

Assumptions based on names and professions: A comparison of form-based and meaning-based gender biases in pronoun resolution

Elsi Kaiser (emkaiser@usc.edu) and Sabrina Soto, University of Southern California

Linguistic expressions can trigger various assumptions. Last-name-only style (e.g. *Jones came in*, *Ramirez was promoted*, see also ex.1a,b) and many professional role nouns (e.g. *mechanic*, *boxer*) often evoke the assumption that the individual is male (e.g.[1,2,4,5,6]). We provide the first systematic comparison of the strength of gender biases triggered by role nouns vs. last-name-only style, through the lens of pronoun interpretation in English.

Meaning-based vs. form-based effects. With **role nouns**, people have expectations about which roles are more often held by which gender(s) (e.g.[6]). Crucially, this is linked to *meaning*: For a role noun to elicit a gender bias in a comprehender's mind, the comprehender needs to know the noun's meaning, i.e., access the noun's lexical entry. In contrast, the male bias of **last-name-only style** is triggered by a *specific linguistic form*, not the meaning of a lexical entry.

Predictions: We expect male-biased role nouns and last-name-only to impact interpretation of the pronouns 'he' and 'she.' However, it is still unclear (i) how robust these effects are, (ii) how they compare to other semantic effects on pronoun interpretation, e.g. effects of implicit causality verb semantics (see e.g. [3,4]), and (iii) how meaning-based vs. form-based biases (role-noun vs. last-name-only) compare to each other.

In **Exp.1 Last-name-only** (n=91 native U.S.-English speakers) and **Exp.2 Male-biased role nouns** (n=54 native U.S.-English speakers), participants wrote continuations for sentence fragments. Targets were two-clause sequences, connected with *because* and ending in a prompt pronoun. Examples for Exp1-2 are given in ex.(2-3).

Design (2x2 Latin Square): We manipulated whether the pronoun prompt is *he/she*. We also manipulated verbs' semantic implicit causality bias: When followed by an explanation signaled by *because*, does the verb favor a subject-referring (IC1) or object-referring (IC2) interpretation of the pronoun prompt? Verbs were using [3]'s norms; positive/negative valence was balanced.

Targets contained one first-name referent, and one last-name-only referent (in **Exp.1**) or one clearly male-biased role noun (in **Exp.2**, rated $\geq 70\%$ male in [6]). Last-name-only/role-noun referents are *in the position favored by verb semantics* (subject/IC1 verbs, object/IC2 verbs). Thus, we ask: **How often do participants follow verb semantics and interpret 'she' (vs. 'he') as referring to last-name-only referents (Exp.1) and male-biased role nouns (Exp.2)?**

Results: Exp.1 Last-name-only. Fig.1 shows how often *he/she* picks out last-name-only referents. **He** overwhelmingly prefers last-name-only referents ($>80\%$, above chance, $p's < 0.001$). However, **she** is used to refer to the last-name-only referent at much lower rates ($40\%-50\%$, at chance). Thus, participants are more reluctant to use *she* for last-name-only referents, even though the implicit-causality verb bias strongly favors that referent.

Results: Exp.2 Male-biased role nouns. Fig.2 shows how often *he/she* picks out the male-biased role noun. **He** prefers male-biased role nouns ($>70\%$, above chance, $p's < 0.003$). But **she** is rarely interpreted as referring to male-biased role nouns ($<40\%$, below chance, $p's < .05$).

Comparing Exp.1-2: **He** patterns alike in Exp.1-2, but **she** is *less likely* to be interpreted as referring to male-biased role nouns than last-name-only referents (glmer, $p < 0.001$).

Both studies show slight, expected subjecthood effects; pronouns favor subjects. Crucially, using IC1/IC2 verbs shows results hold for subjects & objects (not artifacts of grammatical role).

Conclusions. Both form-based (Exp.1, last-name-only) and meaning-based gender biases (Exp.2, role nouns) are powerful enough to overcome otherwise-robust verb-semantic effects on pronoun interpretation. Furthermore, our study provides the first evidence that gender biases rooted in meaning (at least the kind that we tested) can be stronger than those rooted in linguistic form. We suggest this may stem from the mapping between 'linguistic expression' and 'gender association' being *more arbitrary* in the case of last-name-only format. We address ways for theories of pronoun interpretation to incorporate these largely overlooked dimensions.

(1a) Throughout the 1940s, Kahlo participated in exhibitions in Mexico and the United States and worked as an art teacher. (About Frida Kahlo, from Wikipedia)

(1b) Taylor is dating Kelce, the Kansas City Chiefs tight end. (Talking about Taylor Swift and Travis Kelce, CNBC)

(2) *Example target* (**Exp1**: 20 targets with 20 different surnames/last names, **Exp2**: 20 targets with 20 different male-biased role nouns rated $\geq 70\%$ male in [6], 20 verbs, 22 filler items. *In Exp. 1, the first names (e.g. Amanda) matched the gender of the pronoun prompts (e.g. she).*)

- (a) {**Smith/The guard**} impressed **Eric** because **he** [IC1 verb: semantics favor the subject]
 (b) {**Smith/The guard**} impressed **Amanda** because **she** [IC1 verb: semantics favor subject]
 (c) **Frank** promoted {**Mayfield/the engineer**} because **he** [IC2 verb: semantics favor object]
 (d) **Claire** promoted {**Mayfield/the engineer**} because **she** [IC2 verb: semantics favor object]

(3) *Sample continuations*

Smith impressed Eric because **he** ... had done a great job with the project. [he=subject]
 Smith impressed Amanda because **she** ... thought his performance was great. [she=object]
 Claire promoted Mayfield because **she** ... had impressed Claire with her skills. [she=subject]

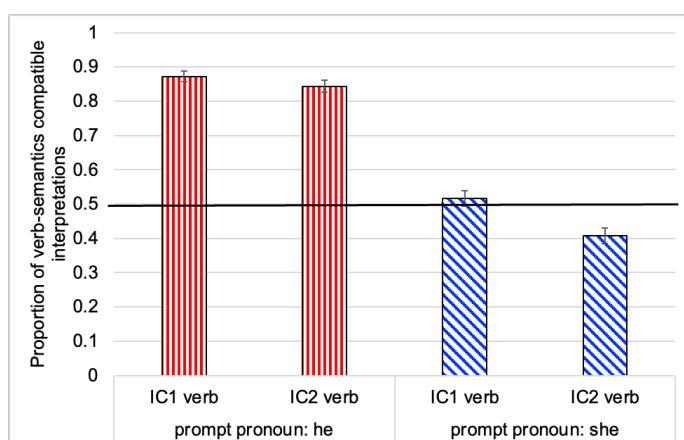


Fig.1 Exp.1: How often do people interpret the pronoun as referring to the **last-name-only** referent? (Y axis shows proportion of subject interpretations for IC1 verbs and object interpretations for IC2 verbs)

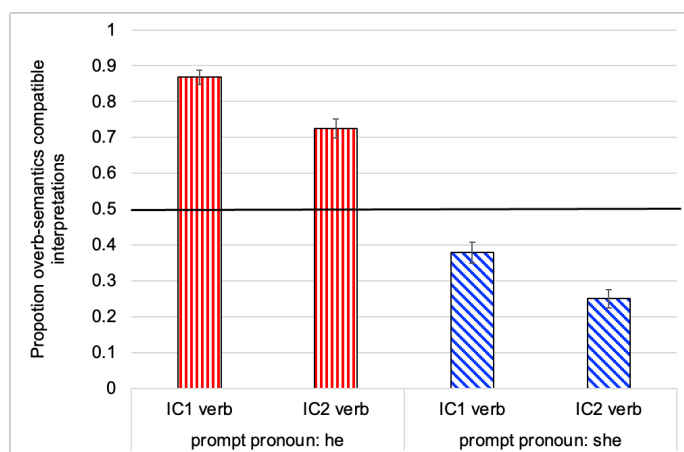


Fig.2 Exp.2: How often do people interpret the pronoun as referring to the **male-biased role noun**? (Y axis is as in Fig.1)

References [1] Atir & Ferguson (2018). How gender determines the way we speak about professionals. *PNAS Proceedings* [2] Gardner & Brown-Schmidt (2024). Biased inferences about gender from names. *Glossa Psycholinguistics*. [3] Hartshorne &

Snedeker (2013). Verb Argument Structure Predicts Implicit Causality, *LCN*. [4] Kaiser et al. (2022). Referring to someone using only their last name: Insights from gender-marked pronouns. Talk at *Linguistic Evidence*. [5] McConnell-Ginet, S. (2003). "What's in a Name?" Social Labeling and Gender Practices. *Handbook of Language and Gender*. [6] Misersky, Gyax, Canal, Gabriel, Garnham, Braun & Sczesny (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *BRM*.