Integrating talker and message in language processing: the influence of talker gender on sentence prediction

Introduction: Language processing often requires integrating the linguistic content of a message with its social context. A major type of social information comes from the speaker's identity, which may activate various stereotypes based on age, gender, *etc.*, and potentially influence subsequent processing. Several studies reported N400-like responses or late positivities when readers/listeners encounter sentences with talker-message incongruency [1,2]. However, the exact nature—or the presence—of the effect is challenged by mixed findings of its temporal and topographic profile and reports of null results [3]. In this paper, we report a series of sentence completion tasks in English that varied in whether and how the speaker's gender is revealed. We aim to address two main questions: (1) Do readers/listeners generate predictions of a sentence differently when there are cues about the likely gender identity of the speaker? (2) Are the effects (if existent) aligned with prevalent gender stereotypes?

Methods: Test materials consisted of 160 self-referring sentence prompts in English with no obvious gender biases (see Table 1 for examples). The sentence prompts were presented in five conditions: a baseline condition (i.e., bare sentence without gender cues) and four gender-cued conditions in a 2×2 design that varied in the type of gender cues (via personal names vs. voices) and trial presentation mode (mixed female/male-cued trials vs. blocking by gender). In the gender-cued conditions, female and male speakers were equally represented: in the name-cue conditions, typical female and male names preceded sentence prompts; in the voice-cue conditions, two computer-generated voices (1F, 1M) read the sentence prompts. 362 native English speakers recruited from Prolific completed the study online, with 16-25 responses per sentence prompt per condition. Participants were asked to complete the sentences in one or two words with their first thought (mean length of responses = 1.37 words; SD = 0.27).

To evaluate if human completions reflect context-driven gender biases, we calculated an approximation of gender bias using semantic embeddings. First, a gender-referential word list with eighteen masculine (e.g., "gentleman") and eighteen feminine (e.g., "schoolgirl") words was taken from [4]. Averaged embeddings were obtained separately for masculine and feminine referential words, which served as the benchmark for gender bias. Second, averaged embeddings of human completions were calculated per sentence prompt per condition, and compared against the average embeddings of masculine and feminine referential words, separately, through cosine similarity. The difference between the two cosine similarity measures (masculine − feminine) was taken as a measure of gender bias, with positive values indicating masculine-oriented responses and negative values indicating feminine-oriented responses. Lastly, paired *t*-tests were conducted to compare the gender bias of human completions across conditions. For better generalizability, the analysis was repeated with two pre-trained models of semantic embeddings, Word2Vec ([5], trained on GoogleNews)[1] and all-MiniLM-L6-v2 ("MiniLM" for short; [6]).[2]

Results: Female names elicited significantly more feminine-oriented responses (i.e., more negative gender bias values) than male names in both presentation modes and supported by both models (*p*s < .001; see Fig 1). Voice cues showed similar but slightly weaker effects: female voice prompted significantly more feminine-oriented responses in the mixed-gender condition (*p*s < .05) supported by both models, and in the blocked condition ($t = -2.69$, $p = .008$) only supported by MiniLM. Overall, responses in the gender-cued conditions were more biased towards the speaker gender than responses in the baseline condition, and the effect was more robust when speaker was male (all *p*s < .05).

Conclusion: This study provides evidence that speaker gender influences sentence prediction. When speaker gender is known, participants tend to generate gendered responses aligned with the speaker gender. The effects seemed more pronounced when gender was presented via names than voices and when speaker was male. Future studies will compare gender biases calculated from pre-trained embeddings with human ratings.

[1] https://www.kaggle.com/datasets/leadbest/googlenewsvectorsnegative300
[2] https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

Table 1. Examples of experimental stimuli

Gender cue	Speaker gender	Example sentence prompts
Nil	Nil	My favorite hobby is ____ .
Name	Female	Helen said: My favorite hobby is ____ .
	Male	Zach said: My favorite hobby is ____ .
Voice	Female	My favorite hobby is ____ . (female voice)
	Male	My favorite hobby is ____ . (male voice)

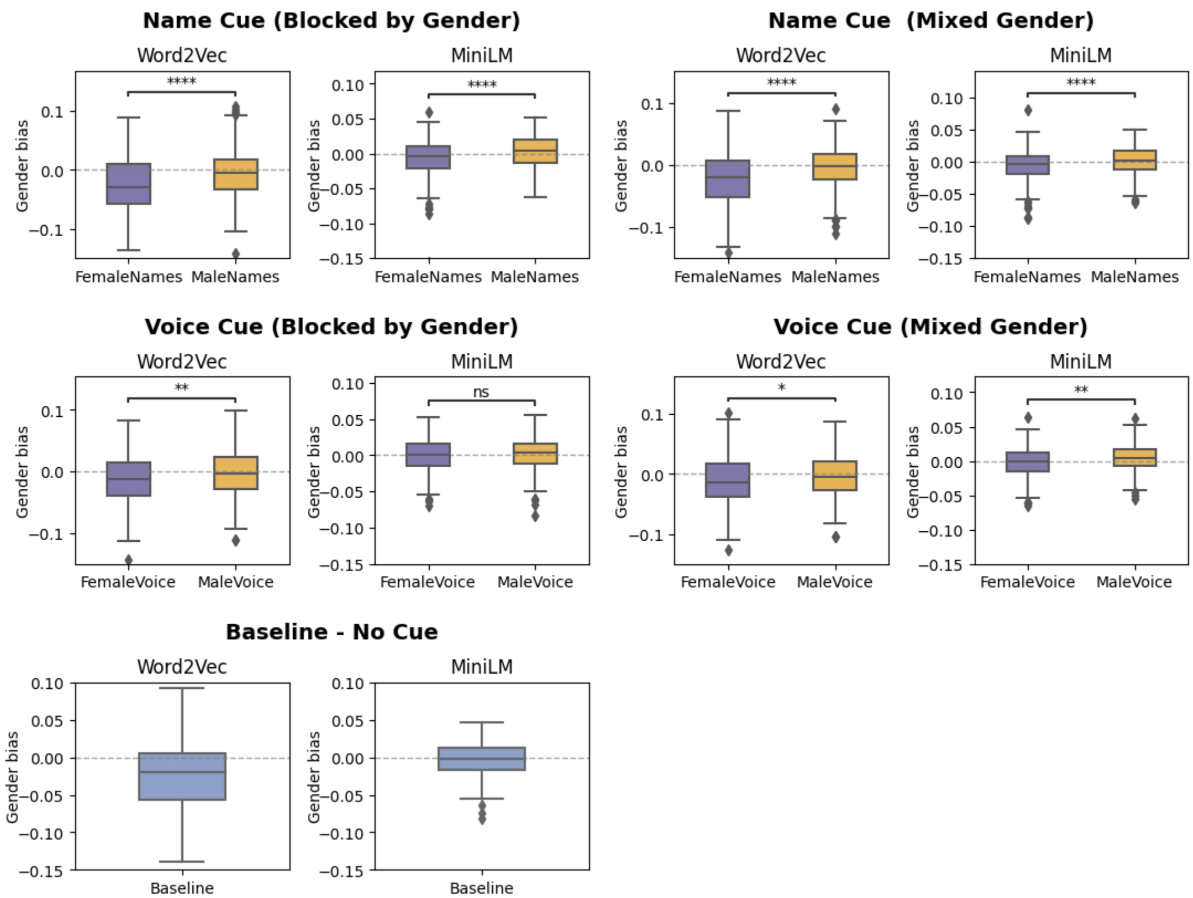

Figure 1. Gender biases from Word2Vec and MiniLM separated by condition.

References

[1] Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing–evidence from event-related brain potentials. *Neuroscience Letters*, *339*(3), 191-194.

[2] Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580-591.

[3] Porkert, J., Siyanova-Chanturia, A., Loerts, H., Schüppert, A., & Keijzer, M. (2024). N400 or P600?—A systematic review of ERP studies on gender stereotype violations. *Language and Linguistics Compass*, 18(5), e12530.

[4] Nadeem, M., Bethke, A., & Reddy, S. (2020). StereoSet: Measuring stereotypical bias in pretrained language models. *arXiv preprint arXiv:2004.09456*.

[5] Mikolov, T. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, *3781*.

[6] Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., & Zhou, M. (2020). Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. *Advances in Neural Information Processing Systems*, *33*, 5776-5788.