1
2
3
4
5
6
7
8
9
10
11
Gauging the impact of gender grammaticization in different languages: Application of a linguistic-visual paradigm
12
13
14
15
16
17
18
19   Sayaka Sato[1,2*], Pascal M. Gygax[1], Ute Gabriel[3]
20   [1] Department of Psychology, University of Fribourg, Fribourg, Switzerland
21   [2]Department of Linguistics and English Language, Lancaster University, Lancaster, U.K.
22   [3] Department of Psychology, Norwegian University of Science and Technology, Trondheim,
23   Norway
24
25
26
27
28
29
30
31
32
33
34
35
36
37   Correspondence:
38   Sayaka Sato
39   University of Fribourg
40   Department of Psychology
41   Rue de Faucigny 2
42   CH-1700 Fribourg, Switzerland
43   sayaka.sato@unifr.ch
44
45
46   Number of words: 10584 words
47   Number of figures: 2
48

Abstract

Employing a linguistic-visual paradigm, we investigated whether the grammaticization of gender information impacts readers' gender representations. French and German were taken as comparative languages, taking into account the male gender bias associated to both languages, as well as the comparative gender biases associated to their plural determiners (French: *les* [generic] vs. German: *die* [morphologically feminine]). Bilingual speakers of French and German had to judge whether a pair of facial images representing *two men* or *a man and a woman* could represent a gender stereotypical role noun prime (e.g., *nurses*). The prime was presented in the masculine plural form with or without a plural determiner. Results indicated that the overt grammaticization of the male gender in the masculine form dominated the representation of the role nouns (though interpretable as *generic*). However, the effect of the determiner was not found, indicating that only gender information associated to a human reference role noun had impacted readers' representations. The results, discussed in the framework of the *thinking-for-speaking* hypothesis, demonstrated that linguistic-visual paradigms are well-suited to gauge the impact of both stereotype information and grammaticization when processing role nouns.

Keywords: gender representation, gender stereotypes, grammatical gender, generic masculine, thinking-for-speaking hypothesis, bilingualism

The ways in which languages organize specific concepts in their linguistic systems have been found to impact how we represent information (e.g., Gennari et al., 2002; Papafragou et al., 2002). This notion, further developed as the *thinking-for-speaking* hypothesis by Slobin (1996) in his work on motion events, proposes that the encoding of concepts and events within a language acts both as a foundational and constraining structure for how verbal information is represented. Processing a specific language therefore imposes speakers to focus on particular concepts that are grammaticized within its structure, resulting in language-bound representations. As will be further discussed in this paper, bilinguals are particularly suited for testing the thinking-for-speaking hypothesis as they offer a platform to examine the extent to which comprehension mechanisms change as a function of the characteristics of the language being used (e.g., Boroditsky et al., 2003; Bylund and Jarvis, 2011; Fausey et al., 2010). In the present study, we focus specifically on the case of gender representation during language comprehension, and argue that processing languages that grammaticize gender information in their linguistic structure will result in heightened biased representations of gender.

Recent psycholinguistic research investigating gender representation during language comprehension has shown that the presence or the lack of gender information in the linguistic structure of a language contributes to shaping distinct gender representations. For example, languages such as English, that do not systematically grammaticize gender information in their linguistic structure, encourage readers to rely on their world knowledge for gender representations (e.g., Carreiras et al., 1996; Kennison and Trofe, 2003; Oakhill et al., 2005; Pyykkönen et al., 2010; Reynolds et al., 2006). Reading about person references such as *nurse* will generate inferences about the possible gender of the depicted person, with gender stereotypes acting as a primary source for representation (e.g., Banaji and Hardin, 1996; Cacciari and Padovani, 2007; Carreiras et al., 1996; Kennison and Trofe, 2003; Kreiner et al., 2008). Banaji and Hardin (1996), for example, showed that participants' judgments to the target stimuli (Experiment 1: judge whether the target was a male or female: *he vs. hers*; Experiment 2: judge whether the target was a pronoun or not: *she vs. do*) following either a gender stereotypical (e.g., nurse, mechanic) or gender definitional (e.g., mother, king) prime was found to be responded to faster when there was a gender congruency between the prime and target stimuli. Oakhill et al. (2005) further substantiated these effects of gender priming with a series of lexical priming experiments. Participants in their study were faster to accept word pairs consisting of a stereotypical role noun (e.g., surgeon) and kinship term (e.g., brother, sister) as referring to the same person in cases when the words were gender congruent. Activating such stereotyped gender inferences has been found to be immediate and robust among English readers, demonstrating that such role nouns may prime a specific stereotypical gender even if morphological or grammatical information may not compel readers to do so (e.g., Carreiras et al., 1996; Kennison and Trofe, 2003).

These representation tendencies however, are not readily generalizable for readers of grammatical gender languages such as French or German, where stereotypical gender is only one of the two possible sources contributing to the construction of gender representations. In these languages, gender is also integrated as part of their grammatical structure. Grammatical gender thus classifies a specific gender category to all nouns (e.g., masculine, feminine, and neuter in the case of German). This gender feature, when marked on person references, commonly corresponds to the biological gender of the referent (i.e., masculine = man, feminine = woman)[1], constraining its language users to consistently monitor gender information at both grammatical and semantic levels. A fundamental claim made by researchers is that the interaction between these two sources of information (i.e., stereotypical and grammatical) during the processing of role nouns is complex, and that the mechanisms for representing gender information are not always straightforward (e.g., Esaulova et al., 2013; Garnham et al., 2012; Gygax et al., 2012; Irmen, 2007).

The complexity of this interaction is rendered by the fact that gender information associated to its surface form does not necessarily coincide with its intended semantic connotations. For instance,

116   when considering the masculine form, there is a discrepancy between form and meaning. Whereas
117   role nouns such as *infirmières*Feminine [nurses] marked in the feminine grammatical form refer
118   unambiguously to female nurses, the masculine form (*infirmiers*Masculine) can refer exclusively to
119   men (i.e., only male nurses) or it may refer to a group composed of both male and female persons
120   (i.e., *generic* interpretation). Readers are presented with a challenge to disambiguate the intended
121   interpretation of the masculine form. It has been argued that its surface forms naturally emphasize
122   the association to the male gender, inevitably prompting a male-specific interpretation (e.g., Gygax
123   et al., 2012). Gygax et al. (2012), for example, adapting Oakhill et al.'s (2005) paradigm in French,
124   found that when participants were instructed to decide whether the person represented by a kinship
125   term in pairs such as *tante* [aunt] – *infirmiers*Masculine [nurses] could belong to a group represented
126   by the second noun (always in the grammatical masculine plural form), they responded positively
127   more often and faster when the kinship term was a man, indicating a male dominant representation.
128   The authors concluded that the generic interpretation could only be activated through active
129   processes, yet the male-specific interpretation was always passively activated (i.e., without control).
130   Most studies using on-line (e.g., Gabriel and Gygax, 2008; Gygax et al., 2008) and off-line (Braun
131   et al., 2005; Stahlberg et al., 2001) tasks concur on the male-specific impact of the masculine form.
132   Crucially however, this male bias effect persisted even when gender stereotypicality violated the
133   grammatical gender information (as seen in *infirmiers*: female stereotype, masculine grammar),
134   leaving the effect of stereotype information unclear.

135   In German, additional grammatical cues associated to its plural determiner (*die* [the]) and pronoun
136   (*sie* [they]) have been investigated, especially in conjunction with possible female biases. In a study
137   investigating gender representation in German, Rothermund (1998) found an unexpected reduction
138   of the male bias when participants conducted a recognition task after reading texts including plural
139   masculine references (*die*plural *Studenten* [the students]). The male-attenuating effect was attributed
140   as being triggered by the plural determiner *die* which shares the same surface form as the singular
141   feminine determiner *die* [*the* – singular – feminine]. Garnham et al. (2012) also showed a male
142   attenuated effect (or an additive female effect) when presenting the German plural pronoun *sie* (i.e.,
143   *they* – also feminine-equivalent) in a sentence judgment task examining the interpretations of
144   masculine role nouns. When the same was done in French however, the masculine pronoun *ils* [they
145   – masculine *specific* or *generic*] did not have a male amplifying effect despite its male association.
146   The authors argued that although cumulating male grammatical cues does not augment male biases,
147   combinations of male and female-equivalent grammatical cues may distract readers from activating
148   male specific representations. To our knowledge, when looking strictly at determiners, only one
149   study (e.g. Gygax et al., 2008) has generated specific hypotheses as to the impact of the definite
150   plural determiner *die* in German, yet its female-bias effect (as shown by Rothermund, 1998) was
151   never clearly replicated.

152   The studies discussed here demonstrate how grammaticized information influences readers'
153   comprehension processes. Grammatical gender languages work in a top-down manner, constraining
154   their users to consistently monitor gender both on grammatical and semantic levels. If, as suggested
155   by the thinking-for-speaking hypothesis, information grammaticized in languages prompts readers'
156   gender biases, which in turn anchor their representations, these regularities should also become
157   evident on their representations. If this were the case, it is reasonable to assume that readers of more
158   than one language may switch representations as they change languages. This notion is further
159   developed in this study by looking particularly at bilinguals where the language biases of each of
160   the bilingual's languages should surface on their representations. Sato et al. (2013) followed this
161   line of logic and investigated in a sentence-based paradigm, whether English-French bilinguals
162   would construct different representations according to their first (L1) and second (L2) language.
163   They presented English and French bilingual participants with sentence primes including role nouns
164   with stereotypical gender (e.g., female: *nurses*, male: *politicians*, neutral: *pedestrians*). Participants
165   judged the plausibility of target sentences including a gender reference (e.g., *some men, some
166   women*) to be a sensible continuation of the prime. The results indicated that switching language

167 was also accompanied by changes of biases in mental representations of gender, with English
168 eliciting stereotyped representations and French male-biased representations triggered by the
169 masculine form. Importantly, participants' L2 proficiency was found to be a good indicator of the
170 extent of the representation switch between L1 and L2.

171 In the present study, we followed Sato et al.'s (2013) study and investigated the effects of
172 stereotypes and linguistic encodings of gender on the representation of person reference role nouns.
173 French and German were taken as comparative languages, provided that they were both marked
174 with grammatical gender. This made them ideal candidates to test thinking-for-speaking effects, as
175 opposed to English, which lacks systematic grammaticization of gender. Characteristics surfacing
176 on representations when processing French and German should essentially reflect the impact of how
177 linguistic encoding contributes in shaping gender representations. Additionally, despite their
178 common usage of the masculine form to denote a generic interpretation, gender associations linked
179 to the plural determiners differ in the two languages. As argued by Rothermund (1996) and Gygax
180 et al. (2008), the German determiner *die* [the - plural] shares the same surface structure as the
181 singular feminine determiner *die* [the – singular – feminine], and should contribute to a female
182 additive bias when presented with a role noun in the masculine form. In contrast, the French plural
183 determiner *les* [the - plural] corresponds to both feminine and masculine nouns as they have a single
184 morphological realization (i.e., gender syncretism: Corbett, 1991) and therefore should not enhance
185 any additional gender information. If in the present study we are able to observe differences in
186 gender biases between French and German representations, it should provide more compelling
187 evidence as to the impact the grammaticization of language has on our conceptualization of gender
188 information.

189 To test these effects, we employed a combined linguistic-visual paradigm. This paradigm was
190 intended to provide a more sensible experimental framework to address the immediacy of gender
191 activation. While a handful of studies have examined gender representation processes employing a
192 lexical-based paradigm (Banaji and Hardin, 1996; Cacciari and Padovani, 2007; Gygax et al., 2012;
193 Oakhill et al., 2005; Siyanova-Chanturia et al., 2012), none have directly addressed the impact of
194 the use of the masculine form, or of role noun determiners. Studies investigating these effects have
195 approached the issue with a sentence comprehension task, applying anaphor resolution paradigms
196 that were dependent on the detection of semantic and syntactic inconsistencies in comprehension.
197 These tasks therefore did not strictly speak to the immediacy of the activation of such surface-level
198 grammatical cues, and discursive contextual elements may have interfered with stereotype
199 activation or with the accessing of signals during activation. More importantly, some, although
200 moderate, effects of stereotype have been observed, indicating that teasing apart these effects in a
201 linguistic context has been complex.

202 For instance, Esaulova et al. (2014) found a subtle effect of stereotypical gender in German. In their
203 experiment (Experiment 1), participants were presented with sentences composed of an anaphor
204 (e.g., *er* [he]) and a stereotyped role noun (e.g., *der Elektriker* [the electrician]) as an antecedent
205 while their eye movements were recorded. Although comprehension difficulty was most prominent
206 when the anaphor did not agree grammatically with its antecedent, as illustrated by most eye-
207 tracking measures, sentence processing was also influenced by the role nouns' stereotypicality, as
208 demonstrated in the late measures only (e.g. regression path on the pronoun region and total fixation
209 path on the role noun). Following the aforementioned Banaji and Hardin's (1996) experiments,
210 Cacciari and Padovani (2007) and Siyanova-Chanturia et al. (2012) also reported stereotype effects
211 in Italian, a grammatical gender language. They found that when a pronoun (e.g., *lui* [he] or *lei*
212 [she]) was primed by a bi-gender role noun (a noun that can vary in grammatical gender as a
213 function of biological gender, as in *insegnante* [a female / male teacher]), participants were
214 particularly slow to decide whether the pronoun was masculine or feminine when primed by a
215 counter-stereotypical role noun. Additionally, Carreiras et al. (1996), in a self paced reading task
216 (Experiment 2), showed that Spanish participants reading was delayed when a role noun (e.g., the

217 carpenter) was written in a grammatical form that mismatched its stereotypicality (e.g., *La*
218 *carpintera*_{feminine form} [the female carpenter] or *El enfermero*_{masculine form} [the male nurse]).

219 In sum, most studies have shown a strong impact of grammatical gender, with some authors
220 claiming that grammatical gender had only overshadowed stereotype effects (e.g., Esaulova et al.,
221 2013; Irmen, 2007, Reali et al., 2015). Although the impact of grammatical cues seems central in
222 representation processes, the reasons for the overriding effects of grammatical cues over gender
223 representations have not been clearly shown. We therefore explore the possibility that the
224 prevalence of male representations in grammatical gender languages (and the lack of stereotype
225 effects) may have well been prompted by the very nature of the paradigms being employed,
226 provided that both the prime and target stimuli were verbal stimuli. The use of verbal target stimuli,
227 maintaining a close link with its verbal prime, may have resulted in mental representations that
228 reflected only and merely linguistic activations. It could be that processing both prime and target
229 stimuli in a verbal context may constrain readers to over-monitor grammatical and syntactical
230 properties. This monitoring in turn may enhance the signal of a representation based on linguistic
231 cues (i.e., toward a male bias in gender-marked languages). In contrast, linguistic-visual paradigms
232 have been found to be effective in gauging effects of gender priming. Studies in social psychology
233 have shown that gender priming may be observed by presenting gender associated words (e.g.,
234 Kawakami and Dovidio, 2001: stereotypical traits; Lemm et al., 2005: words with gender-specific
235 suffixes and role nouns) followed by picture targets that required participants' judgments. For
236 instance, Lemm et al. (2005) showed that although past studies indicated a weaker priming effect
237 when using cross-modal paradigms, the gender priming effects found in their study were still large.
238 Consequently, this approach may indeed be well-suited to gauge the subtle stereotype effects we
239 seek to explore.

240 In our task, stereotypical role nouns in the masculine plural form, either with or without a plural
241 determiner, served as gender primes in German or in French. Participants had to make judgments as
242 to whether a visually presented pair of faces (male pairs or mixed pairs of faces composed of a
243 woman and a man) that followed could represent the preceding prime. The composition of face
244 pairs represented the possible interpretations that the role noun in the masculine form holds (i.e., a
245 male specific or a generic interpretation). We expected to replicate the male bias demonstrated in
246 previous findings (i.e., facilitated responses to male pairs of faces), and intended to explore the
247 influence of stereotype information. Specifically, an attenuated male bias was expected in the
248 female and possibly the neutral stereotyped conditions. Importantly, we also expected that the
249 determiner *die* in German would attenuate this potential male bias arising from the masculine form
250 of the role noun, whereas French rolenouns would retain the male bias. Finally, the experimental
251 task was carried out in participants' L1 and L2 to examine any representational shift that would be
252 prompted by the regularities of each language. For participants' L2, we also took L2 proficiency
253 into account, as measured by a L2 C-test. We expected shifts of representations to be influenced by
254 L2 proficiency (as in Sato et al., 2013).

Method

256 **Participants**

257 *German-speaking sample*

258 Fifty Caucasian German-speaking students from the University of Fribourg (Switzerland)
259 participated in the experiment for course credits. All participants were native speakers of German
260 whose L2 was French (mean age: 22, mean start age of French acquisition: 9.4 years, mean number
261 of schooling of French as L2: 7.2 years). Forty-one participants were women[2].

262   *French-speaking sample*

263   Fifty-one Caucasian French-speaking students from the University of Fribourg participated in the
264   experiment for course credits. All participants were native speakers of French whose L2 was
265   German (mean age: 22, mean start age of German acquisition: 7.5 years, mean number of schooling
266   of German as L2: 9.2 years). Thirty-nine participants were women.

267   **Materials**

268   *Prime role nouns*

269   Thirty-six gender stereotypical role nouns were selected as primes for the experiment (see Table 1).
270   These role nouns were taken from Gygax et al. (2008), all of which were normed and tested for
271   gender stereotypicality in Gabriel et al. (2008) in both German and French. Role nouns were female
272   (e.g., nurses [*Krankenpfleger / infirmiers*]), male (e.g., bosses [*Arbeitgeber / patrons*]) or neutral
273   (e.g., pedestrians [*Spaziergänger / promeneurs*]) in stereotype. To ensure that both female and male
274   stereotyped role nouns were similarly judged as prototypical exemplars of their respective
275   stereotype, we inverted ratings to female stereotypes (i.e., new rating = 100 – initial ratings), and
276   conducted a *t-test* to ensure that both were similarly judged. As expected, both were similar in both
277   languages, t_{French} (22) = .23, p = .82, and t_{German} (22) = .47, p = .64.

278   *Target face pairs*

279   The face pairs were created with the face modeling software FaceGen© Modeller program version
280   3.1.4 (Singular Inversions Inc., Toronto). A total of 30 male and 30 female Caucasian faces with
281   neutral expressions were created. They all had neutral expressions and the crown area of the faces
282   were removed in order to eliminate possible biases associated with certain hairstyles evoking
283   gender-biased information.

284   Twenty-one participants (14 women and seven men who did not participate in the main experiment)
285   participated in the first norming phase by rating the gender typicality of all 60 faces on a 7-point
286   scale (very masculine = 1, very feminine = 7) on a paper-pencil administered questionnaire.
287   Presentation order of the faces was randomized for each participant. Only faces that were clearly
288   rated as female (i.e., average score > 5) or male (i.e., average score < 3) were selected for the
289   experiment. Twenty-four female faces (M = 5.72, SD = .33, range= 5.43 - 6.3) and all thirty male
290   faces (M = 1.58, SD =. 26, range: 1.23 – 2.47) were retained. The average ratings of the female
291   faces (t(23) = 25.272, p < .001; $M_{difference}$ = 1.717) and male faces (t(29) = -50.173, p < .001;
292    $M_{difference}$ = -2.418) were significantly different from the scale midpoint (i.e., 4), with the difference
293   being bigger for male faces than for female faces. We deemed this imbalance in deviation from
294   midpoint non problematic for the purpose of our study, as our main focus was on assuring to select
295   non-ambiguous faces.

296   The 54 faces were then combined to make *male* and *mixed pairs of faces* (see Figure 1 for an
297   example of a presented pair of faces). Female pairs of faces were not constructed for the experiment,
298   as the interpretation of the presented masculine forms could not be grammatically interpreted as
299   being female-specific (i.e., represented by female pairs of faces). More importantly, these female
300   pairs of faces were avoided based on findings by Gygax and Gabriel (2008) who demonstrated that
301   the presentation of both feminine and masculine forms in the same experiment directs readers
302   towards a stronger male-specific representation of the masculine form. Female faces for mixed pairs
303   were always presented on the left in order to avoid a male preferred response according to a
304   possible left-side bias, illustrated in past studies using response scales in left-to-right languages (e.g.,
305   Gabriel et al., 2008). All pairs of faces were comprised of different faces.

306   A second norming phase was conducted in order to ensure that male and mixed pairs of faces were
307   not processed differently due to perceptual properties that we had not foreseen. In this pilot
308   experiment, our experimental pictures were presented on a computer screen running Experiment

309   Builder (SR Research) to another group of 27 participants (25 women and six men who had not
310   participated in the first norming phase). Their task was to decide, on two blocks of trials, whether
311   the presented pairs of faces were *of the same sex* in one block or *of different sex* in the other block,
312   by indicating their responses with a *yes* or *no* button press. The block order was inversed for half of
313   the participants. A repeated-measure ANOVA on correct response times (i.e., 94% of the data)
314   showed no main effect of *block*, $F(1, 26) < 1$, *ns.,* no main effect of *faces*, $F(1, 26) = 3.18^3$, *ns.*, and
315   no interaction, $F(1, 26) = 1.75$, *ns.*, confirming the homogeneity of our experimental target stimuli
316   in terms of perceptual properties.

317   *L2 proficiency assessment*

318   Participants' L2 proficiency levels were operationalized by their performance scores on a given C-
319   test (as done in Sato et al., 2013). Commonly in a C-test, participants are given several distinct
320   passages in which the second half of every other word is deleted except for the first and last
321   sentences. The task is to restore the blanks in the allocated time. This procedure was developed as
322   an effective measurement substituting cloze tests that were used in earlier years, and in recent years,
323   has been frequently encouraged as a measure for language proficiency (Eckes and Grotjahn, 2006;
324   Grotjahn et al., 2002).

325   In fact, C-tests have been shown to be highly correlated with standardized tests (e.g., Studienkollegs
326   in German: Grotjahn and Allner, 1996; TOEFL in English: Hastings, 2002; the five competencies
327   of the Test de Connissance du Français: Reichert et al., 2010). We employed the German C-test
328   offered by onDaF[4] to test German proficiency. Score ratings on this test are considered equivalent
329   to the Common European Framework of Reference for the levels A2 to C1. French proficiency was
330   evaluated with Coleman's (1994) C-test. Four texts were chosen from each original version and 20
331   minutes were allocated to complete the task.

332   *Role noun translation task*

333   To verify whether participants correctly identified the role nouns presented in L2, a role noun
334   translation task was conducted after the experimental trials. Participants were asked to provide a
335   translation for each presented role noun in their L1.

336   **Design and procedure**

337   The experimental task was conducted first in L1, followed by the task in L2 to minimize any data
338   contamination during the processing of a less dominant language[5]. Two experimental lists were
339   created to ensure that a role noun would not appear in both languages for a given participant. The
340   two lists were symmetrically different, in that if a role noun appeared in French in List 1, in List 2,
341   it would appear in German. To avoid an imbalance of gender stereotypicality between languages,
342   role nouns of similar strength of stereotype were always allocated to each language (see Table 1).
343   Each list consisted of six female, six male and six neutral role nouns per language, resulting in 36
344   critical role nouns per list, with each role noun appearing only in either French or German. Each
345   role noun was presented four times per participant (cf. Gygax et al., 2012 and Oakhill et al., 2005,
346   for a similar procedure): twice with a determiner (once followed by male pairs, once by mixed pairs
347   of faces), and twice without. All experimental items were intended to elicit a *yes* response.

348   To trigger *no* responses, twenty filler role nouns that had a gender association by definition (e.g.,
349   grandmother: *Großmütter* / *grand-mères*) were included. Half of the filler role nouns were male by
350   definition, whereas the other half were female. These filler primes were also presented four times
351   with their respective determiner allocations and face pairs. As these nouns were not ambiguous in
352   terms of gender, including them prevented participants from responding *yes* throughout the
353   experimental task without truly processing the role nouns and the target stimuli.

354 Table 1. Role nouns from Gabriel et al. (2008) and their corresponding gender proportion and
355 standard deviations (in parentheses) for each stereotype. All role nouns are presented in the plural
356 form as was in the experiment.
357

English	German	% (SD)	French	% (SD)
Spies	Spione	67 (15)	Espions	74 (17)
Golfers	Golfspieler	68 (14)	Golfeurs	73 (16)
Politicians	Politiker	69 (11)	Politiciens	72 (13)
Police officers	Polizisten	69 (10)	Policiers	70 (13)
Statisticians	Statistiker	72 (12)	Statisticiens	74 (15)
Bosses	Arbeitgeber	72 (12)	Patrons	74 (16)
Computer specialists	Informatiker	79 (11)	Informaticiens	67 (22)
Surgeons	Chirurgen	75 (12)	Chirurgiens	75 (14)
Technicians	Techniker	78 (14)	Techniciens	75 (14)
Engineers	Ingenieure	78 (11)	Ingénieurs	74 (14)
Physics students	Physikstudenten	81 (11)	Etudiants en physique	67 (28)
Pilots	Flieger	76 (13)	Aviateurs	74 (17)
Mean		74 (5)		72 (3)
Singers	Sänger	45 (8)	Chanteurs	48 (9)
Pedestrians	Spaziergänger	46 (8)	Promeneurs	52 (13)
Cinema goers	Kinobesucher	49 (6)	Spectateurs de cinéma	50 (5)
Concertgoers	Konzert-Zuhörer	47 (7)	Auditeurs de concert	51 (10)
Schoolchildren	Schüler	48 (5)	Ecoliers	53 (13)
Spectators	Zuschauer	41 (8)	Spectateurs	51 (5)
Neighbors	Nachbarn	50 (5)	Voisins	50 (8)
Swimmers	Schwimmer	50 (9)	Nageurs	50 (10)
Tennis players	Tennisspieler	52 (7)	Joueurs de tennis	54 (8)
Authors	Autoren	52 (9)	Auteurs	54 (8)
Musicians	Musiker	50 (9)	Musiciens	59 (13)
Skiers	Skifahrer	53 (8)	Skieurs	55 (9)
Mean		49 (3)		52 (3)
Beauticians	Kosmetiker	11 (8)	Esthéticiens	18 (20)
Birth attendants	Geburtshelfer	11 (19)	Assistants maternels	18 (18)
Fortune tellers	Wahrsager	24 (16)	Diseurs de bonne aventure	28 (27)
Cashiers	Kassierer	27 (16)	Caissiers	24 (15)
Nurses	Krankenpfleger	24 (11)	Infirmiers	30 (11)
Hairdressers	Coiffeure	21 (11)	Coiffeurs	38 (25)
Psychology students	Psychologiestudenten	25 (29)	Etudiants en psychologie	33 (10)
Dieticians	Diätberater	27 (15)	Diététiciens	37 (22)
Dressmakers	Schneider/Näher	23 (13)	Couturiers	40 (32)
Dancers	Tänzer	33 (12)	Danseurs	29 (14)
Sales assistants	Verkäufer	33 (14)	Vendeurs	37 (13)
Social workers	Sozialarbeiter	41 (14)	Assistants sociaux	33 (15)
Mean		25 (9)		30 (7)

358 The study was accepted by the Ethics Committee at the Department of Psychology of the
359 University of Fribourg and conformed to relevant regulatory standards. All participants were
360 granted informed consent. For each experimental trial, participants were first presented with a
361 gender stereotypical role noun prime following a fixation point (1000 ms). The role noun was
362 presented in the masculine plural form either in conjunction with a plural definite determiner (e.g.,
363 *die Ingenieure / les ingénieurs* [the engineers]) or without (e.g., *Ingenieure / ingénieurs* [engineers]).
364 Participants were instructed to press the *yes* button after having read the presented role noun, which
365 prompted the presentation of a picture of a pair of faces. Their task was to judge as quickly as
366 possible with a *yes / no* button press whether the presented target face pairs could represent the
367 prime role noun that appeared prior to the faces (see Figure 1 for the procedure). Filler trials, which
368 were randomized among experimental trials followed the same procedure, and the role nouns within
369 them were also presented either with or without a determiner.

370 The experiment was run on a Power Macintosh 4400 with the Psyscope software (Cohen et al.,
371 1993) connected to a button box to provide millisecond accuracy responses. Two buttons were
372 labeled, one "*Ja*" (yes) and the other "*Nein*" (*no*) for German-speaking participants and "*Oui*" (yes)
373 and "*Non*" (no) for French-speaking participants. Items were presented on a computer screen and
374 the "*Ja / Oui*" button was always pressed by the participant's dominant hand. All participants were
375 individually tested in a quiet room, with instructions being given in their respective native
376 languages. They underwent a practice session in their L1 with four items to familiarize themselves
377 with the task and procedure.

378 After the main experimental task, three paper-based post-tests were conducted. First, participants
379 were given a C-test in their respective L2. Following the C-test, participants were requested to
380 assess their L2 competence in terms of their listening, reading, writing and speaking abilities in the
381 L2 and to indicate the years and age of L2 acquisition by means of a self-administered
382 questionnaire. Finally, the role noun translation task was given to the participants to ensure they had
383 properly processed the critical items.

Results

385 We conducted analyses on both participants' binary responses (*yes/no*) to the facial images and
386 their response times for yes-responses (i.e., accepting the faces). Based on the results of the role
387 noun translation task conducted after the main experimental task, items in the L2 that were
388 frequently unknown to each language group (fewer than 10% of the participants were able to
389 provide a correct translation) were omitted from the analyses (*Schneider* [dress makers] and
390 *Wahrsager* [fortune tellers] were removed from L2-French participants' data and *diseurs de bonne*
391 *aventure* [fortune tellers] from L2-German participants' data). Mixed-effects logistic regression was
392 used to model the binary outcome variable (yes/no responses), and linear mixed-effects regression
393 was used to model participants' positive response times. Mixed-effect models provide a means to
394 perform analyses that account for missing values and to avoid the language-as-a-fixed-effect fallacy
395 (Brysbaert, 2007). All analyses were conducted using the R software (R Core Team, 2013), with the
396 *glmer* and *lmer* function from the *lme4* package (Bates et al., 2014). As suggested by Barr et al.
397 (2013) a model with a maximal random factor structure was adopted. Random intercepts and slopes
398 were varied for participants and items in order to account for the variance in performance created
399 by the factors. Random slopes were eliminated if their removal did not result in a significant
400 amelioration of the model or if the model did not converge. All predictors for fixed effects were
401 sum coded (+1, -1) and were entered by step-wise forward selection to an initial null model. Given
402 that participants' L2 *proficiency* was expected to predict general performance in the L2, the
403 *proficiency* predictor, as measured by C-test scores, was centered and entered as a covariate in the
404 null model, which included only random effects. Analyses for each language group were conducted
405 separately as in Sato et al. (2013), given that we expected different variances in the C-test scores.

406   Indeed, C-test difficulty has been found to vary according to various factors such as the language
407   of the C-test, text type or deletion pattern (Sigott and Köberl, 1996).

408   Log-likelihood ratio tests were used to determine the adequacy of including each predictor in the
409   model. A more complex model including the predictor in question was compared to a simpler
410   model without the inclusion of the predictor. If its integration significantly improved the model, the
411   predictor was retained within the model. The predictors tested in the models were *face pairs* (male
412   vs. mixed pairs of faces), *stereotype* (female vs. male vs. neutral), *task language* (German vs.
413   French) and *determiner* (without determiner vs. with determiner).

414   **Responses to Facial Targets**

415   Participants' binary choices were modeled in a mixed logit model to predict the likelihood that
416   participants would accept a face pair presented to them after a particular role noun prime. For both
417   language groups, the first model that followed the null model tested the effects of the masculine
418   form and of stereotype by introducing *face pairs, stereotype* and their interaction to the null model.
419   For both groups, the inclusion of these predictors significantly improved the model fit (Native
420   German group: $\chi^2 = 205.36$, *df* = 5, *p* < .001; Native French group: $\chi^2 = 150.8$, *df* = 9, *p* < .001).
421   The second model proceeded to test whether the effect of the German determiner impacted the
422   interpretation of the presented prime by adding the main effects of *task language* and *determiner*,
423   and importantly, all interactions between *face pair, task language* and *determiner*. While this lead
424   to a significant improvement of the model for the native German group ($\chi^2 = 75.8$, *df* = 6, *p* < .001),
425   the model failed to converge for the native French group. Therefore, only main effects for *task
426   language* and *determiner*, as well as the interaction between *task language* and *face pairs* were
427   introduced into the model, indicating an improvement ($\chi^2 = 31.69$, *df* = 5, *p* < .001). As for the
428   random structure, the final model for the native German group included random slopes for
429   *determiner* at the item level. The model for the native French group included random slopes for *face
430   pair* at the item level and *stereotype* and *task language* at the participant level. Both models
431   indicated a variance inflation factor less than 1.5, indicating that collinearity was not an issue.

432   *Native German group*

433   The results showed significant main effects of *face pairs* and *stereotype* which were qualified by a
434   significant interaction. Overall, the likelihood of a positive response was substantially higher for
435   male pairs of faces than for mixed pairs of faces *(b = .47, SE= .04, p < .001*, odds ratio = 2.56).
436   Consistent with our predictions, the *face pairs* X *stereotype* interaction revealed that this preference
437   for male pairs of faces was especially pronounced when they followed role nouns with male
438   stereotype, compared to when they followed role nouns with neutral stereotype *(b = .43, SE= .06, p
439   < .001*, odds ratio = 2.36) or female stereotyped role nouns (*b = .68, SE= .2, p* < .001, odds ratio:
440   3.89) (see Figure 2).

441   The model also revealed main effects of *determiner* and *task language*, indicating that the
442   likelihood of receiving positive responses was higher if face pairs were preceded by role nouns with
443   an article than when presented without the article (*b = -.19, SE= .05, p* < .001, odds ratio = .68).
444   Face pairs were also more likely to be responded to positively when they were presented with role
445   nouns in participants' L2 French than when presented with role nouns in their dominant L1 German
446   (*b = -.09, SE= .04, p* < .05, odds ratio = 0.83). Contrary to our predictions, these two predictors did
447   not interact, which would have supported the effect of the German determiner *die*. However, a *face
448   pairs* by *task language* interaction surfaced, indicating that male pairs of faces were more likely to
449   be responded to positively when preceded by a role noun in participants' L1 German than when
450   preceded by a role noun in their L2 French (*b = .24, SE= .04, p* < .001, odds ratio = 1.62).

451 *Native French group*

452 As was the case for the German group's responses, the analysis of the French sample revealed a
453 significant main effect of *face pairs* and a marginal significant effect of stereotype further qualified
454 by their significant interaction. The likelihood of accepting face pairs was again higher for male
455 pairs of faces than for mixed pairs of faces (b = .31, *SE* = .05, p < .001, odds ratio = 1.86). The
456 interaction revealed that the likelihood for participants to accept men pairs of faces was again
457 substantial when they followed male stereotyped role nouns than when they followed neutral (*b*
458 = .42, *SE* = .07, p < .001, p < .01, odds ratio = 2.31) or female stereotyped role nouns (b = .75, *SE*
459 = .12, p < .001, odds ratio = 4.48) (see Figure 2).

460 While the model revealed a main effect of *determiner* indicating that role nouns without a
461 determiner triggered greater positive responses (b = .08, *SE* = .03, p < .05, odds ratio: 1.17), no
462 interactions involving this predictor were observed. Finally, as was the case for the native German
463 group, a significant *face pairs* by *task language* interaction indicated that responses to accept men
464 pairs of faces was greater in the dominant L1 French than in the L2 French (b = -.08, *SE* = .03, p
465 < .05, odds ratio = 0.85).

466 **Positive Response Times to Facial Targets**

467 Overall, both groups responded above chance level to accept facial targets (native German group:
468 *yes* responses = 83%, *no* responses = 17%; native French group: *yes* responses = 75%, *no*
469 responses=25%), although these items were intended to elicit positive responses. Only reaction
470 times to these positive responses were subject to analyses. Response times that were 2.5 standard
471 deviations above or below the participant's mean were replaced by their cut-off values (3.5%).
472 Following the analyses of participants' responses to face pair targets, the effects of the masculine
473 form and stereotype were examined by introducing the main effects of *face pairs, stereotype* and
474 their interaction to the null model. There were significant improvements to the models for each
475 language group (Native German group: χ^2 = 157.67, *df* = 5, p < .001; Native French group: χ^2 =
476 210.94, *df* = 5, p < .001). The second model then added the main effects of *task language* and
477 *determiner* and all interactions between *face pair, task language* and *determiner* in order to test the
478 impact of the German determiner. The additions of these predictors resulted in an improvement of
479 the models for the native German group (χ^2 = 123.25, *df* = 6, p < .001) but not for the native French
480 group. Given that none of the effects introduced in the second model were significant, the initial
481 model was retained as the final model. For the native German group, the random structure for the
482 final model included random slopes for *face pairs, determiner* and *task language* at the item level
483 and *face pairs, stereotype, task language* and *determiner* on the participant level. The model for the
484 native French group included random slopes for item level and *face pairs, stereotypes* and their
485 interaction for participant level. Collinearity was not an issue given that both models indicated a
486 variance inflation factor less than 1.9.

487 *Native German group*

488 Consistent with the analyses for participants' responses, the final model showed significant main
489 effects of *stereotype* and *face pairs,* which were qualified with their significant interaction. Male
490 pairs of faces (825 ms) were responded to significantly faster than mixed pairs of faces (995 ms) (*b*
491 = -108.5 *SE*= 29.04, t = -3.73) confirming the male bias in past studies. This male bias was more
492 prevalent when role nouns preceding facial targets were of male stereotype than when they were
493 neutral (b = -47.64, *SE*= 13.89, t = -3.43) or female (b = -60.47, *SE*= 24.27, t = -2.49) stereotype.
494 No main effects or interaction effects including *determiner* were found, but a significant *task*
495 *language* effect indicated that participants were faster to respond in their L2 French (824 ms) than
496 in their L1 German (991 ms).

497   *Native French group*

498   The model revealed a significant main effect of *stereotype* and *face pairs* that was further qualified
499   by a significant interaction. Participants responded to male pairs (870 ms) of faces significantly
500   faster than to mixed pairs (1009 ms) of faces (b = -75.83, SE= 15.88, t = -4.77) confirming a male
501   bias. This effect was stronger for responses to male pairs of faces following male stereotyped role
502   nouns than when following neutral (b = -52.71, SE= 13.62, t = -3.87) or female (b = -77.92, SE=
503   23.78, t = -3.28) stereotyped ones. Contrary to our initial expectations, no effects including
504   *determiner* were significant.

505   **Discussion**

506   The aims of the present study were twofold. First, we aimed to evaluate how linguistic encoding of
507   gender in different languages shape and shift gender representations. Bilinguals of German and
508   French were tested to assess the rather inconclusive effects of a female bias associated to the
509   German determiner *die* (gender non-specific in the plural, but sharing the same surface form as the
510   feminine singular determiner). Although the activation of a male bias was anticipated, the presence
511   of an additional female association (i.e., *die*) was expected to attenuate this male bias in German.
512   The second goal was to provide more compelling evidence of main and interaction effects when
513   both stereotypical and grammatical gender information are available during the processing of role
514   nouns. It has been argued that the impact of gender stereotype information has often been
515   overshadowed by grammatical gender information in past studies, resulting in some uncertainty as
516   to how stereotype information actually influences the interpretation of the masculine form. While
517   past studies have relied on verbal targets, we argue that these tasks may have reinforced the
518   grammatical and morphological cues being tested. Such an impact may have resulted in strong, yet
519   less generalizable grammatical-based representations. In order to overcome these issues, a new
520   experimental approach using visual targets was suggested.

521   Overall, we found a consistent main effect of *face pairs* for both of our groups, where responses to
522   male pairs of faces were facilitated over mixed pairs of faces. This facilitation reflects the general
523   ease in interpreting role nouns in the masculine form as being male-specific rather than generic.
524   Although the surface form of the masculine grammar can theoretically be detached from its
525   semantic association *masculine=men,* it nonetheless boosted the activation of semantic properties
526   associated to the male gender. This was true even when participants were presented with visual
527   targets. Importantly, this male bias was persistent despite the fact that our pilot experiment on the
528   facial images showed a slightly faster, although not statistically significant (p = .08), tendency to
529   process mixed pairs of faces. Our results therefore suggest that a strong male bias is indeed
530   generated by the grammatical masculine form, and is not simply an artifact of the experimental
531   tasks employed in previous studies.

532   However, for both language groups, participants' responses to facial targets were also influenced by
533   stereotypicality, with male stereotyped role nouns generating processing facilitation of following
534   facial targets. In contrast, both response choices and positive response times indicated that facial
535   targets following role nouns with a female stereotype were more difficult to process. We believe
536   this to be indicative of an interference between the grammatically masculine form and the role
537   noun's female stereotypicality. Namely, both sources of information compete, increasing processing
538   difficulty. In contrast, an advantage was observed (i.e., a greater likelihood of allocating positive
539   responses and an elicitation of faster response times) for targets following male stereotyped role
540   nouns, which suggests that the congruency between the grammatically masculine gender and
541   stereotypical gender facilitated participants' construction of mental representations.

542   Importantly, these main effects were further qualified by a consistent *stereotype* by *face pair*
543   interaction for both the German and French group. This interaction indicated that participants'
544   acceptance to face pairs changed as a function of the stereotypicality of the role noun preceding it.

545 Male stereotyped role nouns triggered the greatest facilitation to accept male pairs of faces. These
546 results support the idea that when reading a gender associated role noun such as *nurses*
547 (*Krankenpfleger*German, *infirmiers*French), or *bosses* (*Arbeitgeber*German, *patrons*French) in a grammatical
548 gender language, gender stereotypical information is immediately activated as part of the
549 information associated with the role noun. As we did not embed our primes within sentences, our
550 results suggest that this activation is made at the lexical access, with discursive text elements not
551 needed to guide the activation of gender stereotypical information. Although we did find evidence
552 that the masculine form was highly influential in guiding the representation toward a male-
553 dominant representation as found in previous studies, we also documented that readers rely on
554 immediate stereotypical information, even in the presence of a masculine grammatical form.

555 Our results, however, do not necessarily speak to whether, and to what extent, grammatical gender
556 or stereotypical information has a greater influence over gender representations, as discussed in
557 some discourse-based studies (Irmen, 2007). They mainly support the idea that both are activated at
558 an early stage (i.e., lexical access), a claim that contrasts those of anaphor resolution studies that
559 suggest an activation at later stages of comprehension (e.g., Esaulova et al., 2013; Irmen, 2007).
560 The absence or weak indications of immediate stereotype effects in past studies could be attributed
561 to several reasons. First of all, past research has frequently relied on verbal primes *and* verbal
562 targets (e.g., Gygax et al., 2012; Gygax and Gabriel, 2008) to substantiate a persistent effect of the
563 masculine form as specifically referring to men, with the effects of stereotype being only modest.
564 The present study however, demonstrated that the apparent lack of stereotype effects could be
565 attributed to the tasks used to investigate these issues. We believe that by using facial images as
566 targets, we went beyond simple language-on-language task effects. Essentially the conceptual
567 nature of stereotypes may have made them better candidates for non-verbal tasks which made it
568 possible to delineate the true and noteworthy interaction between grammar and stereotypes when
569 constructing a representation of gender. Another plausible argument for the absence of stereotype
570 effects in past studies can be accredited to the nature of stereotype information, which dwindles
571 rapidly as readers process discourse. Consequently, its effects did not clearly surface in previous
572 studies on text comprehension. In the present study, the lexical-based paradigm may have allowed
573 stereotype effects to surface before fading away, as they would have in a discursive context. Such a
574 view may also support the reason for grammatical gender information to show a greater impact in
575 most studies on the topic.

576 In terms of the impact of language shaping gender representations, the two language groups showed
577 similar representation regularities in both their L1 and L2. This was rather unexpected given that we
578 had anticipated the male bias to be reduced when participants processed the role nouns in German,
579 due to its female-associated determiner. In fact, the German determiner did not elicit any substantial
580 effects. Although there was a modest trend for mixed pairs of faces to be accepted more often when
581 following female and neutral stereotyped role nouns (proportion of positive responses) when adding
582 the determiner *die* for native German readers in L1, it did not lead to statistically significant effects.
583 One could argue that when readers are faced with determining the grammatical gender of a noun,
584 they will make use of available semantic (i.e., conceptual) and phonological information, which
585 may result in processing facilitation (Schiller et al., 2003). In terms of our study, although both
586 conceptual and masculine grammatical gender information competed to represent a probabilistic
587 gender of the role noun, the association to the female gender of the German determiner did not
588 substantially contribute in the representation process.

589 Although we cannot definitively refute the phenomenon, the male-attenuating effect in German
590 documented by Rothermund (1998) appears to be at best superficial, at least in relation to the male-
591 bias exerted by masculine forms. The fact that Garnham et al. (2012) found an effect of *sie*
592 [they_Female], was most likely due to the fact that they combined *die* and *sie*, both feminine equivalent,
593 which offered a cumulative effect in deterring readers' attention from the role nouns' masculine
594 form. In our German data, we observed only a main effect of *determiner*, whereby role nouns

595 presented with a determiner facilitated responses to targets. These effects could be explained by
596 the different rules associated to German. For instance, in French, although a noun must always be
597 accompanied by a determiner even when a general statement is being made (e.g., *Infirmiers* doivent
598 s'occuper des personnes. [*Nurses* need to care for people.] is grammatically incorrect: An article is
599 always needed), in German, a noun can be presented both with and without a determiner
600 (*Krankenpfleger* müssen sich um Menschen kümmern. [*Nurses* need to care for people.] vs. *Die*
601 Krankenpfleger müssen sich um Menschen kümmern. [The nurses needed to care for people.])
602 which denote different meanings. The presence of *die* more clearly specifies that the role noun
603 refers to *a* (particular) group of people, and not to the general activity represented by the role noun,
604 consequently facilitating subsequent associated targets. In this regard, our German group may have
605 constructed different representations according to whether the role noun was presented with or
606 without a determiner.

607 We thus believe that gender information associated to the determiner appears to be trivial, at least in
608 comparison to the information associated to the gender inflection on the role noun. This gender
609 inflexion might be particularly relevant with person reference nouns, as they integrate conceptual
610 gender as part of their lexical representation (Oakhill et al., 2005). In contrast, information linked to
611 a function word, such as a determiner that connote less content and semantic information, would be
612 less readily associated to any conceptual gender. Nonetheless, these results are in line with the
613 numerous studies suggesting that the male bias exerted in grammatical gender languages is strong
614 and appears to govern the comprehension processes. As such, our results substantiate the idea that
615 language contributes in guiding mental representations. In our study, the grammatical masculine
616 form contributed in shaping male-dominant representations across (more or less) all stereotypes,
617 which is at odds with the idea that the masculine is the *unmarked* gender in grammatical gender
618 languages. Although the impact of *die* in German was not observed, the effects of the masculine
619 form of the role nouns lend support to the idea that grammatical markings may well direct (or bias)
620 our attention to particular categories. The masculine form makes the male concept more accessible
621 to readers. Note that this bias may not extend to less ambiguous cases such as *bigender* nouns in
622 Italian, as investigated, for example, by Cacciari and Padovani (2007) or Siyanova-Chanturia et al.
623 (2012).

624 Interestingly, we also observed a *task language* by *face pair* interaction surfacing in our German
625 group's responses, suggesting that the male bias was more persistent in participants' L1 German
626 than in their L2 French. This is crucial given that their dominant language exerted a greater male
627 bias than their less fluent L2, despite having a better understanding and command of the language
628 and the different interpretations of the masculine form in their L1. These results hint that the male
629 bias stem from L1 for grammatical gender language readers. Such an account is in line with
630 bilingual processing theories proposing that the languages of a bilingual are non-selectively
631 activated even when only one language is being used for language comprehension processes (de
632 Groot et al., 2000; Dijkstra and van Heuven, 1998).

633   Finally, we highlight that our linguistic-visual paradigm served as an effective approach to gauge
634 the effects in question. The male bias and stereotype effects observed in our study were apparent in
635 both the participants' L1 and in their less dominant L2. Importantly, despite the lack of stereotype
636 effects observed in the presence of a strong masculine cue in past studies, our paradigm allowed us
637 to observe stereotype effects. While some researchers argue that mixed-modal paradigms produce
638 less priming effects (e.g., Alario et al., 2000), our studies concurred with the conclusions made by
639 Lemm et al. (2005) that they are still very efficient and powerful.

640 Our results suggest that thinking to speak or read in a grammatical gender language emphasizes
641 gender associations, especially when these two are conceptually bound to each other. Although our
642 cognition of gender itself may not be fully influenced by grammatical gender, and this is an
643 empirical question, our social cognition may well be, given that the concept of gender, especially
644 that of male, is enhanced in grammatical gender language readers. These tendencies may then result

645 in shifting or influencing our social perceptions of gender-stereotyped occupations, guiding
646 readers to integrate a representation that is advantageous for men (Braun et al., 2005; Irmen and
647 Köhncke, 1996).

Conclusion

649 Using a linguistic-visual paradigm, the present study showed that readers automatically activate
650 gender-associated information when reading gender stereotypical human referent role nouns. The
651 activation of such information immediately takes place at a lexical level when readers encounter a
652 role noun. Though morphological markings such as the default masculine form in French and
653 German appear to be central when constructing mental representations of gender rather than
654 superficial surface features, our study demonstrated that stereotype information also plays a role in
655 influencing readers' mental representations. A stereotype effect was particularly apparent in the
656 cumulative effects of stereotype and grammar when readers encounter male stereotyped role nouns.
657 While past studies had not clearly found the effects of stereotype information in the presence of
658 strong masculine effects (e.g., Gabriel & Gygax, 2008; Gygax et al., 2012; Garnham et al., 2012),
659 the adaptation of a lexical and conceptual paradigm (with visual stimuli) was able to effectively
660 gauge these effects. Future studies may want to further examine the possibilities of suppressing
661 such male-dominant properties, though they appear to be relatively robust.

References

Alario, F.-X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. *The Quarterly Journal of Experimental Psychology Section A*, *53*(3), 741–764. http://doi.org/10.1080/713755907

Baayen, R. H. (2013). languageR: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics" (Version R package version 1.4.1.). Retrieved from http://CRAN.R-project.org/package=languageR

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. http://doi.org/10.1016/j.jml.2007.12.005

Banaji, M., & Hardin, C. (1996). Automatic stereotyping. *Psychological Science*, *7*, 136–141. http://doi.org/10.1111/j.1467-9280.1996.tb00346.x

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). http://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4* [R package version 1.1-7]. Retrieved from http://CRAN.R-project.org/package=lme4.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 61–80). Cambridge, MA: MIT Press.

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

Brysbaert, M. (2007). "The language-as-fixed effect fallacy": Some simple SPSS solutions to a complex problem. Royal Holloway, University of London.

Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition*, *14*, 47–59. http://doi.org/10.1017/S1366728910000180

Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics*, 277–293. http://doi.org/10.1017/S0142716407070142

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, *49A*, 639–663. http://doi.org/10.1080/027249896392531

Cohen, J. D., MacWhinney, B., Flatt, M. R., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*, 257–271. http://doi.org/10.3758/BF03204507

Coleman, J. A. (1994). Profiling the advanced language learner: the C-Test in British further and

higher education. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und Praktische Anwendungen [The C-test: theoretical foundations and practical applications]* (Vol. 2, pp. 217–237). Bochum: Brockmeyer.

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

de Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology*, *53A*, 397–428. http://doi.org/10.1080/027249800390547

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, *23*(3), 290–325. http://doi.org/10.1191/0265532206lt330oa

Esaulova, Y., Reali, C., & von Stockhausen, L. (2013). Influences of grammatical and stereotypical gender during reading: eye movements in pronominal and noun phrase anaphor resolution. *Language and Cognitive Processes*, *29*, 781–803. http://doi.org/10.1080/01690965.2013.794295

Fausey, C. M., Long, B. L., Inamori, A., & Boroditsky, L. (2010). Constructing agency: The role of language. *Frontiers in Psychology*, *1*. http://doi.org/10.3389/fpsyg.2010.00162

Gabriel, U., & Gygax, P. (2008). Can societal language amendments change gender representation? The case of Norway. *Scandinavian Journal of Psychology*, *49*, 451–457. http://doi.org/10.1111/j.1467-9450.2008.00650.x

Gabriel, U., Gygax, P., Sarrasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, *40*, 206–212. http://doi.org/10.3758/BRM.40.1.206

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

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*, 49–79. http://doi.org/10.1016/S0010-0277(01)00166-4

Grotjahn, R., & Allner, B. (1996). Studierende an Universitäten in NordrheinWestfalen. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* (Vol. 3, pp. 279–335). Bochum: Brockmeyer.

Grotjahn, R., Klein-Braley, C., & Raatz, U. (2002). C-Tests: An overview. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-Test* (pp. 93–114). Bochum: AKS-Verlag.

Gygax, P., & Gabriel, U. (2008). Can a group of musicians be composed of women? Generic interpretation of French masculine role names in the absence and presence of feminine forms. *Swiss Journal of Psychology*, *67*, 143–151. http://doi.org/10.1024/1421-

0185.67.3.143

Gygax, P., Gabriel, U., Lévy, A., Pool, E., Grivel, M., & Pedrazzini, E. (2012). The masculine form and its competing interpretations in French: When linking grammatically masculine role names to female referents is difficult. *Journal of Cognitive Psychology*, *24*, 395–408. http://doi.org/10.1080/20445911.2011.642858

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

Hastings, A. J. (2002). In defense of C-testing. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* (Vol. 4, pp. 11–29). Bochum: AKS-Verlag.

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

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

Irmen, L., & Köhncke, A. (1996). Zur Psychologie des "generischen" Maskulinums [On the psychology of the "generic" masculine]. *Sprache & Kognition*, *15*, 152–166. http://doi.org/10.1515/zfgl.33.2-3.212

Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin*, *27*(2), 212–225. http://doi.org/10.1177/0146167201272007

Kennison, S. M., & Trofe, J. L. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, *32*(3), 355–378. http://doi.org/10.1023/A:1023543602202

Kreiner, H., Sturt, P., & Garrod, S. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language*, *58*, 239–261. http://doi.org/10.1016/j.jml.2007.09.003

Lemm, K. M., Dabady, M., & Banaji, M. R. (2005). Gender picture priming: It Works with denotative and connotative primes. *Social Cognition*, *23*(3), 218–241. http://doi.org/10.1521/soco.2005.23.3.218

Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition*, *33*, 972–983. http://doi.org/10.3758/BF03193206

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, "n" roll: the representation of motion in language and cognition. *Cognition*, *84*, 189–219. http://doi.org/10.1016/S0010-0277(02)00046-X

Pyykkönen, P., Hyönä, J., & van Gompel, R. P. G. (2010). Activating gender stereotypes during online spoken language processing: Evidence from visual world eye tracking. *Experimental Psychology*, *57*, 126–133. http://doi.org/10.1027/1618-3169/a000002

R Core Team. (2013). *R: A language and environment for statistical computing*. [R Foundation for Statistical Computing]. Vienna, Austria. Retrieved from http://www.R-project.org/

Reali, C., Esaulova, Y., & von Stockhausen, L. (2015). Isolating stereotypical gender in a grammatical gender language: Evidence from eye movements. *Applied Psycholinguistics*, *36*(04), 977–1006. http://doi.org/10.1017/S0142716414000010

Reichert, M., Keller, U., & Martin, R. (2010). The C-test, the TCF and the CEFR: a validation study. In R. Grotjahn (Ed.), *Der C-Test: Beiträge aus der aktuellen Forschung. The C-Test: Contributions from Current Research* (pp. 205–231). Frankfurt, Germany: Peter Lang. Retrieved from http://orbilu.uni.lu/handle/10993/14401

Reynolds, D., Garnham, A., & Oakhill, J. (2006). Evidence of immediate activation of gender information from a social role name. *Quarterly Journal of Experimental Psychology*, *59*, 886–903. http://doi.org/10.1080/02724980543000088

Rothermund, K. (1998). Automatische geschlechtsspezifische Assoziationen beim Lesen von Texten mit geschlechtseindeutigen und generisch maskulinen Text-Subjekten [Automatic gender-specific associations to texts containing gender-specific and masculine generic text subjects]. *Sprache & Kognition*, *17*, 183–198.

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

Schiller, N. O., Münte, T. F., Horemans, I., & Jansma, B. M. (2003). The influence of semantic and phonological factors on syntactic decisions: an event-related brain potential study. *Psychophysiology*, *40*(6), 869–877.

Sigott, G., & Köberl, J. (1996). Deletion patterns and C-test difficulty across languages. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und Praktische Anwendungen [The C-test: theoretical foundations and practical applications]* (pp. 159–172). Bochum: Brockmeyer.

Singmann, H., Bolker, B., & Westfall, J. (2015). *afex: Analysis of Factorial Experiments* [R package version 0.13-145.]. Retrieved from http://CRAN.R-project.org/package=afex

Singular Inversions Inc. (2004). FaceGenModeller (Version Version 3.1.4). Retrieved from http://www.FaceGen.com.

Siyanova-Chanturia, A., Pesciarelli, F., & Cacciari, C. (2012). The electrophysiological underpinnings of processing gender stereotypes in language. *PloS One*, *7*, e48712. http://doi.org/10.1371/journal.pone.0048712

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name your favorite musician. *Journal of Language and Social Psychology*, *20*, 464 –469. http://doi.org/10.1177/0261927X01020004004

Footnotes

[1] Note that grammatical gender does not determine gender for all human references and the relation is mutual. For instance, there are special cases such as bi-gender role nouns where the sex of the referent determines the grammatical gender (e.g., *artiste* [artist - French] can be either masculine or feminine depending on the gender of the person). Additionally, there are also examples of epicenes where a single gender can refer to both the biological sex (e.g., *secrétaire*$_{Feminine}$ is always feminine irrespective of whether it refers to a male or a female).

[2] As past studies on gender representation (e.g., Garnham, Gabriel, Sarrasin, Gygax, and Oakhill, 2012; Gygax and Gabriel, 2008; Gygax, Gabriel, Sarrasin, Oakhill, and Garnham, 2008) did not find effects of participants' gender in reading tasks, we did not balance the gender sample of our participants.

[3] If anything, participants were slightly faster (by 36 ms) to respond to mixed pairs of faces than to male pairs ($p = .08$).

[4] www.ondaf.de

[5] Experiment order for participants' L1 and L2 was not randomized as Sato et al. (2013) did not find any experimental order effects based on language.

Figures

Figure 1. The experimental procedure in each condition.

Figure 2. Mean proportion of positive responses to accept facial images for each native group in each stereotype condition (independent of *task language*). Error bars indicate standard errors.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements