Effects of grammatical gender on gender inferences: evidence from French hybrid nouns

Benjamin Storme and Laura Delaloye-Saillen

Abstract

A growing body of research shows that readers/listeners are biased by the grammatical gender of a noun when making inferences about the gender of its referent. This result is central in debates about gender-fair language but has mostly been established using masculine generics. This paper presents two preregistered studies on French that aim to replicate this result but using a lesser-studied type of nouns: generic hybrid nouns. These nouns can refer to both male and female individuals but are either masculine or feminine, depending on the noun (e.g. un talent ‘a talent’ and une vedette ‘a star’). The availability of both genders for hybrid nouns allows for a more comprehensive test of the effect of grammatical gender than permitted by masculine generics. Overall, the paper replicates the role of grammatical biases in gender inferences, with masculine hybrid nouns being judged as more likely to refer to male individuals as compared to feminine hybrid nouns. However the results did not reveal a symmetric bias for feminine nouns, which were interpreted as gender-neutral. But this latter result should be interpreted with caution as it could be due to uncontrolled effects of gender stereotypes coming from the specific stimuli used in the study.

Keywords: grammatical gender; hybrid nouns; masculine generics; gender-fair language; French

Number of words: 4,116

1 Introduction

In languages with sex-based grammatical gender, the grammatical gender of a word does not always transparently reflect the gender of the corresponding referent. For instance, in many languages, masculine forms do not refer only to male individuals but also have a generic reading: they can be used to refer to mixed-gender groups or individuals whose gender is not known or irrelevant (e.g. Aikhenvald 2016: chapter 7). The use of masculine generics is illustrated in (1) for French: although the speaker uses the masculine pronoun il ‘he’, they actually promise to punish whoever broke their glasses, regardless of gender.

(1) Masculine generics in French

Je ne sais pas qui a cassé mes lunettes mais il va être puni.

‘I don’t know who broke my glasses but he will be punished.’

Although grammatical gender is not a reliable cue to the referent’s gender in (1), a growing body of experimental research suggests that readers/listeners still draw gender inferences that are in line with grammatical gender in this type of context (e.g. Gastil 1990 on English he; Gygax et al. 2008, Xiao, Strickland & Peperkamp 2022 on French; Gygax et al. 2008 on German). For instance, Brauer & Landry (2008) found that French readers were more likely to think of male candidates for Prime Minister when presented with a masculine form intended as generic (candidats ‘candidates. MASC’) than when presented with an explicitly gender-neutral form (candidats/candidates ‘candidates.MASC/candidates.FEM’). These findings have been used as a key argument by proponents of gender-fair language to replace masculine generics with gender-neutral forms. Gender-fair language aims at eliminating linguistic biases that contribute to an unfair representation of gender and therefore might perpetuate gender inequalities (see Sczesny, Formanowicz & Moser 2016 for an overview).

The result that readers/listeners tend to overrely on grammatical gender in their gender inferences has mostly been established using masculine generics such as (1) (see Gygax et al. 2021 for an overview). However masculine generics are not the only case where grammatical gender does not transparently reflect the referent’s gender, with potential social impacts. The present paper examines another such case. In French, some nouns have a generic interpretation but a specific grammatical gender, feminine or masculine, depending on the noun, as illustrated in (2a) and (2b). Following Corbett (2015) and Gygax et al. (2019), these nouns will be referred to as hybrid nouns. The present paper aims to test whether the effect of grammatical gender observed for masculine generics extends to generic hybrid nouns, with masculine hybrid nouns inducing more male representations than feminine hybrid nouns and therefore potentially giving rise to problematic biases from a social perspective.
Hybrid nouns with a generic interpretation in French

a. Feminine nouns: *une personne* ‘a person’, *une vedette* ‘a star’, etc.

b. Masculine nouns: *un individu* ‘an individual’, *un talent* ‘a talent’, etc.

Generic hybrid nouns form a closed lexical class and this might explain why they did not receive as much attention as masculine generics in the literature (Gygax et al. 2019). However these nouns also present some advantages compared to masculine generics. First, the availability of both genders for hybrid nouns allows for a more comprehensive test of the effect of grammatical gender than permitted by masculine generics, which are only found in the masculine. In particular, hybrid nouns provide a way to test whether the male bias induced by masculine nouns is compensated by a symmetric female bias for feminine nouns.

Second, hybrid nouns make it possible to control for competition effects that could play a role in the male-biased interpretation of masculine generics. Words that can be used in the masculine generic can also be inflected in the feminine. The male bias for masculine generics could be partly due, or even entirely according to some,\(^1\) to the presence of this feminine competitor. Under this view, the masculine grammatical gender does not directly trigger male inferences. Instead, the reader/listener tends to discard female interpretations for masculine forms by reasoning that the writer/speaker could have used the more informative feminine competitor if they had a female interpretation in mind. In line with the competition-based hypothesis, Gygax & Gabriel (2008) found that readers were more likely to have a male-biased interpretation of masculine generics when reading a text also including feminine-inflected forms than when reading a text only including masculine generics. By contrast, competition is less likely to play a role for hybrid nouns as these nouns do not inflect for gender and therefore lack a salient morphological alternative in the opposite gender that could reinforce or even entirely drive any gender bias.

Despite these advantages, generic hybrid nouns in French have to our knowledge only been investigated in a single study by Brauer & Landry (2008: Study 3), using a single pair of hybrid nouns (*un individu* ‘an.MASC individual’/*une personne* ‘a.FEM person’). In that study, participants were presented with a role noun in the masculine generic and asked to describe the typical person that does the corresponding job. For half of the participants, this person was referred to in the text with the feminine hybrid noun *personne*. For the other half, the masculine noun *individu* was used instead. The authors found a smaller proportion of female responses for participants in the condition with masculine *individu* (16.9\%) than for participants in the condition with feminine *personne* (30\%), in line with the hypothesis that gender inferences are biased by grammatical gender for hybrid nouns.

The present paper aims to follow up on Brauer & Landry (2008) using a larger set of hybrid nouns (14 pairs) to test whether the effect of grammatical gender generalizes beyond the pair *individu/persone*ne. Two online studies were carried out to test the hypothesis, using judgment data from French-speaking participants in France and in Switzerland. The use of judgment data is common in works evaluating grammatical biases in gender inferences (e.g. Gygax et al. 2008, Xiao, Strickland & Peperkamp 2022). The specific design used in the two studies follows Richy & Burnett (2021), where participants were asked to estimate the likelihood that a sentence refers to a man or a woman using a Likert scale. Section 2 presents the first study (Study 1). Section 3 presents a follow-up study (Study 2) that was run to address a methodological issue that came up in the first study. The preregistration, data and code for both studies are available in Storme & Delaloye-Saillen (2022b,a).

2 Study 1

2.1 Methods

2.1.1 Stimuli

Each grammatical gender (feminine, masculine) was represented by 14 generic hybrid nouns. Hybrid nouns were considered to have a generic interpretation by the authors if they could be predicated of both a man and a woman without contradiction. Feminine and masculine hybrid nouns were paired up together according to semantic similarity (e.g. *individu/persone*ne, *vedette/talent*). The nouns in each pair were included in the same carrier sentence, as shown in (3), in order to control for effects of gender stereotypes that could come from the sentential context. The stimuli are available in Storme & Delaloye-Saillen (2022b).

\(^{1}\)For instance, Jakobson (1971: p. 213) treats the masculine gender as semantically underspecified and therefore without any bias towards a male interpretation.
Examples of experimental items

a. Une vedette de la chanson a été invitée pour présider le jury.
   ‘A FEM pop star was invited to chair the jury.’

b. Un talent de la chanson a été invité pour présider le jury.
   ‘A MASC pop talent was invited to chair the jury.’

However gender stereotypes were not controlled for or balanced across noun pairs and sentential contexts. In other words, the specific noun pairs and sentential contexts chosen for the study might come with their own gender biases whose effects will combine with any effect of grammatical gender. Yet, through the comparison of minimal pairs such as (3a) and (3b), the study still makes it possible to assess the effect of grammatical gender at an equal gender-stereotype strength.

2.1.2 Study design

The study was a repeated-measurement experiment with a Latin square design. Two groups of experimental items were created, each one containing seven feminine hybrid nouns and seven masculine hybrid nouns. Each group featured only one among the two paired-up items. For instance, (3a) belonged to one group of items and (3b) to the other group. Participants were randomly assigned to one of the two groups and therefore saw seven feminine hybrid nouns and seven masculine hybrid nouns each. The order of presentation was randomized for each participant.

Participants were asked to guess three properties of the person referred to in the sentence: (i) their age, (ii) their gender, and (iii) their level of education. The question was carefully worded so as not contain any grammatical clue about the referent’s gender. Questions about the referent’s age and level of education were added in order to make the goal of the study harder to guess, following Richy & Burnett (2021). Participants used a seven-point Likert scale to answer all three questions, as shown in Figure 1. For the gender variable, participants were asked to estimate the likelihood that the sentence refers to a man or a woman, 1 indicating a highly confident ‘man’ response, 7 a highly confident ‘woman’ response, and 4 an equal likelihood of the referent being a man or a woman.

Figure 1: An example of item

Following Richy & Burnett (2021) again, experimental items were interspersed with filler items consisting of proper names, in particular (but not exclusively) gender-neutral proper names like Dominique. These fillers were also meant to divert the attention of participants from the research question. Each participant saw 14 experimental items and 20 filler items.

2.1.3 Participants

100 participants were recruited through mailing lists at French-speaking universities in Switzerland and in France. The study was carried out online using the LimeSurvey platform (LimeSurvey 2012). Each participant was randomly assigned to one of the two groups. Each group included 50 participants. Participants participated on a voluntary basis. They provided their informed consent to participate in the research and agreed to make their data available online. No sensitive information about participants was collected.

2.1.4 Data analysis

The judgment data were modeled using the ordinal cumulative model (Bürkner & Vuorre 2019: pp. 78-79). The cumulative model assumes that the observed ordinal response variable derives from the categorization of a latent
continuous unobserved variable. In the present study, the ordinal variable is the rating of the referent’s gender along the seven-point scale. The latent variable is the underlying continuum corresponding to the participant’s uncertainty about the referent’s gender. To model this categorization in the case of a seven-point Likert scale, the cumulative model assumes that there are six thresholds which partition the latent variable into seven ordered categories (1, 2, ..., 6, 7). The model provides estimates for the mean of the two grammatical genders and for the position of the six thresholds along the latent continuous variable. The reader is referred to Bürkner & Vuorre (2019) for further details.

A Bayesian approach was adopted (rather than a frequentist approach) for inferring the parameters of the ordinal regression. This choice was motivated by the fact that Bayesian inference yields outcomes that are intuitive and easy to interpret. In particular, it provides a posterior distribution for all the model’s parameters and combinations of parameter values given the data. This makes it very easy to test any hypothesis about the parameter values and about differences between parameter values. Also, Bayesian approaches virtually always converge to accurate values of the parameters (Liddell & Kruschke 2018). The model was fit using brms (Bürkner 2017) in R (R Core Team 2020).

The model included the fixed effect Grammatical gender and the maximal random-effects structure justified by the study’s design (Barr et al. 2013). There was:

- a by-participant random intercept and a by-participant random slope for Grammatical gender,
- a by-group random intercept and a by-group random slope for Grammatical gender (there were two groups of participants in the Latin square design),
- a by-item random intercept (corresponding to the effect of a specific sentence and pair of masculine/feminine hybrid nouns on the response variable, e.g. (3)),
- and a by-word random intercept (corresponding to the effect of specific nouns, e.g. vedette, talent, etc., on the response variable, independent of their grammatical gender, whose effect is captured by the variable Grammatical gender).

Feminine gender was set as the baseline level. In the analysis, we focus on the parameter $\beta$ that quantifies by how much the baseline level must be adjusted for masculine nouns. Due to the way the Likert scale was set up (1 corresponds to a highly-confident ‘man’ response and 7 a highly-confident ‘woman’ response), a negative value for $\beta$ indicates a stronger male bias for masculine nouns. Compelling evidence for a difference in the inferences triggered by feminine and masculine nouns was considered to be provided only in case zero was outside of the posterior 95% Credible Interval (CI) for $\beta$. Credible Intervals were obtained using the ETI (Equal-tailed Interval) method and the package bayestestR (Makowski, Ben-Shachar & Lüdecke 2019).

### 2.2 Results

Figure 2 shows the posterior probability of each response along the seven-point Likert scale for feminine and masculine hybrid nouns. The effect of grammatical gender on gender inferences goes in the expected direction, with masculine nouns inducing more male representations than feminine nouns ($\beta = -0.85, CI = [-2.50, 1.25]$). But the 95% CI is too large (it includes zero) to conclude for a significant effect.

To get a better understanding of why the Credible Interval was so large, a post-hoc comparison was carried out within each group separately (Group 1, Group 2). Figure 3 shows the posterior density of gender inferences associated with masculine and feminine nouns within each group separately, using the latent continuous scale to make the interpretation easier. The lower and upper thresholds for the gender-neutral response (vertical solid lines) correspond to the posterior mean of the threshold between responses 3 and 4 along the ordinal scale and to the posterior mean of the threshold between responses 4 and 5, respectively.

This post-hoc study reveals that grammatical gender does have a significant effect on gender inferences within each group separately, with masculine nouns inducing more male representations than feminine hybrid nouns (Group 1: $\beta = -1.05, CI = [-1.63, -0.49]$; Group 2: $\beta = -0.87, CI = [-1.45, -0.27]$). The CIs for the within-group differences are smaller than the CI for the across-group difference. The question why a significant effect was found when considering each group separately but not when considering them together will be taken up in Section 2.3.

Inspection of Figures 2 and 3 reveals that the interpretation bias induced by feminine and masculine nouns is asymmetric. Masculine nouns favored a male interpretation whereas feminine nouns did not favor female interpretations but were interpreted as gender-neutral. This is particularly clear in Figure 3: the posterior distribution for the interpretation of feminine nouns is largely included within the range of gender-neutral values whereas the posterior distribution for
Figure 2: The effect of grammatical gender (feminine, masculine) on gender inferences in Study 1 (1: categorically male response; 4: gender-neutral response; 7: categorically female response). Posterior means with 95% CI.

Figure 3: The effect of grammatical gender (feminine, masculine) on gender inferences in Study 1 (represented along the latent continuous scale). Vertical dashed lines indicate means of the corresponding posterior distributions. The two vertical solid lines correspond to the lower and upper thresholds for the gender-neutral response along the seven-point Likert scale.
the interpretation of masculine nouns is shifted towards male values. Inspection of Figure 3 also reveals a difference between Group 1 and Group 2: gender inferences were less biased towards male interpretations in Group 2 than in Group 1, in particular for masculine nouns.

2.3 Discussion

Grammatical gender was found to affect patterns of gender inferences only when focusing on each group separately but not when considering them together. Also, the Credible Interval for the difference between feminine and masculine nouns was much larger when considering the two groups together than when considering them separately. This suggests that there are substantial differences between the two groups. This was confirmed by the exploratory analysis that revealed overall less male-biased inferences in Group 2 than in Group 1.

In light of the post-hoc comparison of Group 1 and Group 2 (Figure 3), we propose the following interpretation for the absence of overall effect of grammatical gender. Due to the way the random-effects structure was set up, the fixed effect for grammatical gender effectively compared feminine nouns in Group 1 with the corresponding masculine nouns in Group 2 (e.g. *une star* and *un as*) and masculine nouns in Group 1 with the corresponding feminine nouns in Group 2 (e.g. *un talent* and *une vedette*). But because feminine nouns were rated lower (= stronger male bias) in Group 1 than in Group 2 and masculine nouns were rated higher (= less strong male bias) in Group 2 than in Group 1, the difference between feminine nouns in Group 1 and masculine nouns in Group 2 ended up being smaller than the difference between masculine and feminine nouns within each group (see Figure 3). On the other hand, the difference between masculine nouns in Group 1 and feminine nouns in Group 2 ended up being larger than the difference between masculine and feminine nouns within each group (see Figure 3). Overall, the presence of a stronger male bias in Group 1 than in Group 2 therefore resulted in more variability and in a larger Credible Interval across groups than within groups.

There are two potential sources for the difference observed between Group 1 and Group 2. The two groups included different hybrid nouns. By chance, the nouns selected for Group 1 could have been more male-biased than the ones selected for Group 2. The two groups also included different participants. The participants in Group 1 could have been more male-biased by chance than the participants in Group 2. However, the first explanation seems more likely because there were less nouns to choose from (28 nouns) than participants (100 participants). If this interpretation is correct, then the absence of effect of grammatical gender across the two groups should be an artefact of the specific way items were grouped. To test this, a second study with a different assignment of items to participants was run and will be presented in Section 3.

The results also revealed asymmetric biases for masculine and feminine nouns. Only masculine nouns had a grammatically biased interpretation. Feminine nouns were largely interpreted as gender-neutral. This result is reminiscent of Brauer & Landry (2008: Study 3). They also found that masculine *individu* had a male-biased interpretation (16.9% female responses) whereas *personne* was not female-biased (30% female responses).

However this asymmetry should be interpreted with caution in both the present study and Brauer & Landry (2008)’s study. In Brauer & Landry (2008)’s study, hybrid nouns *individu* and *personne* were presented alongside masculine generics. The presence of masculine generics could have favored male interpretations overall. In the present study, gender stereotypes were not controlled for across the set of experimental items, as indicated in Section 2.1.1. If the specific set of hybrid nouns and/or carrier sentences chosen for the study happened to have been male-biased, then this could explain the asymmetry between masculine and feminine gender, with inferences being generally shifted towards male interpretations. In particular, several pairs of hybrid nouns chosen in the study are associated with high social status and prestige (*vedette/talent, staras, sommité/génie, célèbrité/monument*). Since there are typically more men than women that occupy high social positions in France and Switzerland, the choice of such words could result in a male bias. We leave for future research the investigation of how gender stereotypes might interact with grammatical gender for hybrid nouns.

3 Study 2

A follow-up study was run to address the issue of the heterogeneity of the two groups observed in Study 1.

---

2The male bias observed in Brauer & Landry (2008) is unlikely to be due to uncontrolled gender stereotypes because they only included role names corresponding to jobs featuring an equal number of men and women according to the French national institute of statistics (e.g. *enseignants de lycée* "high school teachers.MASC").
3.1 Methods

A single change was made in the methods used in Study 1. Instead of being assigned to one of two preestablished groups of experimental items, participants in Study 2 were assigned to a set of seven masculine and seven feminine nouns that were randomly selected for each participant. This change in the design made it possible to remove the Group variable that was problematic in Study 1. Otherwise the same materials and experimental design were used as in Study 1.

60 participants were recruited through mailing lists in French-speaking universities in France and Switzerland and through the CNRS mailing list RISC. The same methods were used for data analysis as in Study 1. The only difference with Study 1 was the absence of random effects for Group (because participants were no longer assigned to groups in Study 2).

3.2 Results

Figure 4 shows the posterior probability of each response along the seven-point Likert scale for feminine and masculine hybrid nouns. This time, the effect of grammatical gender on gender inferences came out significant according to our inference criteria, with masculine nouns inducing significantly more male representations than feminine hybrid nouns ($\beta = -0.94, CI = [-1.41, -0.45]$).

Figure 5 shows the posterior density of gender inferences associated with masculine and feminine nouns in Study 2, plotted along the latent continuous scale. Inspection of Figures 4 and 5 reveals that the interpretation bias induced by feminine and masculine nouns is asymmetric, as in Study 1. This is particularly clear in Figure 5: the value for the interpretation of feminine nouns is included within the range of gender-neutral values whereas the posterior mean for masculine nouns is outside of this range.

3.3 Discussion

Grammatical gender came out as significant on average when the group variable used in Study 1 was removed in Study 2. This suggests that the grouping of items was responsible for the null result observed across groups in Study 1, as hypothesized in Section 2.3.

---

\[3\] This value is constant and equal to zero because Grammatical gender was contrast-coded, with ‘feminine’ being used as the baseline level.
The results also revealed asymmetric biases for masculine and feminine nouns, with only masculine nouns resulting in a grammatically-biased interpretation. However this asymmetric bias should be taken with a grain of salt as it could be due to uncontrolled effects of gender stereotypes in the stimuli (see Section 2.3).

4 Conclusion

The current article has shown that hybrid nouns with a generic interpretation induce more male representations when their grammatical gender is masculine than feminine. This result confirms and generalizes the conclusion that was reached by Brauer & Landry (2008) based on a single pair of hybrid nouns (personne/individu). More generally, this research adds to a growing body of evidence that grammatical gender biases gender inferences. Crucially, the evidence comes from a type of nouns which do not inflect for gender (contrary to masculine generics) and therefore for which an alternative competition-based account is unlikely.

Only the masculine gender was found to bias gender inferences. Indeed, feminine nouns were interpreted as gender-neutral. However this latter result should be interpreted with caution as it could be due to uncontrolled effects of gender stereotypes coming from the specific stimuli used in the study. In particular, several pairs of hybrid nouns included in the experimental items referred to individuals with high social status and this might have contributed to an overall male bias. Further studies are necessary to establish whether the unbiased interpretation of feminine nouns results from methodological issues or is genuine.

The results also have implications for gender-fair language. Indeed they suggest that the general male bias observed in languages such as French could be attenuated not only by replacing masculine generics by gender-neutral forms but also by using feminine instead of masculine generic hybrid nouns (e.g. une personne instead of un individu).

Acknowledgments

Thanks to Guillaume Guex, members of l’Éprouvette at the Université de Lausanne, and Heather Burnett for helpful discussion on this project. Thanks to joffm (from the Limesurvey forum) and Céline Restrepo Zea for technical help. We are very grateful to the following persons for helping us recruit participants: Claire Blandin, Lise Charles, Isabelle Charnavel, Lucile Davier, Alice Delorme, Letitia Ding, Suzanne Duval, Fiona Friedli, Gilles van Heems, Cecile Heim, Jérôme Jacquin, Adèle Jatteau, Sabine Kradolfer, Marine Lebail, Cédric Margot, Robin Meyer, Cédric Patin, Nicolas Pepin, Gilles Philippe, and Madeleine Séguier.
References


Storme, Benjamin & Laura Delaloye-Saillen. 2022a. *Follow-up: Grammatical Gender in French: Its Impact on the Interpretation of Social Gender in Generic Hybrid Nouns*. [Data set]. https://doi.org/10.17605/OSF.IO/5XDBU.
