

Research Article

Gender at first sight: ERP insights into automatic activation, semantic bias, and inclusive language dynamics in Spanish and Italian

Alba Casado^{a,b,*}, Giuseppe Dario Benigno^c, Francesca Pesciarelli^d, Daniela Paolieri^{a,e}

^a Mind, Brain, and Behavior Research Center, University of Granada, Spain

^b Department of Developmental and Educational Psychology, University of Granada, Spain

^c Department of Spanish, Modern, and Classical Languages, University of the Balearic Islands, Spain

^d Università degli Studi di Modena e Reggio Emilia, Italy

^e Department of Experimental Psychology, University of Granada

ARTICLE INFO

Keywords:

Gender-congruency effect

Gender stereotypes

ERPs

Bilingualism

Language activation

ABSTRACT

The aim of this study was to investigate the activation of grammatical gender and its modulation by semantic information, such as biological sex and associated stereotypes. Using a gender-priming paradigm we explored the presence and modulation of the gender-congruency effect (GCE) through behavioral and electrophysiological measures (e.g., N400 and P300) in Spanish native and Italian-Spanish bilingual speakers. Primes included nouns with biological gender (e.g., *cocineros MAS* – male cooks), arbitrary gender referring to inanimate entities (e.g., *ventanas FEM* – windows); these nouns could be associated with a gender stereotype (e.g., *bomberos MAS* – male firefighters; *faldas FEM* – skirt) or neutral. Targets were gender-marked pronouns (*éstos MAS* / *éstas FEM* – these). Participants decided the gender of the pronoun. A GCE for both biological and arbitrary gender nouns was found, supporting the notion that grammatical gender is active at the lexical level and its activation is automatic, even in absence of relevant semantic information. Stereotypical primes enhanced the GCE, particularly among bilinguals, indicating that gender stereotypes can strengthen lexical-semantic activation in non-native speakers. Moreover, unlike Italian, both male and female Spanish speakers did not process the masculine form as inclusive, suggesting a shift toward more gender-sensitive representations in Spanish. Finally, while bilinguals showed evidence of cross-linguistic gender coactivation, native speakers displayed deeper semantic engagement, as reflected in stronger GCEs behaviorally. Overall, these findings demonstrate that grammatical gender processing in Spanish operates at a pure lexical level, interacts with semantic and sociolinguistic factors, and is sensitive to individual differences.

1. Grammatical gender in formal gender systems

Grammatical gender is a lexical-syntactic feature existing in many languages used to classify nouns in two or more categories (Corbett, 1991). For example, languages like Dutch differentiate between animate vs. inanimate entities through common or neutral grammatical gender (e.g., common gender: *de man* – the man; neutral gender: *het huis* – the house). In Indo-European languages, the most common classification of grammatical gender maps the biological distinction between males and females, which in language translates into feminine (FEM) or masculine (MAS), and sometimes neuter (NEU) (e.g., Corbett, 1991; Luraghi, 2011). The gender distinction can be extended to other linguistic elements apart from nouns/pronouns referring to people. For

instance, formal gender languages, like Romance languages, mark with gender all nouns (Corbett, 1991; Luraghi, 2011). Animate nouns have a direct mapping of gender assignment with sex for biological gender nouns (e.g., *hermana FEM* sister – *hermano MAS* brother), or a grammatical gender assignment independent of biological sex distinction for epicene nouns (e.g., *víctima FEM* victim – *personaje MAS* character). Additionally, inanimate entities have a grammatical gender assignment entirely arbitrary (e.g., *casa FEM* house – *vaso MAS* glass). Thus, in Romance languages, all nouns have grammatical gender, which originates from mapping the biological sex distinction and extends this classification to inanimate entities (Vigliocco et al., 2005).

The classical models of language production posit that grammatical gender feature only gets selected when needed for agreement (see

* Corresponding author at: Mind, Brain, and Behavior Research Center (CIMCYC), Department of Developmental and Educational Psychology, University of Granada, Campus Cartuja. C/ Profesor Clavera s/n, 18011 Granada, Spain.

E-mail address: albacasado@ugr.es (A. Casado).

<https://doi.org/10.1016/j.bandl.2026.105842>

Received 21 January 2026; Received in revised form 24 June 2026; Accepted 9 September 2026

Available online 23 September 2026

0093-934X/© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Caramazza, 1997; Schiller & Caramazza, 2003) despite it can be active at the lemma level (see Levelt, 1999; Stolle et al., 2025; and Huang & Schiller, 2021; Wang et al., 2025 for a similar classifier in Chinese). Still, there is evidence of the automatic activation of grammatical gender in bare noun production in languages with two genders, such as Italian or Spanish (e.g., Cubelli et al., 2005; Paolieri et al., 2010, 2020, but see Sá-Leite et al., 2022). Importantly, languages could be classified as more gender transparent or opaque (Kupisch & Rothman, 2018; Sá-Leite & Lago, 2024; Velnić, 2020), with consequences for how grammatical gender is processed during lexical access (e.g., Gollan & Frost, 2001; Sá-Leite et al., 2023; Sá-Leite & Lago, 2024). That is, the transparent language systems with two genders seem to favor the automatic activation of gender, which indeed could influence how speakers of these languages pay attention to the gender feature of nouns. In particular, the grammatical distinction between feminine and masculine along all entities seems to trigger an attentional bias toward those dimensions of meaning in native speakers of two-gender languages (Vigliocco et al., 2004). In that sense, speakers of Romance languages such as Italian or Spanish may pay more attention to the sex of referents than speakers of languages with other gender systems (like English). This attentional tendency could bias the perception of inanimate objects as being more male- or female-like based on the arbitrary grammatical gender assignment (e.g. Boroditsky et al., 2003; Sera et al., 2002).

In 2005, Vigliocco et al. performed a series of studies to explore the effects of grammatical gender on meaning using a transparent language (Italian, two-gender system) and an opaque language (German, three-gender system). They found gender effects but limited 1. to the transparent language (i.e., Italian), 2. to tasks that required mandatory lexical access (i.e., picture naming), and 3. to nouns referring to animate entities (i.e., humans and animals). They explained the effects under the “sex and gender hypothesis”, which proposes that children first associate the sex of the referents and the gender of nouns referring to humans due to the match between the conceptual feature (biological sex distinction female or male) and the linguistic feature (grammatical gender feminine or masculine). Once this association is established for nouns referring to humans, then it is generalized to the rest of the nouns, resulting in a transfer of female-or male-like features to other animate entities. In a less constrained version of the sex and gender hypothesis, the transfer from sex distinction to the purely grammatical one extends to inanimate entities; although Vigliocco et al. (2005) did not find evidence supporting this less constrained version, other studies have found that gender effects can arise in nouns referring to inanimate entities when sex-related information is active. For instance, in a series of studies, Casado et al. (2018, 2021) found that in Spanish, a transparent language with two-genders, the identification of the sex of the speaker facilitated the processing of nouns whose gender assignment was congruent with that of the sex information (e.g., female voice producing a feminine word), in comparison to when the gender was incongruent (e.g., female voice producing a masculine word). By comparing word repetition, lexical decision, and gender decision tasks, Casado et al. (2018) showed that the grammatical gender feature is accessed even when no mandatory attentional focus is required (e.g., in lexical decision). This confirms that gender is an intrinsic lexical property automatically activated during word processing. Crucially, the gender effects emerged for nouns referring to animate entities with semantic gender (e.g., *hermano* MAS - brother) and for nouns referring to inanimate objects with arbitrary gender (e.g., *casa* FEM - house), although in the latter case with weaker strength. Moreover, in order to experience gender effect, explicit language use is required (Bender et al., 2016; Cubelli et al., 2005; Samuel et al., 2019), for instance, the influence of grammatical gender disappears in conceptual tasks that do not require lexical access such as semantic categorization, or when phonological suppression is enforced (Cubelli et al., 2005). Thus, the scope of gender effect is lexical-semantic and does not extend to the purely conceptual level (Vigliocco et al., 2005).

Altogether, it seems that each language system is structured to direct

the speakers' attention to certain elements, and such specificity could modulate how speakers activate information. For instance, speakers of transparent languages in terms of gender, such as the Romance languages, are biased to categorize into feminine and masculine due to the gender transparency of the language system; the extent of this bias depends on the degree of mapping between the semantic information related to sex and the grammatical gender assignment, and the requirements of the linguistic tasks.

The mapping between the semantic information related to sex and the grammatical gender can vary in different nouns. As previously noted, semantically gendered nouns exhibit a high mapping or correspondence between biological sex and grammatical gender assignment. Still, this mapping can be increased or decreased by adding stereotypical information related to gender. For instance, the mapping increases in the case of “female nurse”, given that the proportion of female nurses is higher than that of male nurses, and moreover, this profession implies caring for others, typically associated with female roles. A case in which the mapping decreases could be, for instance, “female firefighters”, given that the proportion of females in this profession is lower compared to males, and this profession implies risk taking, typically associated with male roles. Thus, the associated stereotypical information can modulate the activation of semantic information related to gender. Still, stereotypes depend on personal experience and are idiosyncratic to each culture, and within cultures, to each individual (Allan, 2007; Beller et al., 2015; Bender et al., 2016, 2018; Nicoladis & Foursha-Stevenson, 2012; Segel & Boroditsky, 2011). Past studies (Cacciari & Padovani, 2007; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012) explored the automatic activation of gender stereotypes and compared it with the automatic activation of grammatical gender in Italian native speakers. They used a priming task in which participants were exposed to a prime noun for a very short time (50 ms in Pesciarelli et al., 2019; 200 ms in Cacciari & Padovani, 2007; 300 ms in Siyanova-Chanturia et al., 2012) followed by a personal pronoun (*lui*-he, *lei*-she); their task consisted of deciding the grammatical gender of the personal pronoun. The primes were either role nouns marked by gender (e.g., *sorella* FEM-sister) or role nouns without a gender mark that could be stereotypically feminine (e.g., *insegnante*-teacher), masculine (e.g., *ingegnere*-engineer), or neutral (e.g., *cantante*-singer). The results showed that native speakers were biased by the gender information carried by the prime: gender decision on the pronoun took longer when preceded by gender incongruent information, both for role nouns marked by gender (e.g., *sorella* FEM – *lui* MAS) and for those unmarked but associated with gender stereotypes (e.g., *insegnante* FEM – *lui* MAS). This gender-congruency effect for grammatical and stereotypical gender indicates that both kinds of information are automatically activated and available at very early stages of processing.

These results were replicated with Spanish native speakers in our previous study (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023) with a slightly different design. As primes, we included plural role nouns marked by gender that could be either stereotypically associated with females (e.g., *enfermeras* FEM – nurses) and males (e.g., *bomberos* MAS - firefighters) or stereotypically neutral (e.g., *cocineras* FEM – female cooks; *biólogos* MAS – male biologists), and also plural epicene nouns, making reference to animals but with an arbitrary gender assignment (e.g., *jirafas* FEM – giraffes; *búhos* MAS - owls). As targets, we used demonstrative pronouns (*éstos* MAS / *éstas* FEM – these). In line with Cacciari and Padovani (2007), Pesciarelli et al. (2019), and Siyanova-Chanturia et al. (2012) we found a gender-congruency evoked by all types of primes with a similar effect-size. That is, our results demonstrated that grammatical gender activation, despite being lexical, influenced the activation of upcoming gender information in a similar manner to stereotypical information, a pragmatic world-knowledge that exerts a lexical-semantic influence during real-time processing. Indeed, Siyanova-Chanturia et al. (2012) and Pesciarelli et al. (2019) corroborated this explanation with electrophysiological data. They explored event-related potentials (ERPs) elicited during the processing of the

target word (the pronoun) focusing in particular on the N400 component, which is associated with the integration of lexical-semantic information (Friederici & Jacobsen, 1999; Paolieri et al., 2020; Van Petten & Kutas, 1990). In line with the behavioral results, they showed that gender-incongruent preceding information induced by both role nouns marked by gender and those unmarked but associated with gender stereotypes, evoked larger N400 components than gender-congruent information, thus demonstrating the semantic bias induced by lexical information. Additionally, they also explored the P300 component, associated with the subjective probability of events, typically observed in priming designs (Barber & Carreiras, 2005). Siyanova-Chanturia et al. (2012) found that the feminine pronouns were not affected by the gender-congruency when preceded by role nouns associated with stereotypical gender. In a more recent study, Serafini and Pesciarelli (2025) also found a similar asymmetry with a gender-priming study in Italian including faces, such that the N400 component marked only the violation of the female-related stereotype by male agents, not the other way around. Additionally, they found a P300 effect for female faces, being the first evidence of P300 as modulated by stereotypes. In both cases, the authors explained this asymmetry like Cacciari and Padovani (2007) and Irmen et al. (2010) who found similar behavioral results: gender stereotypes are more constrained in the feminine roles than on the masculine roles, which indicates that, in general, people are more likely to accept females performing a typically masculine roles than the other way around, probably due to the privileged position of males and their roles in western societies.

Pesciarelli et al. (2019) also found a gender asymmetry with marked role-nouns, such that the gender-congruency disappeared with feminine pronouns (e.g., *passaggero* MAS – *lei* FEM passenger). They explained the results based on the grammatical rule that considers the masculine form as the generic or inclusive form. In most gendered languages, the masculine form is considered the default or generic gender (for Spanish see Harris, 1991 or Prado, 1982, and the recommendation from the Royal Academy of Spanish language- RAE). Not only the masculine form is assigned more often to novel nouns (for instance, the majority of borrowings from foreign languages are masculine in Spanish (e.g., *el* MAS iceberg MAS – the iceberg), even when the direct translation is feminine (e.g., *el* MAS *córner* MAS – *la* FEM *esquina* FEM – the corner in football), but is thought to include the representation of both genders in singular (e.g., “*el* MAS *hombre* MAS” – the men as mankind) and in plural (e.g., “*los* MAS *niños* MAS” – the kids) (see Beatty-Martínez & Dussias, 2019 and Irizarry et al., 2025 for a similar explanation). Coming back to the previous example, the Italian version of passengers would include both the representation of female and male figures. Thus, it is likely plausible to encounter a feminine pronoun after a male and a female antecedent. Further, we found a similar asymmetry with a reduced gender-congruency effect when processing the feminine pronoun, but this time uniquely with epicene nouns (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023). We speculated that the asymmetry may be evoked by the masculine form working as the generic gender, such that the masculine nouns referring to animals (e.g., *pingüino* MAS – penguin) induce the activation of female and male exemplars, such that when encountering a feminine pronoun afterwards it is not surprising. On the other hand, the feminine epicene nouns referring to animals (e.g., *tortuga* FEM – turtle) are likely to induce a more salient activation of the female exemplar, also because the feminine form is the marked one (Harris et al., 1991; Prado, 1982), such that it is less likely to encounter a masculine pronoun afterwards and consequently evoked the gender-congruency effect when followed by a masculine pronoun. However, we did not replicate this asymmetry with role-nouns in Spanish. Probably, because this generic use of the masculine form is quite confusing given that speakers/readers/listeners have to deduce by the context which is the appropriate use: generic or specific of males, and on most occasions exclusively masculine interpretation is activated (Anaya-Ramírez et al., 2022).

Altogether, previous research showed that grammatical gender

information modulates the activation of gender information at the semantic level (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012). However, we cannot be sure that grammatical information per se is inducing the effect, since previous studies exclusively used nouns referring to biologically sexed entities, which allow for a coherent mapping between grammatical gender and biological sex. Therefore, we consider it relevant to explore whether pure lexical information modulates the activation of gender information at the semantic level by exploring the processing of arbitrary gendered nouns referring to inanimate entities (e.g., *casa* FEM – house).

Moreover, previous studies showed that the associated stereotypical information can modulate the activation of semantic information related to gender (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), specifically involving: 1. nouns lacking a direct mapping between the biological sex and the grammatical gender but carrying an associated stereotype (e.g., *insegnante* MAS/FEM – teacher), and 2. nouns with a direct mapping between the biological sex and the grammatical gender combined with an associated stereotype that amplifies the degree of mapping (e.g., *infermera* FEM – nurse). Still, it is unknown whether adding stereotypical gender information would modulate the activation of nouns referring to inanimate entities. Hence, we aim at exploring whether in arbitrary gendered nouns, additional stereotypical information (e.g., *falda* FEM – skirt) can induce a mapping between sex and grammatical gender dimensions that would modulate the activation of semantic information related to gender.

Further, it seems that the masculine form can work as the generic, including both female and male figures in Italian role-nouns (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012). However, this generic use was not replicated in Spanish behavioral data (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023), such that the effect was only present with epicene nouns (see also Anaya-Ramírez et al., 2022 for lack of generic use of the masculine in sentence context). Still, we cannot rule out this interpretation given that no ERPs were measured in the Spanish population. Consequently, it is important to explore whether the processing of the masculine form induces the generic representation of males and females or uniquely the masculine interpretation by observing the ERP activity. In addition, the extent of these effects is unknown beyond native speakers, that is, it is not clear whether the effects are replicated in bilingual speakers. It is well known in the bilingualism literature that both languages are active at the same time even when only one is required to perform the task, an effect known as language co-activation (Kroll & Stewart, 1994). For instance, when performing a picture naming task in the second language (L2), participants take longer to name the pictures when the gender assignment in the native or dominant language (L1) is congruent (e.g., Spanish-Italian *ventana* FEM – *finestra* FEM window) with that of the L2 vs. when incongruent (e.g., Spanish-Italian *mesa* FEM – *tavolo* MAS table) (see also, Bordag, 2004; Klassen, 2016; Morales et al., 2011; Paolieri et al., 2010, 2019). While behavioral data from Italian learners of L2 Spanish exhibited clear gender congruency effects, grammatical gender appeared to be less salient than cognate status in modulating P300 amplitudes (Von Grebmer zu Wolfsturn, Pablos-Robles & Schiller, 2022; see also Von Grebmer zu Wolfsturn, Gupta, Pablos-Robles & Schiller, 2023). In language comprehension, using a lexical decision task, Casado, Ferré & Paolieri (2023) observed that bilinguals made more errors when deciding whether an L2 word was real if the translation equivalent in their L1 was gender incongruent as compared to congruent, an effect also supported by larger N400 components. Furthermore, participants take longer to translate words from their L2 to their L1 when grammatical gender is incongruent between languages as compared to when it is congruent, as shown in Catalan-Spanish bilinguals (Paolieri et al., 2019) and Russian-Spanish bilinguals (Paolieri et al., 2019), demonstrating cross-linguistic gender influence. Altogether, previous studies have shown that grammatical gender is active in

the L1 when using the L2, and that its automatic activation can modulate lexical selection in both production and comprehension. Based on previous results, we speculate that bilinguals may experience a greater gender-congruency effect potentiated by congruency in grammatical gender between languages triggered by coactivation of their native language when using the L2. Still, to our knowledge, no previous research has explored the gender-congruency effect with a priming paradigm in a bilingual population.

1.1. Current study

In the current study, we designed a gender-priming paradigm in Spanish (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012) in which participants saw a prime noun for a short period of time (50 ms) and afterwards a personal pronoun. We chose the unmasked condition because a 50 ms prime duration is sufficient to capture automatic grammatical gender activation while preventing conscious strategic processing (Pesciarelli et al., 2019). At this exposure, the prime remains unavailable for deliberate heuristics, making additional masking unnecessary. The task consisted of deciding the grammatical gender of the pronoun (feminine, masculine) and ignoring the prime noun. We tested two groups of participants: native speakers of Spanish, and native speakers of Italian with Spanish as their second language (bilinguals). While performing the task, we recorded their electrophysiological response.

RQ1: We were interested in exploring whether grammatical gender, as pure lexical information, can induce the gender-congruency effect, thus influencing the activation of semantic information in the upcoming elements; to this end, we selected as primes biological gender nouns (marked role-nouns, as in previous studies e.g., *cocineros* MAS – male cooks) and arbitrary gender nouns referring to inanimate entities (e.g., *ventanas* FEM – windows).

Prediction 1: We expected to find evidence of gender incongruency between the prime and the target, both in biological (e.g., *cocineros* MAS – male cooks / *éstas* FEM – these) and arbitrary gendered nouns (e.g., *ventanas* FEM – windows / *éstos* MAS – these), as reflected by longer reaction times (RT) and larger N400 amplitudes compared to when the prime and target were congruent in gender (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023; Siyanova-Chanturia et al., 2012; Pesciarelli et al., 2019).

RQ2: Furthermore, we wanted to explore whether the gender-congruency effect is potentiated by the addition of gender stereotypes, which increase the mapping between semantic information related to sex and grammatical gender assignment; to pursue this aim, we selected as primes nouns associated with gender stereotypes (e.g., *enfermeras* FEM – nurses, *faldas* FEM – skirts) and neuter nouns (e.g., *camareras* FEM – waitresses, *casas* FEM – houses).

Prediction 2.1: We expected to find an increase in incongruency detection for primes with a stronger mapping between semantic information related to sex and gender assignment; in particular, we expected longer RT and larger N400 amplitudes triggered by incongruent primes with biological gender (e.g., *camareras* FEM – waitresses / *éstos* MAS – these) vs. arbitrary gender (*casas* FEM – houses / *éstos* MAS – these) (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023), and also for incongruent primes with an associated stereotype (*faldas* FEM – skirts / *éstos* MAS – these) vs. neutral (*casas* FEM – houses / *éstos* MAS – these) (Casado et al., 2021).

Prediction 2.2.: Additionally, we expected to find larger P300 amplitudes for biological gender incongruent primes with an associated stereotype, since the expectation of encountering an incongruent pronoun afterwards is even lower for stereotypical nouns (e.g., *enfermeras* FEM – nurses / *éstos* MAS – these) than for neutral nouns (e.g., *camareras* FEM – waitresses / *éstos* MAS – these) (following the results of Serafini & Pesciarelli, 2025 with female faces).

RQ3: Moreover, we wanted to dig into the role of the masculine form as the generic gender in Spanish; if according to the norms of Spanish

language use the masculine form works as the generic inducing the activation of male and female figures, we should observe a decrease in the incongruency detection. In this case, encountering a feminine referent after a masculine noun would be plausible and correct, compared to the incongruency response of the masculine pronoun when preceded by a feminine prime referring to inanimate entities (e.g., *camareras* FEM – waitresses / *éstos* MAS – these), which is neither possible nor correct (Pesciarelli et al., 2019).

Prediction 3: We expected the decrease in the incongruency detection to be reflected by a reduced P300 component for the feminine pronoun when preceded by a masculine prime referring to animate entities (e.g., *cocineros* MAS – male cooks / *éstas* FEM – these).

RQ4: To further examine this phenomenon of the generic masculine, we explored how male and female participants process this form separately.

Prediction 4: We expected some differences, given that males are always included when the masculine form is used (generic or definitional) and excluded when the feminine form is used, whereas females must infer whether they are included in the masculine form depending on the generic or definitional use.

RQ5: Finally, we considered it of great interest to examine whether the gender-congruency effect is experienced by bilingual speakers using their L2, given that this can be considered language experience dependent, and also because language co-activation could potentiate the gender-congruency effect. To this aim, we tested native Spanish speakers and compared their performance to that of Italian-Spanish bilinguals in Spanish. All the selected nouns were gender-congruent between languages (e.g., *enfermeras* FEM – *infermieri* FEM nurses, *faldas* FEM – *gonne* FEM skirts, *cocineros* MAS – *cuochi* MAS male cooks, *ventanas* FEM *fenestre* FEM – windows).

Prediction 5: We expected to find greater incongruent detections as marked by longer RT (Paolieri et al., 2020) and larger N400 and P300 amplitudes in bilinguals vs. native speakers.

This study provides a qualitatively new level of integration by examining how linguistic and social factors converge during the processing of the plural generic masculine. To achieve this, the present research introduces four key methodological and theoretical innovations: 1) By comparing native Spanish speakers with Italian-Spanish bilinguals—and ensuring all primes have gender-congruent translation equivalents across both languages—this design allows us to investigate whether L2 processing makes the lexico-syntactic system more permeable to top-down lexical-semantic influences, such as stereotypes. This population comparison aims to provide novel insights into the interactivity of the bilingual brain that native-only studies cannot capture. 2) Unlike previous studies that often focused on a single category, this manuscript employs a balanced design including animate and inanimate nouns (in both categories half stereotyped, half neutral). This configuration makes it possible to isolate, for the first time within the same experimental paradigm, the effects of purely grammatical gender from biological and stereotypical gender. 3) The shift to plural forms is theoretically critical because the plural represents the most robust and systematic environment for the generic masculine in Spanish. Testing the generic masculine in this specific environment allows us to directly challenge prescriptive “default” rules in a way that studies using singular or epicene forms cannot. 4) Finally, this work integrates electrophysiology (ERP) and behavioral results to measure processing effort via the N400 and P300 components. This approach provides a deeper neurocognitive view, tracking how participants react to gender and stereotype variations in real-time.

The study was preregistered in OSF (<https://osf.io/v3xzd/>).

2. Methods

2.1. Participants

We tested a total of 91 participants, 46 native Spanish speakers and

45 Italian-Spanish bilinguals living in Spain. The Spanish natives had no knowledge of the Italian language. The Italian-Spanish bilinguals did not have knowledge of any other language with a formal gender-system beyond an A2 level. The specific socio-linguistic characteristics of the participants are included in Table 1. All participants received either course credits for the University of Granada or 15€ in compensation for their participation.

From the initial sample (91), we excluded 8 participants due to recording problems (4 of each group). We also excluded 3 bilingual participants who did not reach 60% of accuracy in the translation task. The final sample consisted of 42 native Spanish speakers (20 females, 1 non-binary, 21 males) and 38 Italian-Bilinguals (21 females, 1 non-binary, 16 males).

Additionally, 7 participants, three from the Spanish native group and four from the Italian-Spanish group were excluded from the ERP analysis due to the high proportion of rejected epochs (>25%).

The study met the requirements and received approval from the Ethics Committee of the University of Granada for experimental studies with human subjects (2702/CEIH/2022).

2.2. Materials

Like in our previous study (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023), as targets we used the demonstrative pronouns “*éstos MAS*” and “*éstas FEM*” – these. To select the primes, we programmed an online questionnaire on the platform [jotform.com](https://www.jotform.com). Seventy-three native Spanish speakers were asked to rate 379 nouns based on the associated feminine or masculine perception. Out of 379, 130 of the nouns were animate and presented in their plural masculine and feminine form (e.g., *ciudadanas FEM/ciudadanos MAS* – citizens). We employed the plural form because the generic masculine in Spanish and Italian is much more prominent and systematic in plural nouns than in singular ones (e.g., *los niños*—the children—which encompasses both boys and girls). In contrast, the singular form frequently triggers an unconscious mental representation of a specific male individual. Animate nouns were rated through an 11-point scale, in which 100/0 corresponded to the highest masculine perceived frequency ratio, 50/50 to an equal frequency ratio and 0/100 to the highest feminine perceived frequency ratio, based on Carreiras et al. (1996).

The remaining 249 were inanimate nouns (e.g., *calabaza FEM* – pumpkin), and were rated in a 5-point Likert scale, in which 1 corresponded to the most masculine associated perception and 5 to the most feminine one, following a similar procedure as in Casado et al. (2021).

In the end, we selected 168 nouns, 84 animate and 84 inanimate based on their z-score scaled stereotypical value obtained from the online questionnaire. Half of the nouns were stereotyped (i.e., associated with female or male referents based on the norming scores) and were always gender congruent with their associated semantic value (e.g., *enfermeras FEM* – nurses, associated with females). The remaining half of the nouns were neutral (i.e., not associated with any particular gender, e.g., *camareros MAS* – waiters). Inside each category (stereotypical

/neutral), half of the nouns were feminine and the other half masculine (see Table 2 for more information about stimuli selection, and Appendix A for the complete list of stimuli).

All primes were Spanish nouns for which the direct Italian translation was gender-congruent (e.g., *calabazas SP* – *FEM* – *zucche IT* – *FEM* – pumpkins). All the selected items were controlled for transparency, such as 85% of items had a transparent ending for grammatical gender, i.e., masculine nouns that end in -os in their plural form, and feminine nouns that end in -as in their plural. The remaining opaque nouns were phonologically unmarked both in Spanish and Italian (e.g., *pirámides SP* – *FEM* – *OPAQUE* – *piramidi IT* – *FEM* – *OPAQUE* – pyramids). Within each category (animate; inanimate) we controlled for logarithmic lexical frequency (Spanish and Italian), plural noun length (Spanish and Italian), concreteness (English), valence (Spanish), age of acquisition (English) and NLD orthographic similarity between Spanish and Italian. For the lexical frequency, we used the Italian and Spanish version of “Subtlex” database (Amenta et al., 2021; Cueto et al., 2012 respectively). For concreteness we used Brysbaert et al. (2014) values for English translation equivalents. For valence and arousal we used Stadthagen-Gonzalez et al. (2017). For age of acquisition (AoA) we used Kuperman et al. (2012) values for English translation equivalents. Finally, for “NLD”, the Guasch et al. (2013) web-based stimuli research engine NIM was used in its Spanish adaptation (see Table 3).

2.3. Tasks and procedure

All participants were tested in an isolated room; they were seated in front of a 17” 60 Hz LCD monitor positioned at eye level at a distance of 60 cm. Every participant received oral and written instructions for each of the tasks.

At the beginning of the experimental session, while adapting the electrophysiological (EEG) recording cap, each participant had to fill-in four questionnaires: 1. LEAP-Q (Marian et al., 2007) adapted version to explore the linguistic experience; a summary of the results can be consulted in Table 1; 2. The Bem sex roles inventory (Bem, 1981) which allowed us to explore participants' gender self-perception by

Table 2
Experimental conditions.

Category	Prime type	Grammatical gender	
		Masculine	Feminine
Animate (N = 84)	Neutral (N = 42)	e.g.: <i>Autores</i> [Authors] (N = 21)	e.g.: <i>Pintoras</i> [Painters] (N = 21)
	Stereotypical (N = 42)	e.g.: <i>Mineros</i> [Miners] (N = 21)	e.g.: <i>Bailarinas</i> [Dancers] (N = 21)
Inanimate (N = 84)	Neutral (N = 42)	e.g.: <i>Libros</i> [Books] (N = 21)	e.g.: <i>Hojas</i> [Leaves] (N = 21)
	Stereotypical (N = 42)	e.g.: <i>Tractores</i> [Tractors] (N = 21)	e.g.: <i>Faldas</i> [Skirts] (N = 21)

N = 168

Table 1
Socio-linguistic characteristic of the participants.

Group	Biological sex	BEM.f	Age	Benevolent sexism	Hostile sexism	Inclusive language use	BEM
bili	FEMALE	androgine (13)	23.9 (3.0)	3.6 (1.1)	2.6 (1.2)	5.7 (1.6)	5.2 (0.2)
bili	FEMALE	feminine (4)	25.8 (4.7)	3.9 (1.0)	2.8 (1.3)	4.2 (1.4)	4.2 (0.3)
bili	FEMALE	masculine (7)	23.6 (2.4)	3.0 (0.7)	2.8 (1.9)	4.6 (1.7)	5.9 (0.4)
bili	MALE	androgine (7)	24.5 (3.4)	3.7 (0.6)	3.6 (1.9)	5.1 (2.1)	5.2 (0.3)
bili	MALE	feminine (1)	22.0 (0)	2.9 (0)	2.9 (0)	3.3 (0)	4.5 (0)
bili	MALE	masculine (10)	27.0 (7.5)	4.1 (1.1)	3.7 (2.4)	4.3 (2)	5.9 (0.3)
mono	FEMALE	androgine (15)	19.4 (1.4)	3.1 (0.9)	1.8 (1.0)	5.5 (1.7)	5.0 (0.3)
mono	FEMALE	feminine (1)	18.0 (0)	1.9 (0)	1.9 (0)	3.9 (0)	4.0 (0)
mono	FEMALE	masculine (7)	19.7 (2.2)	3.5 (1.5)	2.9 (1.5)	4.9 (1.6)	6.0 (0.2)
mono	MALE	androgine (10)	24.0 (3.1)	2.9 (0.6)	2.1 (0.9)	5.3 (2.1)	5.1 (0.2)
mono	MALE	feminine (1)	23.0 (0)	3.8 (0)	2.7 (0)	3.7 (0)	4.5 (0)
mono	MALE	masculine (11)	20.6 (2.3)	3.9 (1.2)	3.3 (1.6)	4.5 (1.8)	6.0 (0.3)

Table 3
Nouns characteristics and t-tests.

ANIMATE	GENDER		STEREOTYPE	
	FEMININE	MASCULINE	NEUTER	STEREOTYPE
freq. Log spanish	2.87	3.19	2.91	3.15
	t (83) = -1.859; <i>p</i> = 0.06		t(86) = -1.42; <i>p</i> = 0.16	
freq. Log italian	3.12	3.4	3.014	0.536
	t (92) = -1.404; <i>p</i> = 0.06		t(93) = -2.81; <i>p</i> = 0.006	
length spanish	10.16	9.7	9.94	9.92
	t (95) = 1.229; <i>p</i> = 0.22		t(97) = -0.05; <i>p</i> = 0.95	
length italian	9.86	8.94	9.28	9.52
	t (97) = 2.599; <i>p</i> = 0.01		t(97) = -0.66; <i>p</i> = 0.51	
stereotypical score	0.92	1.01	0.25	1.69
	t (95) = -0.512; <i>p</i> = 0.57		t(55) = -17.326; <i>p</i> = 0.00001	
Age of Acquisition	8.87	9.23	9.75	8.34
	t (93) = -0.633; <i>p</i> = 0.53		t(93) = 2.54; <i>p</i> = 0.01	
Concreteness	4.2	4.23	4.11	4.32
	t (85) = -0.34; <i>p</i> = 0.73		t(84) = -2.25; <i>p</i> = 0.02	
Valence	285.87	299.43	303.76	282.28
	t (78) = -0.22; <i>p</i> = 0.82		t(78) = -0.352; <i>p</i> = 0.72	
Arousal	218.24	304.56	291.52	232.09
	t (78) = -1.669; <i>p</i> = 0.09		t(78) = 1.142; <i>p</i> = 0.26	
INANIMATE	GENDER		STEREOTYPE	
	FEMININE	MASCULINE	NEUTER	STEREOTYPE
freq. Log spanish	3.58	3.56	3.66	3.48
	t (91) = 0.082; <i>p</i> = 0.93		t(92) = 1.323; <i>p</i> = 0.19	
freq. Log italian	3.95	3.98	4.01	3.92
	t (93) = -0.282; <i>p</i> = 0.78		t(87) = 0.59; <i>p</i> = 0.55	
length spanish	7.5	7.68	7.56	7.62
	t (92) = -0.48; <i>p</i> = 0.63		t(96) = -0.161; <i>p</i> = 0.87	
length italian	6.8	7.08	7.24	6.64
	t (97) = -0.765; <i>p</i> = 0.44		t(97) = 1.658; <i>p</i> = 0.10	
stereotypical score	0.85	0.69	0.198	1.351
	t (95) = 1.10; <i>p</i> = 0.27		t(55) = -14.126; <i>p</i> = 0.00001	
Age of Acquisition	5.6	6.46	5.82	6.24
	t (98) = -2.07; <i>p</i> = 0.04		t(96) = -0.994; <i>p</i> = 0.32	
Concreteness	4.2	4.23	4.11	4.32
	t (85) = -0.34; <i>p</i> = 0.73		t(84) = -2.25; <i>p</i> = 0.02	
Valence	421.8	326.54	371.97	376.47
	t (90) = 1.742; <i>p</i> = 0.08		t(89) = -0.08; <i>p</i> = 0.93	
Arousal	330.06	226.72	225.02	334.13
	t (89) = 2.24; <i>p</i> = 0.027		t(89) = -2.37; <i>p</i> = 0.02	

Note: Nouns mean on top and t-test below. The comparisons were made by gender (feminine vs. masculine) and by stereotype (stereotypical vs. neutral).

differentiating them in four gender profiles (feminine, masculine, androgynous, indifferent); 3. The Sexist language questionnaire (adapted from Jiménez Rodrigo et al., 2011) to explore the participants' use of inclusive language; 4. The Ambivalent sexist inventory (Cárdenas et al., 2010), to assess participants' sexist behavior.

Spanish native speakers performed three tasks: verbal fluency in Spanish, gender-priming, and picture-word interference. Italian-Spanish bilinguals also performed three tasks: verbal fluency in Italian, gender-priming, and a translation task. All tasks were programmed in E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA, USA). The results from the verbal fluency task and picture-word interference are presented in Casado et al., (in prep.).

The session always started with a verbal fluency task in the native language of the participants (Spanish or Italian). This task was included specifically to promote participants' native language activation before the subsequent tasks. Participants were given 30 s to produce the maximum number of words belonging to a specific phonological category (words beginning with "a" or "m") or semantic category ("animals" or "fruits"). The four categories were counterbalanced across participants.

The gender-priming task was divided into eight blocks, counterbalanced across participants. Each block was composed of 42 counterbalanced trials, 21 containing definitional gender nouns (e.g. queen, girl, where gender is part of the definition of the word) and 21 stereotypical nouns. A total number of 336 trials were included in the task.

Animate and inanimate stimuli were presented in separate blocks. Each prime condition (e.g., *mineros*_{MAS} – miners) was followed by either a masculine or a feminine target demonstrative pronoun in its plural form (e.g., *éstos*_{MAS}/ *éstas*_{FEM} – these) and was presented twice across the task, once followed by the masculine and once by the feminine pronoun or vice versa, always in different blocks. This step was particularly important to make sure that each noun was presented within congruent (same grammatical gender between the prime and the target) and incongruent conditions (different grammatical gender) (see Fig. 1). Moreover, animate and inanimate stimuli were presented in separate blocks to prevent the semantic and conceptual gender information inherently present in animate stimuli from influencing or biasing the processing of inanimate stimuli. Additionally, separating them into distinct blocks allowed us to better isolate and control for differences in lexical variables between categories, ensuring a cleaner assessment of the electrophysiological and behavioral responses within each category.

The structure of the task was designed following Pesciarelli et al. (2019) and Casado, Sá-Leite, Pesciarelli and Paolieri (2023). A fixation cross (500 ms) preceded a blank screen (500 ms), after that the prime stimulus appeared for 50 ms, followed by another blank screen (100 ms) and the target stimulus (1500 ms). Trials were separated from each other by a blank screen of 1000 ms. Participants were asked to respond exclusively to the target stimulus by assigning the grammatical gender of the pronoun as quickly and accurately as possible. When a masculine pronoun appeared, they were instructed to press the "M" key, whereas when it was feminine, they had to press the "F" key. Before the experimental phase, participants completed a practice phase with 12 trials, during which they received automatic feedback on their accuracy and reaction times.

Italian-Spanish bilinguals ended the experimental session with a translation task in which participants had to write the translation equivalent in Italian of the nouns presented in the gender-priming task.

2.4. Electrophysiological acquisition and pre-processing

The EEG signal was acquired during the experimental phase of the gender-priming task. The signal was amplified by a Scan NuAmps system (Compumedics Ltd., VIC, Australia) and recorded through 36 Ag/AgCl electrodes placed on the participants' scalp. Two additional electrodes were placed on the left and right mastoids. Impedances of electrodes were kept below 20 kΩ and the sampling rate was set to 1000 Hz.

EEG data were preprocessed through MATLAB 2022b and the toolboxes EEGLAB 2023 and ERPLAB 2023. The EEG signal was downsampled to 512 Hz and re-referenced to the average signal. An offline band-pass filter (0.1–30 Hz) was applied, and noisy channels were manually removed. Epochs (–250–1200 ms) containing the target conditions were automatically selected and filtered for excessive noise and eye blinking through Independent Component Analysis (ICA). A 200 ms baseline was selected before the target onset.

We selected the time windows for the N400 component (from 250 to 400 ms after target onset) and for the P300 component (400–550 ms after target onset). This timing is especially expected when the N400 is followed by a later positivity/P300 component, as in the present paradigm and it was based on previous studies using closely related priming paradigms, including Siyanova et al. (2012), Pesciarelli et al. (2019), and more recent work by Serafini and Pesciarelli (2025, 2026). We selected nine representative electrodes across frontal (F3, FZ, F4), central (C3, CZ, C4), and posterior (P3, PZ, P4) regions for analysis. This selection ensures consistency with previous paradigms studying N400 and P300 scalp distributions (Pesciarelli et al., 2019; Serafini & Pesciarelli, 2025, 2026; Siyanova et al., 2012). Furthermore, limiting our analysis to these specific sites enables a targeted and comparable test of our hypotheses, thereby preventing Type I error inflation. See Fig. 2.

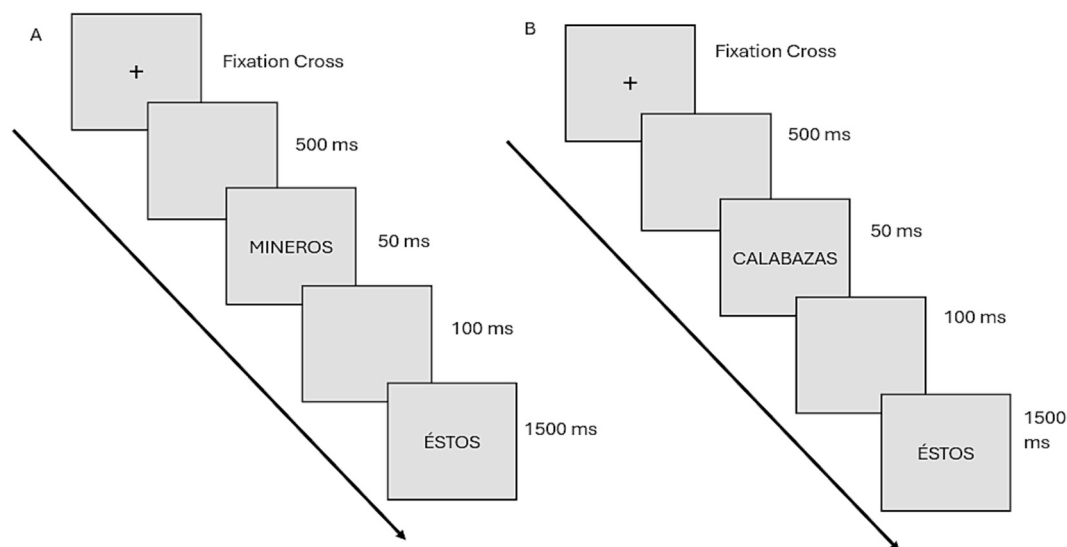


Fig. 1. Schematic depiction of the gender priming task. A: Matched trial with a masculine stereotypical animate prime followed by a masculine target. B: Mismatched trial with a feminine definitional inanimate prime followed by a masculine target.

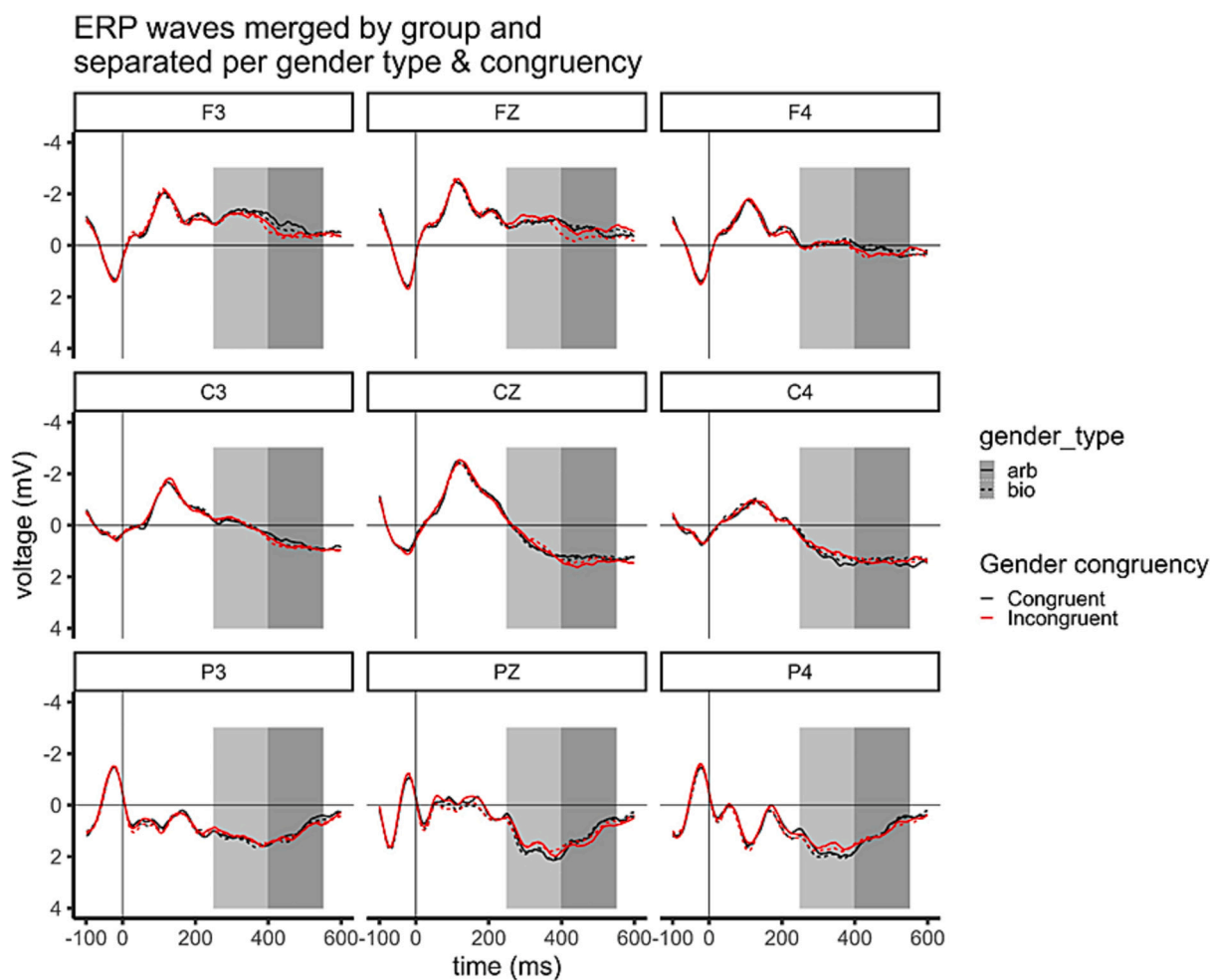


Fig. 2. ERP waves merged by group and separated per gender type & congruency. The color of the line indicates the congruency (red = incongruent; black congruent). The line type represents the gender type (continuous = arb; dotted line = bio). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5. Statistical analysis

The plan for the statistical analyses can be found in the OSF pre-registration (<https://osf.io/pnueq>).

The statistical analyses were performed in R. For the behavioral data we exclusively selected correct trials, discarding 5.74% of the previous data. The remaining data were filtered for outliers, excluding reaction times (RTs) higher than 800 ms, and higher or lower than 2.5 standard deviations which led to discarding an additional 6.24% of the data. The statistical analyses for electrophysiological data were conducted exclusively on trials included in the RT analyses.

Data were analyzed with linear mixed-effects models using the lme4 library and treating RTs and ERP voltages as the dependent variables (DV). Given that the distribution of the DV was positively skewed in the RTs (ms), we transformed the variable to square root to reach a normal distribution.

In the **main models** as fixed effects we included group (Spanish natives –monolinguals–, Italian natives –bilinguals–), congruency (between target and prime gender –congruent, incongruent), gender type (biological, arbitrary), stereotype (stereotypical, neutral), and the interaction between them.

To further explore the role of the masculine as the generic gender we created **additional models** exclusively performed for the biological gender nouns separated by female and male participants. As fixed effect we included group (Spanish natives, bilinguals), target gender (feminine, masculine), prime gender (feminine, masculine), stereotype (stereotypical, neutral), and the interaction between them.

As **control factors** we included the trial number, the block order, the length of the prime, and the lexical frequency of the prime; in the ERP voltages models we also included the horizontality (anterior, central, posterior), laterality (right, central, left), and the reaction times (log transformed). All categorical variables were coded with sum contrast, and all the continuous variables were transformed (if necessary) to reach normal distribution and centered.

The maximal model included participant id and prime id as random intercepts, with congruency, gender type, and stereotype as random slopes in the participant id intercept and group in the prime id intercept. The voltage models also included the electrodes as random intercept.

The maximal model was fitted first and was gradually simplified until non-convergence and singularities were not encountered following Bates et al. (2021) recommendations. The maximal and reduced models were then compared through Akaike Criterion (AIC) and the one with the lowest AIC was selected (Barr et al., 2013; Bates et al., 2021):

Main model : $DV \sim \text{group}^* \text{congruency}^* \text{gender_type}^* \text{stereotype} + (\text{horizontal} + \text{lateral} + \text{rt.}) + \text{trial number} + \text{block} + \text{length} + \text{freq} + (1 + \text{congruency} + \text{stereotype} + \text{gender_type} | \text{ParticipantID}) + (1 + \text{group} | \text{PrimeID}) + (1 | \text{electrode})$

Additional model : $DV \sim \text{group}^* \text{prime_gender}^* \text{target_gender}^* \text{stereotype} + (\text{horizontal} + \text{lateral} + \text{rt}) + \text{trial number} + \text{block} + \text{length} + \text{freq} + (1 + \text{prime_gender} + \text{target_gender} + \text{stereotype} | \text{ParticipantID}) + (1 + \text{group} | \text{PrimeID}) + (1 | \text{electrode})$

All the analysis information can be found in OSF (<https://osf.io/v3xzd>).

3. Results

3.1. Main models

3.1.1. Response time

The model revealed a main effect of congruency, showing that congruent nouns (475 ms) were processed faster than incongruent nouns (506 ms). The congruency effect was modulated by group, and by gender type. Interestingly, there was a three-way interaction between congruency, group, and gender type. The results showed that although both groups experience the gender-congruency effect, as reflected by faster RT for congruent vs. incongruent conditions, for bilingual speakers the difference between congruent and incongruent conditions was larger for arbitrary primes (effect size = -0.344) than for biological primes (effect size = -0.215). No differences were found for Spanish native speakers (effect size arbitrary primes = -0.413 ; effect size biological primes = -0.428) (see Fig. 3 and Table 4).

Crucially, there was also an interaction between congruency and stereotype, showing that the difference between congruent and incongruent conditions was smaller for neutral words (effect size = -0.320) than for stereotypical (effect size = -0.384), see Fig. 4.

3.1.2. N400

The model showed a main effect of gender type, such that arbitrary primes evoked more negative amplitudes (0.116 uV) compared to biological primes (0.176 uV).

There was an interaction between congruency and gender type, showing that the congruency effect was significant for arbitrary primes (con = 0.1368 uV; inc = 0.0959 uV), but not for biological primes (con = 0.1746 uV; inc = 0.1768 uV).

The three-way interaction between group, congruency, and stereotype was significant. The results showed that when the prime is stereotypical, bilinguals triggered more negative amplitudes in the incongruent condition (0.0827 uV) than in the congruent condition (0.1334 uV); however, no significant difference was found for bilinguals when the prime was neutral (congruent = 0.0819 uV; incongruent = 0.0821 uV). No differences were observed for Spanish native speakers (see Fig. 5 and Table 5).

3.1.3. P300

The model revealed a main effect of congruency, with more positive amplitudes evoked in the incongruent (0.269 uV) than in the congruent condition (0.233 uV). Additionally, there was a main effect of gender type, with more positive amplitudes evoked for biological (0.276 uV) compared to arbitrary primes (0.226 uV) (see Fig. 6 and Table 6).

There was a three-way interaction between group, gender type, and congruency. The results showed that bilinguals triggered more positive amplitudes in the incongruent (0.269 uV) than in the congruent (0.196 uV) condition for arbitrary primes, whereas this effect was not signifi-

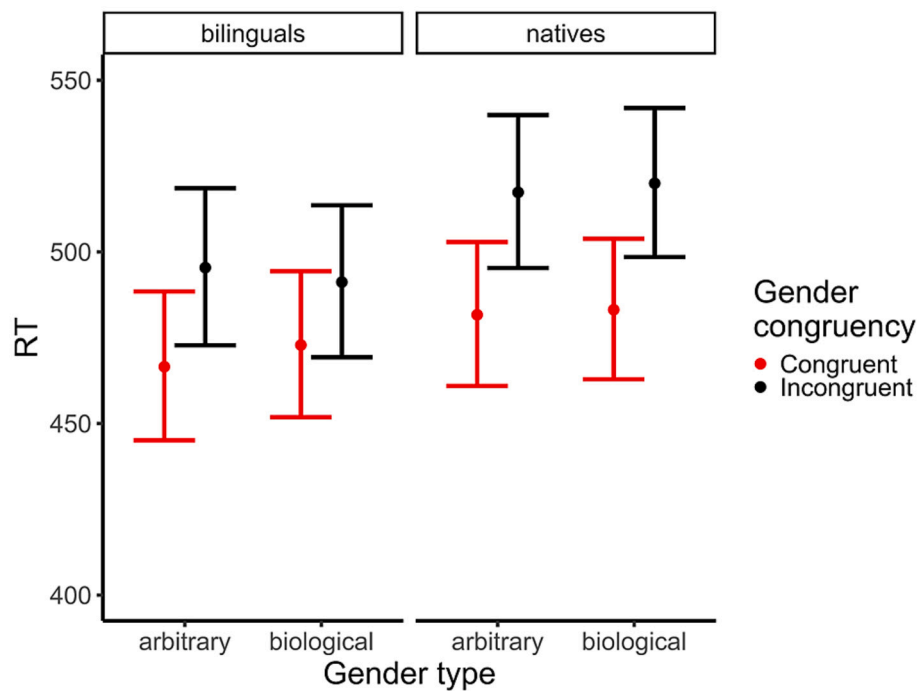


Fig. 3. RT analysis. Interaction between group, gender type, and congruence.

Table 4

Fixed and random effects from the LME of RT.

RT	Estimate	t.value	p value	by-participant	By-item
(Intercept)	22.98	92.11	0.00	1.53	0.11
group1	−0.21	−1.24	0.22		
gender_type1	−0.02	−1.00	0.32	0.12	
congruency1	−0.34	−11.78	0.00	0.23	
stereotype1	0.02	1.51	0.13		
block	−0.04	−6.40	0.00		
trial.s	0.07	5.45	0.00		
freq.s	−0.04	−2.29	0.02		
length.sq	−0.23	−3.77	0.00		
group1:gender_type1	0.01	0.42	0.68		
group1:congruency1	0.07	2.37	0.02		
gender_type1:congruency1	−0.03	−2.19	0.03		
group1:stereotype1	−0.01	−0.68	0.50		
gender_type1:stereotype1	0.00	−0.02	0.99		
congruency1:stereotype1	0.03	2.42	0.02		
group1:gender_type1:congruency1	−0.04	−2.77	0.01		
group1:gender_type1:stereotype1	0.01	0.71	0.48		
group1:congruency1:stereotype1	−0.01	−0.93	0.35		
gender_type1:congruency1:stereotype1	0.02	1.41	0.16		
group1:gender_type1:congruency1:stereotype1	−0.02	−1.26	0.21		

cant for biological primes (con = 0.251 uV; inc = 0.276 uV). Spanish native speakers did not show P300 modulation relative to congruency and gender type interaction, although amplitudes seem larger for

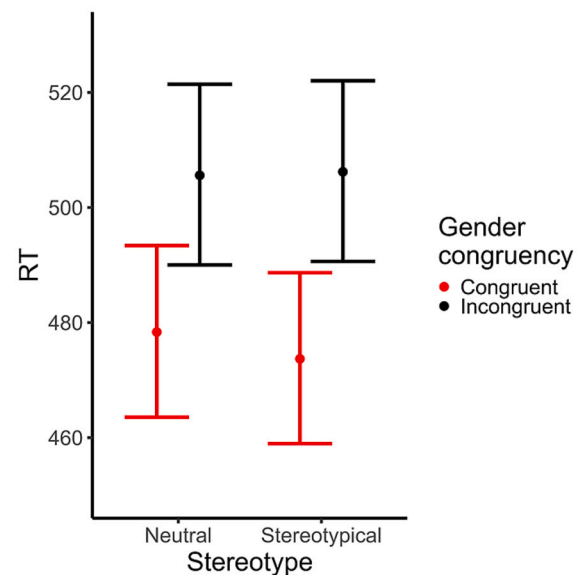


Fig. 4. RT analysis. Interaction between stereotype and congruence.

biological than for arbitrary primes.

3.2. Female participants – only biological nouns¹

3.2.1. Rt

For the female participants, we observed a congruency effect, such that the masculine target was processed faster when preceded by a masculine prime (488 ms) than when preceded by a feminine prime (520 ms). Similarly, the feminine target was processed faster when preceded by a feminine prime (497 ms) than when preceded by a

¹ The model outputs are presented in Appendix B

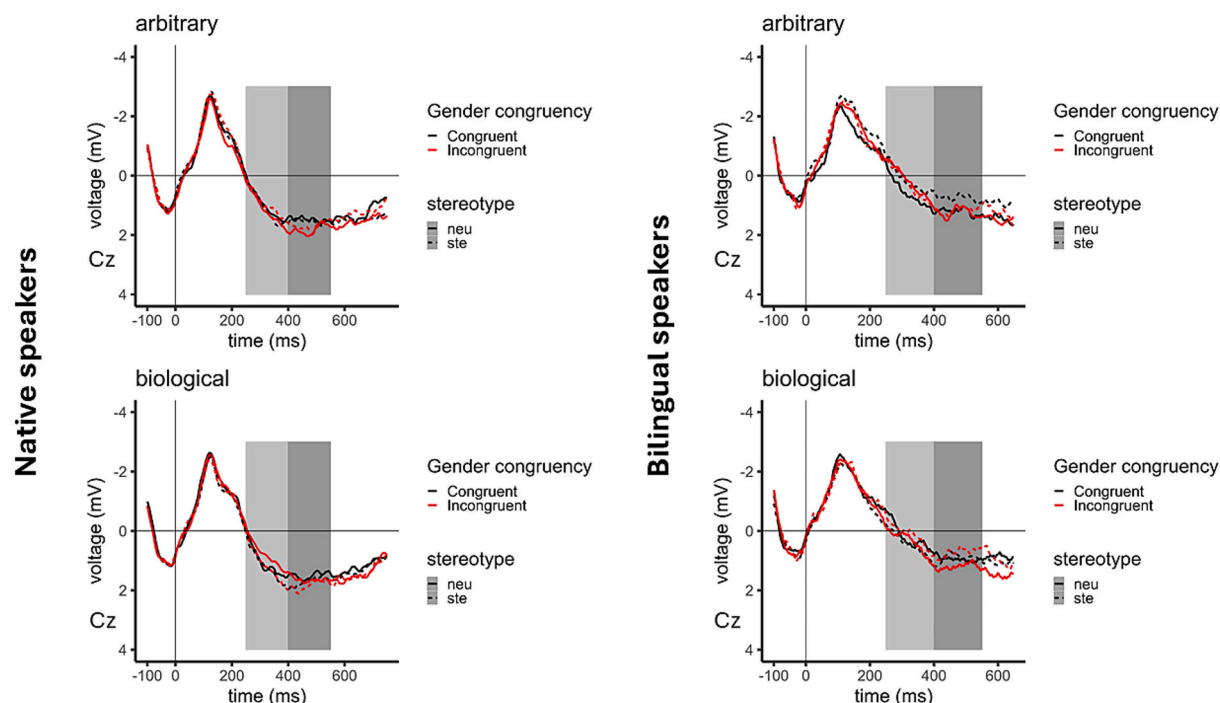


Fig. 5. ERP waves separated by group and gender type; the color represents the gender congruency (red = incongruent; black = congruent); the line type represents the stereotype (continuous = neutral; dotted line = stereotypical). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5

Fixed and random effects from the LME of N400 TW.

N400	Estimate	t.value	p value	by-participant	By-item	by-channel
(Intercept)	0.251	2.442	0.015	0.29	0.06	0.07
group1	-0.051	-1.534	0.129			
gender_type1	-0.03	-4.007	0			
congruency1	0.01	1.862	0.063			
stereotype1	-0.005	-0.79	0.431			
horizontal1	0.042	1.179	0.304			
horizontal2	-0.241	-6.811	0.002			
lateral1	-0.03	-0.845	0.446			
lateral2	-0.146	-4.126	0.015			
rt.sq	-0.005	-1.795	0.073			
freq.s	0.011	1.529	0.128			
length.sq	0	0	1			
block.s	0.007	1.413	0.158			
trial.s	-0.01	-2.006	0.045			
group1:gender_type1	0.004	0.742	0.458			
group1:congruency1	0.003	0.581	0.561			
gender_type1:congruency1	0.011	2.112	0.035			
group1:stereotype1	-0.008	-1.488	0.137			
gender_type1:stereotype1	-0.002	-0.245	0.807			
congruency1:stereotype1	-0.002	-0.477	0.633			
group1:gender_type1:congruency1	0.004	0.726	0.468			
group1:gender_type1:stereotype1	-0.002	-0.306	0.759			
group1:congruency1:stereotype1	-0.01	-2.026	0.043			
gender_type1:congruency1:stereotype1	0	-0.091	0.927			
group1:gender_type1:congruency1:stereotype1	0	-0.062	0.951			

masculine prime (529 ms). Interestingly, this interaction is modulated by group, showing that the effect size was bigger in Spanish natives compared to bilinguals, especially for the masculine target (see Fig. 6).¹

3.2.2. N400

In the N400 we can observed that the congruency effect is modulated by the stereotype; in particular, female participants evoked more negative amplitudes when the masculine target was preceded by a feminine prime with an associated stereotype (0.122 uV) compared to

when the masculine target was preceded by a masculine prime with an associated stereotype (0.225 uV). This congruency effect did not appear with the feminine target (fem prime ste = 0.238 uV; masculine prime ste = 0.225 uV), nor when the prime was neutral, for both masculine targets (fe prime neu = 0.219 uV; mas prime neu = 0.145 uV) or feminine targets (fe prime neu = 0.147 uV; mas prime neu = 0.143 uV). Interestingly, the interaction between stereotype and congruency was modulated by group. We observed that it was the bilingual speakers who showed modulations of the congruency effect by target gender

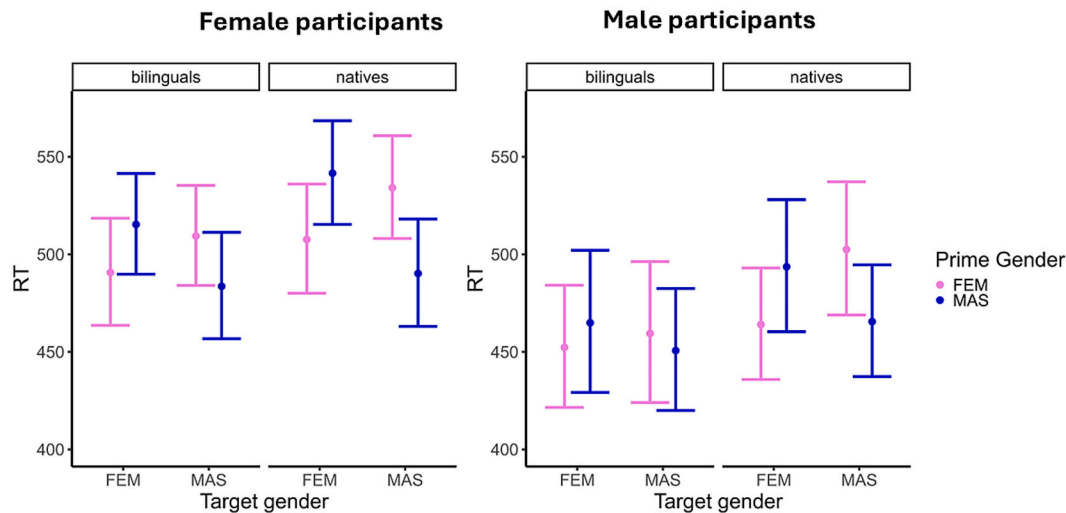


Fig. 6. RT analysis for male and female participants. Interaction between group, target gender, and prime gender.

Table 6
Fixed and random effects from the LME of P300 TW.

P300	Estimate	t.value	p value	by-participant	By-item	by-channel
(Intercept)	0.297	2.813	0.005	0.32	0.06	0.06
group1	−0.003	−0.082	0.935			
gender_type1	−0.025	−3.221	0.002			
congruency1	−0.018	−3.399	0.001			
stereotype1	0	0.005	0.996			
horizontal1	0.196	7.039	0.002			
horizontal2	−0.184	−6.602	0.003			
lateral1	−0.036	−1.307	0.261			
lateral2	−0.053	−1.89	0.132			
rt.sq	−0.002	−0.861	0.389			
freq.s	0.002	0.244	0.807			
length.sq	0.002	0.074	0.941			
block.s	−0.008	−1.49	0.136			
trial.s	−0.008	−1.506	0.132			
group1:gender_type1	0.009	1.715	0.086			
group1:congruency1	−0.006	−1.218	0.223			
gender_type1:congruency1	−0.002	−0.366	0.714			
group1:stereotype1	0.007	1.311	0.19			
gender_type1:stereotype1	0.014	1.851	0.066			
congruency1:stereotype1	0.004	0.773	0.439			
group1:gender_type1:congruency1	−0.01	−1.919	0.055			
group1:gender_type1:stereotype1	0.005	0.921	0.357			
group1:congruency1:stereotype1	−0.004	−0.748	0.455			
gender_type1:congruency1:stereotype1	−0.006	−1.173	0.241			
group1:gender_type1:congruency1:stereotype1	0.002	0.321	0.748			

depending on the stereotype. With the feminine target, the amplitudes triggered were similar for the feminine and the masculine prime, independently on the stereotype (fe prime neu = 0.059 uV; mas prime neu = 0.065 uV; fe prime ste = 0.176 uV; mas prime ste = 0.113 uV). However, with the masculine target, the amplitudes triggered by the masculine target were more negative for the feminine prime when stereotypical (0.003 uV) than for the masculine prime stereotypical (0.124 uV). Crucially, the congruency effect changed polarity with the masculine target when the prime was neutral: masculine targets evoked more negative amplitudes (−0.039 uV) than feminine primes (0.169 uV). Native speakers did not show any significant modulation of the congruency effect.

3.2.3. P300

There were no significant effects related to the target gender or prime

gender.

3.3. Male participants – only biological nouns²

3.3.1. Rt

For the male participants we observed a congruency effect: the masculine target was processed faster when preceded by a masculine prime (458 ms) than when preceded by a feminine prime (480 ms). Similarly, the feminine target was processed faster when preceded by a feminine prime (458 ms) than when preceded by a masculine prime (480 ms). Interestingly, this interaction was modulated by group, with the congruency effect reaching significance in Spanish natives but not in bilinguals (see Fig. 6).²

² Model outputs are presented in Appendix B

3.3.2. N400

In the N400 analysis we observed an interaction between congruency and stereotype, such that the feminine target elicited similar N400 amplitudes for both feminine and masculine primes, independently of stereotype (fe neu = 0.2308 uV; mas neu = 0.2129 uV; fe ste = 0.1904 uV; mas ste = 0.2195 uV). Nonetheless, the masculine target appeared to follow the opposite effect when preceded by neutral information: amplitudes were more negative when preceded by a masculine prime with an associated stereotype (0.0959 uV), compared to a feminine prime with an associated stereotype (0.1916 uV), although this difference was not statistically significant ($z = 1.734$; $p = 0.0829$).

3.3.3. P300

The analysis of the P300 component showed that the congruency effect was modulated by group. In particular, Spanish natives showed a congruency effect, with more positive amplitudes when the feminine target was preceded by feminine primes (0.204 uV) compared to masculine primes (0.292 uV), and when the masculine target was preceded by masculine primes (0.201 uV) compared to feminine primes (0.290 uV), although these comparisons were not significant (feminine target: $z = -1.672$; $p = 0.09$; masculine target: $z = 1.700$; $p = 0.0892$). Bilinguals did not show differences between conditions.

4. Discussion

In the present study, we investigated the automatic activation of grammatical gender in natives and bilingual speakers of Romance languages, specifically Italian and Spanish. Moreover, we explored whether the semantic information related to gender, such as biological sex or associated gender stereotypes, could modulate gender activation. To this aim, we designed a priming task, in which participants were shown a grammatically gendered prime noun for 50 ms, followed by a personal pronoun (*these*) that could be either feminine (*éstas*) or masculine (*éstos*). The participants had to decide the gender of the pronoun. Following previous studies (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), we expected to observe a gender-congruency effect (GCE), which we believe originates at the lexical level (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023). Moreover, we expected that biological gender nouns would elicit stronger GCE than arbitrary gender nouns, in line with the results of our previous study using epicene nouns (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023); at the same time, we expected nouns with an associated gender stereotype to evoke stronger GCE than neutral nouns, at least for arbitrary nouns (Casado et al., 2021). Furthermore, if the masculine works as the generic in Spanish, we should not expect to observe GCE when processing feminine targets, since the masculine prime should activate both masculine and feminine representations; therefore, the upcoming feminine pronoun should not be surprising (as found by Pesciarelli et al., 2019 in Italian native speakers). Additionally, we expected bilingual participants to experience a stronger GCE than natives, due to L1 gender coactivation (Paolieri et al., 2020).

4.1. Gender-congruency effect (GCE)

In line with our initial hypothesis, the GCE occurred for biological gender nouns, replicating previous findings (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012). Crucially, for the first time we demonstrated that the GCE also occurred for arbitrary gender nouns referencing inanimate entities. In both cases, we found a main effect of congruency in RTs, indicating that target nouns are processed more easily when congruent with the prime than when incongruent.

This main effect was also reflected at the electrophysiological level in the P300 component. In particular, we observed that the incongruent condition elicited larger P300 amplitudes compared to the congruent condition, reflecting the surprise of encountering gender mismatching

information for both biological and arbitrary primes. The P300 component appears to be sensitive to the detection of grammatical gender incongruency conflict when using a priming task, in contrast to what was observed by Von Grebmer zu Wolfsturn et al. (2022) in naming tasks with similar populations.

Essentially, this evidence is crucial in supporting our proposal that grammatical gender is active at the lexical level and that its activation is automatic, even when it does not carry semantic information or is not required for agreement (Casado et al., 2018, 2021; Casado, Ferré & Paolieri, 2023).

4.2. GCE modulation by prime biological information

Secondly, we predicted that due to the activation of semantic information related to sex, the GCE will be stronger for nouns with a direct correspondence between biological sex and gender than for nouns with arbitrary gender assignment. Contrary to our prediction, none of the analyses showed this tendency. We observed the opposite pattern in the RT, N400, and P300 models. Specifically, RTs demonstrated that bilinguals experienced a greater GCE when the prime was arbitrary, which was corroborated by the P300 amplitudes. Regarding the N400 amplitudes, we observed that, in general, arbitrary primes elicited more negative amplitudes than biological primes. Considering that the N400 component usually reflects the effort involved in integrating or processing lexical-semantic information, this suggests that arbitrary primes are generally harder to integrate with the target word. Indeed, within the N400 time window, we observed that the GCE was significant only when the prime was arbitrary. The first intuition is to attribute the observed difference to the animacy of the referents; while arbitrary primes refer to inanimate entities, biological primes refer to animate ones. However, based on previous studies, animacy does not appear to play a role in integrating semantic information into noun-pronoun processing during the N400 time window; rather, its effects emerge later, such as in the P600 time window (e.g., for sentence violation Szwedczyk & Schriefers, 2011; for classifier-noun violation Zhang et al., 2012). Therefore, other aspects should be taken into consideration. Despite our best efforts to create a fully matched list of stimuli, it was impossible to find the same number of words for all our conditions (gender type, stereotype, grammatical gender) while maintaining transparent gender endings and matching lexical frequency and length. Consequently, we decided to present the animate and inanimate primes in different blocks, and to account for lexical frequency and length differences statistically. Nevertheless, differences in lexical frequency may have modulated our data; in line with classical studies (e.g., Van Petten & Kutas, 1990), higher-frequency primes evoke more negative N400 amplitudes, especially when incongruent with the target. This pattern was replicated in the P300 time window, such that bilinguals showed more positive amplitudes when arbitrary primes appeared in the incongruent condition compared to the congruent condition. In order to check this possibility, we performed further analyses (presented in the supplementary analyses section of OSF file), including a new model including the interaction between lexical frequency and gender congruency. The results indicate that the congruency effect was modulated by frequency, such as the higher the frequency, the higher the facilitation for gender-congruent word pairs. Thus, the results may be explained by the higher lexical frequency of arbitrary primes compared to biological primes. Another explanation could be related to the differences in the length of the primes between biological and arbitrary primes. Given that the prime only appeared for 50 ms, it may be the case that bilingual participants did not have enough time to process completely the longer words in their L2 (the biological primes). This may explain the fact that the GCE on bilinguals only appeared for arbitrary primes, as reflected by RT and P300. However, for native speakers, in the P300 time window we observed more positive amplitudes when biological primes appeared in the incongruent condition compared to the congruent condition, although this effect was not statistically significant

and was in line with our prediction. It may be the case that native speakers, due to their greater experience with their native language, overcome lexical frequency differences by giving more weight to the semantic information carried by the prime (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023). Still, despite these indications, the design imbalance prevents us from confirming our second hypothesis that the GCE would be greater for biological than for arbitrary primes.

4.3. GCE modulation by prime stereotype information

Similarly, we expected the GCE to be stronger for stereotypical nouns than for neutral nouns. We observed this pattern in the RT analyses, in line with our previous study (Casado et al., 2021): RTs in the congruent condition were faster when the prime was stereotypical compared to when it was neutral. Moreover, the effect was marked at the electrophysiological level in the N400 component. In particular, we observed that bilingual speakers experienced a stronger GCE (as shown by more negative amplitudes in the incongruent vs. congruent condition) with stereotypical primes compared to neutral primes, whereas no such differences appeared for native speakers. This result is quite surprising, given that, based on recent studies (Castro et al., 2025; Jankowiak et al., 2025), one may expect bilinguals in their L2 to be less sensitive to stereotypical information compared to native speakers. These findings demonstrate that native and L2 speakers utilize distinct cue-weighting strategies during comprehension. In native speakers, deeply entrenched and highly automated lexico-syntactic representations take precedence; during early processing windows (e.g., a 50 ms prime), this automated structural processing is activated prior to the integration of higher-level information, rendering L1 processing temporarily shielded from top-down, extralinguistic influences like social stereotypes. Conversely, L2 speakers possess weaker, less automated grammatical representations. Under the inherently higher cognitive load of non-native processing (McDonald, 2006), the L2 system becomes more permeable to top-down semantic influences (Clahsen & Felser, 2006; Clahsen & Felser, 2018; Grüter et al., 2019). In our study, this semantic reliance was likely amplified by the linguistic and cultural overlap shared by our Italian-Spanish bilinguals. This dual-source activation made social expectations exceptionally salient during lexical retrieval, serving as a compensatory semantic crutch to bootstrap grammatical gender processing. Consequently, the stereotype effect emerged exclusively in L2 speakers, demonstrating that real-time processing in a second language is significantly more integrated with, and influenced by, semantic and social world knowledge than in a native language. Still, one possibility is that bilinguals find it easier to recognize words that activate stereotypical information, similar in their L1, and base their congruency effect on purely semantic information. To further explore possible explanations, in the next section (4.4 GCE modulation by masculine as the generic in biological gendered nouns) we examine how stereotypical information is processed by female and male speakers separately when the prime is biological.

4.4. GCE modulation by masculine as the generic in biological gendered nouns

Following previous results in Italian (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), we predicted an absence of GCE for feminine targets in biological nouns, due to the grammatical rule whereby the masculine form includes all gender representations and functions as the default or generic gender. We found differences according to participants' gender, due to the automaticity and strength of the lexico-syntactic retrieval process.

First, male participants showed the congruency effect in RTs, with no differences between masculine and feminine targets. At the electrophysiological level, males exhibited comparable N400 amplitudes for feminine targets regardless of the gender of the prime. However, the congruency effect for masculine targets showed only a non-significant

tendency with neutral primes, and the opposite pattern when the prime was stereotypical. Overall, the N400 amplitudes were not highly informative about the semantic activation of the masculine as generic versus gender marked. In the P300 time window, we observed an interaction between group and congruency, indicating that native speakers were sensitive to the GCE with both the masculine and the feminine targets, independent of stereotype (although not statistically significant). Altogether, these results suggest that native Spanish males do not seem to activate the generic representation when processing the masculine form, in line with our previous findings (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023).

Second, female participants showed the GCE in the RT analyses for targets, although it was greater for the masculine target than for the feminine target, indicating an asymmetry previously reported in Italian (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012). Interestingly, this asymmetry was also reflected in the N400 component and was dependent on the stereotype. Specifically, female participants showed similar N400 amplitudes for feminine targets regardless of the gender of the prime. However, females elicited larger N400 amplitudes when the masculine target was preceded by a feminine prime versus a masculine prime, but only when the prime was stereotypical. This difference was significant only for bilinguals, who showed the GCE pattern exclusively with masculine targets when the prime was stereotypical and reversed the pattern with neutral primes. Our findings suggest that for female participants, the retrieval process is more permeable to top-down information, such as social stereotypes, whereas for males, the response appears more uniform and less sensitive to these variables. This is consistent with evidence that females consistently exhibit a more pronounced sensitivity toward grammatical and stereotypical violations, attributed to a general superiority in language-related skills, such as grammar and verbal fluency (Canal et al., 2015; Osterhout et al., 1997; Pesciarelli et al., 2019).

Altogether, the results revealed that, contrary to prescriptive grammatical rules and previous findings in Italian (e.g., Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), neither group of female and male Spanish natives in the current study process the masculine plural as a truly inclusive or neutral category; instead, both treated it as a marked masculine form. The fact that a reduced GCE was observed in Italian native speakers (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012) but not in Spanish native speakers (Casado, Sá-Leite, Pesciarelli & Paolieri, 2023, current study) may indicate a difference in language use that could be cultural. In the literature, there are studies recommending the use of alternative inclusive forms in Italian (Baiooco et al., 2023; Sulis & Gheno, 2022), while in Spanish, studies are already exploring the impact of such alternative forms in society (e.g., Bonnin & Zunino, 2024; Diaz, Mejía & Villamizar, 2022; Stetie & Zunino, 2022, 2024). These findings suggest that Spanish speakers are further ahead than Italian speakers in adopting alternative forms of generic/inclusive language, probably due to the greater diversity of Spanish usage across many different countries and cultures (e.g., Latin-American vs. Europe). Thus, we advocate for continuing to explore the use of alternative forms that encompass the full spectrum of gender diversity (Anaya-Ramírez et al., 2022; Bonnin & Zunino, 2024; Diaz, Mejía & Villamizar, 2022; Stetie & Zunino, 2022, 2024).

4.5. GCE modulation by L1 gender coactivation in bilinguals

Finally, we predicted that bilingual participants would experience a stronger GCE compared with natives due to the gender co-activation in their L1 when using L2. We did find an interaction between group and congruency in the RT analysis, but in the opposite direction: the GCE was stronger for native speakers than for bilinguals. In particular, the effect size of the difference between congruent and incongruent conditions was higher for natives compared with bilinguals, a pattern that seems to be driven mostly by the incongruent condition. This may suggest deeper processing in natives than in bilinguals, such that natives

activate a greater amount of semantic information from the noun prime, making it harder to inhibit when accessing the target in the incongruent compared to the congruent condition. Still, this interpretation is speculative, as this effect was not reflected in the ERP components: N400 and P300 responses to gender incongruity were not stronger for natives compared to bilinguals. Further research is needed to disentangle the process underlying this native advantage.

4.6. Summary

In summary, the current study showed that the GCE has a lexical nature that can be enhanced by the activation of additional semantic information related to gender. Furthermore, we showed that, in contrast to Italian, the masculine form does not function as a generic in Spanish. Finally, we observed differences in the GCE experienced by natives and bilinguals: although gender coactivation occurs in bilinguals, native speakers appear to engage in deeper semantic processing, and consequently, the expected enhancement of the GCE in bilinguals was not observed. Ultimately, this study offers a comprehensive model of gender processing that accounts for the complex interaction between formal grammar and social associations.

5. Supplementary – control variables

5.1. Main model

5.1.1. RT

Additionally, there was a main effect of block, showing that participants became faster with each block. There was also a main effect of trial, such that within each block, participants became slower toward the end. Moreover, there was a main effect of the length of the prime, such that the longer the words, the slower the RT. Finally, there was a main effect of frequency of the prime, such that the higher the lexical frequency, the faster the RT.

5.1.2. N400

Additionally, there was a main effect of horizontal topography, and lateral topography, such as more negative amplitudes were registered in frontal and left sides. The trial number was also significant, showing that as the trial number rises, the voltage decreases.

5.1.3. P300

Additionally, there was a main effect of horizontal topography, such as more positive amplitudes were registered in central sites.

CRediT authorship contribution statement

Alba Casado: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Formal analysis, Data curation, Conceptualization. **Giuseppe Dario Benigno:** Writing – review & editing, Writing – original draft, Software, Investigation, Formal analysis, Data curation. **Francesca Pesciarelli:** Writing – review & editing, Methodology. **Daniela Paolieri:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research was possible thanks to the grant PID2022-142345NB-I00 from the Spanish Ministry of Science and Universities (MICIU/

AEI/10.13039/501100011033) and FEDER to D.P., grant CEX2023-001312-M/ funded by MCIN/AEI/10.13039/501100011033 and Grant UCEPP2023-11 funded by University of Granada. We gratefully acknowledge the help of all members of the Memory and Language Laboratory, who contributed to the research project by supporting the data collection and providing helpful feedback.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bandl.2026.105842>.

Data availability

Data is available at OSF: <https://osf.io/v3xzd/overview>

References

- Allan, K. (2007). The pragmatics of connotation. *Journal of Pragmatics*, 39(6), 1047–1057.
- Amenta, S., Badan, L., & Brysbaert, M. (2021). LexITA: A quick and reliable assessment tool for Italian L2 receptive vocabulary size. *Applied Linguistics*, 42(2), 292–314.
- Anaya-Ramírez, A., Grinstead, J., Rivera, M. N., 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.
- Baiocco, R., Rosati, F., & Pistella, J. (2023). Italian proposal for non-binary and inclusive language: The schwa as a non-gender-specific ending. *Journal of Gay & Lesbian Mental Health*, 27(3), 248–253.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, 17(1), 137–153.
- 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), 255–278.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2021). *Parsimonious mixed models*. *arXiv preprint arXiv:1506.04967*.
- Beatty-Martínez, A. L., & Dussias, P. E. (2019). Revisiting masculine and feminine grammatical gender in Spanish: Linguistic, psycholinguistic, and neurolinguistic evidence. *Frontiers in Psychology*, 10, 751.
- Beller, S., Brattebø, K. F., Lavi, K. O., Reigstad, R. D., & Bender, A. (2015). Culture or language: What drives effects of grammatical gender? *Cognitive Linguistics*, 26(2), 331–359.
- Bem, S. L. (1981). *Bem Sex Role Inventory: Professional manual*. Palo Alto, CA: Consulting Psychologists Press.
- Bender, A., Beller, S., & Klauer, K. C. (2016). Crossing grammar and biology for gender categorisations: Investigating the gender congruency effect in generic nouns for animates. *Journal of Cognitive Psychology*, 28(5), 530–558.
- Bender, A., Beller, S., & Klauer, K. C. (2018). Gender congruency from a neutral point of view: The roles of gender classes and conceptual connotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(10), 1580.
- Bonnin, J. E., & Zunino, G. M. (2024). Using inclusive language at school: Reported and perceived use among teachers in Buenos Aires. In *Inclusiveness beyond the (non) binary in romance languages* (pp. 9–32). Routledge.
- Bordag, D. (2004). Interaction of L1 and L2 systems at the level of grammatical encoding. *EUROSLA Yearbook*, 4.
- Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. *Language in mind: Advances in the study of language and thought*, 22(61–79), 1.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46, 904–911.
- Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics*, 28(2), 277–293.
- Canal, P., Garnham, A., & Oakhill, J. (2015). Beyond gender stereotypes in language comprehension: Self sex-role descriptions affect the brain's potentials associated with agreement processing. *Frontiers in Psychology*, 6, 1953.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14(1), 177–208.
- Cárdenas, M., Lay, S. L., González, C., Calderón, C., & Alegría, I. (2010). Inventario de sexismo ambivalente: adaptación, validación y relación con variables psicosociales. *Salud & Sociedad*, 1(2), 125–135.
- 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 Section A*, 49(3), 639–663.
- Casado, A., Ferré, P., & Paolieri, D. (2023). Gender congruency between languages influence second-language comprehension: Behavioral and electrophysiological evidence. *Journal of Neurolinguistics*, 68, Article 101156.
- Casado, A., Palma, A., & Paolieri, D. (2018). The influence of sex information on gender word processing. *Journal of Psycholinguistic Research*, 47(3), 557–583.
- Casado, A., Palma, A., & Paolieri, D. (2021). The scope of grammatical gender in Spanish: Transference to the conceptual level. *Acta Psychologica*, 218, Article 103361.

- Casado, A., Sá-Leite, A. R., Pesciarelli, F., & Paolieri, D. (2023). Exploring the nature of the gender-congruency effect: Implicit gender activation and social bias. *Frontiers in Psychology*, 14, 1160836.
- Castro, S., Kalamala, P., Bukowski, M., & Wodniecka, Z. (2025). Relationship between bilingual experiences and social biases: The moderating role of motivation to respond without prejudice. *Bilingualism: Language and Cognition*, 28(2), 535–549.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42.
- Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, 40(4), 693–706.
- Corbett, G. G. (1991). *Gender*. Cambridge University Press.
- Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, 53(1), 42–59.
- Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2012). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica*, 33(2), 133–143.
- Díaz, A. R., Mejía, G., & Villamizar, A. G. (2022). Gender neutral and non-binary language practices in the Spanish language classroom: Tensions between disciplinary and societal changes. *L2 Journal: An Open Access Refereed Journal for World Language Educators*, 14(3).
- Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, 28(5), 467–484.
- Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, 30(6), 627–651.
- Grüter, T., Lau, E., & Wenyi Ling, W. (2019). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, 34(9), 1067–1082.
- Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, 45, 765–771.
- Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22(1), 27–62.
- Huang, S., & Schiller, N. O. (2021). Classifiers in mandarin Chinese: Behavioral and electrophysiological evidence regarding their representation and processing. *Brain and Language*, 214, Article 104889.
- Irizarry, A. R., Beatty-Martínez, A. L., Torres, J., & Kroll, J. F. (2025). “Todes” and “Todxs”, linguistic innovations or grammatical gender violations? *Cognition*, 257, Article 106061.
- Irmen, L., Holt, D. V., & Weisbrod, M. (2010). Effects of role typicality on processing person information in German: Evidence from an ERP study. *Brain Research*, 1353, 133–144. <https://doi.org/10.1016/j.brainres.2010.07.018>
- Jankowiak, K., Naranowicz, M., Pawelczyk, J., Dąrkowski, D., & Gruszecka, J. (2025). Bilingual speakers are less sensitive to gender stereotypes in their foreign language. *Bilingualism: Language and Cognition*, 28(2), 457–468.
- Jiménez Rodrigo, M. L., Román Onsalo, M. L., & Traverso Cortes, J. (2011). Lenguaje no sexista y barreras a su utilización. Un estudio en el ámbito universitario. *Revista de investigación en educación*, 2(9), 174–183.
- Klassen, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus*, 28(1), 9–28.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44, 978–990.
- Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564–582.
- Levitt, W. J. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223–232.
- Luraghi, S. (2011). The origin of the proto-Indo-European gender system: Typological considerations. *Folia linguistica*, 45(2), 435–464.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (Aug 2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940–967. D American Speech-Language-Hearing Association 1092-4388/07/5004-0940.
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late L2 learners. *Journal of Memory and Language*, 55(3), 381–401.
- Morales, L., Paolieri, D., & Bajo, T. (2011). Grammatical gender inhibition in bilinguals. *Frontiers in Psychology*, 2, 284.
- Nicoladis, E., & Foursha-Stevenson, C. (2012). Language and culture effects on gender classification of objects. *Journal of Cross-Cultural Psychology*, 43(7), 1095–1109.
- Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition*, 25, 273–285.
- Paolieri, D., Demestre, J., Guasch, M., Bajo, T., & Ferré, P. (2020). The gender congruency effect in Catalan–Spanish bilinguals: Behavioral and electrophysiological evidence. *Bilingualism: Language and Cognition*, 23(5), 1045–1055.
- Paolieri, D., Lotto, L., Morales, L., Bajo, T., Cubelli, R., & Job, R. (2010). Grammatical gender processing in romance languages: Evidence from bare noun production in Italian and Spanish. *European Journal of Cognitive Psychology*, 22(3), 335–347.
- Paolieri, D., Padilla, F., Koreneva, O., Morales, L., & Macizo, P. (2019). Gender congruency effects in Russian–Spanish and Italian–Spanish bilinguals: The role of language proximity and concreteness of words. *Bilingualism: Language and Cognition*, 22(1), 112–129.
- Pesciarelli, F., Scorolli, C., & Cacciari, C. (2019). Neural correlates of the implicit processing of grammatical and stereotypical gender violations: A masked and unmasked priming study. *Biological Psychology*, 146, Article 107714.
- Prado, M. (1982). El género en español y la teoría de la marcaje. *Hispania*, 258–266.
- Sá-Leite, A. R., Flores, C., Eira, C., Haro, J., & Comesaña, M. (2023). Language balance rather than age of acquisition: A study on the cross-linguistic gender congruency effect in Portuguese–German bilinguals. *Bilingualism: Language and Cognition*, 26(5), 1079–1092.
- Sá-Leite, A. R., & Lago, S. (2024). The role of word form in gender processing during lexical access: A theoretical review and novel proposal in language comprehension. *Psychonomic Bulletin & Review*, 31(5), 1934–1953.
- Sá-Leite, A. R., Luna, K., Tomaz, A., Fraga, I., & Comesaña, M. (2022). The mechanisms underlying grammatical gender selection in language production: A meta-analysis of the gender congruency effect. *Cognition*, 224(105060), 10–1016.
- Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review*, 26(6), 1767–1786.
- Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, 48(1), 169–194.
- Segel, E., & Boroditsky, L. (2011). Grammar in art. *Frontiers in Psychology*, 1, Article 244.
- Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General*, 131(3), 377.
- Serafini, L., & Pesciarelli, F. (2025). The influence of gender stereotypical primes on the neural processing of words and faces. *Social Cognitive and Affective Neuroscience*, 20(1), Article nsaf031.
- Serafini, L., & Pesciarelli, F. (2026). Neural gender stereotype asymmetry in bidirectional word-face priming. *Biological Psychology*, 205, Article 109240.
- Siyanova-Chanturia, A., Pesciarelli, F., & Cacciari, C. (Mar 2012). The electrophysiological underpinnings of processing gender stereotypes in language. *PLoS One*, 7(12), Article e48712.
- Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., et al. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 49, 111–123. <https://doi.org/10.3758/s13428-015-0700-2>
- Stetie, N. A., & Zunino, G. M. (2022). Non-binary language in Spanish? Comprehension of non-binary morphological forms: A psycholinguistic study. *Glossa: A Journal of General Linguistics*, 7(1). <https://doi.org/10.16995/glossa.6144>
- Stetie, N. A., & Zunino, G. M. (2024). Do gender stereotypes bias the processing of morphological innovations? The case of gender-inclusive language in Spanish. *Psychology of Language and Communication*, 28.
- Stolle, S. I., Hanulíková, A., & Schiller, N. O. (2025). Grammatical gender in Slovak word production: An event-related potential study. *Language, Cognition and Neuroscience*, 1–20.
- Sulis, G., & Gheno, V. (2022). The debate on language and gender in Italy, from the visibility of women to inclusive language (1980s–2020s). *The Italianist*, 42(1), 153–183.
- Szewczyk, J. M., & Schriefers, H. (2011). Is animacy special? ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Research*, 1368, 208–221.
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency-related brain potentials. *Memory & Cognition*, 18(4), 380–393.
- Velnić, M. (2020). Acquisition of a transparent gender system: A comparison of Italian and Croatian. *Frontiers in Psychology*, 11, Article 571674.
- Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48(4), 422–488.
- Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General*, 134(4), 501.
- Von Grebmer zu Wolfsturn, S., Gupta, A. I., Pablos-Robles, L., & Schiller, N. O. (2023). When left is right: The role of typological similarity in multilinguals’ inhibitory control performance. *Bilingualism: Language and Cognition*, 26, 165–178.
- Von Grebmer zu Wolfsturn, S., Pablos-Robles, L., & Schiller, N. O. (2022). Does your native language matter? Neural correlates of typological similarity in non-native production. *Lingua e Linguaggio*, 21, 143–169.
- Wang, Y., Witteman, J., & Schiller, N. O. (2025). Activation of classifiers in word production: Insights from lexico-syntactic probability distributions. *Language, Cognition and Neuroscience*, 1–15.
- Zhang, Y., Zhang, J., & Min, B. (2012). Neural dynamics of animacy processing in language comprehension: ERP evidence from the interpretation of classifier–noun combinations. *Brain and Language*, 120(3), 321–331.