

Failures of Gricean reasoning and the role of stereotypes in the production of gender marking

Céline Pozniak and Heather Burnett

May 14, 2020

Abstract

We replicate von der Malsburg et al. (2020)’s recent experiments investigating the relationship between speaker beliefs, gender stereotypes and language use in English on a grammatical gender language: French. The results of our experiment show how the linguistic particularities of the English and French gender marking systems interact with speaker expectations and stereotypes to create different patterns of gender marking production. They also raise a puzzle for current theoretical and computational frameworks that formalize Gricean pragmatics, particularly those in which informativity (Gricean Quantity) is assumed to play a driving role in linguistic production.

Keywords: Language use, grammatical gender, French, Gricean reasoning, formal pragmatics, production experiment

1 Introduction

This squib has two aims. The first one is empirical: we compare the results of von der Malsburg et al. (2020)’s recent experiments investigating the relationship between speaker beliefs, English pronoun semantics, gender stereotypes and pronoun production with a new experiment investigating this relationship in a grammatical gender language: French. von der Malsburg et al. (2020) present a unique study aimed at relating speakers’ beliefs about the social gender of a referent and their pronoun use within the context of the 2016 US and 2017 UK elections. We attempt to replicate a portion of this study using data from the first round of the 2020 French municipal elections.

This squib’s second aim is theoretical: we argue that our results raise a puzzle for current theoretical and computational frameworks that formalize Gricean pragmatics (Grice, 1975), such as the *Rational Speech Act* (RSA) (Frank and Goodman, 2012), particularly those in which informativity (Gricean Quantity) is assumed to play a driving role in linguistic production (see also Franke, 2009). This puzzle is already implicit in von der Malsburg et al. (2020)’s work; however, we argue that it becomes even clearer and more pressing with the comparison to French.

The squib is laid out as follows: In section 2, we give an overview of von der Malsburg et al. (2020)’s studies on English in the context of the US and UK elections, and we outline the pragmatic puzzles that their results raise. Then, in section 3, we present our replication on the 2020 French municipal elections, and show how the puzzles also arise, albeit in a slightly

different form. Section 4 concludes with a discussion of implications of our results for formal pragmatics and the study of ideologies and gender systems.

2 Pronoun use in the US and UK elections

von der Malsburg et al. (2020) studied the relationship between participants' expectations for the next US president and UK prime minister, and their use and interpretation of English gender marked pronouns. In the days and weeks leading up to the 2016 and 2017 elections, they developed measures of participants' degrees of belief that the next president or prime minister would be female vs male. More specifically, in a belief estimation task, American participants were asked to report their expectations about who would win the election among the pool of candidates, and their degree of belief that a woman would win was calculated by comparing their expectation that Hilary Clinton would win with their expectations that Donald Trump or Bernie Sanders would win. At the same time, other participants completed sentences designed to elicit a pronoun referring to the next president/prime minister. In another self-paced reading task, other participants read sentences including a pronoun referring to the future president/prime minister. von der Malsburg et al. (2020) then studied how participants' expectations about the gender of the next leader of the US/UK from the belief estimation task mapped onto their use and interpretation of gendered pronouns from the production and self-paced reading tasks.

The authors of this study were interested in finding out whether the pronouns speakers use when referring to a woman in a traditionally male job could be affected by the ideological representations, such as stereotypes, that speakers associate with this job. If speakers' language use is guided by stereotypes in addition to (or instead of) their beliefs about the social gender of the referent, this could result in a usage pattern that is biased in favour of *he*. From a political perspective, this biased pattern would be problematic since it would contribute to further reinforcing the male stereotypes associated with political leadership, and, in doing so, ultimately contribute to reproducing gender inequalities through language. From the perspective of linguistic theory, this biased pattern would be puzzling, since having language use be determined by stereotypes, rather than guided by Gricean Maxims like Quality (truth telling) and Quantity (informativity) would appear to make the use of gender marking qualitatively different than the use of other kinds of linguistic expressions, like scalar items (Grice, 1975; Horn, 1984), color terms (Frank and Goodman, 2012), gradable adjectives (Lassiter and Goodman, 2017), action verbs (Bergen et al., 2016), among many others.

Based on the discussion in von der Malsburg et al. (2020), we can therefore oppose two competing hypotheses: (what we call) the **transparent hypothesis**, which states that "linguistic preferences transparently reflect event expectations" (p.2), and (what we call) the **stereotype hypothesis**, which proposes that gender-based stereotypes associated with the nouns *president* and *prime minister* play a role in addition to event expectations. The transparent hypothesis boils down to the idea that gendered pronoun use is Gricean (guided by truth, informativity and rationality); whereas, the stereotype hypothesis boils down to the idea that gender stereotypes can override truth and informativity, at least in some cases.

Testing the predictions of these two hypotheses requires making assumptions about the semantics of English pronouns, namely how gender marking in English is related to social gender. Although they briefly explore other options (p.12), von der Malsburg et al. (2020)

assume that the semantics of English pronouns is as in (1), where, following Cooper (1983); Heim and Kratzer (1998), we represent the semantic contribution of gender marking as a presupposition on the pronoun: pronouns denote functions who return an individual, provided that individual has the property stated in the gender presupposition.

- (1) $\llbracket \text{he} \rrbracket = \lambda x : x \in \llbracket \text{male} \rrbracket. x$
 $\llbracket \text{she} \rrbracket = \lambda x : x \in \llbracket \text{female} \rrbracket. x$
 $\llbracket \text{they} \rrbracket = \lambda x : x \in \llbracket \text{female} \rrbracket \cup \llbracket \text{male} \rrbracket. x$
 Inclusive forms (*(s)he, he or she, he/she* etc.): $\llbracket \text{she/he} \rrbracket = \lambda x : x \in \llbracket \text{female} \rrbracket \cup \llbracket \text{male} \rrbracket. x$

The predictions of the transparent hypothesis, given (1), can be made explicit using the architecture of the *Rational Speech Act* model (Frank and Goodman, 2012; Scontras et al., 2018)¹. Within this framework, it is easy to show that the transparent hypothesis predicts that participants should use the most informative (in the sense of Shannon (1948)) pronoun to communicate their belief about the social gender of the referent. More specifically, since the semantic denotation of *he* includes only men, *he* is the most informative pronoun for male social gender. Therefore, participants are predicted to favour *he* when they are (almost) certain that the future president/prime minister will be male. Since the denotation of *she* includes only women, participants are predicted to favour the feminine pronoun when they are (almost) certain that the referent will be female. Since the semantic denotations of *they* and inclusive forms include both men and women, depending on how the transparent hypothesis is formalized, participants may be predicted to favour *they* or inclusive forms (which von der Malsburg et al. (2020) call *gender hedged forms*) when they are uncertain².

The stereotype hypothesis, on the other hand, predicts that participants should take into account something else besides truth and informativity when choosing which pronouns to use. This hypothesis predicts that they should deviate from the informativity-driven transparent pattern in a way that corresponds to the stereotypes that they associate with *president* or *prime minister*. Following previous work on gender stereotypes in language (Duffy and Keir, 2004; Foertsch and Gernsbacher, 1997; Misersky et al., 2014; Garnham et al., 2015, among others), von der Malsburg et al. (2020) further assume that speakers’ gender stereotypes are not random, but are a product of experience. This experience can be direct, i.e. one can develop gendered mental representations associated with a noun like *nurse* through observing that the majority of nurses that one interacts with are female (Garnham et al., 2015). Alternatively, this experience can be discursive, i.e. one can develop gendered mental representations from listening to the way that others talk about members of professions or other social categories³. Given that stereotypes are assumed to be grounded in experience, the stereotype hypothesis

¹Formal implementations of the transparent hypothesis are made available on the OSF repository: https://osf.io/9fn2j/?view_only=317da82f8d504daaabcfe474cb723df2

²The predictions for the use of *they* and inclusive forms depend a bit on how speaker uncertainty is modelled: if “uncertain” is treated as a separate category from “male belief” and “female belief”, then the transparent hypothesis indeed predicts that *they* should be used under uncertainty. On the other hand, if uncertainty is treated as hesitation between “male belief” and “female belief”, then, since *they* is always less informative than *he* or *she* for a particular social gender, *they*/inclusive forms are predicted to be very rare. These points are outlined in the code on OSF.

³This can sometimes result in stereotypes that have little basis in reality (see Cameron (2007); Cameron and Shaw (2016) for discussion).

also predicts that participants who have direct or discursive experience with both male and female leaders will develop different stereotypes from those whose leaders have always been male. By this logic, since the US has never had a female president, participants in this country probably have a strong male stereotype for *president*, and deviation from the transparent pattern is predicted to be in favour of *he* in this country. The UK, on the other hand, has had a high profile female prime minister (Margaret Thatcher), and the incumbent, Theresa May, is also a woman. So UK participants probably have less strongly male stereotypes for their country's leaders than American participants. Therefore, the stereotype hypothesis predicts that deviation from the transparent pattern should be less favourable to *he* in the UK than in the US, and possibly include gender neutral forms like *they*.

The predictions of the stereotype hypothesis were largely borne out in the results. In the UK study, von der Malsburg et al. (2020) did not do a belief estimation task, since they assumed that everyone would have a high degree of belief (say at least 0.8) that Theresa May would win. With the British participants, as predicted by the transparent hypothesis, the proportion of *she* is higher than *he* at ≥ 0.8 degree of belief. This result gives suggests that gendered pronoun use is, at least at some level, guided by Gricean principles. However, von der Malsburg et al. found that the form that is favoured at the highest degree of female belief is *they* (the least gender informative expression in the English pronominal paradigm); whereas, the transparent hypothesis predicts it should be *she*, since *she* is the most informative for female social gender. On the other hand, the deviation from the transparent pattern in the UK data to the benefit of *they* could be in line with the stereotype hypothesis, on the assumption that the stereotype UK participants associate with *prime minister* is gender neutral, thanks to prime ministers like Margaret Thatcher and Theresa May. Thus, von der Malsburg et al. (2020)'s results from the UK election present a first puzzle for formal pragmatics: **What are the pragmatic mechanisms that allow gender stereotypes to override informativity in English pronoun use?**

The results of the US experiment further complicate the puzzle. On the one hand, von der Malsburg et al. (2020) found that American participants' beliefs about the social gender of the next president did play a large role in their pronoun use: as participants' belief that the next president will be female rose, the proportion of their use of *he* declined. This again suggests that, at some basic level, informativity-based reasoning does underly English pronoun production. However, the authors also found that, as participants' expectations of a female president rose, the proportion of *she* did not rise. In fact, unlike in the UK study, *he* remained a productive choice for participants even when they thought that it was likely that the next president would be female. Indeed, at all degrees of belief, use of *he* remained higher than *she* in the American study, something which violates both Gricean Quantity and Quality. The increased use of *he* in the US compared to the UK is predicted by the stereotype hypothesis: since Americans have a strong male stereotype for president, this stereotype can result in the production of *he* even at high degrees of female belief. This raises a puzzle for our pragmatic models which is parallel to the one raised by the UK data: *What are the mechanisms that allow gender stereotypes to override informativity (and perhaps even truth)?*

The US study showed an additional pattern that complicates the puzzles even more: as expectation in a female president rose, the pronoun whose rate increased was not *she*, but rather *they*. This is unexpected under both the transparent hypothesis and a simple stereotype hypothesis. If (1) is correct, *they* is the least informative expression in the English pronoun

system, and yet it appears to be the optimal way in which American participants choose to express beliefs about a future female president. This is unlikely to be due to a stereotype effect, since the strong male stereotype is presumably what generates the high rate of *he* in the data. Instead what seems to be going on here is that the rise of *they* is an interaction between stereotypes and beliefs: when participants have a high degree of belief that the referent is female with a strongly male stereotype noun, they need to combine these ‘conflicting’ beliefs together. And apparently *they* is the pronoun that participants find optimal to resolve this conflict. This introduces the additional puzzle: **What are the pragmatic mechanisms that allow the interaction between beliefs and stereotypes to override informativity?**

In summary, von der Malsburg et al. (2020) have shown us a number of interesting generalizations and puzzles about the relation between beliefs, stereotypes and English pronoun use:

- Participants’ beliefs about a referent play a role in determining use of gender marked expressions. This is straightforwardly predicted within Gricean pragmatics and its formalizations such as the RSA.
- Stereotypes also play a role in determining use of gendered expressions, and people sometimes use the pronoun associated with the noun’s stereotype, rather than their belief. This is not straightforwardly predicted by Gricean pragmatics, and requires some other mechanism allowing ideological constructs like stereotypes to override informativity.
- When beliefs and stereotypes are in conflict (eg. female belief and male stereotype), participants choose a gender neutral form. This is not straightforwardly predicted by Gricean pragmatics, nor is it predicted if one simply adds stereotypes. Accounting for this pattern requires some other mechanism allowing the combination of beliefs and stereotypes to override informativity.

von der Malsburg et al. (2020)’s study strongly suggests that we need to integrate gender stereotypes into our models of gendered pronoun use. The question is how this should be done. Are the patterns that we see the product of general reasoning, or are they language specific?

In order to investigate these questions, we turn to our replication of this experiment on French.

3 Gender marking in the French elections

French is a grammatical gender language; therefore, the mapping relations between grammatical gender and social gender are a bit different than in English. The first thing to note is that the question of whether there is a reliable relationship between grammatical gender and social gender is, itself, rather controversial. The view of traditional grammarians, such as the Académie Française (1984, 2004) and Grevisse-Goosse (2008), has been that grammatical gender does not reliably indicate social gender. On the other hand, feminist qualitative researchers have argued the contrary: that, in the vast majority of cases, noun phrases with masculine marking refer to men, while those with feminine marking refer to women (Violi, 1987; Michard, 1996; Houdebine, 1998, among others).

In the past 10 years, a significant body of research in psycholinguistics has developed investigating this question, and the consensus that emerges from this work is that feminine marking

reliably maps to female gender and masculine marking maps to male gender; however, this mapping is probabilistic (see Brauer, 2008; Gygax et al., 2008, 2012, among others). Building on the psycholinguistic work, we will therefore assume that the mapping between French grammatical and social gender (for human nouns) is as in (2): masculine gender maps to male social gender with a high (but not total) probability; whereas, feminine gender maps more consistently to female social gender. Written French also has a wide variety of inclusive forms (*écriture inclusive*) for noun phrases (*le/la maire*, etc.) and pronouns (*il ou elle*, *il/elle*, etc.). Although (to our knowledge) there has been no psycholinguistic research on the interpretation of these forms, they presumably map to both men and women.

- (2) Let N be a (pro)noun phrase,
- a. $\llbracket N_M \rrbracket = \begin{cases} \lambda x : x \in \llbracket \text{male} \rrbracket. \llbracket N \rrbracket(x) & p > 0.5 \text{ of the time} \\ \lambda x : x \in \llbracket \text{female} \rrbracket \cup \llbracket \text{male} \rrbracket. \llbracket N \rrbracket(x) & 1 - p \text{ of the time} \end{cases}$
 - b. $\llbracket N_F \rrbracket = \lambda x : x \in \llbracket \text{female} \rrbracket. \llbracket N \rrbracket(x)$
 - c. $\llbracket N_{INC} \rrbracket = \lambda x : x \in \llbracket \text{female} \rrbracket \cup \llbracket \text{male} \rrbracket. \llbracket N \rrbracket(x)$

Given (2), we can now formulate the predictions of the transparent and stereotype hypotheses for French, set within the *Rational Speech Act* model. The transparent hypothesis predicts that masculine should be used when the participants believe the referent is male, and that the proportion of feminine should rise as belief in a female referent also rises. (2) does allow for the possibility that an expression with masculine marking can be used to refer to a woman, so it is consistent with masculine being used when participants are (almost) certain that the referent is female. However, because masculine grammatical gender is male-biased, this means that feminine grammatical gender is a more informative signal for female social gender. Therefore, even though the semantics of the French gender marking system in (2) does allow for some masculine to be used at high degrees of female belief, the transparent hypothesis predicts that the rate of masculine should not exceed the rate of feminine in contexts where participants are certain that the referent is female.

Similar to English, the stereotype hypothesis predicts that gender stereotypes associated with mayors should play a role in the production of gender marked expressions on top of participants' beliefs. Again, since these stereotypes are based on experience, participants from different cities with different electoral histories are expected to have different stereotypes. In Marseille, the current mayor, Jean Claude Gaudin, is not seeking reelection because he is retiring after 25 years as the city's mayor. His former depute, Martine Vassal, was favoured to win leading up to the first round of the election⁴. In Paris, the incumbent is a woman: Anne Hidalgo. She was favoured to win leading up to the election⁵. She has also made gender and sexuality a very salient aspect of her first term, holding public consultations with LGBT activist groups, hosting the 2019 *Gay Games* in Paris, and using inclusive forms on official signs in city hall. Given these different histories, Parisians are well placed to have a less strongly male stereotype for the leader of their city than Marseillais, so the stereotype hypothesis predicts more deviation from the transparent pattern in favour of the masculine in Marseille than in Paris.

Finally, we would like to know whether the French inclusive forms show the behaviour shown

⁴<https://www.lesechos.fr/elections/municipales/municipales-2020-ce-que-disent-les-derniers-sondages-1171347>

⁵<https://www.lesechos.fr/elections/municipales/municipales-2020-ce-que-disent-les-derniers-sondages-1171347>

by English *they* at high degrees of female belief: if using a gender neutral or inclusive expression is a general cognitive strategy to reconcile female expectation with male stereotype, we should expect to find an increase in inclusive forms as belief in a female reference increases, particularly in Marseille. This being said, French *écriture inclusive* and English *they* have very different histories: gender neutral singular *they* has been widely used in English since the 14th century (Curzan, 2003) and, in the syntactic context studied in this paper, has little social meaning. French inclusive forms, on the other hand, appear to be innovations of the second half of the 20th century, becoming more widespread at the beginning of the 21st (see Abbou et al. (2018) for more information). They are still not universally accepted in written French and often communicate the political orientations of those who use them (Abbou, 2017). So it remains to be seen whether we will find the same pattern in French as von der Malsburg et al. (2020) found in English.

3.1 Experiment

We replicated the continuation experiment from von der Malsburg et al. (2020) on the first round of the 2020 municipal elections in Paris and Marseille. French elections usually have two rounds: a first one with all the candidates, and then, if no single candidate gets more than 50% of the vote, a second run-off round with the top two candidates is held. The first round of municipal elections was held on March 15th, 2020. Because of the coronavirus pandemic, the second round was postponed and, at the time of writing, it has not yet been announced when it will be held.

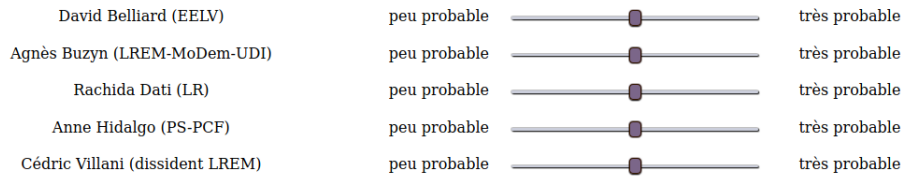
3.1.1 Design & materials

The experiment was the same in Paris and in Marseille (only the names of the candidates and the city were changed). Inspired by von der Malsburg et al. (2020), our experiment consisted of two parts: a completion task (a context sentence ((3) or (4)), followed by a sentence to complete, (5)-(9)), and an estimation of the probability of winning the municipal elections for the five most popular candidates on a 11-point level slider (figures 1 and 2).

- (3) Les élections municipales de mars 2020 vont déterminer qui dirigera la ville de Paris.
'The municipal elections of March 2020 will decide who will govern the city of Paris.'
- (4) Les élections municipales de mars 2020 vont déterminer la personne qui dirigera la ville de Paris.
'The municipal elections of March 2020 will decide the person who will govern the city of Paris.'
- (5) Même si son pouvoir n'est pas absolu,
'Even if his/her power is not absolute,'
- (6) Une fois au poste,
'Once in office'
- (7) Pendant sa première année,
'During his/her first year,'

- (8) Le lendemain de son élection,
‘The following day of his/her election,’
- (9) Pour remercier son équipe politique,
‘To thank his/her political team,’

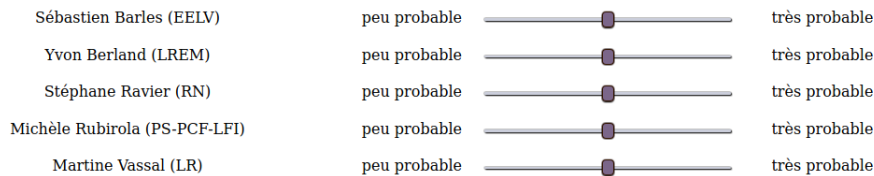
À votre avis, pour les élections municipales de Paris 2020, quelle est la probabilité que les personnes suivantes gagnent les élections ?



[– Cliquez ici pour continuer](#)

Figure 1 – Probability estimation presented in the experiment for Paris

À votre avis, pour les élections municipales de Marseille 2020, quelle est la probabilité que les personnes suivantes gagnent les élections ?



[– Cliquez ici pour continuer](#)

Figure 2 – Probability estimation presented in the experiment for Marseille

3.1.2 Participants

144 participants (mean age: 34 years old, $\sigma=13$) from the region of Paris (Île-de-France) and 68 participants (mean age: 40 years old, $\sigma=12$) from the region of Marseille (Provence-Alpes-Côte d’Azur) did the experiment before the end of March 15th (8pm) on IbexFarm (Drummond, 2013)⁶. They were recruited via the Crowdpanel platform whose cost for one participant is 0.32 euro per minute (www.crowdpanel.io) and via social media (for 32 participants for the experiment about the Paris elections ⁷)⁸.

⁶8 participants from Paris and 6 from Marseille did the experiment twice, so we excluded their second participation.

⁷For these participants, there was no restriction about living in the region of Paris.

⁸Analyses with participants from the Crowdpanel platform only are available on the OSF repository: https://osf.io/9fn2j/?view_only=317da82f8d504daaabcfe474cb723df2

3.1.3 Procedure

Participants read one sentence and completed another one as they wished. Then, they estimated the probability of winning the election for five candidates in Paris (3 women, 2 men) or Marseille (2 women, 3 men). The experiment lasted around 4 minutes.

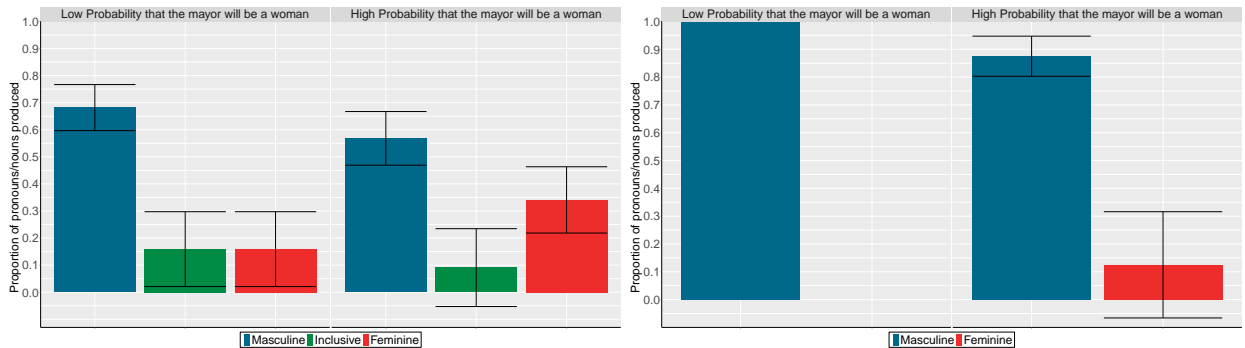
3.1.4 Results

Completion task

We excluded completions that weren't a sentence (only one word, *le maire*, *oui*, etc N=14 for Marseille and N=35 for Paris), and completions that didn't directly express candidate's gender (*la personne*, *la ville est forte*, etc N=6 for Marseille and N=29 for Paris). We then took into account completions about the mayor in the three possible grammatical gender forms (2), either DPs (*le*, *la maire*, even the candidate's name) or pronouns (*il*, *elle*). This led us to 48 tokens for Marseille and 88 tokens for Paris. Our final dataset is therefore unfortunately smaller than we originally expected; however, these results are still informative as to how participants' beliefs are related to their language use.

Figure 3 shows the proportion of the grammatical forms depending on the probability that the mayor will be a woman (by taking the median)⁹. It shows that masculine grammatical gender is dominant, especially in Marseille, with inclusive forms only appearing in Paris¹⁰.

Figure 3 – Production depending on grammatical gender for Paris (left panel) and Marseille (right panel)



Bayesian binomial regression model

We did Bayesian binomial regression models (Carpenter et al., 2017; Bürkner, 2017; Bürkner and Charpentier, 2018). The dependent variable, grammatical gender, was coded as 1 for

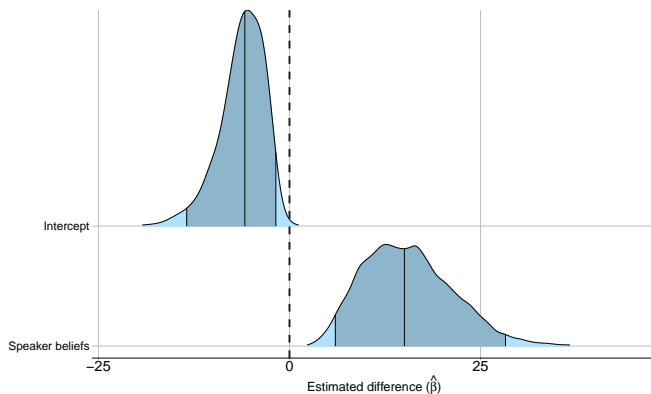
⁹Probability was calculated by adding female candidates' probability (3 for Paris and 2 for Marseille) divided by the total probability of all candidates.

¹⁰Results depending on context sentences are available on the OSF repository: https://osf.io/9fn2j/?view_only=317da82f8d504daaabcfe474cb723df2. Context sentences with *la personne qui* (the person that) might lead to more feminine grammatical gender forms, but it wouldn't expect the masculine grammatical gender dominance.

feminine and 0 for masculine ¹¹ Since the city variable and the speaker beliefs variable are highly correlated ($t = 11.087$, $p < 2.2e-16$, $cor = .73$), we only kept the speaker beliefs variable in the model and we applied mean centered coding. The model was run with 4 chains with 3000 iterations by chain. Weakly informative priors were used (normal(0,10)). Random variables were Participants and Items.

Besides the dominance of the masculine form, Bayesian binomial regression models showed an effect of speaker expectations ($\hat{\beta} = 15.56$, 95% CrI=[6, 28.26], $P(\hat{\beta}) > 0 = 1$, Figure 4): the more participants think that the mayor will be a woman the more they will use the feminine form. Furthermore, there was a correlation between city and speaker expectations: participants think a woman is more likely to win in Paris, so they use more feminine forms.

Figure 4 – Posterior distributions for the grammatical gender condition depending on speaker beliefs (95 % CrI)



4 Discussion and conclusion

Our study on French replicated a number of results found by von der Malsburg et al. (2020); however, we also found some differences. Our first main result is that, like in the English studies, we find an effect of speaker expectation of the social gender of the next mayor on their use of grammatical gender. In both Paris and Marseille, higher degree of belief in a female referent translates at least somewhat to more use of the feminine (and less use of the masculine). These results suggest that Gricean reasoning does underly use of gender marking in French.

However, our most striking result is the dominance of masculine gender, regardless of degree of belief. Like the US study, not only is masculine often used when participants think it's likely that the next mayor will be female, its rate exceeds that of the feminine even when participants are certain or almost certain that a woman will win. This pattern is not predicted by the transparent hypothesis, and actually even suggests that Gricean reasoning has been suspended here: even if French masculine gender can be used to refer to women, feminine grammatical gender is so much more informative to signal female social gender that any Gricean/rational/informativity-based theory of language use presumably predicts that

¹¹Because of the very number of Inclusive forms present only in Paris (N=11), we excluded this variable level.

feminine gender should at some point overtake masculine. However, this does not even come close to happening. To the extent that this ‘overuse’ of the masculine is driven by stereotypes associated with French mayors, it is predicted by the stereotype hypothesis. Coming back to the puzzles raised in von der Malsburg et al. (2020) (i.e. the pragmatic mechanisms allowing stereotypes to override informativity), results from our experiment in French provide an additional argument in favour of including gender stereotypes into our pragmatic models of the use of gender-marked expressions, although, in this squib, we leave how exactly to do this as a puzzle for future work. The fact that we find strong stereotype effects in both English and French opens the door to the possibility that the stereotype effect is cognitively general. However, of course this needs to be further investigated with crosslinguistic research.

Evaluating the predictions of the stereotype hypothesis brings us to the consideration of the differences between Paris and Marseille. As discussed in the previous section, we find that Parisians think that it is more likely that a woman will win than Marseillais¹². Because higher female belief is related to more feminine grammatical gender, this translates into more feminine in Paris than in Marseille, which is predicted by the transparent hypothesis. The stereotype hypothesis also predicts that, at equal degrees of belief, there should be more masculine in Marseille than in Paris since Marseillais presumably have a more male stereotype for mayor than Parisians. The differences in expectation between Paris and Marseille render evaluating this prediction difficult. If we look at the degrees of belief between 0.5-0.75 (i.e. where it is more likely than not that a woman will win, and for which we have data in both Marseille and Paris), there indeed appears to be slightly more masculine in Marseille than in Paris, as shown in Figure 5. However, these results should be taken with a grain of salt since the data within this slice are sparse.

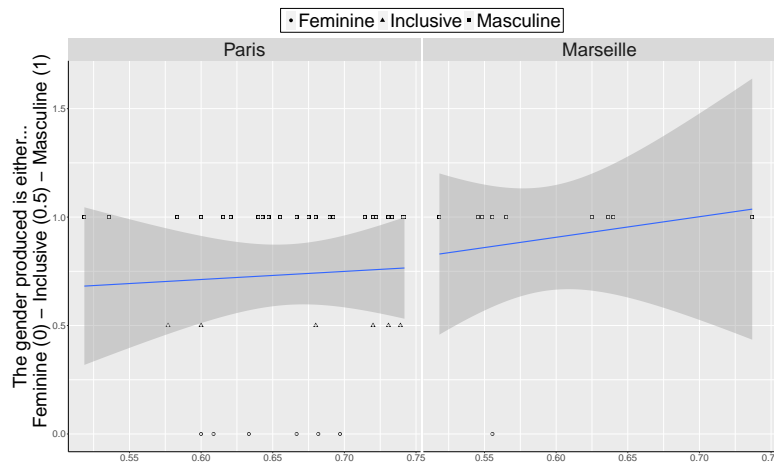


Figure 5 – Production of grammatical gender depending on belief that the mayor will be a woman (bw. 0.5 et 0.75, continuous variable)

Besides the differences in expectation, the most striking difference observed between Paris and Marseille is the inventory of forms used. While inclusive forms are absent from Marseille, we find a small amount of *écriture inclusive* in our Paris data. In line with the other puzzle (the

¹²Note that the results of the first round put female candidates in the top two spots in both Paris (Anne Hidalgo and Rachida Dati) and Marseille (Michèle Rubirola and Martine Vassal), so the current real probability that Paris and Marseille will have a female mayors is 1.

pragmatic mechanisms allowing the interaction between beliefs and stereotypes to override informativity), this pattern shows that French inclusive forms are not playing the same role in the gender marking system as English *they*: they do not increase as participants' expectations in a female mayor increase, meaning that inclusive forms are not a strategy used by French speakers to resolve a clash between a high degree of female belief and a strong male stereotype. Our study therefore suggests that, perhaps unlike the stereotype effect itself, the special way that beliefs and stereotypes interact to produce more gender inclusive/neutral forms is language specific.

Why *écriture inclusive* should be used in Paris and not Marseille is unclear. One possibility may be that Parisians have a more gender neutral stereotype for mayor and that the inclusive forms arise because of this stereotype. Another possibility has to do with the salience of the inclusive forms themselves in Paris. As discussed in section 3, the incumbent Anne Hidalgo is very vocal about issues related to gender and sexuality, and, under her direction, city hall uses inclusive forms in many official contexts. Studies of other French cities (for example, when the second round of the municipal election happens) would be desirable to see whether *écriture inclusive* is a Parisian exception.

In conclusion, while some of the generalizations found by von der Malsburg et al. (2020) also characterize our French results, our cross-linguistic comparison highlights how the linguistic particularities of the English and French gender marking systems interact with speaker expectations and stereotype mental representations to create different patterns of production of gender marked expressions. Our study also makes the theoretical puzzles raised in von der Malsburg et al. (2020)'s English results more pressing since we show that the ability of stereotypes interfere with Gricean reasoning is not specific to English. More generally, our study and von der Malsburg's together argue in favour of incorporating ideological concepts like gender stereotypes into our pragmatic models and having language specific patterns arise from the combining stereotypes, Gricean reasoning and language specific gender marking inventories. However, how exactly this should be done is left to future work.

Supplementary files

Code, scripts and data from the production experiment are available on the OSF repository: https://osf.io/9fn2j/?view_only=317da82f8d504daaabcfe474cb723df2.

References

- Abbou, J. (2017). (Typo) graphies anarchistes. ou le genre revele l'espace politique de la langue. *Mots. Les langages du politique*, 1:53–72.
- Abbou, J., Arnold, A., Candea, M., and Marignier, N. (2018). Qui a peur de l'écriture inclusive? entre délire eschatologique et peur d'émascation entretien. *Semen. Revue de sémio-linguistique des textes et discours*, 44. <https://doi.org/10.4000/semen.10800>.
- Bergen, L., Levy, R., and Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, 9.

- Brauer, M. (2008). Un ministre peut-il tomber enceinte? l'impact du générique masculin sur les représentations mentales. *L'Année psychologique*, 108(2):243–272.
- Bürkner, P.-C. (2017). Advanced bayesian multilevel modeling with the r package brms. *arXiv preprint arXiv:1705.11123*.
- Bürkner, P.-C. and Charpentier, E. (2018). Modeling monotonic effects of ordinal predictors in regression models.
- Cameron, D. (2007). *The myth of Mars and Venus*. OUP Oxford.
- Cameron, D. and Shaw, S. (2016). *Gender, power and political speech: Women and language in the 2015 UK General Election*. Springer.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1). doi:110.18637/jss.v076.i01.
- Cooper, R. (1983). *Quantification and Semantic Theory*. Reidel, Dordrecht.
- Curzan, A. (2003). *Gender shifts in the history of English*. Cambridge University Press.
- Drummond, A. (2013). Ibex farm. *Online server: <http://spellout.net/ibexfarm>*.
- Duffy, S. A. and Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, 32(4):551–559. <https://doi.org/10.3758/BF03195846>.
- Foertsch, J. and Gernsbacher, M. A. (1997). In search of gender neutrality: Is singular they a cognitively efficient substitute for generic he? *Psychological science*, 8(2):106–111. <https://doi.org/10.1111/j.1467-9280.1997.tb00691.x>.
- Frank, M. C. and Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.
- Franke, M. (2009). *Signal to act: Game theory in pragmatics*. Institute for Logic, Language and Computation Amsterdam.
- Garnham, A., Doehren, S., and Gygax, P. (2015). True gender ratios and stereotype rating norms. *Frontiers in Psychology*, page 1023. <https://doi.org/10.3389/fpsyg.2015.01023>.
- Grevisse-Goosse, A. (2008). Le bon usage. *Bruzelles: De Boeck-Duculot*.
- Grice, P. (1975). Logic and conversation. *Syntax and Semantics*, 3:41–58. https://doi.org/10.1163/9789004368811_003.
- Gygax, P., Gabriel, U., Lévy, A., Pool, E., Grivel, M., and Pedrazzini, E. (2012). The masculine form and its competing interpretations in french: When linking grammatically masculine role names to female referents is difficult. *Journal of Cognitive Psychology*, 24(4):395–408. <https://doi.org/10.1080/20445911.2011.642858>.
- Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., and 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>.
- Heim, I. and Kratzer, A. (1998). *Semantics in generative grammar*, volume 1185. Blackwell Oxford.
- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Schiffrin, D., editor, *Georgetown University Roundtable on Languages and Linguistics*, pages 11–42. Georgetown University Press, Washington DC.
- Houdebine, A.-M. (1998). La féminisation des noms de métiers. *Paris: Harmattan*.
- Lassiter, D. and Goodman, N. D. (2017). Adjectival vagueness in a bayesian model of interpretation. *Synthese*, 194(10):3801–3836.
- Michard, C. (1996). Genre et sexe en linguistique: les analyses du masculin générique. *Mots. Les langages du politique*, 49:29–47.
- Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., Chiarini, T., Englund, K., Hanulíková, A., Öttl, A., et al. (2014). Norms on the gender perception of role nouns in czech, english, french, german, italian, norwegian, and slovak. *Behavior research methods*, 46(3):841–871. <https://doi.org/10.3758/s13428-013-0409-z>.
- Scontras, G., Tessler, M.-H., and Franke, M. (2018). Probabilistic language understanding: An introduction to the Rational Speech Act framework. Retrieved 2020-5-12 from <https://www.problang.org>.
- Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Violi, P. (1987). Les origines du genre grammatical. *Langages*, 21(85):15–34.
- von der Malsburg, T., Poppels, T., and Levy, R. P. (2020). Implicit gender bias in linguistic descriptions for expected events: The cases of the 2016 US and 2017 UK election. *Psychological Science*. <https://doi.org/10.1177/0956797619890619>.