Pretty Plumbers to the Rescue: Adjectives Aid in Gender Mismatch Recovery

Previous research has found that gender bias impacts coreferential dependency formation, finding e.g., a slowdown in reading times (RTs), termed Gender Mismatch Effect (GMME), when a grammatically accessible noun phrase is judged by the comprehender as "mismatching" a pronoun in gender[1-5]. This suggests a role for not only grammatical constraints (e.g., morphological gender) but also non-grammatical factors such as social information. English has nouns that are stereotypically gendered (e.g., *nurse*) as they are not morphologically (e.g., *waitress*) or definitionally (e.g., *mother*) marked for gender but still carry a conceptual gender bias grounded in world knowledge. As such, it has been argued that stereotypical gender information is more susceptible to surrounding contexts than morphological or definitional gender. [1-3] have found that certain discourse contexts can successfully eliminate a GMME for these stereotypical nouns; however, this work relied on using morphological/definitional gender information for context, such as contexts that explicitly state the gender of a noun (e.g., *The electrician*$_{MASC}$ *was a cautious* <u>*woman*</u>…) and finding that this prior disambiguating context eliminated a GMME on a later reflexive (e.g., *The electrician*$_{MASC}$ *taught herself*) [2]. It is an open question how other forms of contextual information interact with gender biases of nouns. The present study uses attributive adjectives (e.g., *The muscular*$_{MASC}$ *nurse*$_{FEM}$) as prior gender context for a noun, to further probe if the gender of a stereotypical noun can be overridden, and the GMME eliminated, by context that only carries stereotypical gender information. Adjectives are an interesting test case as their gender bias – if it exists – tends not to be morphological/definitional. While adjectives in English are argued to carry gender biases based on corpus data [6-11], their impact on processing has not been tested; thus, this study also more broadly asks if and how the potential gender bias of adjectives impacts coreferential dependency formation during sentence processing.

Methods: The experiment used a 2 x 2 design (item N=24): *Adjective Bias* (adjective & reflexive match vs. mismatch) x *Noun Bias* (noun & reflexive match vs. mismatch), reflexives *himself/herself* balanced across items – see (1). Gendered adjectives were selected based on prior corpus studies; stereotypical nouns were selected from prior GMME studies. Participants (N=120) saw items word-by-word using a G-Maze paradigm[13] in which participants needed to distinguish the correct continuation from a distractor.

Results. A linear mixed effects regression model was fit, predicting RTs at the reflexive (*himself/herself*) region from *Adjective Bias, Noun Bias*, and their interaction. We found a main effect of *Noun Bias* ($p<0.001$), as well as an interaction of *Adjective Bias* and *Noun Bias* ($p<0.001$), such that in the noun-mismatch conditions (1c&d) RTs are faster when the adjective matches the reflexive ($p<0.001$). In noun-match conditions (1a&b), there is no effect of adjectives ($p=0.116$) – Fig. 2. RTs on the noun were also analyzed, with the predictor *Adjective-Noun Match*, reflecting whether the noun and adjective gender bias match (1a&c coded as *match* and 1b&d as *mismatch*). We found a main effect of *Adjective-Noun Match*, such that RTs on the noun are higher in mismatch conditions ($p<0.001$) – Fig. 1.

Implications. RTs on the noun show that comprehenders are sensitive to a mismatch in gender bias between the adjective and noun. Most interestingly, on the reflexive, the effect of adjectives was restricted to noun-mismatch conditions. This suggests that when the stereotypical noun mismatches the reflexive in gender, comprehenders are able to take into account the gender biases of adjectives in order to reinterpret the gender information of the noun, which helps to "repair" coreferential dependency formation between the noun and the reflexive pronoun, resulting in faster RTs for condition (1d) than (1c). However, a context containing gender information in the form of biased adjectives does not completely override the gender of the stereotyped noun (cf. context manipulations of [1-3]), as the GMME from the noun was not completely eliminated, and information associated with the noun overall had greater influence on coreferential dependency formation (as evidenced by no RT difference at the reflexive between (1a) vs. (1b)).

References:
[1] Carrieras et al., 1996 [2] Duffy & Keir, 2004 [3] Kreiner et al., 2008 [4] Reynolds et al., 2006 [5] Sturt, 2003 [6] Moon, 2014 [7] Turrentine et al., 2019 [8] Adams et al., 2022 [9] Khan et al., 2023 [10] Raphael, 2023 [11] Evans et al., 2024 [12] Hess et al., 1995 [13] Boyce et al., 2020

Figure 1: G-Maze RTs on Noun Region

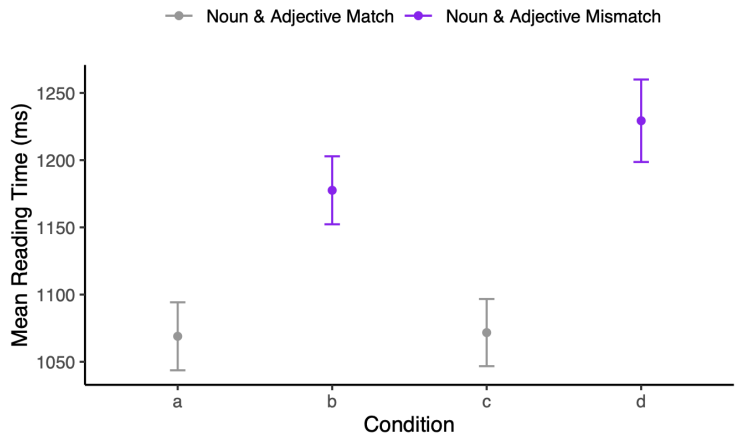
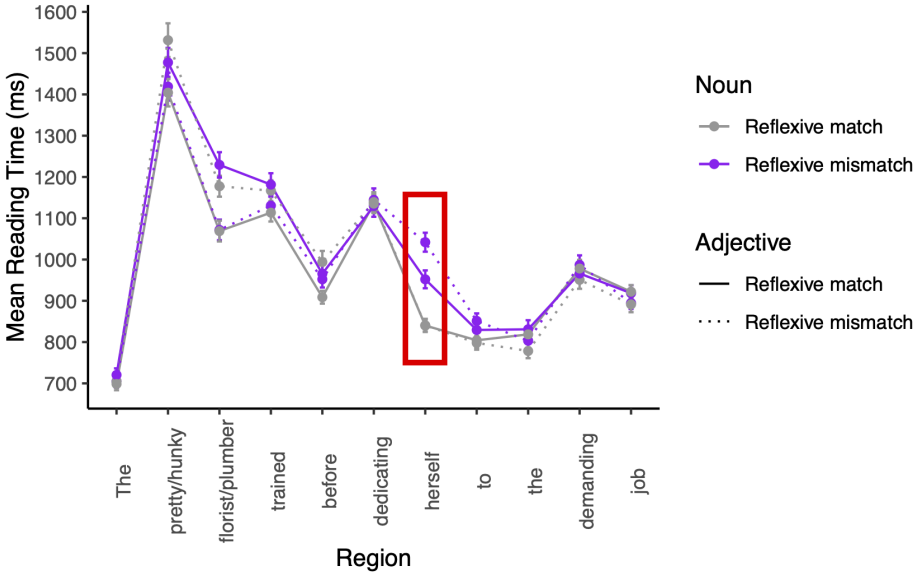


Figure 2: G-Maze Results



(1) Example Items:

a.The *pretty*$_{FEM}$ *florist*$_{FEM}$ trained before dedicating *herself* to the demanding job. **[A & N match]**

b.The *hunky*$_{MASC}$ *florist*$_{FEM}$ trained before dedicating *herself* to the demanding job. **[A mismatch, N match]**

c.The *hunky*$_{MASC}$ *plumber*$_{MASC}$ trained before dedicating *herself* to the demanding job. **[A & N mismatch]**

d.The *pretty*$_{FEM}$ *plumber*$_{MASC}$ trained before dedicating *herself* to the demanding job. **[A match, N mismatch]**