Applied Psycholinguistics (2026), **47**, e4, 1–34
doi:10.1017/S0142716426100484

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Recognized but harder to integrate: An eye-tracking study of French gender-fair forms during reading

Julia Tibblin[1], Pascal Gygax[2], Joost van de Weijer[3] and Jonas Granfeldt[1]

[1]Centre for Languages and Literature, Lund University, Sweden, [2]Department of Psychology, University of Fribourg, Switzerland and [3]Lund University Humanities Lab, Lund University, Sweden
Corresponding author: Julia Tibblin; Email: julia.tibblin@rom.lu.se

(Received 22 January 2025; revised 19 November 2025; accepted 23 January 2026)

Abstract
Over the last few decades, linguistic gender-fair forms have become increasingly used by individuals and official institutions. In the French-speaking sphere, this has led to heated discussions among politicians and other stakeholders, some of whom claim that these forms render texts illegible and inaccessible to the general public. However, the processing of gender-fair forms in reading has been the topic of a few empirical studies. In the present paper, we add to this small body of research by reporting results from a pre-registered eye-tracking study where 58 native French-speakers read short texts which included a masculine form (*voisins*), complete double form (*voisines et voisins*), or contracted double form (*voisin·es*). Consistent with previous findings, the complete double forms were not more costly to process. In contrast, contracted double forms led to increased processing costs in intermediate and late stages of processing, but had no effect on the early stages of processing. Our data also indicate that the processing of contracted double forms becomes easier over time, and that it is facilitated by positive attitudes towards gender-fair language. These findings provide important insights that enlighten the current debate and should therefore be considered in the elaboration of official guidelines regarding gender-fair language.

Keywords: eye movements; French; gender-fair language; gender-inclusive language; processing cost; reading

Introduction

In grammatical gender languages, such as French or German, masculine forms can be interpreted in two ways. They can be interpreted specifically, just like the feminine form, in which case the referent is a male individual or a group consisting only of men. However, the masculine form can also be employed generically, when

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

referring to mixed-gender groups or to individuals whose gender identity is unknown or irrelevant. Thus, speakers of French cannot be completely certain of the gender composition of the group in a sentence like *Avant la cérémonie, les acteurs défilaient sur le tapis rouge* [Before the ceremony, the actors.$_{\text{MASC}}$ walked down the red carpet].

For almost two decades, a large body of research has investigated the generic use of the masculine form in various grammatical gender languages. The findings of these studies suggest that even when the masculine form is used with a generic intention, it is often the specific interpretation that is activated and prevailing (e.g., Gygax et al., 2013, 2019, 2021). Thus, the masculine form does not appear to be the most suitable option to achieve balanced mental representations of gender. In this field of research, it has been pointed out that the effects of the masculine form could actually have consequences on a societal level. For example, adolescents hearing job descriptions with only the masculine form showed greater gender bias in the ascription of success of the persons occupying the jobs, compared to adolescents hearing the same descriptions with both the feminine and masculine forms (Vervecken et al., 2015). In parallel to the development of research on the gender biases associated with the masculine form, users of grammatical gender languages, such as French, have started to employ alternative forms to avoid the masculine form when referring to mixed-gender groups or to individuals whose gender is unknown (Díaz & Heap, 2020; Flesch & de Beaumont, 2023; Simon & Vanhal, 2022). These forms are often referred to as gender-fair forms.

In French, common gender-fair forms include contracted (*acteur·rices* [actor.$_{\text{MASC·FEM.PL}}$]) or complete double forms (*actrices et acteurs* [actors.$_{\text{FEM}}$ and actors.$_{\text{MASC}}$]), as opposed to the masculine form (*acteurs* [actors.$_{\text{MASC}}$]). As gender-fair forms have become more frequent, researchers have begun to shift their focus from studying the masculine form in isolation to comparing its impact to that of gender-fair forms. Globally, they found that gender-fair forms increase the mental representations of women (e.g., Brauer & Landry, 2008; Kim et al., 2023; Pozniak et al., 2024; Richy & Burnett, 2021; Spinelli et al., 2023; Tibblin, Granfeldt, et al., 2023; Tibblin, van de Weijer, et al., 2023; Xiao et al., 2022).

As a consequence of the increased use of these new forms, researchers and stakeholders have expressed a note of caution, fearing that some gender-fair forms, such as the contracted double form, may be cognitively costly to process, and that, as a result, texts employing these forms might be less intelligible and less accessible to the general public. Some critics have been more cautious in their claims (e.g., Klinkenberg, 2019), whereas others have declared—yet without supporting data—that the implementation of some particular gender-fair forms could pose severe problems in reading, especially for second-language learners or readers with language disorders (Dister, 2021; Dister & Moreau, 2020; Grinshpun et al., 2020). However, only a few studies have looked at the processing of these forms in reading, especially using methods that access reading as the text is being processed.

Given that more and more gender-fair forms are being used in French, by everyday users and by institutions (see Simon & Vanhal, 2022), it is crucial to understand whether they actually cause increased processing costs for readers. Some studies addressed this question on the basis of self-paced reading tasks (e.g., Gygax & Gesto, 2007; Liénardy et al., 2023; Zami & Hemforth, 2024) or eye-tracking data,

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

but in the latter case, only exploratory studies were conducted (e.g., Girard et al., 2022). Since the processing of these forms was not studied in detail (i.e., distinguishing between different stages of the reading process, see section "*Language processing in reading*"), there is an important gap to fill in the research on the processing of gender-fair forms in French.

In this pre-registered[1] study, we address this gap by analyzing L1 French speakers' eye movements while reading short texts with masculine, complete, or contracted forms, allowing us to investigate the processing of these forms in detail. We also analyze how exposure to and attitudes towards gender-fair language impact the processing of these forms. As far as we know, this study is the first controlled, experimental study to do so.

Background

Gender-fair language in French

When French speakers choose not to use the masculine grammatical gender as a generic form, they typically resort to feminization and neutralization strategies. The feminization strategy consists of explicitly marking the presence of both women and men by using complete (*voisines et voisins* [neigbors.$_{FEM}$ and neighbors.$_{MASC}$]) and contracted (*voisin·es* [neighbor.$_{MASC·FEM.PL}$]) double forms. In contrast, the neutralization strategy includes all gender identities by employing forms that do not explicitly mark a referent's gender, like collective nouns (*le voisinage* [the neighborhood]), common gender nouns (*les locataires* [the tenants]), and epicene nouns (*les personnes* [the people]). Common gender nouns differ from epicene nouns in the way grammatical gender is marked on adjacent elements like determiners. While the former can take either feminine or masculine agreement, the latter can only agree with one grammatical gender (Corbett, 1991). Thus, one can say either *un* or *une locataire* [a.$_{MASC}$ or a.$_{FEM}$ tenant], but only *une personne* [a.$_{FEM}$ person] or *un individu* [an.$_{MASC}$ individual].

It is important to note that although the contracted double forms were originally proposed as an alternative to the complete ones when writers lacked space (Haut Conseil à l'Égalité entre les femmes et les hommes, 2016), they appear to be more common than their complete equivalents in everyday written French (Díaz & Heap, 2020; Kamblé-Bagal & Tatossian, 2022). Since they are also used by non-binary individuals to refer to themselves (Bolter, 2022), although a further *·x* is sometimes added, as in *un·e·x étudiant·e·x*, the contracted double forms could be considered as a hybrid strategy to be placed somewhere in between feminization and neutralization.

The effects of both strategies on mental representations of gender have been the topic of several empirical investigations over the last couple of years. Some studies have focused solely on comparing double forms to the masculine form (e.g., Brauer & Landry, 2008; Pozniak et al., 2024; Xiao et al., 2022), while others have compared gender-neutral forms to the masculine form (e.g., Kim et al., 2023; Richy & Burnett, 2021) or grammatically masculine or feminine epicene nouns to each other (e.g., Storme & Delaloye Saillen, 2024). Finally, some researchers have examined both strategies and have compared them either to each other or to the masculine

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

form (e.g., Spinelli et al., 2023; Storme & Storme, 2025; Tibblin, Granfeldt, et al., 2023; Tibblin, van de Weijer, et al., 2023). Taken together, these studies indicate that all gender-fair forms give rise to more balanced mental representations of gender compared to the masculine form.

In terms of official guidelines, the French High Council for Gender Equality promotes the complete double forms, but contracted double and gender-neutral forms are encouraged to some extent (Haut Conseil à l'Égalité entre les femmes et les hommes, 2022). Despite these recommendations, the French Senate recently adopted a law proposal intending to prohibit the use of gender-fair language—at least of contracted double forms—in certain official documents (see Proposition de loi visant à protéger la langue française des dérives de l'écriture dite inclusive, 2023). In both Québec and in Belgium, a combined use of gender-neutral and complete double forms is advised (Direction de la langue française, 2024; Office québécois de la langue française, n.d.). The differences across linguistic areas go beyond official recommendations, as research has shown such differences both in regard to attitudes towards gender-fair language (e.g., Tibblin, 2020) and to the use of gender-fair forms (e.g., Dumais, 2021; Kamblé-Bagal & Tatossian, 2022).

In the above-mentioned guidelines and in the public debate, it is often claimed that French gender-fair language, especially the double forms, would render the reading process difficult. It also constitutes the main counterargument used by the French Senate. In order to understand to what extent, and in what way, the gender-fair forms under investigation could influence reading processes, we must first grasp how language processing in reading operates. This is the aim of the following section.

Language processing in reading

Reading is a complex activity that encompasses several processes, such as extracting visual information, activating stored phonological and semantic representations, engaging in inference processing, and planning when and where to move the eyes next (Traxler, 2012). Although researchers have devoted over a century to the study of reading, few theoretical models aim to explain the reading process in its entirety (Rayner & Reichle, 2010). However, the Über-Reader model (Reichle, 2021; Veldre et al., 2020) is a recent model that aims to account for all perceptual, cognitive, and motor processes involved in reading. It consists of four components, each corresponding to a central process in reading: word identification, sentence processing, discourse processing, and the visual, attention, and oculomotor systems that program eye movements during reading (Reichle, 2021).

A crucial step of successful reading is identifying each word presented in a sentence, before integrating them into a coherent overall message (Cortese & Balota, 2012). Consequently, if the word identification process fails, the access of form- and meaning-related information in readers' mental lexicon is compromised, or at least more cognitively demanding. However, if the word-identification process is successfully completed, readers have access to the meaning of the word and can undertake syntactic parsing (i.e., the analysis of the grammatical structure of the sentence and the syntactic relationship between the words) and semantic interpretation (i.e., the integration of the meaning of a word into what has previously

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

been comprehended from the sentence and text) (Staub, 2015). In the present study, we are interested in whether two types of French gender-fair forms (i.e., complete, *voisines et voisins* [neigbors.$_{FEM}$ and neighbors.$_{MASC}$], and contracted, *voisin·es* [neighbor.$_{MASC·FEM.PL}$], double forms) lead to increased processing costs in reading and, if that is the case, at which stage(s) of the reading process these costs emerge.

A well-suited method to investigate the different stages of processing during reading is the use of eye-tracking, as its high spatio-temporal resolution can provide researchers with detailed insights into reading processes. Decades of studies of eye movements in reading have shown that the processing of words in reading is reflected in the time readers spend fixating on them (e.g., Rayner, 1998, 2009; Schotter & Rayner, 2015; Starr & Rayner, 2001). More specifically, different eye-movement measures reflect different stages of the reading process. Early measures are commonly taken as an indication of the word identification and lexical access processes, and late measures reflect syntactic parsing and semantic integration of the words into a coherent mental representation of the sentence and/or the text (Boston et al., 2008, Tables 3 and 4; Conklin et al., 2018, Chapter 3). While this *early vs. late measures* distinction might seem binary, these two categories should be used somewhat cautiously, as there is no strict one-to-one relationship between these measures and the different processes involved in reading (Kuperman & Van Dyke, 2011). Some authors even suggest adding a third, intermediate category (e.g., Conklin et al., 2018). Still, eye-tracking is advantageous over other more modest methods such as self-paced reading, as it offers researchers a richer picture of the different stages of the reading process. For example, analyses of *regressions* (i.e., backward eye movements to a previously viewed area of text) and *first fixation durations* (i.e., how long the eye stays on an area of the text at the first encounter) document very different stages of the reading process. It is thus a highly suitable method to test the assumption that gender-fair forms, especially contracted double forms, may hinder reading fluency.

In terms of the stages of the reading process, some gender-fair forms may constitute an initial obstacle to a smooth reading process by hindering lexical access. Given that the complete double forms include nouns that appear in standard French, it seems unlikely that they would impede such an early process. However, this seems more likely for contracted double forms, which clearly stand out from standard French, making it more difficult to access the mental lexicon. Interestingly, specific assumptions as to the timing of lexical access have been put forward on the basis of eye-tracking and EEG reading studies. One such assumption is the timeline proposed by Sereno and Rayner (2003), which specifies that this process occurs between 100 and 200 ms from stimulus onset. According to this timeline, the programming of the eye movements is initiated long before the brain signals the eyes to move forward, something we will come back to when discussing our results.

Individual differences and processing in reading

In addition to intra-linguistic factors (e.g., orthographic length, regularity, and word frequency), extra-linguistic, participant-related factors could also influence reading processes. In fact, recent models of reading increasingly account for readers' background knowledge, including their motivation and attitudes (e.g., Kim, 2020).

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Some have also suggested that *subjective frequency* (i.e., how familiar a person is with a word) could be a stronger predictor of lexical processing than *objective frequency* (i.e., the occurrence of a word in a reference corpus) (e.g., Cortese & Balota, 2012). In the context of our study, attitudes towards gender-fair language are a particularly interesting variable, as positive attitudes have sometimes been found to influence the effects of gender-fair forms on mental representations (e.g., Anaya-Ramírez et al., 2022; Braun et al., 2005; Tibblin, van de Weijer, et al., 2023) and on processing cost (e.g., Román Irizarry & Guzzardo Tamargo, 2025; Steiger-Loerbroks & von Stockhausen, 2014). More specifically, participants with positive attitudes towards gender-fair language may be more exposed to and may use these forms more frequently, which could facilitate their processing, while participants holding negative attitudes may process them more reluctantly, especially the contracted double forms.

In the following section, we address the studies that have investigated the processing of gender-fair forms in French and in other grammatical gender languages, some of which included attitudes towards gender-fair language as a factor.

Previous research: processing gender-fair forms in reading

The number of studies on the interpretation of certain French gender-fair forms has increased quickly during the last couple of years (see section "*Gender-fair language in French*"). However, studies investigating the processing of these forms in reading remain scarce. To our knowledge, for French, there have been three studies using a self-paced reading task (Gygax & Gesto, 2007; Liénardy et al., 2023; Zami & Hemforth, 2024) and one exploratory eye-tracking study (Girard et al., 2022). In this section, we begin with an overview of the results of these studies, briefly address similar studies conducted in other languages, and end with a summary taking both overviews into account.

Processing gender-fair forms in French

In a self-paced reading task, Gygax and Gesto (2007) let 39 Swiss French speakers read five texts containing three occurrences of a role noun. The role nouns were presented either in their masculine (*maçons* [bricklayers.$_{\text{MASC}}$]) or feminine (*maçonnes* [bricklayers.$_{\text{FEM}}$]) form, or as a complete (*maçons ou maçonnes* [bricklayers.$_{\text{MASC}}$ or bricklayers.$_{\text{FEM}}$]) or contracted double (*maçon-ne-s* [bricklayer.$_{\text{MASC-FEM-PL}}$]) form. The form of the role nouns varied between participants. The results, based on orthogonal contrast analyses which controlled for the differences in length between the target items, showed significantly longer reading times at the first encounter of feminine and contracted double forms. However, the difference between the four conditions lessened by the second and third encounter. The authors thus concluded that the contracted double and feminine forms may initially slow down reading, but that there is a strong habituation effect. They also concluded that the complete double forms did not slow down the reading at any time.

Over a decade later, Liénardy and colleagues (2023) used a similar design as Gygax and Gesto (2007), but only compared the influence of complete (*autrices et*

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

auteurs [authors.$_{FEM}$ and authors.$_{MASC}$) and contracted double (*auteur·rices* [author.$_{MASC·FEM.PL}$) forms to that of the masculine form (*auteurs* [authors.$_{MASC}$]). Using a between-participants design, they had 256 Belgian French speakers read descriptions of 18 different professions, each consisting of six sentences. In these descriptions, the manipulated role noun appeared in sentences 1, 3, and 5. The results showed no significant effects of the form on reading times of the manipulated role noun regardless of its position in the text. This means that no habituation effect could be observed, as no increased reading times were found in the first place.

Since the results of this study notably differed from those of 2007, it is worth highlighting a few differences between the two. First, the study took place online, giving the researchers less control over participants' behavior during the reading task. In other words, there is a possibility that some participants did not fully focus on the task (if, for example, they were interrupted by their phone or other notifications), in which case the reading times would not be indicative of their actual reading behavior. Second, once each text was read, the participants were asked to estimate the percentage of women present in the described group. One could imagine that this additional task influenced the reading strategies adopted by the participants, making the task slightly less ecologically valid than studies that focus on reading only. Lastly, the debate and practices concerning French gender-fair forms have changed considerably from 2007 to 2023, and it is likely that French speakers are more used to reading these forms in 2023 than in 2007.

Finally, Zami and Hemforth (2024) let 30 French speakers, living in France, read 24 sentences, each split into six segments, in which a role noun appeared in the masculine (*les conducteurs* [the drivers.$_{MASC}$]) or the contracted double form (*les conducteur·rice·s* [the driver.$_{MASC·FEM·PL}$]). Half of the participants read either the masculine or the contracted double form, whereas the other half were exposed to both forms. The analyses indicated that the contracted double forms led to significantly longer reading times of the critical words, while no difference was observed in the analyses of the spillover region. Descriptive data also showed a stronger habituation effect for the contracted double forms compared to the masculine form. However, this interaction was not tested inferentially.

In the only eye-tracking study to date investigating some specific French gender-fair forms, Girard and colleagues (2022) had 18 French speakers from Quebec read 40 sentences. Each sentence contained a role noun presented in one of four conditions, to which all participants were exposed: masculine form (*étudiant* [student.$_{MASC}$]), contracted double forms using either a full stop (*étudiant.e* [student.$_{MASC.FEM}$]), an interpunct (*étudiant·e* [student.$_{MASC·FEM}$]), or parentheses (*étudiant(e)* [student.$_{MASC(FEM)}$]). Total fixation durations on the critical word, as well as the probability of regressing into the critical region, were analyzed. The analyses, although exploratory given the rather low number of participants, showed no significant differences between the conditions for any outcome measure.

As the studies presented in this overview concern different populations in the francophone world, one could argue that it may be difficult to compare them to each other and to the present study. However, although scarce, previous studies seem to suggest high homogeneity across different francophone samples in terms of the mental representations from gender-fair and unfair forms (e.g., Kim et al., 2023). In fact, what mainly distinguishes speakers from different areas from each other may

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

be their attitudes towards gender-fair language (as hinted by Tibblin, 2020), a variable that we specifically include in the present study.

Processing gender-fair forms in other languages

As mentioned in the introduction, gender-fair language is not a phenomenon limited to French, but the way new gender-fair forms take shape varies across languages. To mention a few examples, the most prominent German gender-fair form is the so-called *gender star form* (e.g., *Biolog*in* [biologist.$_{\text{MASC*FEM}}$]) (Waldendorf, 2024), while speakers of Spanish tend to replace the binary suffixes *-a* (fem.) or *-o* (masc.) with a non-binary neo-morpheme like *-x* or *-e* (e.g., *todxs* or *todes* [all.$_{\text{NON-BINARY}}$]) (Slemp, 2021), mimicking the Italian *schwa* which functions as a non-binary morpheme (e.g., *avvocatə* [lawyer.$_{\text{NON-BINARY}}$]) (Abbondanza et al., 2025). In addition to these novelties, complete double forms as well as different neutralization strategies are available in all these languages.

Importantly, despite superficial resemblances (*Biolog*in* in German vs *étudiant·e* in French), similar gender-fair forms may at times have different connotations across languages. For example, the German gender star form is framed as an explicit marker of non-binarity, while the French contracted double forms were, at least originally, put forward as an abbreviation of the binary complete double forms. Even within a language, different typographical signs can have different ideological connotations (see Burnett & Pozniak, 2021), and their use can even vary across linguistic areas.

Empirical studies on these linguistic innovations have often focused on their interpretation (i.e., the effects of gender-fair forms on mental representations of gender), but some have taken a closer look at their cognitive processing. For example, a recent eye-tracking study on Spanish non-binary pronouns (e.g., *todxs* [all.$_{\text{NON-BINARY}}$]) found these novel pronouns to increase processing costs in all stages of the reading process when compared to the masculine form (Román Irizarry & Guzzardo Tamargo, 2025). In German,[2] the question of whether the gender star form may impede lexical access was recently investigated by Zacharski and colleagues (2025) by means of a lexical decision task. Their results showed high acceptance rates (over 95%) for the star form overall, but in their older, non-student sample, participants were slower to judge words composed by the gender star form as real words compared to both feminine and masculine forms.

Interestingly, both the above-mentioned studies, as well as an eye-tracking study on gender-neutral nominalizations in German (e.g., Steiger-Loerbroks & von Stockhausen, 2014), included participants' attitudes towards gender-fair language as an additional predictor of the processing of the gender-fair forms. The findings of these studies are somewhat inconclusive, but they suggest that the processing of gender-fair forms could be facilitated by positive attitudes towards them. In addition, Zacharski and colleagues (2025) found a similar habituation effect to the one reported in some studies on French gender-fair forms (see the previous section).

Taken together, these two overviews indicate that the French contracted double forms are likely to be more costly to process, but that they may not necessarily affect the early stages of the reading process (i.e., lexical access), and this may depend on attitudes towards gender-fair language and exposure to these forms. Furthermore,

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

the complete double forms may not be more costly to process, at least when controlling for their length. These overviews also underline the scarcity of studies using eye-movement data to investigate the processing of gender-fair forms in general, and French forms in particular. The present study, therefore, fills an important gap in the current literature.

The present study

The overarching aim of the present study is to investigate the processing cost in reading of two types of gender-fair forms: complete and contracted double forms. Given that the masculine and the gender-fair forms differ both in length and in structure (as shown in Example 1), comparing these three forms to each other is somewhat like comparing apples and oranges.

1. a) *voisins* [neighbors.$_{\text{MASC}}$]
   b) *voisin·es* [neighbor.$_{\text{MASC·FEM.PL}}$]
   c) *voisines et voisins* [neighbors.$_{\text{FEM}}$ and neighbors.$_{\text{MASC}}$]

First, the masculine form (1a) is a simple NP containing only the noun. Second, in the contracted double forms (1b), the NP also contains only one noun, but includes a typographical sign that is not used in standard French, namely the interpunct. In addition, the interpunct separates the inflectional morphemes *-e* and *-s*, marking the feminine and the plural, from the stem, respectively. Finally, the complete double form (1c) contains two nouns conjoined by the conjunction *et* [and].

These structural differences make it more challenging to informatively compare these forms using standard processing measures, such as reading times. Namely, it would not be surprising, and thus not very informative, to observe longer reading times with the conjoined NP in (1c) than with the simple NP in (1a), simply because it contains more characters. Alike, the NP in (1b) also contains more characters than the simple one, and in addition, it contains a typographical sign that splits the word in two. Since the interpunct is not used in standard French, and since typographical signs in French are typically not found between a stem and an inflectional morpheme, one may expect that reading times will be longer, given that infrequent and unexpected words typically require more processing (see section "*Language processing in reading*").

To make the comparison between these different forms as informative as possible, we included a control condition next to the experimental condition, which will be explained in further detail in the Methods section. Thus, in the present study, we manipulated two variables that varied within-participants (NP structure [Simple, Hyphenated, or Conjoined] and Condition [Experimental or Control]) and included two additional predictor variables: exposure to the gender-fair forms throughout the experiment and participants' attitudes towards gender-fair language. Our hypotheses relating to these variables were the following[3]:

(H_1) Gender-fair forms will lead to increased processing costs as compared to the masculine form, even when controlling for differences in length and NP structure.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Based on the previous literature (e.g., Gygax & Gesto, 2007; Zami & Hemforth, 2024), it might be the case that only the contracted double forms lead to increased processing costs ($H_{1\text{-}ALT}$).

(H_2) Increased exposure throughout the experiment will facilitate the processing of the contracted double forms as compared to the masculine form, in line with the habituation effect found in previous studies (e.g., Gygax & Gesto, 2007; Zacharski et al., 2025; Zami & Hemforth, 2024).

(H_3) Positive attitudes towards gender-fair language will facilitate the processing of the gender-fair forms under investigation, as found in one previous study (e.g., Román Irizarry & Guzzardo Tamargo, 2025).

Method

This section, as well as the following ones, adheres, as closely as possible, to the best practices for reporting eye-tracking research recommended by Carter and Luke (2020). All materials, the frequency values of the NPs, and all code used to run the experiment can be found on the OSF repository of the project: https://osf.io/6hxm3/.

Research design

As previously mentioned, we included a control condition next to the experimental one in an attempt to control for the differences in length and NP structure between the forms under investigation. The items used in the control condition were NPs of similar length and structure to the experimental ones (see Figure 1 for an overview of the research design and example items). With such a design, we do not directly compare the processing cost of the gender-fair forms to that of the masculine form, but we are interested in the difference in processing cost between the experimental and control items for each NP structure, i.e., the comparisons represented with solid lines in Figure 1. Therefore, we ensured that our comparison of gender-fair forms is as accurate as possible by matching them with NPs of similar length and structure. The selection of control items is further explained and problematized in the Materials section below.

For the second hypothesis, we used trial number as an indication of increased exposure throughout the experiment. To measure the participants' attitudes towards gender-fair language, relevant to the third hypothesis, we employed an eight-item scale used in previous studies (e.g., Sauteur et al., 2023; Tibblin, 2020) on a five-point Likert scale ranging from *Pas du tout d'accord* [I don't agree at all] to *Tout à fait d'accord* [I agree completely]. As such, a high score on the scale reflects positive attitudes. This scale draws on items from previous, more comprehensive scales (e.g., Parks & Roberton, 2000; Prentice, 1994; Sczesny et al., 2015). Four items were reverse-scored, for example, L'*usage du masculin générique est une propriété de la langue française et n'a rien à voir avec les inégalités entre les sexes dans la société* [The

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

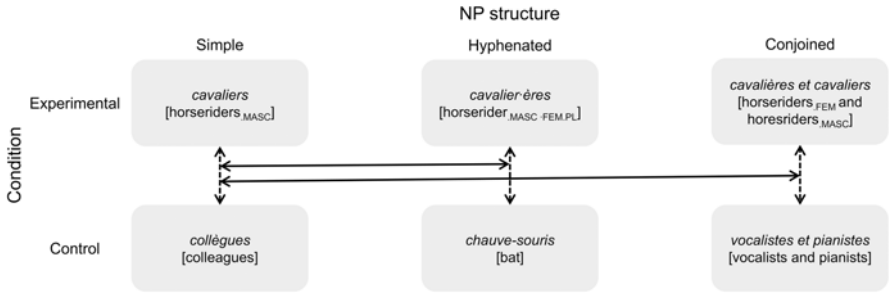


Figure 1. Overview of the research design with example NPs in each condition. Dashed lines indicate the direct comparisons within each NP structure, whereas the solid lines indicate the indirect comparisons between NP structures, i.e., the comparisons of interest.

use of masculine forms to refer to generic nouns is a property of the French language and has nothing to do with social inequalities between males and females].

Participants

Seventy-five participants with normal or corrected-to-normal vision underwent the experiment. However, data from seventeen participants (23%) were excluded (see section "*Data screening*" for details). Thus, the final sample consisted of 58 L1 French speakers (median age = 22 years, IQR = 3 years, range = 18–57 years), 90% of whom had lived the majority of their lives in France. Other responses included French-speaking parts of Switzerland (3 participants) and Belgium (1 participant), or a non-French-speaking country (2 participants). Forty-two participants (72%) self-identified as women, 14 as men, and two participants indicated another gender identity. The median length of stay in Sweden prior to the experiment was one month (IQR = 1.5 months, range = 0–60 months[4]).

Apparatus

Binocular eye movements were recorded with a desk-mounted Tobii Pro Spectrum eye-tracker at a sample rate of 600 Hz. This eye-tracker has a reported median accuracy of 0.30° and a precision of 0.06° RMS in optimal conditions (Tobii AB, 2018). A head- and chinrest was used to minimize head movements, and separator screens were attached to three sides of the desk to avoid possible distractions. The distance between the participants' eyes and the screen was approximately 63 cm. The stimuli were presented in PsychoPy2 (Peirce et al., 2019) on a 23.8″ EIZO FlexScan EV2451 screen at 1920 × 1080 resolution. We used the Titta toolbox (Niehorster et al., 2020) to connect the eye-tracker to the experiment in PsychoPy.

Materials

The target NPs in the experimental condition consisted of 18 role nouns, all of which were non-stereotyped in terms of gender (see Tibblin, van de Weijer, et al., 2023). They were presented either in the masculine (*joggeurs* [joggers.$_{\text{MASC}}$]),

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

complete double (*joggeuses et joggeurs* [joggers.$_{FEM}$ and joggers.$_{MASC}$]), or contracted double form (*joggeur·euses* [jogger.$_{MASC·FEM.PL}$]). These role nouns include both nouns where the feminine form is obtained by adding a suffix to the masculine form, e.g., *voisin·es*, ($n = 6$), and nouns where it is obtained by replacing the masculine suffix with its feminine counterpart, e.g., *skieur·euses*, ($n = 12$). In the control condition, the target NPs were of similar length as the experimental equivalents and included either common gender nouns (e.g., masculine: *juristes* [legal experts]; complete double: *cyclistes et arbitres* [cyclists and referees]) or hyphenated compound nouns (e.g., contracted double: *rouge-gorges* [robins]). As such, all the nouns included as control items were nouns that cannot be inflected for grammatical gender. These were selected by the following search procedures in the Lexique2 database (New et al., 2004).

First, control NPs in the masculine form were found by searching for common gender nouns (defined as nouns marked neither as feminine nor masculine in their plural form) of the same length as the experimental NPs, and that, if possible, shared their initial and final letters. Second, control NPs to the complete double forms were constructed by scanning the list of common gender nouns for nouns that could compose semantically suitable pairs (e.g., *enfants et adults* [children and adults] or *artistes et peintres* [artists and painters]) containing approximately (+/− one character) the same number of characters as the experimental NPs. By semantically suitable, we mean that the two nouns were related in meaning in such a way that they could be easily integrated into the context and naturally function as sentence subjects. This was verified by multiple L1 French speakers who were explicitly instructed to signal if any word seemed unnatural or unexpected. Finally, control NPs to the contracted double forms were selected by extracting hyphenated compound nouns of similar length as the experimental NPs from the database. We then explored this list searching for concrete nouns (e.g., preferring *chauve-souris* [bats] to *contre-mesure* [countermeasure]) that we deemed as possible topics of a short text.

While the control items of the gender-fair forms may not be fully comparable to their experimental equivalents, we believe that the selected research design allows for a more accurate comparison than a simple comparison between the gender-fair and masculine forms (as previously done). It is true that the complete double forms repeat the same lexical item within the NP, unlike their control equivalents, but we chose this comparison due to a lack of more suitable options. Along the same lines, we are aware that one cannot expect contracted double forms to be processed the same way as hyphenated compound nouns, given, among other factors, their novelty. However, French hyphenated common gender nouns are extremely scarce,[5] which is why we included nouns not referring to people. It is also true that there are semantic differences between the control and experimental NPs, as the meaning of *joggeur·euses* is, at least roughly, equivalent to the sum of its two constituents, whereas that is not the case for compound nouns (i.e., the meaning of *chauve-souris* [bat] is not the sum of *chauve* [bald] and *souris* [mouse]). Nevertheless, we believe this was the best option as we did not want to introduce neologisms.

Once the full list of target NPs was finalized, they were each embedded in a three-sentence-long text, see example (2) below (NP structure: hyphenated, Condition: experimental, i.e., the contracted double form):

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Tous les deux ans, un très grand concours de violon s'organise en Angleterre.
Pour ce concours, les joueur·euses doivent se préparer minutieusement.
Chaque année, une nouvelle ville anglaise accueille la compétition.

Every second year, a very big violin contest takes place in England.
For this contest, the player.$_{\text{MASC·FEM.PL}}$ must prepare themselves carefully.
Each year, a new English city hosts the competition.

In the experimental condition, the texts only differed regarding the form of the role noun they contained, whereas a new text was written for each control item. Thus, the complete materials consisted of 72 texts (18 for the experimental items and 54 for the control items). The texts were presented with double spacing on a gray background in black, bold Courier New font at a size of 25 px, giving an x-height of approximately 11 px or 0.27° of visual angle. For the experiment, a Latin-square design was employed to create three lists in which each participant read six role nouns in each of the three experimental forms (masculine, complete double, and contracted double), and the control-condition equivalent of each experimental item, giving a total of 36 trials per participant.

Procedure

The data collection took place in the Digital Classroom at the Lund University Humanities Lab. Participants were recruited through different channels, including mailing lists for exchange students, social media, and posters in university buildings, and were told they would engage in a reading study in French. A maximum of four participants could undertake the experiment simultaneously, but most participants (55%) underwent the experiment alone or in pairs. All participants received both oral and written instructions and gave their consent before beginning the experiment. Once they had adjusted the desk and chinrest to a comfortable position, a 9-point calibration and 4-point validation procedure began. This procedure was repeated until the accuracy error was approximately 0.3° or until it had been repeated three times. The mean accuracy and precision error for the entire sample were 0.32° ($sd = 0.09$) and 0.06° RMS ($sd = 0.03$), respectively.

Upon completion of the calibration procedure, the experiment started with a brief repetition of the instructions to remind the participants to read the texts silently at their normal reading speed, as if they were reading a magazine. At the start of each trial, a fixation cross was presented where the first letter of the text would appear. After 1000 ms, the text appeared and remained on the screen until the space key was pressed, or for a maximum of 25 seconds. Following this, the participants were asked to rate the text on a 5-point Likert scale ranging from *Pas du tout* [Not at all] to *Tout à fait* [Absolutely] (for example, *A quel point avez-vous trouvé cette information improbable?* [To what degree did you find this information unlikely?]). The question, whose purpose was to keep the participants attentive, remained on the screen until an answer was recorded. The participants were given three practice trials before the real experiment began.

When all 36 texts were read, participants could take a short break before they underwent a second reading experiment, unrelated to the one reported here, which

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

did not focus on gender-fair language. Subsequently, we measured participants' attitudes towards gender-fair language, and asked questions relating to their knowledge and personal uses of gender-fair language, and the frequency with which they encounter gender-fair forms. Finally, the participants underwent the Brussels Orthographic Quality Scale (Chetail et al., 2019) as it was relevant for the second reading experiment, and answered a battery of socio-demographic questions, including parts of the French version of the Language Experience and Proficiency Questionnaire (Marian et al., 2007). Upon completion, participants were debriefed and rewarded with a cinema ticket (worth approximately €15) for their participation. On average, participants spent 9 minutes on the reading task, and 30 minutes on the entire session.

Analyses

Due to limited space, the descriptions of some procedures in this section are slightly abbreviated. Additional details, as well as the codes used to calculate and analyze the measures, are available on the OSF repository of the project.

Data preparation

The transformation of the raw data to fixations was done in line with Titta, a toolbox for creating PsychToolbox and Psychopy experiments with Tobii eye trackers (Niehorster et al., 2020).[6] Thus, fixations were identified by means of a two-means clustering (I2MC) algorithm (see Hessels et al., 2017 for a detailed account of the algorithm). We used the Python-implemented version of this algorithm with the default settings, meaning that fixations with a duration shorter than 40 ms were removed upon completion of the fixation detection procedure, and that no maximum fixation duration was defined. Once fixations had been identified, they were mapped to AOIs. Next, we visually inspected each participant's scanpaths and identified a slight vertical offset between the recorded gaze position and the text for some participants. Since these data were reliable in all other aspects, we compensated for this offset by adjusting the fixations between 10 and 30 px vertically for these participants. This is not an optimal solution (see Holmqvist et al., 2011, p. 314), which is why we reran the analyses with a subset of the sample including only participants whose data had not been adjusted ($n = 44$). The results did not differ from the entire sample, and all participants were therefore kept in the analyses reported below.

The analyzed eye-movement measures are listed and defined in Table 1. As specified in the pre-registration, we initially set out to analyze both local (i.e., with the critical NP as AOI) and global (i.e., with the entire text as AOI) measures, but for two reasons, we report only the local measures in the present paper. First, the majority of the global measures were significantly correlated with each other ($p < .05$), and second, the fixed effects of the linear mixed-effects models only explained between 0.6 and 4% of the variance. Therefore, we refer to the OSF repository of the project for model outputs of the global measures (i.e., mean fixation duration, percentage of regressions, reading time per character, and total number of fixations).

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Table 1. All analyzed, local eye-movement measures, the cognitive process they are hypothesized to reflect (see Boston et al., 2008; Conklin et al., 2018), and their definition

Measure	Reflected process	Definition
First fixation duration	Early: Word identification	The duration of the first fixation in the AOI.
Gaze duration	Early: Word identification, text integration	The sum of all fixations in the AOI from the first fixation until the AOI is exited to the right or to the left.
Regression-path duration	Intermediate: Word integration difficulty, syntactic processing	The sum of all fixations in the AOI and regressions to earlier parts of the text before the AOI is exited to the right.
Regressions-in	Intermediate: Regressions are difficult to classify as they can be caused by several different problems (Conklin et al., 2018)	Whether the fixation landing on the AOI was preceded by a regression or not.
Total fixation duration	Late: General comprehension difficulty, syntactic processing	The sum of all fixations made in the AOI.

Data screening

Data screening was conducted first on a participant level and then on a trial level. In the by-participant screening process, we excluded 17 participants. Seven participants were excluded because they did not fit the required profile (i.e., they had lived in Sweden for over five years or did not indicate French as their most dominant language), and ten due to poor calibration or data quality. Binocular eye movements were recorded for all participants, but if, for any eye, the calibration accuracy error was above 0.7° or the mean percentage of data loss across all trials was above 20%, only data from the other eye were kept. Consequently, only data from one eye was analyzed for six participants.

The by-trial screening was done in several steps. First, 22 trials with a total reading time below 1000 ms were removed. Second, 15 trials where the participant had accidentally pressed the space key too early, as indicated by the visual inspection of the scanpaths, were removed. Third, five trials were excluded because the reading time reached the maximum threshold of 25,000 ms. Finally, we would exclude trials with a percentage of data loss above 20%. However, this was not the case for any trial. In total, the by-trial screening led to the removal of 2% of the data.

Inferential analyses

All analyses were conducted with R (R Core Team, 2021) and are in keeping with the pre-registration except for the regressions-in variable, which was analyzed as a binary variable (i.e., whether the fixation on the NP was preceded by a regression or not) rather than as a proportion of regressions. We made this adjustment because in many trials, very few regressions were made and, consequently, these trials had a proportion of regressions of 0.5 or 0.33. Since this adjusted analysis provides one

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

data point per fixation on the NP rather than one per trial, it allows for a more fine-grained analysis. In accordance with the pre-registration, the frequency value of each NP refers to lemma frequency from the written corpus of the Lexique2 database (New et al., 2004), except for the conjoined control NPs whose values reflect the mean frequency of the two NPs.

Prior to the analyses, we log-transformed the four continuous outcome variables (i.e., all except for the regressions-in which was binary) to reduce skewness of the data, centered the variables *Attitudes towards gender-fair language*, *Length in characters*, and *Frequency* to their means ($M_{\text{ATTITUDES}} = 3.8$ ($sd = 0.8$), $M_{\text{LENGTH}} = 14.8$ ($sd = 6.4$), and $M_{\text{FREQUENCY}} = 18.4$ ($sd = 42.3$)), and dummy coded the variables *NP structure* and *Condition* with Simple and Control as reference levels. Thus, in all models reported below, the intercept refers to NPs with a simple NP structure in the control condition (e.g., *collègues* [colleagues]).

All continuous outcome variables were analyzed separately with linear mixed-effects modeling using the *lmer* function of the lme4 package (Bates et al., 2021), while the binary variable was analyzed through generalized linear mixed-effects modeling with binomial distribution using the *glmer* function of the same package. For each outcome variable, a maximal model was created containing a fixed-effects structure consisting of two three-way interactions (one including the variables *NP structure, Condition, Trial number,* and one including *Attitudes* instead of *Trial number*) as well as their main effects, with the addition of Length and Frequency as main effects. The interaction between NP structure and Condition is central for all three hypotheses, while the interaction including Trial number concerns H_2, and the one including Attitudes relates to H_3. The random-effects structure consisted of random intercepts of Participant and Lexeme and by-participant random slopes for NP structure, Condition, and Trial number. This model was entered into the *fitLMER.fnc* function of the LMERConvenienceFunctions package (Tremblay & Ransijn, 2020), which 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.

Next, we created a second model, which included any interaction kept by the *fitLMER.fnc* function, or the one between NP structure and Condition if no interaction was kept, the main effects of Trial number, Length, and Frequency, and the random intercepts of Participant and Lexeme. Due to convergence errors, the generalized linear mixed-effects model did not contain the main effects of Trial number, Length, and Frequency. For all models, we attempted to add by-participant random slopes for NP structure, Condition, and Trial number, but only random intercepts were kept since the inclusion of any random slope, regardless of which, led to singular fits in all analyzed measures. We then compared the Akaike information criterion (AIC)s of these two models, and since the AIC of the manually created model was lower than that of the one created by the *fitLMER.fnc* function for all outcome variables, the manual model was kept as the final model. This model-selection procedure deviates slightly from the pre-registration, in which we stated that we would keep the model output from the *fitLMER.fnc* function and only update it with the interaction between NP structure and Condition. However, given that both the marginal R^2 and the AIC values indicated that the manual models

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

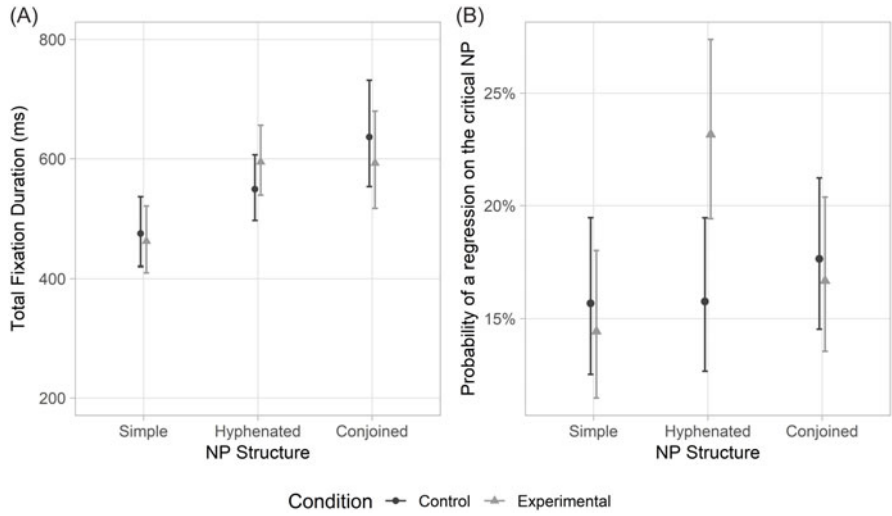


Figure 2. Total fixation durations (ms) (A) and probability of regression-in (B) by NP structure and Condition as predicted by the best model of fit. The vertical lines indicate the 95% confidence interval.

seemed to better fit the data, we chose to prefer them over the models produced by the *fitLMER.fnc* function. In addition, this procedure allowed us to have the same random-effect structure for all outcome variables and to always control for frequency, length, and trial number.

Finally, we used the *tab_model* function from the sjPlot package (Lüdecke, 2024) to output model summaries. This function calculates marginal and conditional R^2 values for the model, values which represent the proportion of variance explained by the fixed effects alone (marginal R^2) and by both the fixed and random effects (conditional R^2) (Nakagawa & Schielzeth, 2013). We also calculated effect sizes for the linear mixed-effects models in the form of Cohen's *d* by dividing the difference in predicted values by the square root of the sum of the variance components (Brysbaert & Stevens, 2018). When referring to the size of these effects, we use the empirically derived benchmarks for psychological studies suggested by Sommet et al. (2023), namely smaller = 0.20, median = 0.35, and larger = 0.50. Furthermore, we used the *ggpredict* function from the ggeffects package (Lüdecke, 2018) to obtain the predicted values visualized in Figures 2–4.

Results

In the following section, the results for each hypothesis are presented separately. We provide model outputs for gaze duration (Table 4), regression-path duration, total fixation duration, and regressions-in (Table 3). Since the final model of the first fixation durations was not very informative (no significant effects were found, and the amount of variance explained was below 20%), it is not presented here but available on the OSF repository of the project. Descriptive plots of the three-way interactions hypothesized by H_2 and H_3 are also available on the repository.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

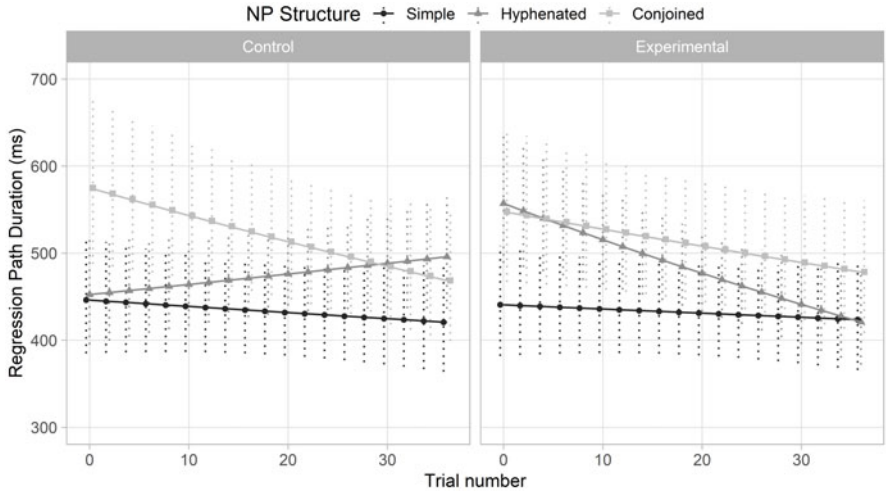


Figure 3. Predicted regression-path durations by Trial number, grouped by Condition and NP structure. The vertical lines indicate the 95% confidence interval.

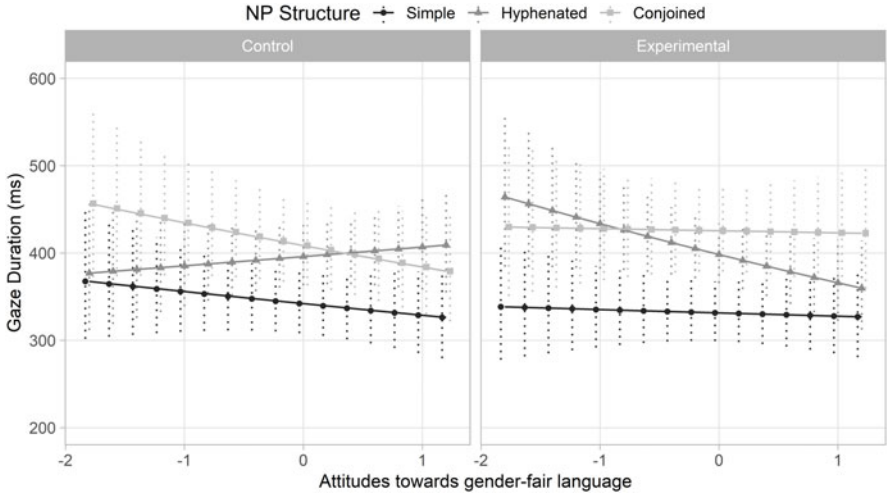


Figure 4. Predicted gaze durations by Attitudes towards gender-fair language, grouped by Condition and NP structure. The vertical lines indicate the 95% confidence interval.

H1: Gender-fair forms lead to increased processing costs

The first hypothesis concerned the difference between the two gender-fair forms and the masculine one. For this hypothesis to be confirmed, we should observe a significant interaction effect between the variables NP structure and Condition. More precisely, the difference between the control and experimental conditions should, for each outcome variable, be larger within a hyphenated or conjoined NP structure than within a simple NP structure (H_1). Since previous studies

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Table 2. Mean (standard deviation) first fixation duration (FFD), gaze duration (GD), regression-path duration (RPD), total fixation duration (TFD), and actual percentage of regressions out of all fixations on the critical word (reg-in). Values are in milliseconds except for Reg-in

NP Structure	Simple		Hyphenated		Conjoined	
Condition	Control (n = 339)	Exp (n = 339)	Control (n = 339)	Exp (n = 338)	Control (n = 339)	Exp (n = 340)
FFD	251 (101)	247 (88)	232 (84)	240 (83)	251 (95)	251 (97)
GD	311 (149)	297 (133)	415 (195)	445 (306)	612 (335)	649 (368)
RPD	394 (293)	387 (278)	496 (285)	542 (439)	829 (535)	822 (467)
TFD	452 (323)	427 (262)	583 (341)	660 (455)	984 (550)	920 (521)
Reg-in (%)	17	16	17	25	19	18

(Gygax & Gesto, 2007; Zami & Hemforth, 2024) only found an effect of contracted double forms, and not of complete ones, we also presented an alternative version of H_1 in which only the contracted double forms would lead to increased processing costs ($H_{1\text{-ALT}}$).

As seen in Table 2, the descriptive statistics of all measures show greater differences between the two conditions for hyphenated NP structures compared to simple ones (in line with $H_{1\text{-ALT}}$), whereas only the gaze durations show such a difference for conjoined NP structures (in keeping with H_1).

In the inferential analyses, the interaction of interest (i.e., the one between Condition and NP structure) was significant for regression-path durations, in which case it was medium-sized, and for regressions-in (Table 3, line 8). For total fixation durations, this interaction did not reach significance, but the increase predicted by the final model was around 50ms (see Figure 2). As expected, the difference between the experimental and the control condition within the hyphenated NP structure (e.g., *cavalier·ères* [horse rider.$_{\text{MASC·FEM.PL}}$] vs. *chauve-souris* [bat]) differed either significantly (in the case of regression-path durations and regressions-in) or visually (in the case of total fixation durations) from the difference observed within a simple NP structure (e.g., *cavaliers* [horse riders.$_{\text{MASC}}$] vs. *collègues* [colleagues]). This interaction is visualized in Figure 2, which shows the total fixation durations (plot A) and chances of a regression on the critical NP (plot B) as predicted by the final models. The fixed effects of the regression-path duration and total fixation duration models account for about a third of the variance in the data (31%), as indicated by the marginal R^2-values, whereas the fixed effects of the regressions-in model only account for 1% of the variance. This indicates that there are other factors, which we have not taken into account, that could explain the regressions-in variable. Table 3 also shows that the three-way interaction between NP structure, Condition, and Trial number was significant for regression-path durations (line 13). This result will be discussed in the following section.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Table 3. Models of best fit for regression-path durations, total fixation durations, and regressions-in. Asterisks indicate *p*-values (* = <0.05, ** = <0.01, *** = <0.001). Note that the output for regressions-in varies since it is a generalized linear mixed-effects model and not a linear mixed-effects model. The number of observations also differs since in the first two models each data point represents one trial, whereas in the third each data point represents one fixation on the critical NP

	Regression-path duration (log)				Total fixation duration (log)				Regressions-in		
Predictors	*Est.*	*95% CI*	*Statistic (t)*	*Cohen's d*	*Est.*	*95% CI*	*Statistic (t)*	*Cohen's d*	*Odds Ratio*	*95% CI*	*Statistic (z)*
(Intercept)	6.10	5.95–6.25	81.11		6.24	6.12–6.37	96.41		0.19	0.14–0.24	−12.49
$NP_{.HYPH}$	0.01	−0.16–0.18	0.16	0.03	0.15	0.03 –0.26	2.50*	0.27	1.01	0.72 –1.41	0.04
$NP_{.CONJ}$	0.25	0.01–0.50	2.06*	0.49	0.29	0.09–0.50	2.79**	0.55	1.15	0.84–1.58	0.87
$Cond_{.EXP}$	−0.01	−0.17–0.14	−0.15	−0.02	−0.03	−0.13–0.08	−0.51	−0.05	0.91	0.64–1.28	−0.56
Trial N	−0.00	−0.01–0.00	−0.70	0.00	−0.00	−0.01–−0.00	−4.53***	−0.01			
Length	0.04	0.03–0.06	6.87***	0.08	0.04	0.03–0.05	6.19***	0.07			
Frequency	−0.00	−0.00–0.00	−0.06	0.00	0.00	−0.00–0.00	0.43	0.00			
$NP_{.HYPH} \times Cond_{.EXP}$	0.22	0.01–0.43	2.03*	0.43	0.11	−0.02–0.23	1.69	0.20	1.78	1.13–2.80	2.47*
$NP_{.CONJ} \times Cond_{.EXP}$	−0.04	−0.25–0.17	−0.34	−0.07	−0.04	−0.17–0.08	−0.68	−0.08	1.03	0.66–1.62	0.14
$NP_{.HYPH} \times$ Trial N	0.00	−0.00–0.01	1.22	0.01							
$NP_{.CONJ} \times$ Trial N	−0.00	−0.01–0.00	−1.21	−0.01							
$Cond_{.EXP} \times$ Trial N	0.00	−0.01–0.01	0.16	0.00							
$(NP_{.HYPH} \times Cond_{.EXP}) \times$ Trial N	−0.01	−0.02–−0.00	−2.27*	−0.02							
$(NP_{.CONJ} \times Cond_{.EXP}) \times$ Trial N	0.00	−0.01–0.01	0.29	0.00							
Random effects components											
Residual variance (σ^2)		0.20				0.21				3.29	

(*Continued*)

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Table 3. (*Continued*)

	Regression-path duration (log)				Total fixation duration (log)				Regressions-in		
Predictors	*Est.*	*95% CI*	*Statistic (t)*	*Cohen's d*	*Est.*	*95% CI*	*Statistic (t)*	*Cohen's d*	*Odds Ratio*	*95% CI*	*Statistic (z)*
Random intercept variance (τ_{00})		0.01_{lexeme}				0.01_{lexeme}				0.01_{lexeme}	
		0.05_{pcp}				0.06_{pcp}				0.15_{pcp}	
Observations		2034				2034				3307	
Marg. R^2		0.31				0.31				0.01	
Cond. R^2		0.48				0.48				0.06	

Table 4. Model of best fit for gaze durations. Asterisks indicate *p*-values (* = <0.05, ** = <0.01, *** = <0.001)

Gaze duration (log)				
Predictors	*Est.*	*95% CI*	*Statistic (t)*	*Cohen's d*
(Intercept)	5.87	5.76–5.98	104.82	
NP $str_{.HYPH}$	0.15	0.05–0.24	3.05**	0.29
NP $str_{.CONJ}$	0.18	0.01–0.35	2.05*	0.35
$Condition_{.EXP}$	−0.03	−0.12–0.05	−0.72	−0.06
Attitudes	−0.04	−0.13–0.05	−0.83	−0.08
Length	0.03	0.02–0.04	6.21***	0.06
Trial number	−0.00	−0.00–0.00	−2.14*	0.00
Frequency	−0.00	−0.00–0.00	−0.26	0.00
NP $str_{.HYPH}$ × $Condition_{.EXP}$	0.04	−0.07–0.15	0.67	0.07
NP $str_{.CONJ}$ × $Condition_{.EXP}$	0.07	−0.04–0.18	1.28	0.14
NP $str_{.HYPH}$ × Attitudes	0.07	−0.02–0.15	1.54	0.13
NP $str_{.CONJ}$ × Attitudes	−0.02	−0.11–0.06	−0.51	−0.04
$Condition_{.EXP}$ × Attitudes	0.03	−0.06–0.11	0.66	0.06
(NP $str_{.HYPH}$ × $Condition_{.EXP}$) × Attitudes	−0.14	−0.26–−0.02	−2.29*	−0.28
(NP $str_{.CONJ}$ × $Condition_{.EXP}$) × Attitudes	0.03	−0.09–0.15	0.46	0.05
Random effects components				
Residual variance (σ^2)	0.20			
Random intercept variance (τ_{00})	0.01_{lexeme}			
	0.05_{pcp}			
Observations	2034			
Marginal R^2/conditional R^2	0.240/0.405			

Taken together, the data support $H_{1\text{-}ALT}$, since only contracted double forms led to increased processing costs compared to the control condition[7]. However, this hypothesis was confirmed inferentially only in the intermediate (regression-path duration and regressions-in), but visually in the late (total fixation duration) measures. This suggests that interference did not occur at early stages of processing (i.e., word identification), but at later stages (i.e., syntactic integration or general comprehension processes).

H2: Increased exposure facilitates the processing of gender-fair forms

Our second hypothesis was concerned with the habituation effect discovered in previous studies (Gygax & Gesto, 2007; Zami & Hemforth, 2024). If increased exposure of contracted double forms facilitates the processing of the contracted double forms

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

compared to the masculine form, there should be a significant interaction between the variables Condition, NP structure, and Trial number. More precisely, in the experimental condition, we expect a larger effect of Trial number with a hyphenated NP (i.e., a contracted double form) than with a simple NP (i.e., a masculine form).

Despite being present in descriptive plots of all measures except for first fixation duration, the hypothesized effect was confirmed inferentially only for regression-path durations. The model of best fit, presented in Table 3, contained a significant interaction effect between Trial number, Condition: experimental, and NP structure: hyphenated (line 13). The predicted values are visualized in Figure 3, which clearly shows that regression-path durations of NPs like *voisin·es* [neighbor.$_{\text{MASC·FEM.PL}}$] resemble those of *voisines et voisins* [neighbors.$_{\text{FEM}}$ and neighbors.$_{\text{MASC}}$] at the start of the experiment but are almost identical to those of NPs like *voisins* [neighbors.$_{\text{MASC}}$] by the end. This visualization also shows, unexpectedly, that in the control condition, the predicted regression-path durations of hyphenated NPs increase throughout the experiment. Given this effect, and the fact that the three-way interaction was significant in only one of all analyzed measures, H_2 is only partially confirmed.

H3: Positive attitudes towards gender-fair language mediate processing costs

Finally, we hypothesized that positive attitudes towards gender-fair language would facilitate the processing of gender-fair forms. If that were the case, we should observe a significant interaction between the variables Condition, NP structure, and Attitudes towards gender-fair language. More precisely, there should be a significant negative effect of positive attitudes with experimental, hyphenated NPs.

Plots of the data show that this was the case in gaze durations, regression-path durations, and total fixation durations. However, the three-way interaction of interest was significant only in the analyses of gaze durations, as reported in Table 4. Indeed, the interaction effect between a hyphenated NP structure, the experimental condition, and attitudes towards gender-fair language was significant, such that gaze durations for experimental, hyphenated NPs decreased as the attitudes became increasingly positive (line 14). In short, our data support a weak version of H_3, given that the hypothesized effect was significant only in one measure. However, these results should be interpreted with caution, since, as Figure 4 shows, positive attitudes towards gender-fair language unexpectedly influenced gaze durations in the control condition. This unexpected finding is further problematized in the discussion.

General discussion

In brief, our descriptive data showed increased processing costs for the contracted double forms in all measures compared to the control condition. In the inferential analyses, this effect was significant in the intermediate eye-movement measures (regression-path durations and regressions-in), but visually quite important also for the late measure (total fixation durations). We also replicated the previously reported habituation effect (Gygax & Gesto, 2007; Zami & Hemforth, 2024) on the processing of gender-fair forms, but only in one of the analyzed measures. Finally,

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

the predicted effect of attitudes towards gender-fair language was found in one measure. In the remainder of this section, we discuss these results in relation to our initial hypotheses and to the current understanding of the reading process as outlined in section 2.2.

The influence of gender-fair forms on the reading process

First of all, contrary to our initial hypothesis, the difference between the experimental and control conditions within conjoined NP structures was not greater than that within simple NP structures in any of the analyzed measures. This indicates that when other relevant factors are controlled for (length, structure, etc.), complete double forms do not lead to increased processing costs in reading. Considering that complete double forms make use of nouns that are part of standard French, it is not surprising that they require similar processing costs as other conjoined NPs of similar length. In contrast, the contracted double forms led to significantly longer regression-path durations and were significantly more likely to attract regressions than the compound nouns used as controls. In brief, $H_{1\text{-ALT}}$ was confirmed inferentially in the intermediate measures.

As stated in Table 1, early measures are assumed to reflect word identification and lexical access processes, while late measures typically are taken as indications of word integration and syntactic processing (Conklin et al., 2018). Regressions are somewhat more difficult to classify, since they can reflect both early and later processes (Clifton et al., 2007), which is why we categories regression-path duration and regressions-in as intermediate measures, in line with Conklin et al. (2018). Against this background, it does not seem to be the case that the contracted double forms hinder the processes of word identification and lexical access. Although, as previously discussed, French and German forms are not fully equivalent to each other, it is interesting to note that our findings are in line with the results of Zacharski et al. (2025), previously discussed. In their younger, student sample, similar to ours, gender star forms were accepted to the same extent and as quickly as feminine and masculine forms, leading the researchers to conclude that this novel form does not hinder lexical access for this group.

Furthermore, the timeline of lexical access described in section *Language processing in reading* (Sereno & Rayner, 2003) provides an additional interesting ground for discussion of our results on the contracted double forms. More specifically, if the programming of eye movements begins before the eyes are prompted to move forward, it could be that, at least for some participants, the threshold for lexical access was reached before the interpunct in the contracted double forms was registered. Given that first fixations often land between the beginning and the middle of the word (White & Liversedge, 2006) and that the interpunct in most experimental items appeared in the second half of the NP (only in two cases out of 18 did it appear exactly in the middle of the NP), the interpunct might not have been registered during the first fixation on the NP. Consequently, one could imagine a scenario where the oculomotor systems programmed the eyes to move forward (i.e., to the subsequent verb phrase) immediately after a successful lexical access of the NP. Thus, the unfamiliarity that the interpunct constitutes might have been registered only after this saccadic programming had been executed.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Such a scenario could explain the descriptively longer total fixation durations and significantly increased probability of regressions for the contracted double forms. However, it would not explain the increased regression-path durations, since these are caused by regressions to earlier parts of the text. Bearing in mind that regression-path durations and total fixation durations are considered to reflect different processes (see Table 1), it is not unlikely that there could be two distinct explanations for these results. As such, the scenario outlined above could explain the increased total fixation durations and regressions to the NP. In contrast, it could be the case that in trials with increased regression-path durations, the interpunct was registered during the initial reading of the critical NP. This appears to have caused a disruption of the reading process, which forced the reader to return to earlier parts of the text before feeling ready to continue forward. It could also be that the contracted double form was initially read as a singular form, given that the plural marking appears only after the feminine suffix. This incongruency between the plural article *les* [the.$_{PL}$] and the masculine stem might have caused regressions to the article, leading to increased regression-path durations, at least when compared to the hyphenated compound noun. In sum, our results indicate that contracted double forms do not affect early stages of processing, such as word identification and lexical access, but rather intermediate and late stages of the reading process.

Our finding that the contracted double forms attracted more regressions than the other NPs is somewhat similar to previous findings reporting that first fixations on words with orthographically irregular word beginnings landed closer to the beginning of the word compared to words with regular word beginnings (White & Liversedge, 2006). Thus, the authors suggested that "it is possible that very infrequent individual letters may be so visually unfamiliar (rather than just linguistically infrequent) that they draw saccades directly towards them." (White & Liversedge, 2006, p. 769). Extending this statement to include typographical signs in addition to individual letters, and to hold for entire words and not only their beginnings, could explain our results. Future studies might want to compare contracted double forms using different typographical signs (such as a hyphen) to shed further light on this question.

When comparing our results to those of previous studies on the processing of gender-fair forms, we notice a few differences. Out of the four previous studies on the processing of French gender-fair forms, two reported increased processing costs of contracted double forms (Gygax & Gesto, 2007; Zami & Hemforth, 2024). However, it is difficult to compare the results of these two studies to ours, since they did not study the processing of these forms in different stages of reading and compared the different gender-fair forms directly to the masculine one (accounting for word length in their statistical models). Therefore, we focus our discussion on the one study that analyzed the processing of French contracted double forms throughout the reading process (Girard et al., 2022), but also compare our results to a similar study on Spanish neopronouns (Román Irizarry & Guzzardo Tamargo, 2025).

Girard and colleagues (2022) did not find any differences in total fixation durations or regressions-in between masculine and contracted double forms. This stands in stark contrast to our results, since these two measures, in addition to regression-path durations, were the ones that showed the most important

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

differences in our data. A few notable differences between the studies could explain this. First, the study by Girard et al. (2022) did not include a control condition, but the contracted double forms were instead directly compared to the masculine form (and controlled for length). Second, in the texts presented to the participants, the position of the critical word varied as it was embedded in the NP constituting the direct or indirect object. In our study, the critical word was always the subject and followed by a verb phrase. Given the important role the subject of a sentence occupies, especially in a language with subject-verb agreement like French, participants possibly paid more attention to contracted double forms in subject position than when they were embedded in the direct or indirect object. Third, the texts used by Girard et al. (2022) consisted of one sentence, as opposed to three in the present study. It is possible that readers are more likely to make regressions in a longer text, which in turn would lead to longer total fixation durations. Finally, their sample consisted of Quebec French speakers, whereas ours mainly included French speakers from France. Despite contracted double forms not being officially recommended in Quebec, it is possible that the participants' use and knowledge of gender-fair language, as well as their attitudes towards it, differed between the two samples.

In the previously introduced study on Spanish non-binary pronouns (Román Irizarry & Guzzardo Tamargo, 2025), increased processing costs of the non-binary form were found in all measures when these forms were compared to the masculine one. These results are more similar to ours than those of Girard and colleagues (2022), yet still differ, given that the Spanish gender-fair forms were found to increase processing costs across the entire reading process. However, in the materials used in the Spanish study, the target NPs appeared as anaphoric expressions of a common gender noun antecedent (i.e., *élèves . . . celleux* in French), a structure quite different from the one used in our materials.

Against this background, it is hard to fairly compare the results of previous studies to ours. Nevertheless, there seems to be a general trend indicating that gender-fair nouns, unlike pronouns, may not hinder lexical access but could interfere with the reading process at later stages. In sum, these previous studies, and ours, provide important insights and call for further investigations on the processing of contracted double forms in reading. Future studies might want to contrast French speakers from different geographic areas (as done in Kim et al., 2023), as well as contracted double forms appearing in different syntactic roles.

The influence of exposure on the processing of gender-fair forms

In our data, H_2 was confirmed only in the analyses of regression-path durations. These analyses showed that the predicted regression-path durations of contracted double forms were similar to those of the complete double forms at the beginning of the experiment but resembled those of the masculine form by the end. This finding is in line with previous research showing that the frequency effect, i.e., higher processing costs for infrequent words, is mitigated as words are repeated in a short passage (Clifton et al., 2007). It is also in keeping with the results of Zacharski and colleagues (2025, Figure 3) and with two of the previous studies on French gender-fair forms (Gygax & Gesto, 2007; Zami & Hemforth, 2024).

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

The fact that the predicted effect was significant only in the analyses of regression-path durations might be a consequence of our research design. Given that this effect was present in descriptive plots of gaze durations and total fixation durations, the lack of a significant effect in these two measures might be caused by a lack of power, since three-way interactions require a larger sample size than two-way interactions to observe a significant effect (Heo & Leon, 2010). It is also easier to obtain a significant difference between slopes in opposing directions (a reversed interaction) than between a null slope and a positive or negative one (a fully attenuated interaction) (Sommet et al., 2023). In other words, if the slope for hyphenated NP structures in the control condition (see the left panel in Figure 3) had been negative or null, the interaction effect would most likely not have been significant. To conclude, there seems to be a habituation effect such that the processing of contracted double forms becomes easier over time, but with the current research design it is difficult to pinpoint the strength of this effect.

The influence of attitudes towards gender-fair language on the processing of gender-fair forms

The results concerning H_3, according to which positive attitudes towards gender-fair language facilitate the processing of gender-fair forms, largely follow the same pattern as the results discussed in the previous section. That is, we observed the hypothesized effect in descriptive plots for gaze durations, regression-path durations, and total fixation durations, but it was only confirmed inferentially for gaze durations. Similarly, the slope representing the influence of attitudes on gaze durations for hyphenated NPs in the control condition was positive (see the left panel in Figure 4), as opposed to a negative slope in the experimental condition. Thus, the statistical explanations provided in the previous section might account for the lack of a significant three-way interaction in regression-path durations and total fixation durations. Nevertheless, the predicted effect was observed, in line with some previous studies (Román Irizarry & Guzzardo Tamargo, 2025; Steiger-Loerbroks & von Stockhausen, 2014), and it is therefore possible that H_3 would have been confirmed with another research design. In the future, it would be of great value to have a wider range of attitudes towards gender-fair language in the sample (as we might have a ceiling effect in our sample), or to measure the attitudes more implicitly, for example, with an implicit association test (Greenwald et al., 1998).

Finally, we unexpectedly found positive attitudes towards gender-fair language to slightly increase gaze durations of the hyphenated control NPs. This indicates that this variable might reflect something else in addition to the participants' attitudes towards gender-fair language, or could question the relevance of our hyphenated control condition. Also, the participants had rather positive attitudes towards gender-fair language on average ($M = 3.8$, $sd = 0.8$, range = 2.1–5), which may not reflect a more general variance present in the population.

Limitations of the present study

Before concluding, we will discuss a few limitations of the present study. First, the effect we found on total fixation durations was non-significant, which indicates that

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

it is less robust than the ones we found on regressions-in and on regression-path durations. In this light, a more high-powered study would be welcome to either strengthen or disprove our findings. This discrepancy in the results also shows, as has been indicated by previous researchers, the difficulty of classifying regressions as a reflection of early or late stages of processing.

Furthermore, our sample mainly consisted of students or non-students with a high level of education. In addition, the sample had generally positive attitudes towards gender-fair language, and no participant had any language disorders. It is therefore important to bear in mind that our results are not representative of all French speakers. Given the relatively increased presence of contracted double forms (Díaz & Heap, 2020; Flesch & de Beaumont, 2023; Simon & Vanhal, 2022), a replication of the present study with a more heterogeneous sample would be highly beneficial both for the scientific community and the general public, as well as determining the factors that may help better processing of these contracted forms.

Another limitation relates to our materials, which consisted of highly controlled, short texts in which the critical word only appeared once. Future studies may want to use more ecological texts in which the gender-fair forms appear more than once (as in Pozniak et al., 2024) and not only as nouns, to see how the appearance of several gender-fair forms in a text influences reading patterns. They might also want to include explicit comprehension questions to see how these forms affect reading comprehension.

Finally, while eye-tracking technology provides detailed data on reading patterns, it does not tell us why participants return to a certain word. It could be that they had trouble understanding what it represents, that they were positively surprised to see it, or something else. Therefore, we cannot truly account for how contracted double forms influence later stages of the reading process. Extending the present study with analyses of Event-Related Potentials (ERP)s instead of eye movements could yield further insights into the effects of gender-fair forms on language processing. On the basis of our results, we hypothesize that contracted double forms affect later components such as the N400 or the P600 rather than early components.

Concluding remarks

To conclude, the study reported in this paper has contributed to the scant body of research on the processing of French gender-fair forms in reading by analyzing L1 French speakers' eye movements in a reading task. First of all, we found that complete double forms did not lead to increased processing costs in any of the analyzed measures. Reading a complete double form, such as *voisines et voisins* [neigbors.$_{FEM}$ and neighbors.$_{MASC}$] is not harder than reading pairs such as *artistes et peintres* [artists and painters]. However, contracted double forms (e.g., *voisin·es* [neighbor.$_{MASC·FEM.PL}$]) led to increased processing costs in the measures reflecting intermediate and late stages of processing, at least when compared to common compound nouns of similar length (for example, *chauve-souris* [bat]). Given that no effect was found on early measures, our results suggest that contracted double forms do not hinder lexical access but might constitute an obstacle in syntactic or discourse-level processing, or in general comprehension processes. Since we cannot know the exact

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

reason behind the increased durations and regressions, nor how they influence reading comprehension, we call for future studies to investigate the matter more closely.

While our results show that contracted double forms increase processing costs, at least to some extent, it is important to underline that this would be the case for other types of words as well. Some less common feminine forms, such as those for certain stereotypically male professions (e.g., *pompière* [firefighter.$_{FEM}$] or *charpentière* [carpenter.$_{FEM}$], see van Compernolle (2008)), would possibly require initial higher processing costs without that being a reason for avoiding them. Given the robust, positive effect that contracted double forms have on representation of women (e.g., Spinelli et al., 2023; Tibblin, Granfeldt, et al., 2023; Tibblin, van de Weijer, et al., 2023; Xiao et al., 2022), and that the processing of these forms might become easier as they grow more frequent, one could argue that the benefits outweigh the risks. Finally, it should not be forgotten that for some speakers of French, the contracted double form is the only form that truly corresponds to their gender identity.

Acknowledgments. The authors gratefully acknowledge Lund University Humanities Lab for use of the Digital Classroom. The authors also thank Diedrick Niehorster and Marcus Nyström for their thoughtful advice throughout the process, as well as Roger Johansson for his early input on the research design. Finally, they thank Fil dr Uno Otterstedt's foundation for financing payment to the participants.

Author contributions according to the CRediT taxonomy. JT: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing—original draft preparation, Writing—review & editing.

PG: Conceptualization, Methodology, Supervision, Writing—review & editing.

JW: Conceptualization, Formal analysis, Methodology, Writing—review & editing.

JG: Conceptualization, Methodology, Project administration, Supervision, Writing—review & editing.

Competing interests. The authors declare none.

Replication package. All materials, the frequency values of the NPs, and all code used to run the experiment can be found on the OSF repository of the project: https://osf.io/6hxm3/.

Notes

1. The pre-registration can be consulted at the following link: https://osf.io/28ck3/.

2. We are aware of the large strand of research that has investigated the comprehensibility of German gender-fair forms using subjective measures (e.g., Friedrich et al., 2021, 2024; Friedrich & Heise, 2019; Pabst & Kollmayer, 2023) and one eye-tracking study investigating the processing of nominalized (i.e., gender-neutral) forms (e.g., Steiger-Loerbroks & von Stockhausen, 2014), but choose to focus on the processing of novel forms here.

3. In this section, we present the hypotheses in a more general, less methodology-dependent manner than in the pre-registration. We therefore refer readers interested in the hypotheses regarding the statistical analyses to the pre-registration.

4. Three participants had lived in Sweden for 60 months (5 years). Since they were all L1 French speakers and indicated French as their most dominant language, we chose to retain them in the final sample.

5. For instance, out of the over 500 nouns included in a psycholinguistic norming study (e.g., Bonin et al., 2021), only a handful would have met our criterion, such as *porte-parole* [spokesperson] or *sans-emploi* [unemployed], without even taking NP length into consideration.

6. https://github.com/marcus-nystrom/Titta/tree/master/demo_analyses.

7. As rightfully pointed out by one of the reviewers, a comparison between the hyphenated and conjoined NP structures using reverse Helmert coding was in order to confidently confirm $H_{1\text{-}ALT}$ rather than H_1. Since such an analysis did not change the interpretation of our data, we refer interested readers to the OSF repository of the project for a full report of these analyses.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

References

Abbondanza, M., Galimberti, V., Bonomi, V., Reverberi, C., Durante, F., & Foppolo, F. (2025). Neutralizing gender in role nouns: investigating the effect of ə in written and oral Italian. *Frontiers in Communication*, **9**, 1530778. https://doi.org/10.3389/fcomm.2024.1530778

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*, **43**(6), 1257–1274. https://doi.org/10.1017/S0142716422000352

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* (Version 1.1-27.1) [Computer software]. https://github.com/lme4/lme4/

Bolter, F. (2022). « Le masculin l'emporte »: Stratégies linguistiques et politiques de genre dans les associations LGBT+ en France. In V. Swamy, & L. Mackenzie (Eds.), *Devenir non-binaire en français contemporain* (pp. 21–43). Éditions le Manuscrit.

Bonin, P., Laroche, B., & Méot, A. (2021). Psycholinguistic norms for a set of 506 French compound words. *Behavior Research Methods*, **54**(1), 393–413. https://doi.org/10.3758/s13428-021-01601-w

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: an evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, **2**(1), 1–12. https://doi.org/10.16910/jemr.2.1.1

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.

Braun, F., Sczesny, S., & Stahlberg, D. (2005). Cognitive effects of Masculine generics in German: an overview of empirical findings. *Communications*, **30**(1), 1–21. 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. 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. https://doi.org/10.1111/josl.12489

Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, **155**, 49–62. https://doi.org/10.1016/j.ijpsycho.2020.05.010

Chetail, F., Porteous, F., Patout, P. A., Content, A., & Collette, E. (2019). BOQS : une échelle de qualité orthographique pour l'adulte. *Revue de neuropsychologie*, **11**(3), 214–226. https://doi.org/10.1684/nrp.2019.0518

Clifton, C., Staub, A., & Rayner, K. (2007). Eye Movements in Reading Words and Sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: a window on mind and brain* (pp. 341–371). Elsevier. https://doi.org/10.1016/B978-008044980-7/50017-3

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: a guide for applied linguistics research* (1 ed.). Cambridge University Press. https://doi.org/10.1017/9781108233279

Corbett, G. G. (1991). *Gender*. Cambridge University Press.

Cortese, M. J., & Balota, D. A. (2012). Visual Word Recognition in Skilled Adult Readers. In M. Spivey, M. F. Joanisse, & K. McRae (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 159–185). Cambridge University Press.

Díaz, Y., & Heap, D. (2020). Variation dans les accords du français inclusif. In A. Hernández, & E. Butterworth (Eds.), *Actes du congrès annuel de l'Association canadienne de linguistique 2020* (p. 15) Association canadienne de linguistique. https://cla-acl.ca/actes/actes-2020-proceedings.html

Direction de la langue française. (2024). Quand dire, c'est inclure. Pour une communication officielle et formelle non discriminatoire quant au genre. Ministère de la Fédération Wallonie-Bruxelles. https://www.languefrancaise.cfwb.be/index.php?eID = tx_nawsecuredl&u = 0&g = 0&hash = 9a12f122a383c2a2953af0a1ade5101351b1754e&file = fileadmin/sites/lff/upload/lff_super_editor/lff_editor/documents/2024/Quand_dire-brochure-20240422-vf.pdf

Dister, A. (2021, March 5). L'enjeu démocratique de la lisibilité des textes. La Revue Parlamentaire. https://www.larevueparlementaire.fr/articles-revue-parlementaire/3679-anne-dister-l-enjeu-democratique-de-la-lisibilite-des-textes

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Dister, A., & Moreau, M.-L. (2020). *Inclure sans exclure: Les bonnes pratiques de rédaction inclusive*. Ministère de la Fédération Wallonie-Bruxelles, Direction de la Langue française.

Dumais, G. (2021). "What do they say in Quebec?": non-binary gender expression in informal spoken Quebec French. *Toronto Working Papers in Linguistics*, **43**(1), 1–11. https://doi.org/10.33137/twpl.v43i1.35934

Flesch, M., & de Beaumont, É. (2023). Usages informels du français inclusif: Étude des doublets abrégés et complets sur twitter, reddit et youtube. *Langue Française, N°*, **220**(4), 59–78. https://doi.org/10.3917/lf.220.0059

Friedrich, M. C. G., Drößler, V., Oberlehberg, N., & Heise, E. (2021). The influence of the gender asterisk ("Gendersternchen") on comprehensibility and interest. *Frontiers in Psychology*, **12**, 760062. https://doi.org/10.3389/fpsyg.2021.760062

Friedrich, M. C. G., Gajewski, S., Hagenberg, K., Wenz, C., & Heise, E. (2024). Does the gender asterisk ("Gendersternchen") as a special form of gender-fair language impair comprehensibility?. *Discourse Processes*, **61**(9), 439–461. https://doi.org/10.1080/0163853X.2024.2362027

Friedrich, M. C. G., & Heise, E. (2019). Does the use of gender-fair language influence the comprehensibility of texts?: An experiment using an authentic contract manipulating single role nouns and pronouns. *Swiss Journal of Psychology*, **78**(1–2), 51–60. https://doi.org/10.1024/1421-0185/a000223

Girard, G., Foucambert, D., & Le Mené, M. (2022). Lisibilité de l'écriture inclusive: Apports des techniques d'oculométrie. In P. Tarahomi (Ed.), *Actes du Congrès de l'ACL 2022* (p. 15) Association canadienne de linguistique. https://cla-acl.ca/actes/actes-2022-proceedings.html

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. https://doi.org/10.1037/0022-3514.74.6.1464

Grinshpun, Y., Neveu, F., Rastier, F., & Szlamowicz, J. (2020, September 18). Une 'écriture excluante' qui 's'impose par la propagande': 32 linguistes listent les défauts de l'écriture inclusive. Marianne. https://www.marianne.net/agora/tribunes-libres/une-ecriture-excluante-qui-s-impose-par-la-propagande-32-linguistes-listent-les

Gygax, P., Gabriel, U., & Zufferey, S. (2019). Le masculin et ses multiples sens: un problème pour notre cerveau . . . Et notre société. *Savoirs en Prisme*, **10**(e-publication), 24.

Gygax, P., & Gesto, N. (2007). Féminisation et lourdeur de texte. *L'Année psychologique*, **107**(02), 239. https://doi.org/10.4074/S0003503307002059

Gygax, P., Sarrasin, O., Lévy, A., Sato, S., & Gabriel, U. (2013). La représentation mentale du genre pendant la lecture: État actuel de la recherche francophone en psycholinguistique. *Journal of French Language Studies*, **23**(2), 243–257. https://doi.org/10.1017/S0959269512000439

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. https://doi.org/10.1016/j.langsci.2020.101328

Haut Conseil à l'Égalité entre les femmes et les hommes. (2016). *Pour une communication publique sans stéréotype de sexe. Guide pratique (1ère)*. La documentation francaise.

Haut Conseil à l'Égalité entre les femmes et les hommes. (2022). *Pour une communication publique sans stéréotype de sexe. Guide pratique (2ème édition)*. La documentation francaise. https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/guide_egacom_sans_stereotypes-2022-versionpublique-min.pdf

Heo, M., & Leon, A. C. (2010). Sample sizes required to detect two-way and three-way interactions involving slope differences in mixed-effects linear models. *Journal of Biopharmaceutical Statistics*, **20**(4), 787–802. https://doi.org/10.1080/10543401003618819

Hessels, R. S., Niehorster, D. C., Kemner, C., & Hooge, I. T. C. (2017). Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC). *Behavior Research Methods*, **49**(5), 1802–1823. https://doi.org/10.3758/s13428-016-0822-1

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: a comprehensive guide to methods and measures*. Oxford university press.

Kamblé-Bagal, N., & Tatossian, A. (2022). Étude comparative sur l'usage de l'écriture inclusive dans deux médias écrits français et québécois. *SHS Web of Conferences*, **138**, 12003. https://doi.org/10.1051/shsconf/202213812003

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Kim, J., Angst, S., Gygax, P., Gabriel, U., & Zufferey, S. (2023). 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*, **33**(1), 1–26. https://doi.org/10.1017/S095926952200014X

Kim, Y. S. G. (2020). Toward integrative reading science: the direct and indirect effects model of reading. *Journal of Learning Disabilities*, **53**(6), 469–491. https://doi.org/10.1177/0022219420908239

Klinkenberg, J. M. (2019). Quelle écriture pour quelle justice? « Écriture inclusive » et politique linguistique. *Le discours et la langue. Revue de linguistique française et d'analyse du discours*, **11**(1), 15–26.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, **65**(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Liénardy, C., Tibblin, J., Gygax, P., & Simon, A. C. (2023). Écriture inclusive, lisibilité textuelle et représentations mentales. *Discours*, **33**, 1–27. https://doi.org/10.4000/discours.12636

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

Lüdecke, D. (2024). *sjPlot* (Version 2.8.16) [Computer software]. https://CRAN.R-project.org/package=sjPlot

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, **50**(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

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. https://doi.org/10.1111/j.2041-210x.2012.00261.x

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behavior Research Methods, Instruments, & Computers*, **36**(3), 516–524. https://doi.org/10.3758/BF03195598

Niehorster, D. C., Andersson, R., & Nyström, M. (2020). Titta: a toolbox for creating PsychToolbox and Psychopy experiments with Tobii eye trackers. *Behavior Research Methods*, **52**(5), 1970–1979. https://doi.org/10.3758/s13428-020-01358-8

Office québécois de la langue française. (n.d.). *Rédaction épicène*. Retrieved 19 September 2024, from https://vitrinelinguistique.oqlf.gouv.qc.ca/banque-de-depannage-linguistique/la-redaction-et-la-communication/feminisation-et-redaction-epicene/redaction-epicene

Pabst, L. M., & Kollmayer, M. (2023). How to make a difference: the impact of gender-fair language on text comprehensibility amongst adults with and without an academic background. *Frontiers in Psychology*, **14**, 1234860. https://doi.org/10.3389/fpsyg.2023.1234860

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.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, **51**(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pozniak, C., Corbeau, E., & Burnett, H. (2024). Contextual dilution in French gender inclusive writing: an experimental investigation. *Journal of French Language Studies*, **34**(2), 273–292. https://doi.org/10.1017/S0959269523000236

Prentice, D. A. (1994). Do language reforms change our way of thinking. *Journal of Language and Social Psychology*, **13**(1), 3–19.

Proposition de loi visant à protéger la langue française des dérives de l'écriture dite inclusive, 16 Session ordinaire de 2023-2024 (2023). https://www.senat.fr/leg/tas23-016.pdf

R Core Team. (2021). *R: A Language and Environment for Statistical Computing* (Version 4.1.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**(3), 372–422.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, **62**(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *WIREs Cognitive Science*, **1**(6), 787–799. https://doi.org/10.1002/wcs.68

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Reichle, E. D. (2021). *Computational models of reading: formal descriptions of the mind in action*. Oxford University Press. https://academic.oup.com/book/39768

Richy, C., & Burnett, H. (2021). Démêler les effets des stéréotypes et le genre grammatical dans le biais masculin: Une approche expérimentale. *GLAD!*, **10**, 1–31. https://doi.org/10.4000/glad.2839

Irizarry, A. R., & Tamargo, R. E. G. (2025). Processing of pronouns with gender-inclusive –x in Spanish: an eye-tracking study. *Social Psychological Bulletin*, **20**, e13611. https://doi.org/10.32872/spb.13611

Sauteur, T., Gygax, P., Tibblin, J., Escasain, L., & Sato, S. (2023). « L'écriture inclusive, je ne connais pas très bien . . . mais je déteste ! »: Liens entre connaissances linguistique et historique, orientation politique et attitudes envers l'écriture inclusive. *GLAD!*, **14**, 1–22. https://doi.org/10.4000/glad.6400

Schotter, E. R., & Rayner, K. (2015). The Work of the Eyes During Reading. In A. Pollatsek, & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 44–60). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199324576.001.0001

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. https://doi.org/10.1177/0146167215585727

Sereno, S., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences*, **7**(11), 489–493. https://doi.org/10.1016/j.tics.2003.09.010

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. https://doi.org/10.3917/lf.215.0081

Slemp, K. (2021). Attitudes towards varied inclusive language use in Spanish on Twitter. *Working Papers in Applied Linguistics and Linguistics at York*, **1**, 60–74. https://doi.org/10.25071/2564-2855.6

Sommet, N., Weissman, D. L., Cheutin, N., & Elliot, A. J. (2023). How Many participants do i need to test an interaction? Conducting an appropriate power analysis and achieving sufficient power to detect an interaction. *Advances in Methods and Practices in Psychological Science*, **6**(3), 1–21. https://doi.org/10.1177/25152459231178728

Spinelli, E., Chevrot, J.-P., & Varnet, L. (2023). Neutral is not fair enough: testing the efficiency of different language gender-fair strategies. *Frontiers in Psychology*, **14**, 1256779. https://doi.org/10.3389/fpsyg.2023.1256779

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: some current controversies. *Trends in Cognitive Sciences*, **5**(4), 156–163. https://doi.org/10.1016/S1364-6613(00)01619-3

Staub, A. (2015). Reading Sentences: Syntactic Parsing and Semantic Interpretation. In A. Pollatsek, & R. Treiman (Eds.), *The Oxford Handbook of Reading* (pp. 202–216). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199324576.013.15

Steiger-Loerbroks, V., & von Stockhausen, L. (2014). Mental representations of gender-fair nouns in German legal language: an eye-movement and questionnaire-based study. *Linguistische Berichte*, **237**, 57–80.

Storme, B., & Delaloye Saillen, L. (2024). Effects of grammatical gender on gender inferences: evidence from French hybrid nouns. *Linguistics Vanguard*. **10**(1), 297–308. https://doi.org/10.1515/lingvan-2022-0064

Storme, B., & Storme, M. (2025). Feminization is more gender-fair than neutralization: evidence from gender-stereotyped contexts. *Journal of Language and Social Psychology*. 45(1), 147–172. https://doi.org/10.1177/0261927x251349151

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. https://doi.org/10.1051/shsconf/20207813006

Tibblin, J., Granfeldt, J., van de Weijer, J., & Gygax, P. (2023). The male bias can be attenuated in reading: on the resolution of anaphoric expressions following gender-fair forms in French. *Glossa Psycholinguistics*, **2**(1), 1–33. https://doi.org/10.5070/G60111267

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. https://doi.org/10.1017/S0959269522000217

Tobii AB. (2018). *Tobii Pro Spectrum Data Quality Test Report. Accuracy, precision, and detected gaze under optimal conditions – controlled environment*. Tobii AB.

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press

Traxler, M. J. (2012). *Introduction to psycholinguistics: understanding language science* (1 ed). Wiley-Blackwell.

Tremblay, A., & Ransijn, J. (2020). *Package 'LMERConvenienceFunctions': Model Selection and Post-Hoc Analysis for (G)LMER Models* (Version 3.0) [Computer software]. https://cran.r-project.org/web/packages/LMERConvenienceFunctions/LMERConvenienceFunctions.pdf

van Compernolle, R. A. (2008). « Une pompière? C'est affreux! » Étude lexicale de la féminisation des noms de métiers et grades en France. *Langage et société*, **123**, 107–126.

Veldre, A., Yu, L., Andrews, S., & Reichle, E. D. (2020). Towards a complete model of reading: simulating lexical decision, word naming, and sentence reading with über-reader. *Proceedings of the Annual Meeting of the Cognitive Science Society*, **42**, 151–157. https://escholarship.org/uc/item/92p992b5

Vervecken, D., Gygax, P. M., Gabriel, U., Guillod, M., & Hannover, B. (2015). Warm-hearted businessmen, competitive housewives? Effects of gender-fair language on adolescents' perceptions of occupations. *Frontiers in Psychology*, **6**, 1–10. https://doi.org/10.3389/fpsyg.2015.01437

Waldendorf, A. (2024). Words of change: the increase of gender-inclusive language in German media. *European Sociological Review*, **40**(2), 357–374. https://doi.org/10.1093/esr/jcad044

White, S. J., & Liversedge, S. P. (2006). Linguistic and nonlinguistic influences on the eyes' landing positions during reading. *Quarterly Journal of Experimental Psychology*, **59**(4), 760–782. https://doi.org/10.1080/02724980543000024

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*, **42**(1), 82–106. https://doi.org/10.1177/0261927X221084643 (Original work published 2023).

Zacharski, L., Kruppa, A., & Ferstl, E. C. (2025). The readability of the non-binary gender star in German: evidence from a lexical decision task. *Social Psychological Bulletin*, **20**, e13719. https://doi.org/10.32872/spb.13719

Zami, J., & Hemforth, B. (2024). Intelligibilité de l'écriture inclusive: Une approche expérimentale. *SHS Web of Conferences*, **191**, 10004. https://doi.org/10.1051/shsconf/202419110004

Cite this article: Tibblin, J., Gygax, P., van de Weijer, J., & Granfeldt, J. (2026). Recognized but harder to integrate: An eye-tracking study of French gender-fair forms during reading. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100484

https://doi.org/10.1017/S0142716426100484 Published online by Cambridge University Press