The male bias can be attenuated in reading: on the resolution of anaphoric expressions following gender-fair forms in French

Julia Tibblin, Centre for Languages and Literature, Lund University, Lund, Sweden, Julia.tibblin@rom.lu.se

Jonas Granfeldt, Centre for Languages and Literature, Lund University, Lund, Sweden, jonas.granfeldt@rom.lu.se

Joost van de Weijer, Humanities Lab, Lund University, Lund, Sweden, joost.van_de_weijer@humlab.lu.se

Pascal Gygax, Department of Psychology, University of Fribourg, Fribourg, Switzerland, pascal.gygax@unifr.ch

Despite the increased use of different types of gender-fair forms in French, studies investigating how they are interpreted when presented in a sentence remain few. To fill this gap, we conducted a pre-registered study using a timed sentence evaluation task to examine the possibility of speakers' establishing an anaphoric relationship between a gendered anaphoric expression (*femmes* 'women' or *hommes* 'men') and non-stereotyped role nouns as antecedents. The antecedents were presented in their masculine form or in one out of three different gender-fair forms (complete double forms: *les voisines et voisins* 'the neighbours.$_{FEM}$ and neighbours.$_{MASC}$', contracted double forms: *les voisin·es* 'the neighbours.$_{MASC·FEM}$', or gender-neutral forms: *le voisinage* 'the neighbourhood'). In line with previous findings, the masculine form led to a male bias in the participants' mental representations of gender. All three examined gender-fair forms resolved this bias, but comparisons of the different forms to each other revealed differences between them. The results show that complete double forms lead to equally balanced mental representations of gender while contracted double forms slightly favour representation of women. Finally, gender-neutral forms result in a small male bias, although significantly smaller than the one produced by the masculine form. The results are discussed in relation to the mental models theory and provide new and important insights on how gender-fair forms in French are interpreted.

1. Introduction

Over the past decades, speakers of different languages with grammatical gender, such as French, are increasingly using so called gender-fair language (Simon & Vanhal, 2022). By using gender-fair language, these speakers aim to include all genders when writing or speaking without using the masculine form as a default value. This is done by means of gender-fair forms (GFFs), which can either be explicitly inclusive using both the feminine and masculine forms, *les voisines et voisins* 'the neighbours.FEM and neighbours.MASC', or implicitly inclusive using collective nouns, *le voisinage* 'the.MASC neighbourhood', or common gender nouns, *les locataires* 'the tenants' (Simon & Vanhal, 2022). The problem that users of GFFs aim to resolve springs from the ambiguity of the masculine form in grammatical gender languages. In these languages, the masculine form often has two distinct interpretations, one specific and one generic, while the feminine form only has a specific interpretation. Thus, when a group consists of both women and men, or when one does not know or want to specify the gender of a person, it is the masculine form that should be used in French according to standard grammars (Melis & Godard, 2021). For example, one would use the masculine *les voisins* to generally refer to neighbours. The interpretation of this intended generic use of the masculine form[1] has been extensively investigated in numerous psycholinguistic studies conducted in several different languages (see Gygax et al., 2021, for an overview). The results consistently show that the masculine form, even if intended as generic, is more often interpreted as referring to men, even when the intention is to refer to both women and men.

While the use of French GFFs has increased (Simon & Vanhal, 2022), the body of literature investigating how these forms are interpreted remains sparse. Moreover, the few studies (Brauer & Landry, 2008; Richy & Burnett, 2021) that have investigated GFFs in discourse have mainly done so using questionnaires. In addition, out of the existing studies on the non-contextualized effect of French GFFs on the representation of women (Kim et al., 2022; Tibblin et al., 2023; Xiao et al., 2022), only Tibblin et al. (2023) included both double and gender-neutral forms. These authors did not find any differences between the representations of women that different GFFs led to, but since they studied the GFFs as isolated words with an untimed technique, whether such an effect would also surface spontaneously with a response timing method remains open. In sum, there is currently a lack of studies that both look at how different GFFs are processed in discourse and use timed tasks that can provide more fine-grained measures.

Against this backdrop, the aim of this pre-registered[2] study is to investigate the influence of GFFs (*les voisines et voisins, les voisin·es,* or *le voisinage*) on the resolution of the anaphoric

---

[1] The masculine form is sometimes referred to as 'the generic masculine' when it is used generically. Since the form of the noun does not differ depending on the two different uses of the masculine form, we refrain from this terminology and will instead use the term *the masculine form* henceforth.

[2] The pre-registration is available at the following link: https://osf.io/p4vkw.

expressions *les femmes* 'the women' or *les hommes* 'the men' as compared to the masculine form (*les voisins*). Anaphora and anaphora interpretation have been extensively studied both from a linguistic perspective (e.g. Ariel, 1990; Kamp, 1981; van Eijck & Kamp, 1997) and from a psycholinguistic perspective (e.g. Garnham, 2001; Garnham et al., 2012), which makes anaphora resolution a good testing ground for studies on the interpretation and processing of GFFs in reading. To this aim, we tested native French-speakers on a timed sentence evaluation task.

2. Background

2.1 French GFFs and their effects on mental representations of gender

In French, two main strategies are used to form GFFs: (a) the re-feminisation strategy that explicitly includes women and men by using double forms, and (b) the neutralisation strategy that implicitly includes all genders as it makes use of forms that lack gender markings corresponding to a person's gender (Elmiger, 2008; Simon & Vanhal, 2022). Double forms can be either complete, *les voisines et voisins* 'the neighbours.$_{FEM}$ and neighbours.$_{MASC}$' or contracted by means of different typographical signs, e.g. *les voisin·es* 'the neighbours.$_{MASC·FEM}$'. The neutralisation strategy mainly uses three different types of nouns: collective nouns, epicene nouns, and common gender nouns. Collective nouns, like *le voisinage* 'the neighbourhood', implicitly refer to all members of the group independently of their gender. Epicene and common gender nouns resemble each other but differ in one important aspect, namely, in how they handle grammatical gender (Corbett, 1991). On the one hand, epicene nouns can only take one grammatical gender, as *une personne* 'a.$_{FEM}$ person' or *un individu* 'an.$_{MASC}$ individual'. On the other hand, common gender nouns accept both masculine and feminine agreement (one can say *une locataire* 'a.$_{FEM}$ tenant' as well as *un locataire* 'a.$_{MASC}$ tenant'), but the form of the noun does not change. As French plural articles are not marked for grammatical gender, common gender nouns can be considered gender-neutral in their plural form, *les locataires* 'the tenants'.[3]

Both strategies have been studied empirically using different experimental methods with the aim of investigating their effects on speakers' mental representations of gender. As for the re-feminisation strategy, studies have repeatedly found that complete and contracted double forms increase representation of women in German (e.g., Braun et al., 2005; Hansen et al., 2016; Irmen & Roßberg, 2004; Körner et al., 2022; Schunack & Binanzer, 2022) and in French (e.g., Brauer & Landry, 2008; Tibblin et al., 2023; Xiao et al., 2022). The few existing studies

---

[3] In France and in Québec, the officially recommended strategy is a combined use of complete double forms and gender-neutral forms. However, the French Haut Conseil à l'Égalité entre les Femmes et les Hommes recommends the use of the contracted double forms in some cases, but solely as an abbreviation of the complete forms. See https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/guide_egacom_sans_stereotypes-2022-versionpublique-min.pdf and https://vitrinelinguistique.oqlf.gouv.qc.ca/banque-de-depannage-linguistique/la-redaction-et-la-communication/feminisation-et-redaction-epicene/redaction-epicene for more detailed recommendations for the use of each GFF in both written and spoken French.

on French gender-neutral forms indicate that these forms are successful in attenuating the male bias induced by the masculine form, at least as long as there is no interference with stereotype information (e.g., Kim et al., 2022; Richy & Burnett, 2021; Tibblin et al., 2023).

The studies that have examined German gender-neutral forms are more numerous. However, the results of these studies are mixed, as some studies found gender-neutral forms to resolve the male bias (Sato, Gabriel, et al., 2016; Stahlberg et al., 2001), while other studies found that these forms reproduce the male bias (Braun et al., 2005; Irmen, 2007; Irmen & Roßberg, 2004). There could be several explanations for this discrepancy. First, different methods (e.g., eye-tracking, self-paced reading, sentence evaluation task, questionnaires) and materials have been used. Second, the number of participants and the statistical analyses conducted do vary. This variation could lead to differences in statistical power, and it might therefore be difficult to compare the results of these studies.

Considering the varying results of the studies on gender-neutral forms, one could expect the double forms to be more effective in increasing representation of women compared to the gender-neutral forms. To the best of our knowledge, only three studies (Irmen & Roßberg, 2004; Stahlberg et al., 2001 (Exp. 1); Tibblin et al., 2023) have compared double and gender-neutral forms. While Irmen and Roßberg (2004) found significant differences between complete double forms and gender-neutral forms using common gender nouns in German, the two other studies did not find any statistically significant differences between double and gender-neutral forms, meaning that both GFFs were equally effective in attenuating the male bias induced by the masculine form (Stahlberg et al., 2001 (exp. 1); Tibblin et al., 2023).

2.1.1 Individual differences: Attitudes towards gender-fair language

Discrepancies across studies could also be caused by sample characteristics. In other words, some individual differences may be at the very base of different representational mechanisms. For example, one factor that could interact with the effect of GFFs on the representation of women is participants' attitudes towards gender-fair language. Generally, attitudes are considered to be a person's global evaluations towards an object (Perloff, 2003). A common method to measure attitudes towards gender-fair language is to use questionnaires including statements regarding its use and beliefs about it, and ask participants to indicate to which degree they agree or disagree with each statement (e.g., Parks & Roberton, 2000; Prentice, 1994; Sczesny et al., 2015; Tibblin, 2020).

Previous studies investigating how participants' attitudes towards gender-fair language influence their mental representations of gender are few, and the results are inconclusive. Braun et al. (2005), for example, found that the masculine form produced a stronger male bias among participants with positive attitudes towards gender-fair language compared to those having negative ones. The authors argued that participants having positive attitudes towards gender-fair language are more likely to use GFFs to refer to mixed-gender groups. Consequently, the masculine

forms would nearly entirely lose their generic sense, and more often and strongly be interpreted as specifically referring to men. However, other researchers found that participants' attitudes towards gender-fair language had a somewhat negative effect on the representation of women regardless of whether the participants were presented with a masculine form or a GFF. For example, in Tibblin et al. (2023), participants holding more negative attitudes towards gender-fair language generally perceived higher proportions of women in the occupations presented. According to the tentative explanation put forward by the authors, the participants with negative attitudes might believe that women are already sufficiently represented in society (i.e., estimate higher proportions of women), and thus think that gender-fair language is unnecessary (i.e., have more negative attitudes towards gender-fair language). Finally, Anaya-Ramírez et al. (2022) explored the influence of Spanish speakers' attitudes towards gender-fair language on the possibility of establishing an anaphoric relationship between a feminine anaphor and a stereotypically masculine role noun presented in its masculine form (e.g., *los camioneros* 'the lorry drivers'). Interestingly, they found that positive attitudes facilitated the establishment of this anaphoric relationship.

2.2 Mental representations and text processing

When considering anaphoric relationships, a broad and common assumption in text processing is that as we read, we construct mental models of the situation described in the text to maintain local and global coherence (e.g., Garnham, 2001; Graesser et al., 1994; Gygax & Gabriel, 2011). A mental model is a partial representation of a real or imaginary world that readers build while reading (Garnham, 2001). When reading, there are two main sources of information that influence the construction of our mental models: one is grounded in the text itself (bottom-up), and the other in our knowledge of the world[4] (top-down) (Gygax & Gabriel, 2011). Over the course of reading, the mental model will be updated as new information is presented in the text. These updates can be based on new information in the text itself, or on inferences made about the world represented in the mental model.

Thus, when constructing a mental model of *the nurses* in a sentence like *The nurses came out of the hospital*, both information present in the lexeme *nurses* (such as its lexical meaning and the plural mark signalling the presence of several nurses) and information drawing on background knowledge contribute to the mental model. Based on their background knowledge, readers might infer that the majority of the nurses are women, given that they live in a society where nurses are – or are believed to be – women more often than men. In a natural gender language like English, the surface form of the role noun *nurses* will be unlikely to influence the construction of its mental representation with respect to the nurses' gender, but this is not the case in grammatical gender languages like French. In these languages, the form of the role noun will influence how the group is mentally represented

---

[4] In the present paper we use, in line with Garnham (2001), the terms *knowledge of the world* and *background knowledge* interchangeably.

also with respect to gender (Carreiras et al., 1996). Reading *infirmières* 'nurses._{FEM}' should lead to a model representing only female nurses, while *infirmiers* 'nurse._{MASC}' could either lead to a model representing only male nurses (specific interpretation) or to a model representing a group of female and male nurses (generic interpretation). When comparing English speakers to French and German speakers, researchers (Gygax et al., 2008) found that, in French and German, the surface form of a role noun had a stronger influence than the stereotype information conveyed by the role noun, a factor that is based on the participants' knowledge of the world. In other words, when constructing their mental models, English readers relied on their background knowledge – or belief – of nurses often being women, while French and German-speaking readers were more influenced by the masculine form – resulting in a male bias in their mental representations of gender.

To further understand how anaphoric relationships work within mental models, the notational instruments from Reboul and Moeschler (1998) are particularly useful. Although it is generally assumed that mental models do include all elements pertinent to the situations depicted in the text (Garnham, 2001), Reboul and Moeschler (1998) offer tools and a terminology that enable us to make clear predictions as to the smaller entities – such as objects and people – included in readers' mental representations (MRs). These notions are useful when discussing anaphora resolution and provide a solid ground for understanding how gender information works in anaphoric resolution, which is central in the present paper. As such, two operations discussed by Reboul and Moeschler (1998) are particularly relevant to the current study, namely the *grouping* and the *extraction* operations. These operations are exemplified with the following sentences:

(1)   (Reboul & Moeschler, 1998, p. 138, translated from French)
   a.   [_{NP1} Un homme et   une femme] entrèrent. Ils   allèrent s'asseoir au
           A   man   and a   woman entered.   They went   sit   in + the
        fond du   bar.
        back of + the bar
        '[_{NP1} A man and a woman] came in. They sat down in the back of the bar.'
   b.   Jean avait [_{NP2} neuf billes].   Il   les   a   laissé tomber. Il   n'   en   a
        Jean had   [_{NP2} nine marbles]. He them have let   drop.   He _{NEG} some have
        retrouvé que huit. La   dernière avait roulé sous le   canapé.
        found   only eight. The last   had   rolled under the sofa.
        'John had [_{NP2} nine marbles]. He dropped them. He only found eight of them. The last one had rolled under the sofa.'

When reading NP1 in example (1a), two MRs are created: [@woman][5] and [@man], but when reading the pronoun *they*, a new MR [@woman&man] is created by the process of grouping the two former ones. Inversely, an extraction process starts with one MR from which two or more new MRs are extracted. As such, after reading NP2 in (1b), the MR [@marbles] is created. But

---

[5]   The brackets and the at sign are used by Reboul and Moeschler (1998) to label MRs.

as the reading continues, the reader must distinguish the marbles that were found from the one that was not. Therefore, two new MRs are extracted from the existing one: [@8marbles] and [@1marble]. Both grouping and extraction operations are involved when interpreting anaphoric relations, to which we will turn in the following section.

2.3 The effect of mental representations of gender on the possibility of resolving anaphoric expressions in French

Since anaphoric expressions take their meaning from an entity previously mentioned in the text, the creation and subsequent manipulations of MRs play an important role when processing and interpreting anaphoric expressions (Garnham, 2001). The investigation of anaphoric expressions, such as *she* in (2) below, can therefore provide us with useful insights when studying mental representations of gender.

(2)   Sally admired Bill's jacket and then she got one for Christmas (adapted from Garnham, 2001, p. 41).

In order to successfully resolve an anaphoric expression, readers must go through three stages. First, they must identify the anaphor, *she*. Second, they must determine the meaning of the anaphor with the help of linguistic cues about, for example, gender (feminine) and number (singular). After these two stages, readers can match the anaphoric expression to the correct antecedent, *Sally* (Garnham, 2001).

In psycholinguistic studies, the resolution of a gendered anaphoric expression with a possible antecedent presented in the masculine form has often been studied to investigate how the masculine form is interpreted by speakers of gendered languages. For example, sentence pairs like (3) have been used:

(3)   (adapted from Gygax et al., 2008, p. 472)

    a. $[_{NP3}$ Les voisins]   marchaient dans la gare.
        The   neighbours.$_{MASC}$ walked   in   the   station
        '$[_{NP3}$ The neighbours] were walking through the station.'

    b. Du   beau temps   étant prévu $[_{NP4}$ plusieurs femmes] n' avaient pas
        Some good   weather being forecast several   women $_{NEG}$ had $_{NEG}$
        de veste.
        $_{ART}$ coat
        'Since sunny weather was forecast $[_{NP4}$ several of the women] were not wearing a coat.'

The task for the participants is to evaluate whether sentence (3b) is a possible continuation from sentence (3a) or not. As we will see, this judgment relies on the possibility of readers' establishing an anaphoric relation between the subjects of the two sentences, a possibility that ultimately relies on the gender included in the MRs that are created during reading.
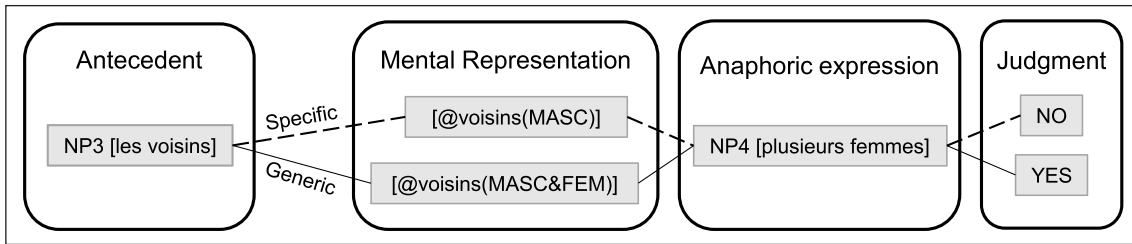
Figure 1: The process of interpreting the masculine form. The dashed line reflects a specific interpretation and the full line represents a generic interpretation.

Figure 1 shows how the two different interpretations of the masculine form would lead to two different MRs and, in turn, to two different outcomes in a sentence evaluation task with sentence pairs such as (3). As we can see, when reading (3), readers should, on the basis of NP3, first create a MR of a group of neighbours. Subsequently, when encountering the anaphoric expression (i.e., NP4) in (3b), they must recall and confront the MR created while reading NP3 in (3a) in order to assess if there is a possible antecedent. Since the two possible interpretations of the masculine form, the specific and the generic, differ in the gender-related information they contain, the interpretation of NP3 selected will have consequences for the possibility of resolving the anaphoric expression in NP4. If the readers have made a specific interpretation of NP3, the MR [@voisins(MASC)] should be activated. Since all the neighbours in this MR are men, it would not be possible to extract [@voisines(FEM)] from it, which would be needed in order to interpret NP4 as an anaphoric expression with NP3 as its antecedent. However, a generic interpretation of NP3 would activate the MR [@voisins(MASC&FEM)]. Since this MR contains both women and men, it should be possible to extract both [@voisines(FEM)] and [@voisins(MASC)] from it. In this case, when the reader encounters NP4 and is forced to create a new MR of the neighbours, it should be possible to make the anaphoric link back to NP3. In brief, a generic interpretation of *voisins* should allow for a quick and easy extraction of [@voisines(FEM)], requiring little cognitive effort. Inversely, a specific interpretation would not allow for a successful extraction process and, consequently, not allow for an establishment of an anaphoric link back to NP3.

(4)   [$_{NP5}$ Les voisines   et voisins]/   [$_{NP6}$ Les voisin·es]
        The neighbours.$_{FEM}$ and neighbours.$_{MASC}$/   The neighbours.$_{MASC·FEM}$
        /[$_{NP7}$ Les locataires] marchaient dans la gare.
        /   The tenants   walked   in   the station
        '[$_{NP5}$ The neighbours and neighbours]/[$_{NP6}$ The neighbours]/[$_{NP7}$ The tenants] were walking through the station.'

If GFFs are used instead of the masculine form, the anaphora resolution process can be predicted to be rather different. Using complete double forms, NP5 in example (4) should activate the

two separate MRs [@voisines(FEM)] and [@voisins(MASC)], analogous to example (1) above. Since one of the MRs created on the basis of NP5 corresponds perfectly to the information presented in the anaphoric expression, no operation needs to be executed on the MR and the anaphoric expression can be resolved effortlessly. As for the contracted double forms, these are morphologically abbreviations of the complete forms, and this is also in line with the use officially recommended in France (Haut Conseil à l'Égalité entre les Femmes et les Hommes, 2022). Therefore, the processing and interpretation of the contracted double forms, NP6 in example (4), should be identical to the interpretation of NP5 just described.

The last type of GFF examined in this study are the gender-neutral forms. These forms included common gender nouns like NP7 in example (4), collective nouns, and epicene nouns (see **Table 1** for examples). Since NP7 contains no explicit information regarding the referents' gender, the MRs created when reading NP7 should include both genders, and thus resemble the generic interpretation of the masculine form described above. Consequently, NP7 should activate the MR [@locataires(MASC&FEM)]. The process of extracting the MR [@locataires(FEM)] should therefore always be possible, similar to the case in which the masculine form is generically interpreted. If epicene or collective nouns are used instead of common gender nouns, the process should be the same, seeing as they all implicitly include referents of all genders.

Table 1: Overview of the research design.

Form of the antecedent (between-participant)	Type of strategy	Example	Anaphor gender (within-participant)
Masculine	Non-GF: masculine	*Les voisins* 'the neighbours.$_{MASC}$'	Women
			Men
Complete double	GF: re-feminisation	*Les voisines et voisins* 'the neighbours.$_{FEM}$ and neighbours.$_{MASC}$'	Women
			Men
Contracted double	GF: re-feminisation	*Les voisin·es* 'the neighbours.$_{MASC·FEM}$'	Women
			Men
Gender-neutral	GF: neutralisation	*Les locataires* 'the tenants'*, le voisinage* 'the.$_{MASC}$ neighbourhood' or *les personnes* 'the people'	Women
			Men

3. The present study

The present study aims to investigate the mental representations of gender that are activated when reading role nouns in a GFF and compare those MRs to ones generated by the masculine form. To this end, we used a sentence evaluation task (initially elaborated by Tanenhaus &

Carlson, 1985, 1990), including sentence pairs like (3) as presented above. Thus, we investigated how anaphoric expressions were processed when the antecedents were presented in a GFF compared to when the antecedents were presented in their masculine form.

With the aim of focussing on the influence of the surface form of the antecedent, and minimising the influence of the participants' background knowledge, only role nouns representing non-stereotypical professions or activities were included (Tibblin et al., 2023). This choice allowed us to focus more precisely on the influence of linguistic information and decreased the influence of the gender-related information represented in the role nouns.

We used a 4×2 mixed design with Form of the antecedent and Anaphor gender as the two main predictor variables (cf. **Table 1** for an overview of the research design). Since previous studies on the effects of attitudes towards gender-fair language on mental representations of gender are rare and have shown inconclusive results, we measured these attitudes and included them as a predictor variable in our study. Thereby, we hope to shed some light on the interaction between these attitudes and mental representations of gender.

3.1 Hypotheses

The sentence evaluation task provides two outcome variables: the participants' judgments of the second sentence (yes or no) and the response times (RTs). The judgments reflect whether it was at all possible for participants to establish an anaphoric relation between the anaphoric expression and its antecedent. Positive judgment times were analysed (in line with previous research) and give us insight as to how easy the process of establishing the anaphoric relationship was.[6]

First, we hypothesised that there would be a significant two-way interaction between the variables Form and Anaphor gender, such that when *women* was used as an anaphor, the GFFs would lead to more positive judgments and faster RTs, compared to the masculine form. When *men* was used anaphorically, we did not expect any differences between the GFFs and the masculine form (Hypothesis 1).

Second, we hypothesised that when comparing the GFFs to each other, there would be significant differences between them, such that when *women* was used as an anaphor, the double forms would lead to significantly more positive judgments and faster RTs compared to the gender-neutral forms (as in Irmen & Roßberg, 2004). Again, we did not expect such differences with *men* as an anaphor (Hypothesis 2).

---

[6] Due to limited space, the hypotheses are slightly abbreviated here. However, they can be found in their full form in the preregistration, https://osf.io/p4vkw. The pre-registration also contains one hypothesis on expected main effects of Form and Anaphor gender which assumes that no interactions were found. Since we found a significant interaction, this hypothesis is automatically falsified and therefore not included here.

Finally, we expected a significant three-way interaction between the variables Form, Anaphor gender, and Attitudes towards gender-fair language, such that when the role noun was presented in its masculine form, positive attitudes would lead to fewer positive judgments and slower RTs when the anaphor *women* was used (as in Braun et al., 2005). We did not expect any effects of attitudes towards gender-fair language in the GFF conditions or when *men* was used as an anaphor (Hypothesis 3).

In brief, we expected (i) that the GFFs, compared to the masculine form, would increase the chances of resolving the anaphoric expression *women* and speed up this process, (ii) that the double forms would increase the chances of a successful anaphoric resolution of *women* as well as accelerate this process compared to the gender-neutral forms, and (iii) that positive attitudes towards gender-fair language would decrease the chances of resolving the anaphoric expression *women* with a masculine antecedent and slow down this process.

4. Method

4.1 Participants

One hundred and sixty-three right-handed native French-speakers participated in the experiment. They were recruited on Prolific, an online participant pool that provides high quality data (Peer et al., 2022), requires fair payments to the participants, and allows researchers to do fine pre-screening. The pre-screening questions, whose exact wording can be found in the pre-registration, concerned the participants' handedness, minimum age, language disorders and linguistic background. All participants were paid £2.45. Prior to analyses, 10 participants were removed from the sample due to high error rates concerning the filler items (see 4.3 below). The final sample therefore consisted of 153 participants (mean age = 28 years, *SD* = 9 years; 41% students). 84 of these were women, 64 men, 4 of another gender identity and 1 did not wish to state their gender. The majority (72%) of the sample lived in France, but some lived in the French-speaking regions of Belgium (10%), Switzerland (5%) or Canada (3%). The remaining 10% lived in a non-French-speaking country or region.

4.2 Materials and procedure

The materials consisted of 22 experimental and 44 filler sentence pairs. These were based on the materials used by Gygax et al. (2008) and followed the structure of (3) above. Thirty-three of the filler sentence pairs were incongruent, requiring a negative response, and eleven were congruent, requiring a positive response. All sentence pairs were presented in a randomized order.

The participants' task was to read and understand the first sentence, then to press the space bar to make the first sentence disappear and the second one appear. Once they had read the second sentence, they should decide, as quickly and accurately as possible, whether it was a

possible continuation of the first sentence. It was explicitly stated that the group described in the second sentence was part of the one mentioned in the first sentence. The participants were told to keep their thumbs on the space bar and their index fingers on the keys corresponding to *yes* and *no* throughout the experiment. To minimise the risk of including trials where participants might not have been paying attention, the experiment automatically continued if no key was pressed within 10 seconds. Before each sentence pair, a screen asking whether the participants were ready to continue was presented. This screen was presented until the space bar was pressed, and participants were then allowed to take a break if needed. The experiment started after six familiarisation trials.

The 22 role nouns used as antecedents in the experimental sentences were selected on a criterion of non-stereotypicality based on a previous study that collected stereotypicality norms of role nouns in their masculine and GF forms (Tibblin et al., 2023), such as *joueurs de tennis* 'tennis players' or *militants écologistes* 'climate activists'. Thus, the average estimated percentage of women in each role noun ranged from 39% (*les athlètes* 'the athletes') to 64% (*les personnes qui montent à cheval* 'the people who go horseback riding') ($M = 49\%$, $SD = 13\%$). When the role nouns were presented in the complete double form, the feminine form always preceded the masculine form (for example *les voisines et voisins* 'the neighbours.$_{FEM}$ and the neighbours.$_{MASC}$').

We created two lists for each level of the variable Form. In each list, half of the target sentences contained *women* and the other half *men*. If a role noun was followed by *women* in the first list, it was followed by *men* in the second list, and vice versa. The context-setting content of the sentences did not change between the two lists, and all lists included all 44 filler items. Each participant was randomly assigned to one of the four forms and within each form, to one experimental list.

Upon completing the experiment, participants were presented with questions regarding GFFs, an eight-item scale measuring attitudes towards gender-fair language (Tibblin, 2020; Tibblin et al., 2023) ($\alpha = .92$, 95% CI[.90, .94]), and a set of socio-biographical questions. Complete lists of both filler and experimental items, details regarding their content, as well as further procedure details, the attitudes scale, and all questions presented to the participants can be found in the pre-registration.

4.3 Data preparation

The data were analysed with R (R Core Team, 2021) and all analyses described are in line with the pre-registration unless otherwise stated. The judgments and the RTs were analysed separately, and the preparations are presented accordingly in this section. First, the percentage of correct responses for the two types of filler items (congruent and incongruent) was calculated for each participant. Before doing so, filler responses were coded as incorrect if the RT of the first or second sentence was below the set minimum threshold (300 ms) or hit the maximum time

limit (10,000 ms). After this procedure, we removed all trials of the participants ($n = 10$) who responded incorrectly to more than 50% of the congruent or the incongruent fillers. Following this first data screening, we removed all trials in which the RT of the first or second sentence was above or below the set threshold (i.e., all analysed trials had a RT of the first and second sentence between 300–10,000 ms). This second data screening led to a removal of 0.95% of the data.

Before analysing each outcome variable, we calculated a mean score reflecting each participant's attitudes towards gender-fair language. Missing values (i.e., No option suits me) were excluded and the scores of items 1–3 and 6–7 were inversed prior to the calculation. This mean score served as a predictor in the upcoming analyses. This variable and the Sentence length in characters variable were centred to their mean ($M = 3.4$ and $M = 70.1$, respectively).

4.3.1 Analyses of the judgment data

The participants' judgments of the second sentences were analysed through generalized linear mixed-effects modelling with binomial distribution using the *glmer* function of the lme4 package (Bates et al., 2021). To minimise convergence errors, we used a bound optimization by quadratic approximation (BOBYQA) with a set maximum of 200,000 iterations. The best model of fit was selected using the *fitLMER.fnc* function of the LMERConvenienceFunctions package (Tremblay & Ransijn, 2020). This function takes a maximal model as an input, then finds an optimal fixed-effects structure through backwards elimination of the fixed effects, before finding an optimal random-effects structure through forward selection, and finally refits the fixed-effects structure through further backward elimination. The maximal model contained a three-way interaction between the variables Form, Anaphor gender, and Attitudes towards gender-fair language and their main effects as fixed effects, and random intercepts of Participant and Role noun. The back- and forward fitting threshold was set to a z statistic of 2[7].

4.3.2 Analyses of the response time data

The RTs reflect the time taken from the onset of the second sentence to the moment any of the two response keys was pressed. Only the RTs of the positive judgments were analysed, which led to a removal of 11% of the data. The remaining data were analysed with linear mixed-effects modelling using the *lmer* function of the lme4 package (Bates et al., 2021). All RTs were log-transformed prior to the analyses to reduce the skewness of the data. The maximal structure was identical to that of the generalized linear mixed-effects model, but also included main effects of

---

[7] As pointed out by one of the reviewers, model selection based on stepwise regression can be criticised for producing alpha inflation. We therefore investigated the maximal models of both outcome variables to see whether the main and interaction effects of interest remained significant in the maximal model. Since this was the case, we only present the final models in the paper and refer to the OSF repository of the project for summaries of the maximal models: https://osf.io/5gc8f/.

the variables Trial number and Sentence length in characters, and a random slope of Trial number by Participant. The final model was selected through the same procedure as the judgment data, with the exception that the threshold was set to a *p*-value of .05.

4.3.3 Further details on the data analyses

In this section, we present details of the analyses that were not specified in the pre-registration. For both models, we used the R^2 measures suggested by Nakagawa and Schielzeth (2013) to quantify the variance explained by each model. These measures were calculated with the *r2* function of the performance package (Lüdecke et al., 2021) and include both the marginal R^2, taking only the fixed effects into account, and the conditional R^2, accounting for both fixed and random effects. In both models, all variables were dummy coded with Form: *masculine* and Anaphor gender: *men* as reference categories. Furthermore, we used the *ggpredict* function from the ggeffects package (Lüdecke, 2018a) to predict the values visualised in **Figures 3** and **5**.

We report effect sizes in forms of partial η^2 for the main and interaction effects of the generalized linear mixed-effects model (thresholds: small = .01 and medium = .06). These values were calculated with the *anova_stats* function from the sjstats package (Lüdecke, 2018b). Cohen's *d* is used to report the effect sizes of the linear mixed-effects model (thresholds: small = 0.2, medium = 0.5, and large = 0.8). These values were calculated by dividing the difference in predicted values by the square root of the sum of the variance components, as suggested by Brysbaert and Stevens (2018).

Within each anaphoric expression, all levels of Form were compared against each other, and within each form, the two levels of Anaphor gender were compared against each other. These comparisons were conducted using the *emmeans* and *pairs* functions from the emmeans package (Lenth, 2022). Within each anaphoric expression, *p*-values were adjusted with the Tukey method for multiple comparisons. These analyses were validated with the *glht* function of the multcomp package (Hothorn et al., 2008), using a single-step method for *p*-value adjustment. As both analyses showed the same significant comparisons, only the output using the Tukey adjustment method is reported in Section 5 below. Visualisations and summaries of all comparisons are available on the OSF repository of the project: https://osf.io/5gc8f/.

5. Results
5.1 Judgment data

Before presenting the results of the inferential analyses, we present descriptive data grouped by the form of the role noun and the anaphoric expressions used. As the leftmost bars in **Figure 2** show, the proportion of positive judgments in the masculine condition (*les voisins*) is lower when *women* was used as an anaphor compared to when *men* was used. However, when complete

double forms (*les voisines et voisins*) were used, there is barely any difference between the anaphoric expressions. Interestingly, the contracted double forms (*les voisin·es*) produced the opposite pattern compared to the masculine form. In fact, the proportion of positive judgments is lower when *men* was used as an anaphor compared to *women*. Finally, the gender-neutral forms (*le voisinage, les locataires*) led to fewer positive judgments overall and to a slightly lower proportion of positive judgments when *women* was used anaphorically compared to *men*. In sum, these descriptive results suggest that the GFFs attenuate the male bias induced by the masculine form, but not all in the same way.[8]
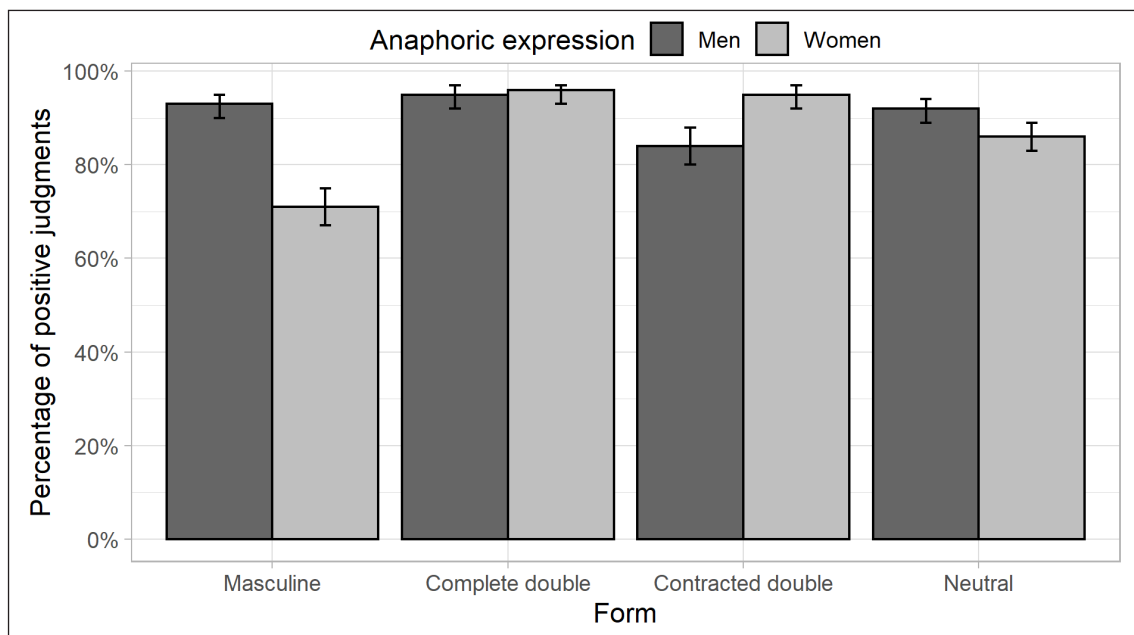
Figure 2: Percentage of positive judgments grouped by form and anaphoric expression. The error bars indicate 95% confidence intervals.

The descriptive results previously described are to a large extent confirmed by the inferential analyses, summarised in **Table 2**. The intercept in this table refers to the odds ratio, i.e., the exponent of the model's coefficients, of the following values: *men* as Anaphor gender, *masculine* as Form, and the Attitudes variable centred to its mean. An odds ratio of 1 indicates that the odds of a positive response are the same as the odds of a negative response, while a value greater than 1 indicates higher odds of a positive response. Inversely, an odds ratio less than 1 indicates lower odds of a positive response. The model of best fit included a significant interaction between the variables

---

[8] As suggested by one of the reviewers, we visually inspected the judgments of each role noun depending on form and anaphor gender to investigate the noun-by-noun variation. Since the number of observations per condition is very low, and because these figures did not reveal any new insights, we do not include them in the current paper but refer to the OSF repository of the project.

Form and Anaphor gender, as well as a significant main effect of the variable Attitudes towards gender-fair language. Due to convergence errors, the final random-effects structure provided by the analyses described above only included random intercepts of the variable Participant. However, we compared this final model with an identical one including random intercepts of both Participant and Role noun. Since the AIC was lower when two random variables were included ($AIC = 1975$, compared to 2099), and since it was significantly better according to an analysis of variance ($p < .001$), we included random intercepts of both Participant and Role noun in the final model. The significant main and interaction effects were the same in both models.

Table 2: Summary of the best generalized linear mixed-effects model of fit.

Predictors	Odds Ratios	95% CI	Statistic (z)	p
(Intercept)	26.75	14.45–49.53	10.46	<**0.001**
Form $_{Complete\ double}$	1.28	0.57–2.90	0.59	0.552
Form $_{Contracted\ double}$	0.30	0.15–0.62	–3.24	**0.001**
Form $_{Neutral}$	0.80	0.37–1.70	–0.59	0.557
Anaphor $_{Women}$	0.13	0.08–0.21	–8.38	<**0.001**
Attitudes $_{centred}$	1.26	1.02–1.54	2.19	**0.029**
Form $_{Complete\ double}$ * Anaphor $_{Women}$	9.80	4.26–22.56	5.37	<**0.001**
Form $_{Contracted\ double}$ * Anaphor $_{Women}$	31.49	15.15–65.44	9.24	<**0.001**
Form $_{Neutral}$ * Anaphor $_{Women}$	3.69	1.88–7.24	3.80	<**0.001**
Random effects				
Residual variance	3.29			
Random intercept variance $_{role\ noun}$	0.34			
Random intercept variance $_{participant}$	1.20			
ICC	0.32			
Observations	3334			
Marginal R^2	0.134			
Conditional R^2	0.410			

The variable Form was qualified by a small main effect (partial $\eta^2 = .02$). More precisely, the simple effect of Form: *contracted double* (line 3) shows that the odds of a positive judgment after the anaphor *men* (the reference value) are significantly lower when the antecedent is presented in a contracted double form (*les voisin·es*) compared to when it is presented in the masculine form (*les voisins*), but not when complete double or gender-neutral forms are used (lines 2 and 4).

Furthermore, the variable Anaphor was qualified by a close to small main effect (partial $\eta^2 = .009$) and the simple effect of Anaphor: *women* (line 5) indicates that when the antecedent is presented in the masculine form, the odds of obtaining a positive judgment after the anaphor *women* are significantly lower than after the anaphor *men*. The very small main effect (partial $\eta^2 = .003$) of Attitudes towards gender-fair language (line 6) shows that positive attitudes increased the odds of a positive judgment, regardless of the anaphor and the form used. Thus, the more positive a participant's attitude towards gender-fair language was, the more likely they were to respond *yes*. This holds for both *women* and *men* anaphors, and for all forms of the antecedent. Finally, the effect size of the interaction between Form and Anaphor gender was almost medium-sized (partial $\eta^2 = .053$). In fact, all the GFFs in combination with Anaphor: *women* led to significant interaction effects (lines 7–9), and all coefficients are larger than 1. Thus, when a GFF is used, the likelihood of a positive judgment following *women* increases significantly if one of the GFFs is used instead of the masculine form. The predictions made by this model are visualised in **Figure 3**. The marginal R^2 value indicates that the fixed effects of this model account for 13% of the variance, while the fixed and random effects together account for 41% of the variance.
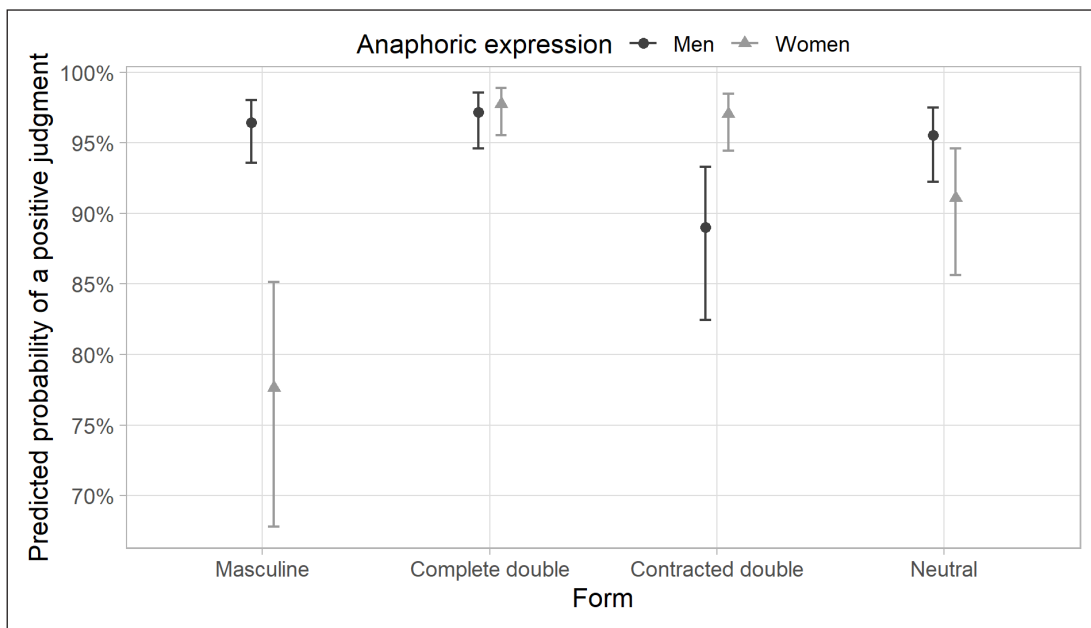
Figure 3: Probability of positive judgments as predicted by the final model grouped by form and anaphoric expression. The error bars indicate 95% confidence intervals.

When the different forms were compared against each other within each anaphoric expression, contrast analyses revealed that when *men* was used as an anaphor, the contracted double forms led to a significantly lower probability of a positive judgment compared to all the other forms

(*ps* < .05). When *women* was used anaphorically, the masculine form led to a significantly lower probability of a positive judgment compared to all GFFs (*ps* < .01), and both double forms led to a probability that was significantly higher compared to the gender-neutral form (*ps* < .05).

The comparisons between the anaphoric expressions within each form revealed significant differences between the two anaphoric expressions in the masculine, gender-neutral, and contracted double forms. Within the masculine and gender-neutral forms, the probability of a positive judgment was significantly lower when *women* was used compared to *men* (*ps* < .01). Within the contracted double form, *men* led to a significantly lower probability compared to *women* (*p* < .001).

5.2 Response times of the positive judgments

The mean RTs of the positive judgments, grouped by form and anaphor gender, are presented in **Figure 4**. These means suggest that when participants were able to establish an anaphoric relation between *women* and an antecedent presented in the masculine form, it took them a longer time to do so. As this difference is not observed when GFFs are used, GFFs seem effective in neutralising this effect. Interestingly, it seems to require slightly more cognitive effort to establish an anaphoric relationship between *men* and an antecedent presented in a contracted double form than *women*. Taken together, these descriptive results reveal patterns similar to the judgment data.
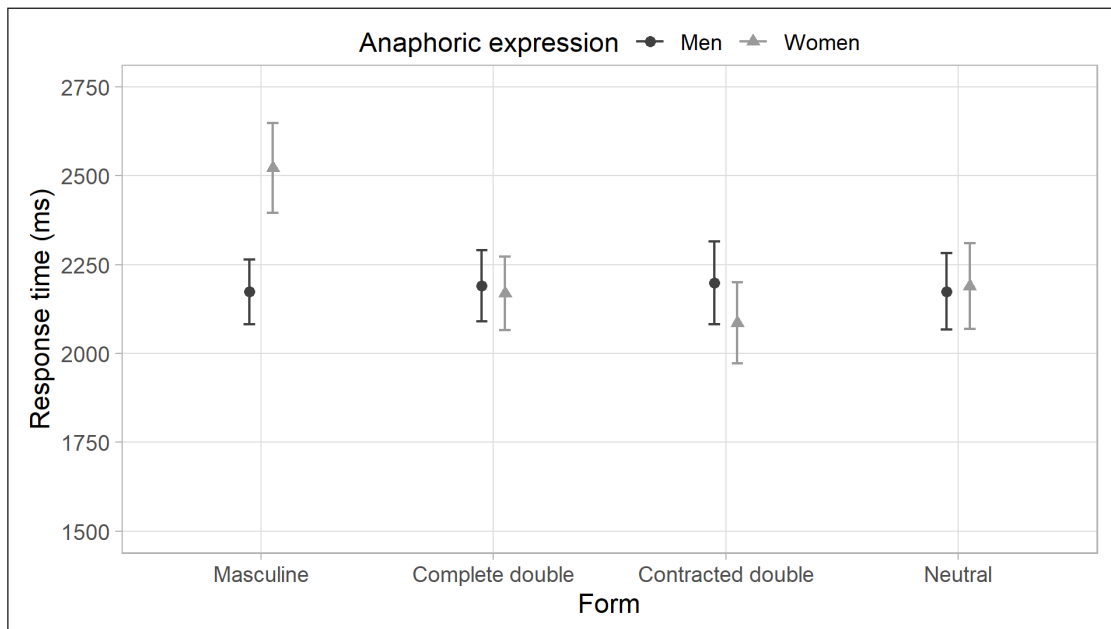
Figure 4: Mean response times (in ms) of the positive judgments grouped by form and anaphoric expression. The error bars indicate 95% confidence intervals.

A summary of the final model that best fit the RT data is presented in **Table 3**. The intercept in this summary refers to the log-transformed RT of the following values: *men* as the anaphoric expression, *masculine* as the form, the Attitudes and the Sentence length variables are centred around their means, and trial number is set at 1 (i.e., the first trial of the experiment). The fixed-effects structure of the model of best fit produced by the analyses, described in Section 4.3, contained an interaction between the variables Form and Anaphor gender, and main effects of the variables Attitude, Sentence length and Trial number. It also included Participant and Role noun as random intercepts.

Table 3: Summary of the best linear mixed-effects model of fit.

Predictors	Estimates	95% CI	Statistic (t)	p	Effect size (d)
(Intercept)	7.76	7.65 – 7.86	146.89	<**0.001**	18.021
Form ₍Complete doublets₎	0.01	–0.13 – 0.15	0.16	0.871	0.027
Form ₍Contracted doublets₎	0.04	–0.10 – 0.18	0.52	0.604	0.087
Form ₍Neutral₎	0.00	–0.14 – 0.14	0.01	0.993	0.002
Anaphor ₍Women₎	0.11	0.07 – 0.16	4.74	<**0.001**	0.263
Attitudes ₍centred₎	–0.06	–0.10 – –0.01	–2.51	**0.012**	–0.131
Sentence length ₍centred₎	0.01	0.00 – 0.01	2.49	**0.013**	0.015
Trial number	–0.00	–0.01 – –0.00	–16.61	<**0.001**	–0.012
Form ₍Complete double₎ * Anaphor ₍Women₎	–0.13	–0.19 – –0.06	–3.89	<**0.001**	–0.293
Form ₍Contracted double₎ * Anaphor ₍Women₎	–0.19	–0.26 – –0.13	–5.78	<**0.001**	–0.444
Form ₍Neutral₎ * Anaphor ₍Women₎	–0.10	–0.16 – –0.03	–2.98	**0.003**	–0.225
Random effects					
Residual variance	0.09				
Random intercept variance ₍participant₎	0.09				
Random intercept variance ₍role noun₎	0.00				
ICC	0.50				
Observations	2966				
Marginal R^2	0.081				
Conditional R^2	0.541				

First, none of the simple effects of the GFFs (lines 2–4) were significant, signalling that when *men* was used anaphorically, it did not take more time to judge the sentence as possible with a GFF antecedent compared to a masculine antecedent. However, the simple effect of the anaphor *women* was significant (line 5), indicating that in the masculine condition, RTs were slower when *women* was used as an anaphor compared to when *men* was used. The small negative main effect of attitudes towards gender-fair language (line 6) indicates that the more positive participants' attitudes were, the faster their judgments were. No interaction including attitude was found, suggesting that the attitude effect was roughly constant across both anaphoric expressions and all antecedent forms. Furthermore, significant main effects of Sentence length (line 7) and of Trial number (line 8) were found, signalling that RTs increase as the number of characters in the sentence increases and that RTs decrease over the course of the experiment. However, the sizes of these effects were very small ($d = 0.015$ and -0.012, respectively).

Finally, we found significant interaction effects of all GFFs in combination with the anaphor *women* (lines 9–11). Thus, the difference between the RTs following the different anaphoric expressions is significantly smaller in the GFF conditions compared to the masculine condition. The interaction effects of the complete double and gender-neutral forms in combination with the anaphor *women* were small ($d = -0.293$ and -0.225, respectively), while that of the contracted double forms was close to medium-sized ($d = -0.444$). The response times predicted by the model are visualised in **Figure 5**. As indicated by the marginal and the conditional R^2 values, the fixed effects of this model account for 8% of the variance, while the fixed and random effects together account for 54%.
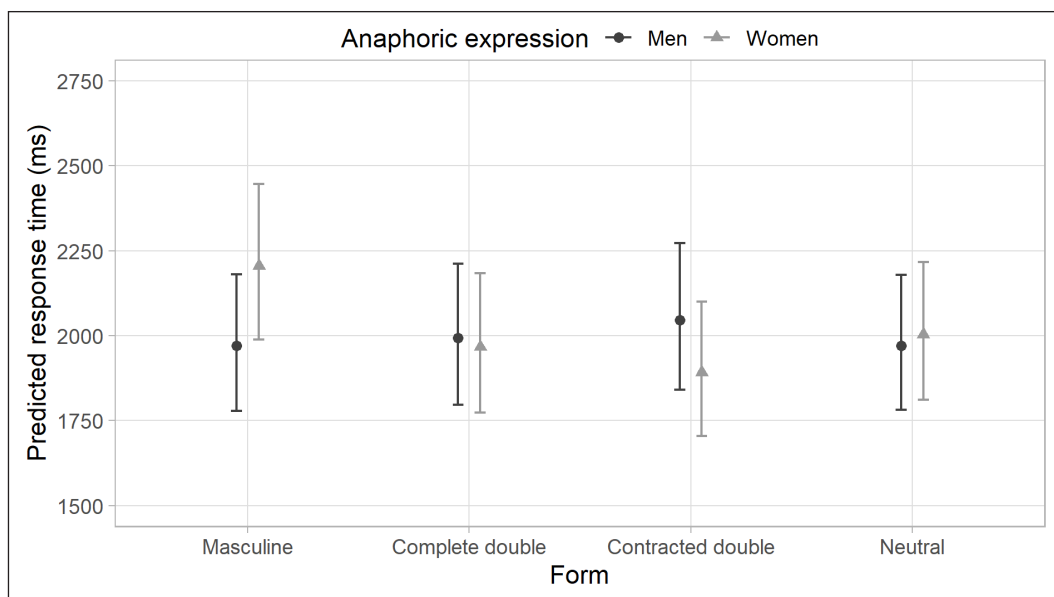
Figure 5: Response times (in ms) as predicted by the final model grouped by form and anaphoric expression. The error bars indicate 95% confidence intervals.

The contrast analyses between each form within each anaphoric expression revealed that no form led to significantly faster or slower RTs when compared to each of the other forms ($ps > .05$). However, the comparisons between the anaphoric expressions within each form revealed that within the masculine form, *women* led to significantly slower RTs compared to *men* ($p < .001$). Inversely, within the contracted double forms, *men* led to significantly slower RTs compared to *women* ($p = .001$). No such differences were found within the complete double and the gender-neutral forms ($ps > .05$).

6. General discussion

In the present study, we investigated the effects of French GFFs on the resolution of the anaphoric expressions *women* and *men*. The GFFs were compared both to the masculine form and to each other. The results indicate that presenting the antecedent in its masculine form made the establishment of an anaphoric relationship between the anaphoric expression *women* and the antecedent more difficult compared to when the anaphoric expression *men* is used. However, when the antecedent was presented in any of the three analysed GFFs, this bias was either attenuated (by the gender-neutral forms) or disappeared completely (double forms). We also compared the different GFFs to each other. When doing so, we only found significant differences when analysing the judgments, and not the RTs. More specifically, the double forms were more effective in resolving the male bias compared to the gender-neutral forms. While the latter forms still led to a slight male bias, it was weaker than the one induced by the masculine form and was, as said, not present when looking at the RTs. We also found that the contracted double forms evoked a small female bias, meaning that it was harder to resolve the anaphoric expression *men* compared to *women* with an antecedent presented in a contracted double form. Finally, we found that positive attitudes towards gender-fair language led to more positive judgments and faster RTs regardless of the form of the antecedent or the gender of the anaphoric expression. In the following sections, we will discuss these findings in relation to our initial hypotheses and to our theoretical framework.

6.1 The effects of the different forms on mental representations of gender

6.1.1 The masculine form

First of all, let us look at how the masculine form was processed. The analyses of the judgments clearly indicate that when a role noun was presented in its masculine form, it was harder to successfully establish an anaphoric relationship between the anaphor *women* and the antecedent compared to when *men* was used as an anaphor. The RTs show that even when participants were able to establish this relationship, and consequently responded *yes* in the judgment task, it required more cognitive effort. These findings are in keeping with previous studies on the masculine form in French (e.g., Garnham et al., 2012; Gygax et al., 2008, 2012; Sato et al., 2013;

Sato, Gygax, et al., 2016), and we can once again conclude that in French, the masculine form leads to a male bias in the mental representations of gender generated when reading.

As discussed above, Reboul and Moeschler (1998) provide interesting and adequate tools to explain our results. As **Figure 1** shows, a specific interpretation of the masculine form would lead to a negative judgment of the anaphoric linkage, while a generic interpretation would lead to a positive judgment. However, there are cases where the participants responded positively but required more time. Since the MR [@voisines(FEM)] of the anaphor must be extracted from the preceding antecedent MR to resolve the anaphoric expression, the preceding MR must be [@voisins(MASC&FEM)]. It therefore seems plausible that the longer RTs reflect an update by the participants of the antecedent MR. In this scenario, the MR of the antecedent changes from being specific [@voisins(MASC)] to generic [@voisins(MASC&FEM)] upon encountering the anaphoric expression. However, the process of updating a specific MR to a generic one requires more cognitive effort, as reflected by the longer RTs, compared to if the specific MR is maintained. Actually, participants who maintained the specific MR (i.e., those who rejected the anaphor *women* following a masculine antecedent) were quicker in their judgments ($n = 122$ observations, $M = 2127$ ms, $SD = 1277$ ms) compared to those who updated their MR (i.e., those who accepted the anaphor *women*, $n = 301$ observations, $M = 2521$ ms, $SD = 1113$ ms).[9] This indicates that readers' mental representations are clearly male, and as such these readers do not allocate any cognitive resources to find a solution to resolve the incongruency induced by *masculine antecedent – female anaphor*.

Taken together, this suggests that the specific interpretation of the masculine form or, rather, the creation of the MR [@voisins(MASC)] can potentially be overridden, maybe depending on its strength. A MR activated by a strong specific interpretation would not be able to update into a generic one, in contrast to a weak specific interpretation that could be susceptible to an update. With the method used in the current study, we cannot study empirically whether and how this process takes place, but future studies may want to look at the conditions granting easy or more complex updates.

It is also possible that different factors could influence the strength of the specific interpretation. Such factors could be based in the text itself, like explicit information on the gender of the people represented. For example, Duffy and Keir (2004) found that, in English, the effect of a mismatch between the stereotypicality of a role noun and the referent's gender was mitigated when it was explicitly stated that the person occupying a stereotypically masculine role was a woman. As such, participants had trouble processing the pronoun *herself* in *The electrician taught herself* but

---

[9] We ran additional analyses on all RTs including the preceding judgment as a predictor variable. However, there are very few negative judgments in several conditions, and we therefore refrain from discussing these results further here. However, a full summary and a visualisation of these analyses are available on the project's OSF repository.

not in *The electrician was a cautious woman who taught herself*. Along the same lines, Gygax et al. (2012) used a word association task to investigate this issue in French. They asked participants to judge whether a person represented by a female (*aunt*) or male (*uncle*) kinship term could be part of a group represented by a role noun presented in its masculine form. Halfway through the experiment, the participants were explicitly instructed to interpret the masculine form generically. Following these instructions, participants were more likely to associate female kinship terms with the role noun. However, the RTs still indicated higher cognitive costs of linking women to the role noun, suggesting that the male bias in French is hard to overcome. It would therefore be interesting for future studies to investigate if other kinds of information could attenuate the male bias, and if that is the case, the strength of such an effect.

6.1.2 The effects of the gender-fair forms on mental representations of gender

The principal aim of this study was to investigate how our MRs of gender are affected when the antecedent is presented in a GFF instead of in its masculine form, with the hypothesis that the use of a GFF would lead to increased representation of women. This hypothesis was confirmed in the analyses of both the judgments and the RTs, showing that GFFs increased the chances of a successful anaphoric resolution with *women*, and facilitated this resolution, as reflected by shorter RTs. These findings are in keeping with those of previous research on French GFFs (Brauer & Landry, 2008; Kim et al., 2022; Richy & Burnett, 2021; Tibblin et al., 2023; Xiao et al., 2022). Taken together, our results indicate that both double and gender-neutral forms are effective when it comes to attenuating the male bias induced by a masculine form in a timed task investigating the processing of French GFFs in context. To our knowledge, the present study is the first to directly compare different GFFs and show that this is the case.

Secondly, we were interested in investigating whether any GFF was more effective than another in attenuating the male bias, with the hypothesis that the double forms would be more effective compared to the gender-neutral forms (Hypothesis 2). As the descriptive statistics in **Figure 2** and **Figure 4** indicate, and as confirmed by the inferential statistics, the complete double forms produced the most gender-balanced MRs. This can be explained by referring to example (1a) above, since a NP like *les voisines et voisins* leads to two separate MRs, analogous to $[_{NP1}$ A man and a woman] in example (1). The two MRs created are then [@voisines(FEM)] and [@voisins(MASC)], and since both anaphoric expressions correspond perfectly to one of the MRs, no extraction operation needs to be executed on the MR in order to resolve them. This can explain the little to no difference in both the proportion of positive judgments and the RTs between the two anaphoric expressions following a complete double form as antecedent.

Furthermore, both the complete and contracted double forms were more effective than the gender-neutral forms in attenuating the male bias, but only in regard to judgments. In other words, when *women* was used anaphorically, it was easier to establish an anaphoric relationship

when the antecedent was presented in any double form compared to if it was presented in a gender-neutral form. Since no GFF led to significantly different RTs when the GFFs were compared to each other within each anaphoric expression, Hypothesis 2 can only be partially confirmed. It should still be stated that even if a small male bias was provoked by the gender-neutral forms, it was significantly smaller than the one provoked by the masculine forms. This is in line with the results of previous studies having found that gender-neutral forms led to an increased representation of women compared to the masculine form in French (e.g., Kim et al., 2022; Richy & Burnett, 2021; Tibblin et al., 2023). The small, but present, male bias produced by the gender-neutral forms might be explained by the fact that these forms do not contain any explicit gender-related information, in contrast to the double forms. In addition, six of the nine collective nouns used to form the gender-neutral forms were grammatically masculine (e.g., *le voisinage* 'the.$_{MASC}$ neighbourhood'), while only three were feminine (e.g., *la clientèle* 'the.$_{FEM}$ customer base'). The fact that the participants were confronted with more grammatically masculine nouns could possibly have influenced their assignment of semantic representations to the collective nouns. Actually, Cacciari et al. (1997, 2011) found that when Italian epicene nouns were used as antecedents, an incongruence between the grammatical gender of the antecedent and the gender represented by the anaphoric expression led to greater processing costs. Further studies on the differences between grammatically feminine and masculine collective and epicene nouns as well as the differences between collective and common gender nouns in French are needed in order to understand these results better.

Finally, the contrast analyses revealed unexpected effects of the contracted double forms on the resolution of *men* as an anaphoric expression. In fact, our results indicated that this form resulted in a slight female bias in the MRs generated when reading, as the participants in the contracted double condition presented lower probabilities of positive judgments and slower RTs following *men* compared to the participants in the other conditions. In addition, it was harder for these participants to establish the link between the antecedent and *men* compared to *women,* and when the link was successfully established with *men,* the RTs were longer. However, it should be underlined that the difference between the two anaphoric expressions is smaller than the one observed when the antecedent is presented in its masculine form. Such a female bias has, to our knowledge, only been observed in one previous study investigating German GFFs (Körner et al., 2022) and never before in French. In line with our results, Körner et al. (2022) found that the female bias induced by the German asterisk form, *student*innen* 'students.$_{MASC*FEM}$', was weaker than the male bias induced by the masculine form. While the experimental paradigm used by Körner et al. (2022) and their results are similar to those of the present study, one should be careful when drawing parallels between the two languages. An important difference is that the asterisk form in German aims at representing non-binary individuals in addition to women and men (Körner et al., 2022), while contracted double forms in French are, according to official

recommendations, at least in France, meant to function as an abbreviation of the complete double forms (Haut Conseil à l'Égalité entre les Femmes et les Hommes, 2022).

When interpreting these results using the tools provided by Reboul and Moeschler (1998), it does not seem as though the contracted double forms lead to two separate MRs, as the complete double forms do. Rather, a NP like *les voisin·es* seems to activate either a specific MR, [@ voisines(FEM)], or a generic MR, [@voisins(MASC&FEM)]. We propose two possible explanations to why such NPs sometimes seem to hinder representation of men. On the one hand, a contracted double form might be processed as a feminine form, thus only having a specific interpretation, and not as a double form. Since the inflectional suffix marking the feminine gender is separated from the stem (ex. *les spectateur·rices* 'the spectators.$_{MASC·FEM}$'), this morpheme is somewhat highlighted, and participants might therefore perceive it as salient. Thus, when the participants are confronted with the anaphoric expression, they might remember reading only the feminine form. Should that be the case, their MR would be [@voisines(FEM)], from which it would be impossible to extract the MR [@voisins(MASC)]. This would result in a negative judgment. If a contracted double form is interpreted as a feminine form, it should be impossible to update the specific antecedent MR to a generic one. Thus, this interpretation could only explain the difference between the two anaphoric expressions in the judgments, but not the longer RTs, given that they reflect an update of the antecedent MR. On the other hand, these results could also be explained by societal factors. Since GFFs represent a highly political topic in the French-speaking world (Abbou et al., 2018) and since contracted double forms with the interpunct often are used to exemplify GFFs, it is possible that the interpunct would call upon reader's background knowledge and be associated with feminism. Actually, a corpus study on different French GFFs suggests that a certain GFF can be used to construct a political identity and to take a stance on gender-related issues (Burnett & Pozniak, 2021). If the interpunct triggers ideas about activism in favour of representation of women, it might be hard for participants to include men in their MRs.

These two explanations have different implications for the processing of contracted double forms. If the first explanation is accurate, there should be a difference between nouns whose feminine form is created by adding an affix onto the masculine form (*les voisin·es*) and those where it is created by replacing the affix marking the masculine gender with an affix marking the feminine gender (*les spectateur·rices*). Since the feminine form of the former type of nouns (*les voisines*) is very similar to the contracted double form, it is possible that these could be interpreted as feminine forms. In a contrary way, it seems unlikely that the latter type of nouns would be read as a feminine form (*les spectatrices*), since readers would then have to skip the affix marking the masculine gender (*-eur*). Such a difference between the two types should not be present if the second explanation is correct. The role nouns used in the present study were not controlled for the way the feminine form was created (seven nouns marked the feminine form by adding a feminine affix while 15 marked it by replacing the masculine affix), and it is therefore

difficult to compare the two types to each other. These are only two suggested explanations, as a different experimental design comparing the two ways to form the feminine form, and perhaps using other typographical signs in addition to the interpunct, would be needed to answer this question. A replication of the present study using balanced contracted forms might be able to shed light on the issue.

6.2 The effect of attitudes towards gender-fair language on mental representations of gender

Finally, we studied the effects of participants' attitudes towards gender-fair language on their MRs of gender. Hypothesis 3, based on Braun et al. (2005), was that the masculine form would produce a stronger male bias among the participants with positive attitudes compared to those with negative attitudes. In contrast to this hypothesis, we found that positive attitudes activated more gender-balanced MRs, as shown by higher odds of positive judgments and faster RTs regardless of the form of the antecedent or the gender represented in the anaphoric expression. These results stand in contrast to two previous studies that have explored the link between attitudes towards gender-fair language and MRs of gender. This difference might be explained by the methodological differences between our study and theirs in how the variable was analysed. While we kept attitudes as a ratio-scaled variable, Braun et al. (2005) transformed it into a categorical variable by dividing the participants into two groups representing positive or negative attitudes. In addition, the authors investigated the difference between two different forms, while we investigated four different forms, and the statistical power of our analyses is therefore not as strong as theirs. Thus, it is possible that the lack of an interaction in our data is partially due to a lack of statistical power in the present study. In contrast to Braun et al. (2005), Tibblin et al. (2023) found that negative attitudes towards gender-fair language led to an increased perceived general percentage of women when participants were asked to estimate the percentage of women in a set of role nouns. Since most studies that have investigated the effects of attitudes towards gender-fair language on MRs of gender have yielded different results, we can only conclude that the role of attitudes is still not clear, and more research is needed. All studies presented so far have measured attitudes explicitly by means of a questionnaire, and it would therefore be very interesting to measure them implicitly, for example, with an implicit association test (Greenwald et al., 1998), or to conduct in-depth interviews (Gerson, 2020) to gather a deeper understanding of what these attitudes actually reflect.

7. Concluding remarks

A few words should also be said about the chosen research design. First, while a sentence evaluation task provides more insight than a questionnaire study, it does not provide as detailed data as ERP or eye tracking methods. An ERP study could tell us whether GFFs influence the

lexical-semantic or syntactic processing of text by measuring N400 and P600 effects (as in Misersky et al., 2019), whereas an eye tracking study could show the regressions made from the anaphoric expression to its antecedent, and show whether the processing of the different GFFs gives raise to different gaze patterns, at different locations such as the role noun itself and its spillover regions.

Secondly, the marginal R^2 values of the models of best fit, i.e., the variance accounted for by the fixed effects alone, were rather low (13% and 8%, respectively). This indicates that there probably are many other variables for which we did not control that could influence the outcome variables. For example, we controlled our items for stereotype by including only non-stereotyped role nouns, but the role nouns may differ in word frequency, familiarity, length, etc. They may also represent different levels of stereotypicality for different participants. In the same fashion, there are possibly other participant-related variables for which we did not control that could influence the judgments and the RTs.

Finally, it should also be noted that the texts used in this study might not be representative of other text genres, such as novels or newspapers. However, we chose this kind of texts to more accurately be able to compare the results with previous studies. In addition, one could argue that some participants may not actually read the second sentences as linked to the first one, and thus may not be involved in any anaphoric resolution processes. Three arguments go against this explanation. First, previous studies in fact pre-tested similar pairs and showed that the antecedents were understood to include the group described by the anaphor in the second sentence (see Garnham et al., 2012; Gygax et al., 2008), Second, if participants were not doing any anaphor-antecedent resolutions, responses should always be positive (i.e., there should be no bias). Third and finally, in previous experiments, as in ours, emphasis is always put on this relationship in the instructions to the participants.

To conclude, the reported study replicated a male bias in the mental representations of gender activated by the masculine form, even when it is generically intended. Fortunately, our results clearly show that the different types of GFFs are viable alternatives to the masculine form as they effectively manage to either completely eliminate (double forms) or mitigate (gender-neutral forms) this bias. Interestingly, we found differences between the GFFs in participants' judgments, but not in the RTs. More specifically, the complete double forms (*les voisines et voisins*) led to almost perfectly balanced mental representations of gender, while the contracted double forms (*les voisin·es*) actually resulted in a slight female bias. We encourage future studies to investigate this form in closer detail to find the source of this interpretation. Finally, the gender-neutral forms (*les locataires* or *le voisinage*) produced a slight male bias, although significantly smaller than the one produced by the masculine form. Taken together, these findings provide strong support in favour of all kinds of GFFs in French when the aim is to increase representation of women.

Abbreviations

ART = article

FEM = feminine

GFF = gender-fair forms

MASC = masculine

MR = mental representation

NEG = negation

Data accessibility statement

The dataset generated and analysed during the current study is available in the OSF repository of the project, https://osf.io/5gc8f, where the consent form, instructions to participants, materials, and the analysis code files are also available.

Ethics and consent

This research was approved by the Swedish Ethical Review Authority (application number 2021-01932). Informed consent was obtained from all individual participants included in the study.

Acknowledgements

Competing interests

The authors have no competing interests to declare.

Author contributions

Julia Tibblin: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Visualisation, Writing – original draft, Writing – review & editing.

Jonas Granfeldt: Conceptualisation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Joost van de Weijer: Formal analysis, Methodology, Writing – review & editing.

Pascal Gygax: Conceptualisation, Methodology, Resources, Supervision, Writing – review & editing.

References

Abbou, J., Arnold, A., Candea, M., & Marignier, N. (2018). Qui a peur de l'écriture inclusive? Entre délire eschatologique et peur d'émasculation. *Semen : Revue de sémio-linguistique des textes et discours*, *44*(1), 133–150. DOI: https://doi.org/10.4000/semen.10800

Anaya-Ramírez, A., Grinstead, J., Nieves Rivera, M., Melamed, D., & Reig-Alamillo, A. (2022). The interpretation of Spanish masculine plural NPs: Are they perceived as uniformly masculine or as a mixture of masculine and feminine? *Applied Psycholinguistics*, 1–18. DOI: https://doi.org/10.1017/S0142716422000352

Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge.

Bates, D., Maechler, M., Bolker, B., Walker, S., Haubo Bojesen Christensen, R., Subgnabb, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., & Krivitsky, P. N. (2021). *Package 'lme4': Linear mixed-effects models using 'Eigen' and S4* (1.1–27.1). https://github.com/lme4/lme4/

Brauer, M., & Landry, M. (2008). Un ministre peut-il tomber enceinte? L'impact du générique masculin sur les représentations mentales. *L'année psychologique*, *108*(22), 243–272. DOI: https://doi.org/10.4074/S0003503308002030

Braun, F., Sczesny, S., & Stahlberg, D. (2005). Cognitive effects of masculine generics in German: An overview of empirical findings. *Communications*, *30*(1), 1–21. DOI: https://doi.org/10.1515/comm.2005.30.1.1

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 1–20. DOI: https://doi.org/10.5334/joc.10

Burnett, H., & Pozniak, C. (2021). Political dimensions of gender inclusive writing in Parisian universities. *Journal of Sociolinguistics*, *25*(5), 808–831. DOI: https://doi.org/10.1111/josl.12489

Cacciari, C., Carreiras, M., & Cionini, C. B. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory and Language*, *37*(4), 517–532. DOI: https://doi.org/10.1006/jmla.1997.2528

Cacciari, C., Corradini, P., Padovani, R., & Carreiras, M. (2011). Pronoun resolution in Italian: The role of grammatical gender and context. *Journal of Cognitive Psychology*, *23*(4), 416–434. DOI: https://doi.org/10.1080/20445911.2011.526599

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, *49A*(3), 639–663. DOI: https://doi.org/10.1080/713755647

Corbett, G. G. (1991). *Gender*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139166119

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, *32*(4), 551–559. DOI: https://doi.org/10.3758/BF03195846

Elmiger, D. (2008). *La féminisation de la langue en français et en allemand. Querelle entre spécialistes et réception par le grand public*. Editions Champion.

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Psychology Press.

Garnham, A., Gabriel, U., Sarrasin, O., Gygax, P., & Oakhill, J. (2012). Gender representation in different languages and grammatical marking on pronouns: When beauticians, musicians, and mechanics remain men. *Discourse Processes*, *49*(6), 481–500. DOI: https://doi.org/10.1080/0163853X.2012.688184

Gerson, K. (2020). Depth interviewing as science and art. In K. Gerson & S. Damaske, *The science and art of interviewing* (pp. 1–29). Oxford University Press. DOI: https://doi.org/10.1093/oso/9780199324286.003.0001

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. DOI: https://doi.org/10.1037/0033-295X.101.3.371

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. DOI: https://doi.org/10.1037/0022-3514.74.6.1464

Gygax, P., & Gabriel, U. (2011). Gender representation in language: More than meets the eye. In R. K. Mishra & N. Srinivasan (Eds.), *Language-cognition interface: State of the art* (p. 321). Lincom Europa.

Gygax, P., Gabriel, U., Lévy, A., Pool, E., Grivel, M., & 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. DOI: https://doi.org/10.1080/20445911.2011.642858

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, *23*(3), 464–485. DOI: https://doi.org/10.1080/01690960701702035

Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: A challenge for our cognitive system. *Language Sciences*, *83*, 101328. DOI: https://doi.org/10.1016/j.langsci.2020.101328

Hansen, K., Littwitz, C., & Sczesny, S. (2016). The social perception of heroes and murderers: Effects of gender-inclusive language in media reports. *Frontiers in Psychology*, *7*, 369. DOI: https://doi.org/10.3389/fpsyg.2016.00369

Haut Conseil à l'Égalité entre les Femmes et les Hommes. (2022). *Pour une communication publique sans stéréotype de sexe. Guide pratique.* https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/guide_egacom_sans_stereotypes-2022-versionpublique-min.pdf

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. DOI: https://doi.org/10.1002/bimj.200810425

Irmen, L. (2007). What's in a (role) name? Formal and conceptual aspects of comprehending personal nouns. *Journal of Psycholinguistic Research*, *36*(6), 431–456. DOI: https://doi.org/10.1007/s10936-007-9053-z

Irmen, L., & Roßberg, N. (2004). Gender markedness of language: The impact of grammatical and nonlinguistic information on the mental representation of person information. *Journal of Language and Social Psychology*, *23*(3), 272–307. DOI: https://doi.org/10.1177/0261927X04266810

Kamp, H. (1981). A theory of truth and semantic representation. In P. Portner & B. H. Partee (Eds.), *Formal semantics—The essential readings* (pp. 189–222). Blackwell. DOI: https://doi.org/10.1163/9789004252882_014

Kim, J., Angst, S., Gygax, P., Gabriel, U., & Zufferey, S. (2022). The male bias and ways to avoid it: A study on epicenes and group nouns in Swiss and Québec French. *Journal of French Language Studies*, 1–26. DOI: https://doi.org/10.1017/S095926952200014X

Körner, A., Abraham, B., Rummer, R., & Strack, F. (2022). Gender representations elicited by the gender star form. *Journal of Language and Social Psychology*, 0261927X2210801. DOI: https://doi.org/10.1177/0261927X221080181

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means* (1.8.1-1). DOI: https://CRAN.R-project.org/package=emmeans

Lüdecke, D. (2018a). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772. DOI: https://doi.org/10.21105/joss.00772

Lüdecke, D. (2018b). *Sjstats: Statistical functions for regression models.* Zenodo. https://doi.org/10.5281/ZENODO.1489175

Lüdecke, D., Makowski, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., Wiernik, B. M., Arel-Bundock, V., & Jullum, M. (2021). *performance: An R package for assessment, comparison and testing of statistical models.* https://easystats.github.io/performance/. DOI: https://doi.org/10.21105/joss.03139

Melis, L., & Godard, D. (2021). La formation, le genre et le nombre des noms. In A. Abeillé & D. Godard (Eds.), *La grande grammaire du français* (1st ed., Vol. 1, pp. 377–397). Actes sud/Imprimerie nationale Éditions.

Misersky, J., Majid, A., & Snijders, T. M. (2019). Grammatical gender in German influences how role-nouns are interpreted: Evidence from ERPs. *Discourse Processes, 56*(8), 643–654. DOI: https://doi.org/10.1080/0163853X.2018.1541382

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. DOI: https://doi.org/10.1111/j.2041-210x.2012.00261.x

Parks, J. B., & Roberton, M. A. (2000). Development and validation of an instrument to measure attitudes toward sexist/nonsexist language. *Sex Roles, 42*(5/6), 415–438. DOI: https://doi.org/10.1023/A:1007002422225

Peer, E., Rothschild, D., Gordon, A., Zak, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*, 1643–1662. DOI: https://doi.org/10.3758/s13428-021-01694-3

Perloff, R. M. (2003). *The dynamics of persuasion: Communication and attitudes in the 21st century* (2 ed.). Lawrence Erlbaum Associates.

Prentice, D. A. (1994). Do language reforms change our way of thinking. *Journal of Language and Social Psychology, 13*(1), 3–19. DOI: https://doi.org/10.1177/0261927X94131001

R Core Team. (2021). *R: A language and environment for statistical computing* (4.1.1). R Foundation for Statistical Computing. https://www.R-project.org

Reboul, A., & Moeschler, J. (1998). *Pragmatique du discours. De l'interprétation de l'énoncé à l'interprétation du discours*. Armand Colin.

Richy, C., & Burnett, H. (2021). Démêler les effets des stéréotypes et le genre grammatical dans le biais masculin: Une approche expérimentale. *GLAD!*, *10*. DOI: https://doi.org/10.4000/glad.2839

Sato, S., Gabriel, U., & Gygax, P. M. (2016). Altering male-dominant representations: A study on nominalized adjectives and participles in first and second language German. *Journal of Language and Social Psychology*, *35*(6), 667–685. DOI: https://doi.org/10.1177/0261927X15625442

Sato, S., Gygax, P. M., & Gabriel, U. (2013). Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals. *Bilingualism: Language and Cognition*, *16*(4), 792–807. DOI: https://doi.org/10.1017/S1366728912000739

Sato, S., Gygax, P. M., & Gabriel, U. (2016). Gauging the impact of gender grammaticization in different languages: Application of a linguistic-visual paradigm. *Frontiers in Psychology*, *7*. DOI: https://doi.org/10.3389/fpsyg.2016.00140

Schunack, S., & Binanzer, A. (2022). Revisiting gender-fair language and stereotypes – A comparison of word pairs, capital I forms and the asterisk. *Zeitschrift für Sprachwissenschaft*, 1–29. DOI: https://doi.org/10.1515/zfs-2022-2008

Sczesny, S., Moser, F., & Wood, W. (2015). Beyond sexist beliefs: How do people decide to use gender-inclusive language? *Personality and Social Psychology Bulletin*, *41*(7), 943–954. DOI: https://doi.org/10.1177/0146167215585727

Simon, A. C., & Vanhal, C. (2022). Renforcement de la féminisation et écriture inclusive: Étude sur un corpus de presse et de textes politiques: *Langue française*, *N° 215*(3), 81–102. DOI: https://doi.org/10.3917/lf.215.0081

Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name your favorite musician: Effects of masculine generics and of their alternatives in German. *Journal of Language and Social Psychology*, *20*(4), 464–469. DOI: https://doi.org/10.1177/0261927X01020004004

Tanenhaus, M. K., & Carlson, G. N. (1985). Processing deep and surface anaphors. *North East Linguistics Society*, *15*(1), 24.

Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of deep and surface verbphrase anaphors. *Language and Cognitive Processes*, *5*(4), 257–280. DOI: https://doi.org/10.1080/01690969008407064

Tibblin, J. (2020). Les attitudes envers le langage inclusif des francophones et leur effet sur l'évaluation d'un texte. *Congrès Mondial de Linguistique Française – CMLF 2020, SHS Web of Conferences*(78), 13006. DOI: https://doi.org/10.1051/shsconf/20207813006

Tibblin, J., van de Weijer, J., Granfeldt, J., & Gygax, P. (2023). There are more women in joggeur·euses than in joggeurs: On the effects of gender-fair forms on perceived gender ratios in French role nouns. *Journal of French Language Studies*, *33*(1), 28–51. DOI: https://doi.org/10.1017/S0959269522000217

Tremblay, A., & Ransijn, J. (2020). *Package 'LMERConvenienceFunctions': Model selection and post-hoc analysis for (G)LMER models* (3.0). https://cran.r-project.org/web/packages/LMERConvenienceFunctions/LMERConvenienceFunctions.pdf

van Eijck, J., & Kamp, H. (1997). Representing discourse in context*. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 179–237). Elsevier. DOI: https://doi.org/10.1016/B978-044481714-3/50006-0

Xiao, H., Strickland, B., & Peperkamp, S. (2022). How fair is gender-fair language? Insights from gender ratio estimations in French. *Journal of Language and Social Psychology*, 0261927X221084643. DOI: https://doi.org/10.1177/0261927X221084643