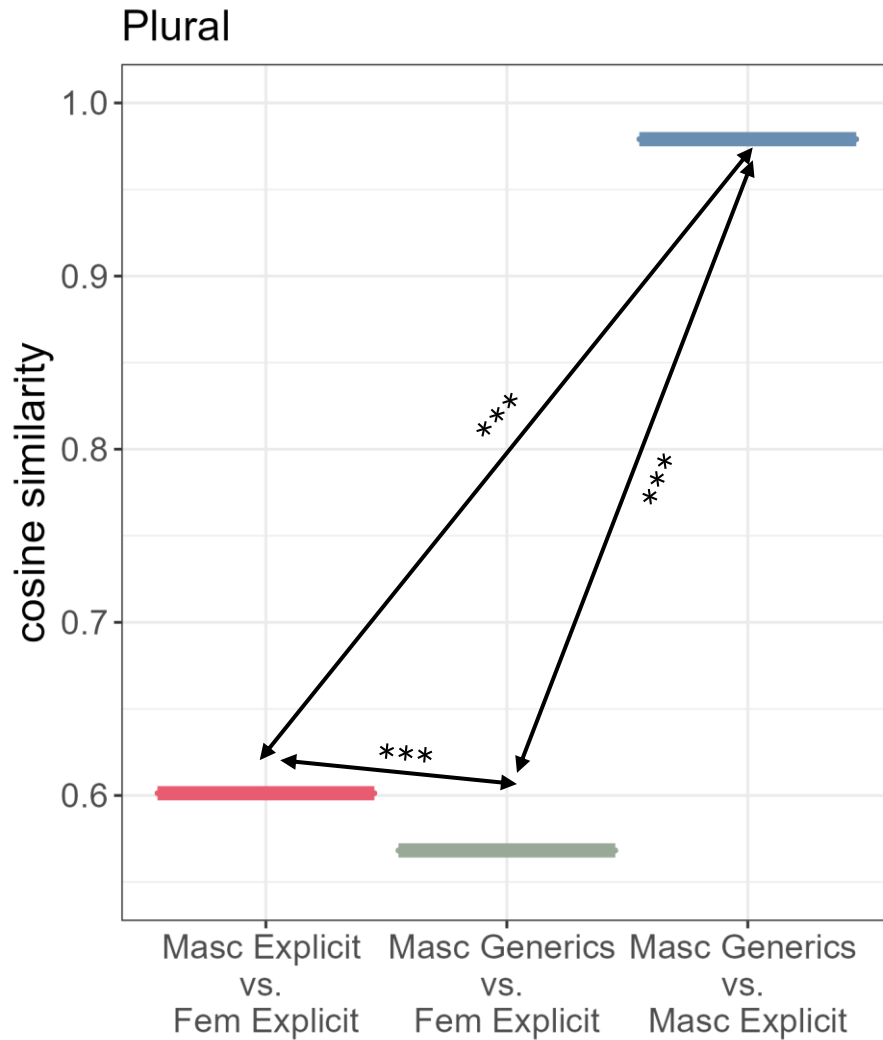
target form	base		number		gram. gender		genericity
Lehrer	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{singular}}$	+	$\overrightarrow{\text{masculine}}$	+	$\overrightarrow{\text{generic}}$
Lehrer	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{singular}}$	+	$\overrightarrow{\text{masculine}}$	+	$\overrightarrow{\text{explicit}}$
Lehrerin	$\overrightarrow{\text{Lehrer}}$	+	$\overrightarrow{\text{singular}}$	+	$\overrightarrow{\text{feminine}}$	+	$\overrightarrow{\text{explicit}}$

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

Method

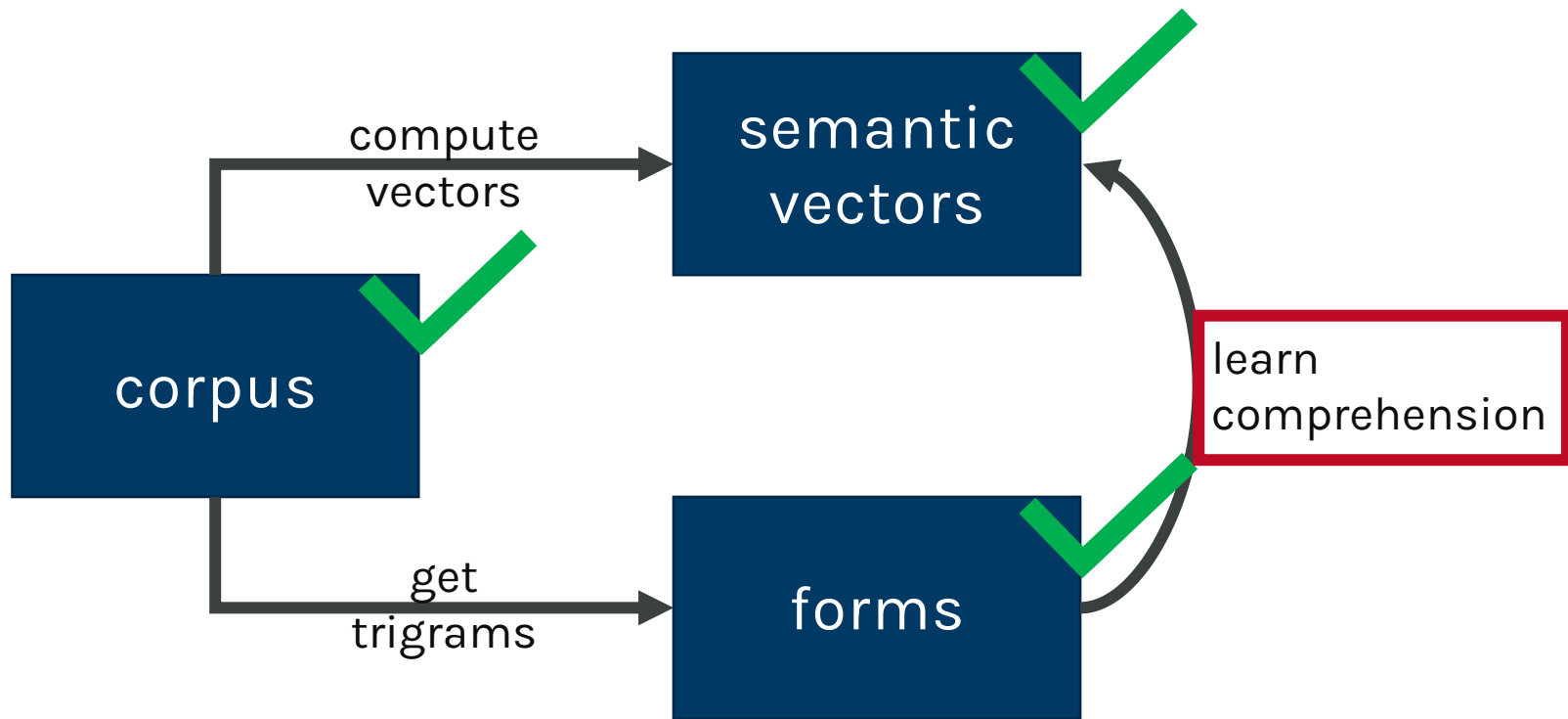
Forms

- we use trigrams as unit for a word's form
- trigrams / triphones have been shown to capture the form variability of words well (e.g. Chuang et al., 2020; Schmitz et al., 2021)

target form	#le:	le:r	e:rA	rA#	ArI	rIn	In#
Lehrer	1	1	1	1	0	0	0
Lehrer	1	1	1	1	0	0	0
Lehrerin	1	1	1	0	1	1	1

Method

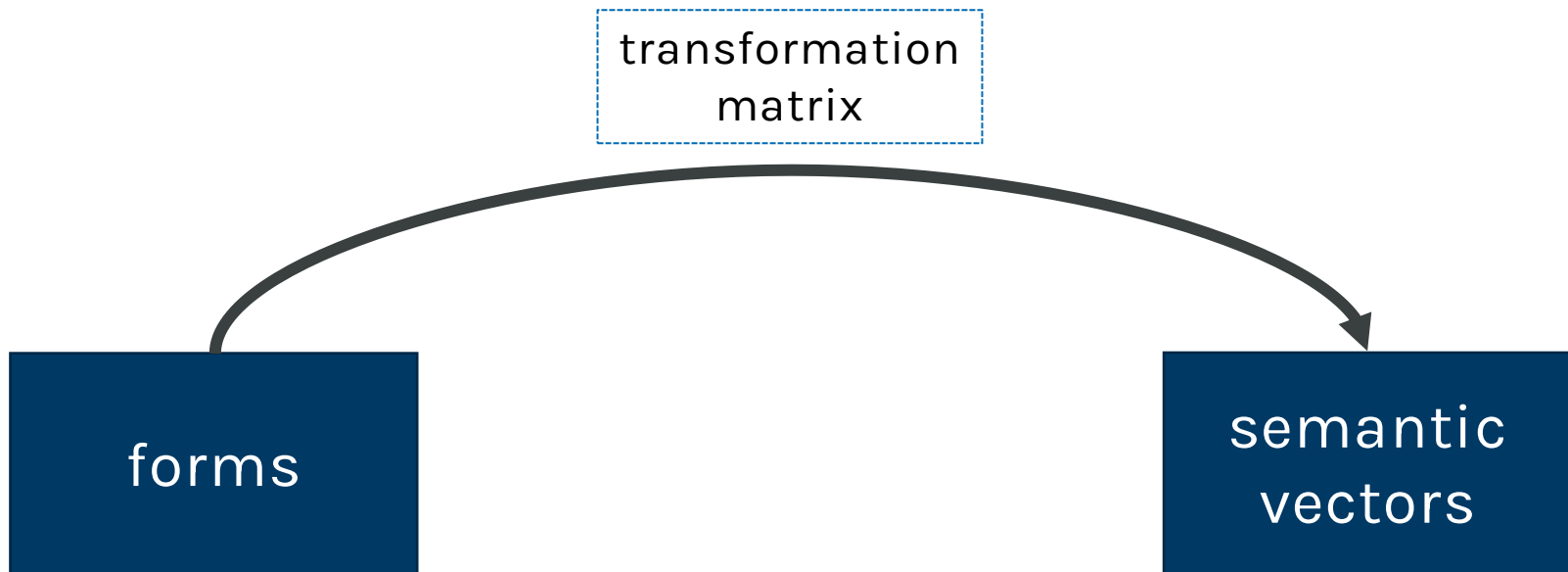
- we simulate an individual's mental lexicon by implementing a linear discriminative learning network (e.g. Baayen et al., 2019)
- for this, semantics and word forms are required as starting points



Method

Learning comprehension

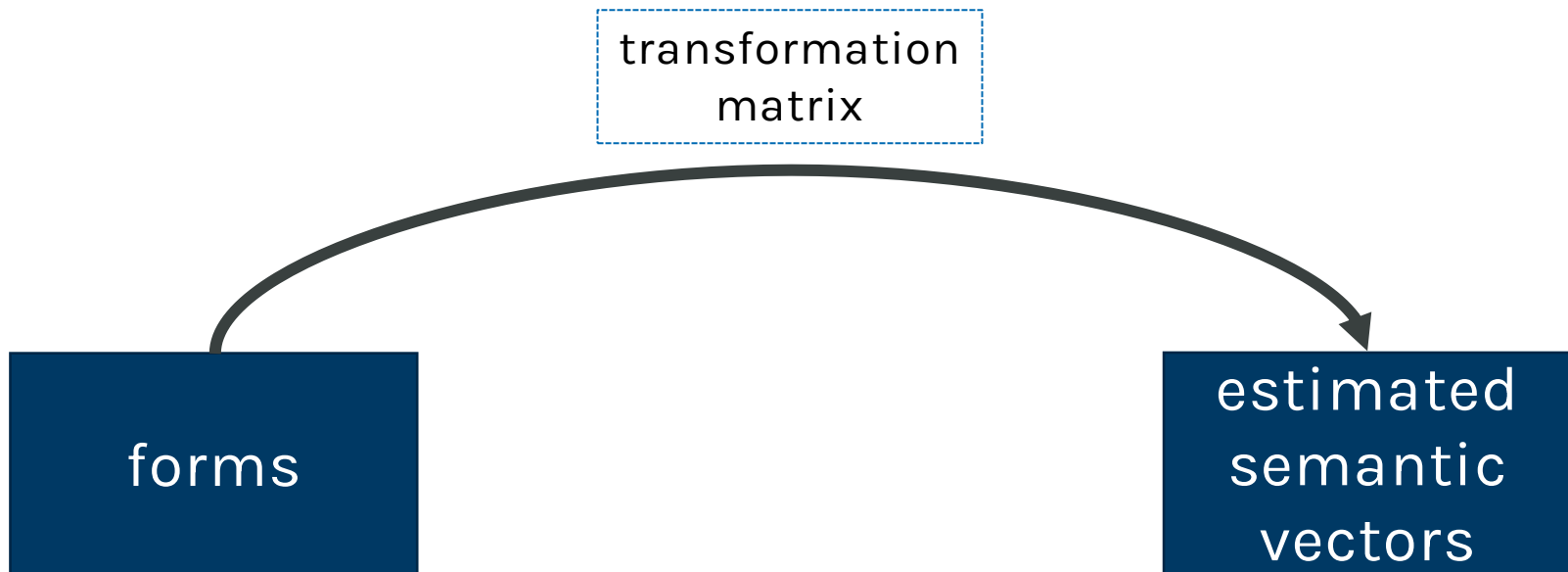
- comprehension is learnt by linearly mapping the matrix of forms onto the matrix of semantic vectors



Method

Learning comprehension

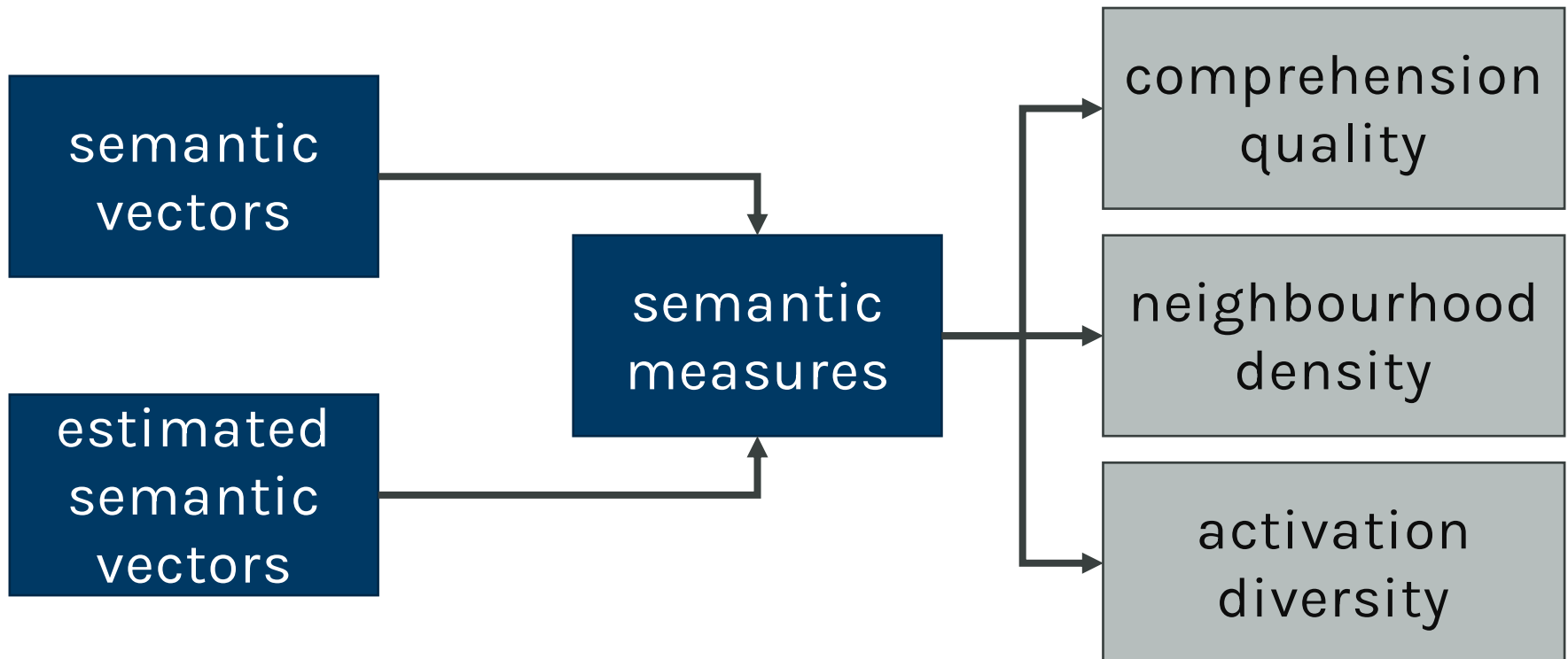
- comprehension is learnt by linearly mapping the matrix of forms onto the matrix of semantic vectors



Method

Learning comprehension

- using the original semantic vectors and the semantic vectors estimated by the comprehension learning, we can extract semantic measures



Analysis

Variables

- measures derived from the LDL implementation
 - COMPREHENSION QUALITY
correlation of a target's original and estimated vectors
 - NEIGHBOURHOOD DENSITY
correlation of a target with its 8 nearest neighbours
 - ACTIVATION DIVERSITY
Euclidian distance of a target's vector
- STEREOTYPICALITY JUDGEMENTS taken from Gabriel et al. (2008)

Analysis

Multinomial logistic regression

- dependent variable: GENERICITY

singular masculine generic; singular masculine explicit; singular feminine explicit

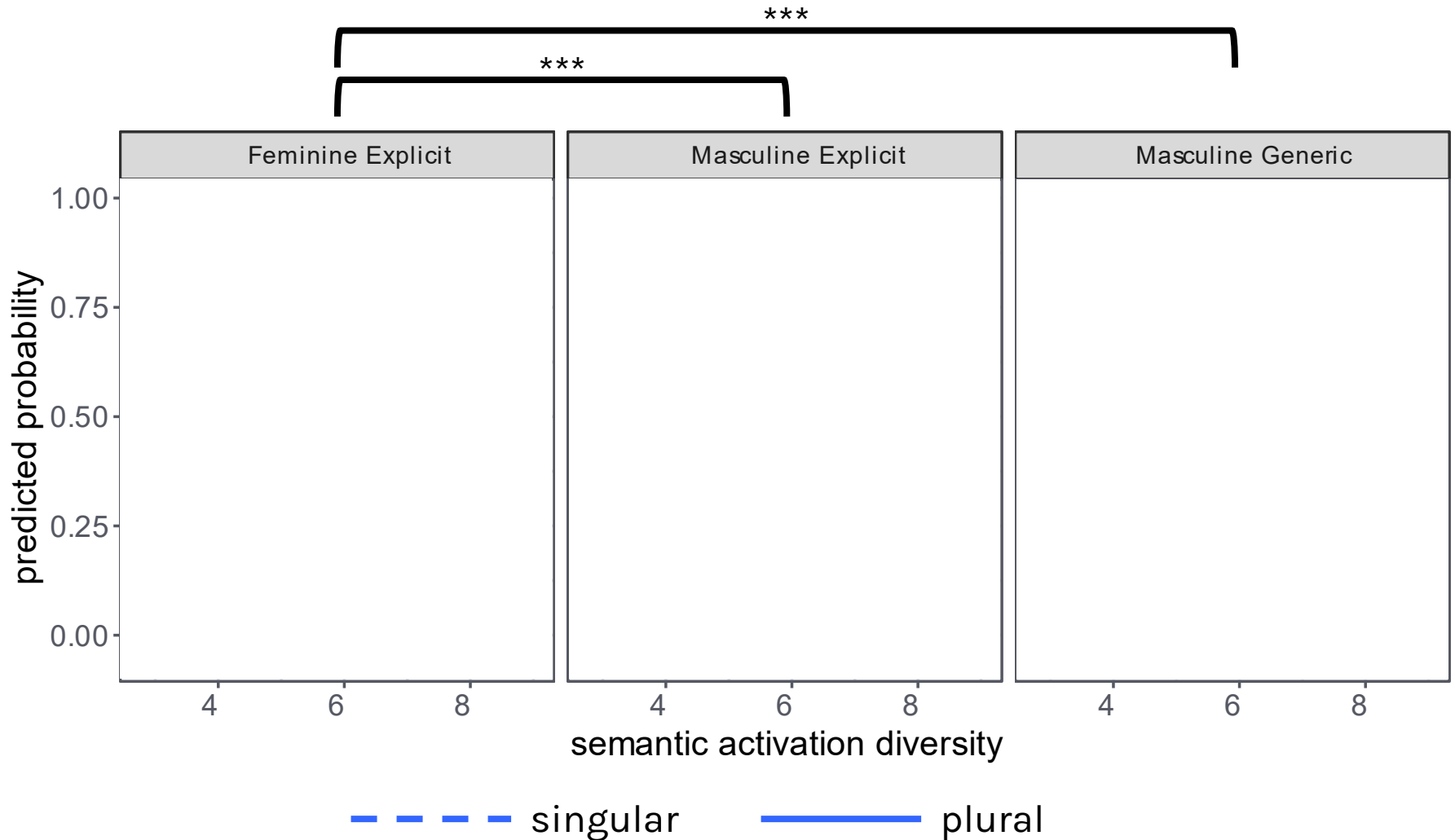
plural masculine generic; plural masculine explicit; plural feminine explicit

- explanatory variables

- ACTIVATION DIVERSITY
- a PC consisting of COMPREHENSION QUALITY & NEIGHBOURHOOD DENSITY
- STEREOTYPICALITY JUDGEMENTS

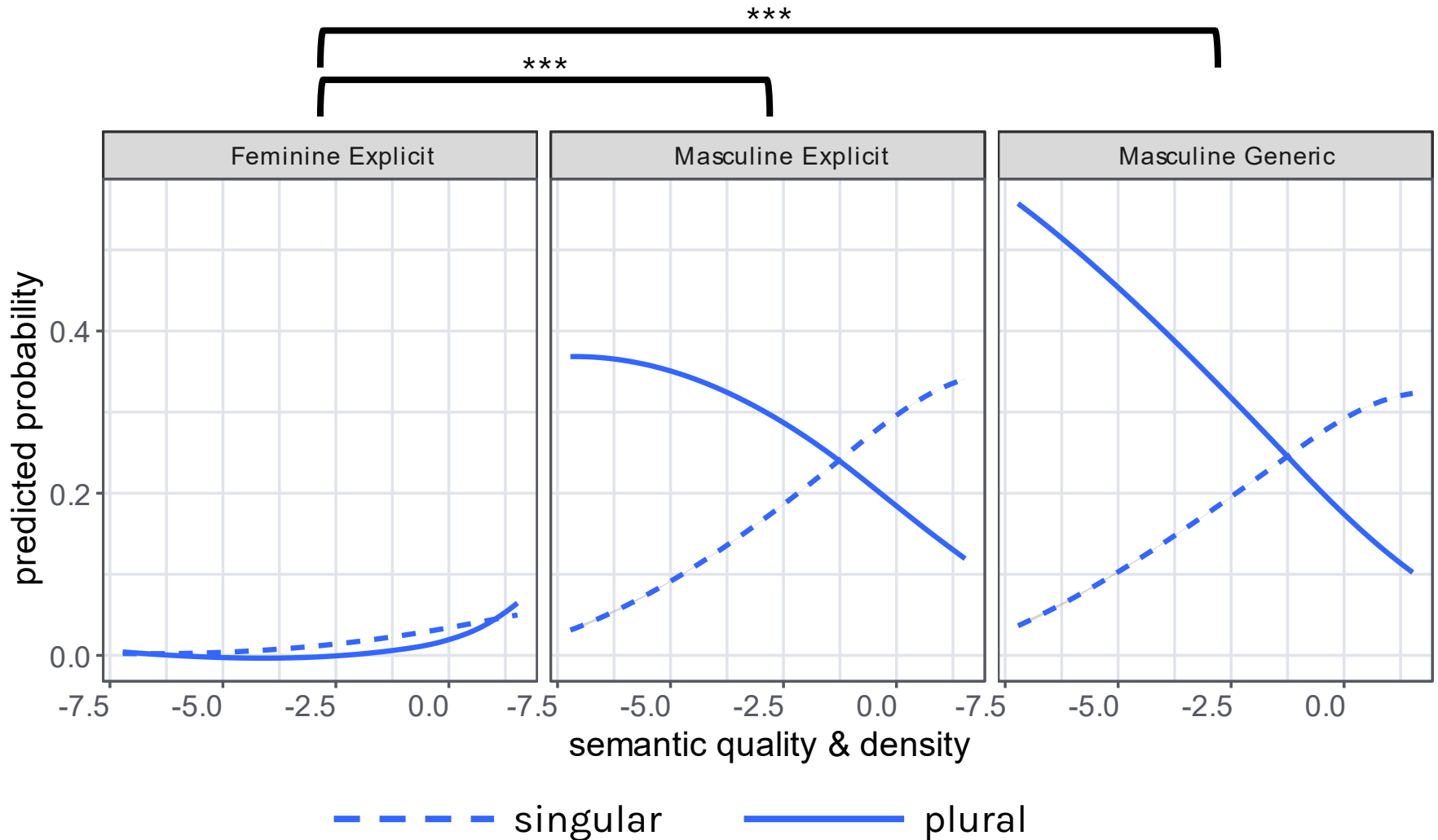
Results

Activation diversity



Results

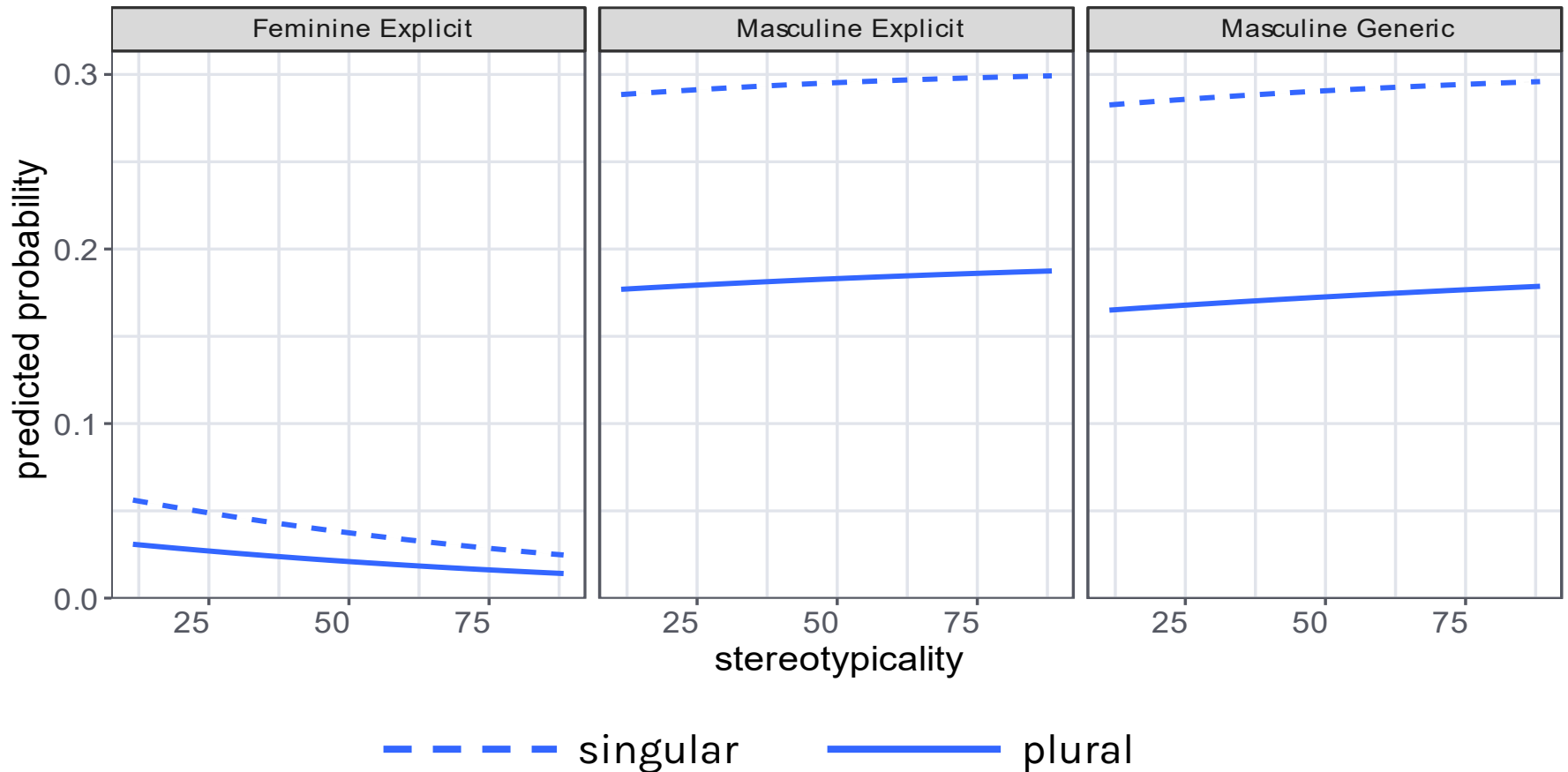
Comprehension quality & neighbourhood density



Results

Stereotypicality judgements

no significant differences



Results

Summary

- masculine generic and masculine explicit forms are highly similar in terms of
 - ACTIVATION DIVERSITY
 - they co-activate entries in the mental lexicon to a similar extent
 - COMPREHENSION QUALITY & NEIGHBOURHOOD DENSITY
 - they are comprehended equally well
 - they live in similarly dense neighbourhoods
- feminine explicit forms are significantly different as compared to masculine forms in regard to all semantic measures
- stereotypicality judgements do not show a significant effect

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 [masculine] generics has a semantic component of “maleness”
 - Irmen & Linner (2005)
semantic similarity of masculine generics and explicit nouns 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 nouns, leading to a semantic activation of masculine explicit nouns, 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)
- future research will show
 - whether the LDL measures computed for our data are predictive of behavioural measures
 - how (new) more neutral forms, e.g. *Lehrer*innen*, *LehrerInnen*, perform

Thank you!

References 1/2

- 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>
- Chuang, Y.-Y., Vollmer, M. L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., & Baayen, R. H. (2021). The processing of pseudoword form and meaning in production and comprehension: A computational modeling approach using linear discriminative learning. *Behavior Research Methods*, 53(3), 945–976. <https://doi.org/10.3758/s13428-020-01356-w>
- 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>

References 2/2

- 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>
- 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
- Schmitz, D., Plag, I., Baer-Henney, D., & Stein, S. D. (2021). Durational Differences of Word-Final /s/ Emerge From the Lexicon: Modelling Morpho-Phonetic Effects in Pseudowords With Linear Discriminative Learning. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.680889>
- 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>