Journal of Neurolinguistics 68 (2023) 101156

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Gender congruency between languages influence second-language comprehension: Behavioral and electrophysiological evidence

Alba Casado [a,*], Pilar Ferré [b], Daniela Paolieri [a]

[a] *Mind, Brain, and Behavior Research Center (CIMCYC), Department of Experimental Psychology. University of Granada, C/ Profesor Clavera, s/n, 18011, Granada, Spain*

[b] *Department of Psychology and Research Center for Behavior Assessment, University Rovira i Virgili, Carretera de Valls, s/n, 43007, Tarragona, Spain*

ARTICLE INFO

Keywords:
Gender congruency
Bilinguals
Lexical access
ERPs
N400

ABSTRACT

In the present study we explore whether gender congruency between languages modulates bilinguals' access to their second language words presented in isolation. We predicted that accessing L2 words that have a different gender across languages (gender-incongruent) would be more costly and require more effort than accessing same-gender words (gender-congruent) due to language co-activation, even when no access to L1 was required to perform the task. Additionally, we intended to shed some light into the mechanism underlying the gender congruency effect. To these aims, we compared the performance of Spanish native speakers with that of Italian-Spanish bilinguals (Italian native speakers) during a lexical decision task. The participants saw Spanish words that were gender-congruent and gender-incongruent between languages while event related potentials were recorded. Moreover, as an additional manipulation, we selected nouns that in both languages could be ambiguous or unambiguous. With the aim to examine whether the underlying mechanism is activation of multiple information during word processing, we focused on the N400 component, related with the effort to integrate lexical-semantic information: higher N400 amplitudes indicate greater effort. According to our prediction, Italian-Spanish bilinguals produced more errors and evoked larger N400 amplitudes when accessing gender-incongruent than gender-congruent words, while no differences appeared for Spanish native speakers between conditions. These results indicate that gender-incongruent words are harder to integrate compared with gender-congruent words, and that bilinguals automatically activate the grammatical gender of both languages during L2 language comprehension. Nevertheless, the results do not seem to support the assumption of a similar mechanism responsible for the gender congruency and the ambiguity effects. In short, the gender-congruency effect seems to originate due to activation of multiple information at the lexical level which generates difficulties to integrate at the semantic level during word access.

* Corresponding author.
E-mail addresses: albacasado@ugr.es (A. Casado), mariadelpilar.ferre@urv.cat (P. Ferré), dpaolieri@ugr.es (D. Paolieri).

https://doi.org/10.1016/j.jneuroling.2023.101156
Received 16 November 2022; Received in revised form 16 June 2023; Accepted 30 June 2023
Available online 19 July 2023
0911-6044/

1. Introduction

1.1. Grammatical gender processing in bilinguals

Bilingual speakers are able to use at least two languages, both of which are activated simultaneously regardless of the speaker's intention to use just one of them. Accordingly, the lexical access in bilinguals is a non-selective process (for reviews see, for example, Bialystok et al., 2012; Kroll & Bialystok, 2013; Kroll et al., 2014) and the two languages constantly interact. The interaction between languages has been observed during reading, speaking, and listening (Dijkstra, 2005; Hoshino & Thierry, 2011; Kroll et al., 2005; Marian & Spivey, 2003), such that the linguistic properties of the non-intended language affect the processing of the intended language at the semantic, lexical, and phonological levels (see, for example, Blumenfeld & Marian, 2007; Bobb et al., 2020; Bordag & Pechmann, 2007, 2008; Colomé, 2001; Costa et al., 2000; Guasch et al., 2017; Hoshino & Thierry, 2011; Iniesta et al., 2021; Ju & Luce, 2004; Lemhöfer et al., 2008; Macizo et al., 2010; Weber & Cutler, 2004).

Regarding the influence of the non-intended language over the processing of the intended language at the lexical level, we can find some evidence coming from the processing of the grammatical gender. Based on the literature we can assume that grammatical gender is an intrinsic lexical property instead of merely a syntactic feature exclusively processed in noun phrase contexts (Alario et al., 2008; Bates et al., 1995; Casado et al., 2018; Cole & Segui, 1994; Cubelli et al., 2005; De Martino et al., 2011; Palma Duràn & Pillon, 2011; Paolieri et al., 2010; Paolieri et al., 2011; see also Wang et al., 2018, for a similar lexico-syntactic classifier feature; and Sá-Leite et al., 2019 for a review). Therefore, observing the effects of grammatical gender during bilingual language processing can give us information about language co-activation at the lexical level.

In this line, previous research already showed that the effects of grammatical gender can be detected during bilingual speech production (Bordag & Pechmann, 2007, 2008; Klassen, 2016; Lemhöfer et al., 2008; Morales et al., 2011; Paolieri et al., 2010, 2018; Salamoura & Williams, 2007; Zu Wolfsthurn et al., 2021) and comprehension (Halberstadt et al., 2018; Lemhöfer et al., 2008; Morales et al., 2016; Paolieri et al., 2020). The co-activation of grammatical gender in both languages induces what is called the *gender congruency effect*, i.e., a facilitation when processing and producing words that match in gender between languages (e.g., the word "bottle" has feminine grammatical gender in both Spanish and Italian— *botella/bottiglia*) compared to words for which the gender assignment is different (e.g., the word "candy" has different grammatical gender assignment in Spanish (masculine, *caramelo*) and Italian (feminine, *caramella*); for a review, see Sá-Leite et al., 2019). Previous research generally suggests that the gender systems of the native or stronger language (L1) and the second or weaker language (L2) interact in the bilingual mental lexicon such that L1 gender is activated when L2 is used. Therefore, when the nouns of both languages have the same grammatical gender assignment (e.g., feminine in L1 and L2), the grammatical gender receives activation from the two languages which turns into a greater level of activation of this representation at the lexical level, which ultimately facilitates lexical selection, especially in L2 (Cubelli & Paolieri, 2008). Moreover, the lexical representations in L1 and L2 are connected at the semantic level, as they share the conceptual representation (Duyck & Brysbaert, 2004; Kroll & De Groot, 2009; Kroll & Stewart, 1994; Sunderman & Kroll, 2006). Given that L1 and L2 representations share information at the lexico-semantic level, those words that have the same grammatical gender assignment in both languages would be more similar between them. What is more, they will be more similar compared with words that have different grammatical gender in L1 and L2; in particular, some authors claim that congruent gender words (same grammatical gender between languages) share more semantic features than incongruent gender words (see, for example, Athanasopoulos & Kasai, 2008; Boutonnet et al., 2012; Konishi, 1993; Saalbach et al., 2012; for a review, see Samuel et al., 2019). In this line, the "thinking-for-speaking hypothesis" (Slobin, 1997, pp. 437–467) posits that people are trained by their native language to think according to certain patterns that are aligned to grammatical and lexical structures. For instance, for Italian-Spanish bilinguals the representation of "bottle" (*bottiglia/botella*, which is feminine in both languages) should be thought of as having more lexico-semantic features in common than the representation of "spoon", which is designated by two nouns with different genders (the Italian noun, *cucchiaio*, is masculine, whereas the Spanish noun, *cuchara*, is feminine), although grammatical gender assignment is arbitrary. To explore this idea, Paolieri et al. (2018) compared bilingual speakers of languages with similar gender systems (Italian-Spanish) and bilingual speakers of dissimilar language systems (Russian-Spanish) who conducted a translation task in which they were asked to translate words with either same or different grammatical gender between the languages. The authors hypothesized that words with the same grammatical gender between languages would be easier to produce than words with different grammatical genders. They also predicted that the gender congruency advantage would be stronger for concrete nouns than for abstract nouns. That would be explained because abstract nouns have fewer shared semantic features than concrete nouns (see, for example, De Groot, 1989; Plaut & Shallice, 1993); therefore, fewer semantic elements are shared between abstract nouns and their translation equivalents than with concrete nouns (Van Hell & De Groot, 1998). Paolieri et al. (2018) showed that the gender congruency effect appeared both in pairs of languages with similar gender systems and pairs of languages with dissimilar gender systems. Moreover, the gender congruency effect was greater when the speakers produced concrete nouns than when they produced abstract nouns, which indicates that there is a close relationship between semantics and grammatical information in bilingual language production (Paolieri et al., 2018).

All in all, in the reviewed literature we found quite some evidence showing that L1 gender is activated automatically during L2 use. Moreover, the shared grammatical gender assignment between languages facilitates the access to lexico-semantic representations. While these findings have been corroborated at the behavioral level, little is known about how the automatic activation of gender is reflected in more fine-grained measures of cognitive processing, such as event-related brain potentials (ERPs). In this study we aim to find the neural correlate of bilingual gender co-activation.

1.2. ERP markers

ERPs can index differences in the level of meaning activation (Barber et al., 2013). Specifically, the N400 component is defined as a negative-going potential around 400 ms that is thought to reflect mainly semantic processing (for a review, see Kutas & Federmeier, 2011). It is known to be a good index of the ease with which information is accessed within long-term semantic memory and the ease with which it is integrated with the local context (Kutas & Federmeier, 2000). Previous research has shown that words with many semantic features are associated with larger N400 amplitudes (Amsel, 2011; Haro et al., 2017; Rabovsky et al., 2012). For example, Haro et al. (2017) conducted a lexical decision task in Spanish in which participants had to decide whether the string of letters they were presented corresponded to a real word or to a pseudoword. They selected ambiguous nouns (with more than one meaning, e.g., *sirena* – mythological creature/alarm) and unambiguous nouns (with only one meaning, e.g., *piano*). The ambiguous nouns were also divided into homonyms and polysemous words. The results showed that the ambiguous words elicited larger N400 amplitudes than the unambiguous words regardless of the kind of ambiguity (polysemy or homonymy). Similarly, using a lexical decision task, Rabovsky et al. (2012) found that words with many semantic features elicited larger N400 amplitudes than those with fewer semantic features. The results of these experiments suggest that accessing words that activate a large amount of semantic information requires greater effort (Holocomb et al., 1999; West & Holcomb, 2000), which is captured by modulations in the N400 component.

In agreement with this interpretation, Barber et al. (2013) found that the lexical status of the stimuli during lexical access modulates the N400 component. In particular, they found that concrete nouns elicited larger N400 amplitudes than abstract nouns in a lexical decision task. They posited that, unlike abstract nouns, concrete nouns activate a greater amount of sensory-motor information, which requires a greater effort to be integrated within a single lexico-semantic system. The greater effort needed to integrate the more semantic features is reflected by greater N400 amplitudes.

Additionally, the N400 component has recently been shown to reflect changes in the amount of activated information at the lexico-semantic level in bilinguals. For instance, Taler and collaborators (2016) designed a lexical decision task to explore the processing of ambiguous and unambiguous words in L1 in L2. Contrary to previous studies, they found lower N400 amplitudes to ambiguous words relative to unambiguous words, especially in L2. In a different line, and in a pioneering study, Paolieri et al. (2020) asked highly proficient early Catalan-Spanish bilinguals to perform a translation-recognition task comprising gender-congruent pairs (*pluja* (FEM)/*lluvia* (FEM)-rain) and gender-incongruent pairs (*genoll* (MAS)/*rodilla* (MAS)-knee). The behavioral results showed that it took longer to translate the gender-incongruent pairs than it did to translate the gender-congruent pairs. Moreover, the processing of gender-incongruent pairs evoked more negative N400 amplitudes than the processing of gender-congruent pairs. This indicates that the automatic activation of gender in both languages when incongruent activates extra information that requires more effort to be integrated at the lexico-semantic level compared to when the grammatical gender is congruent (for similar evidence of a lexico-syntactic classifier analogous to grammatical gender in production, see Wang et al., 2018). Finally, Zappa et al. (2022), using a different task, showed rapid changes in cortical activity, associated with L2 learning. Interestingly, while all L2 words were well learned at a behavioral level, a large post-training N400 effect was observed only for L1-L2 gender congruent nouns. This larger N400 effect was taking as evidence of gender overlap influences during the initial stages of an L2 acquisition. Nevertheless, due to the ecological relevance of the study, all auditory nouns were preceded by the determiner. Therefore, it remains an open question whether the same effect would be observed with the presentation of bare nouns. Actually, some relevant inconsistent findings seem to emerge using tasks with different demands of syntactic activation, such as the processing of bare nouns or noun phrases (i.e., article + noun or adjective + noun) (see Wang & Schiller, 2019; and Sá-Leite, Haro, Comesaña, & Fraga, 2021). In our view, observing a gender effect in bare noun processing is crucial as it constitutes evidence of the obligatory access to grammatical gender information even when it is not explicitly needed for speech processing, supporting the notion that grammatical gender is an intrinsic part of the lexical representation (e.g., Paolieri et al., 2011; and Sá-Leite et al., 2019 for a review).

From the review, therefore, we can assume that the amplitude variations in the N400 component index the effort that is required to integrate all the information that is activated during word processing within the lexico-semantic system, also in bilinguals. The N400 is therefore the ideal component for studying the effects of gender-congruency between languages.

1.3. The present study

In this study we explored whether access to L2 words presented in isolation during comprehension was modulated by gender congruency with their L1 translation equivalents at the behavioral and neural levels. We want to emphasize that as a novelty, we focused on bare noun processing instead of noun phrases, and on word comprehension instead of production. Moreover, we used electrophysiological techniques to explore the underlying mechanisms. To do so, we tested a group of Italian-Spanish bilinguals (Italian native speakers) and a control group of Spanish native speakers. We recorded their electrophysiological response while they conducted a lexical decision task in Spanish (L2 for the Italian native speakers). The participants had to decide whether the items presented were real words or pseudo-words. The real words could be gender-congruent (same grammatical gender in both languages; e.g., *botella/bottiglia* is feminine in both Spanish and Italian), or gender-incongruent (opposite grammatical gender in Spanish and Italian; e.g., the Spanish word for sweet – *caramelo* – is masculine, while the Italian word – *caramella* – is feminine). Moreover, as an additional manipulation we selected real words (nouns) that in both languages could be ambiguous (words with more than one meaning; e.g., *colonia*, which in Spanish and Italian means both perfume and colony) or unambiguous (words with only one meaning; e.g., *casa*, which means house in both Spanish and Italian). In this study, therefore, we manipulated two factors: gender congruency and ambiguity. By comparing the manipulation of gender congruency with that of ambiguity we were able to examine whether the underlying mechanisms of both effects (the gender congruency effect and the ambiguity effect) are the same: activation of multiple information during word processing within the lexico-semantic system. To our knowledge, no studies have compared the effort to integrate the activation of multiple information during word processing using the gender-congruency effect and the lexical ambiguity effect.

We hypothesized that L1 grammatical gender would be activated even when bilinguals process isolated L2 words. Access to gender-incongruent nouns would therefore activate a larger amount of information than access to gender-congruent nouns. This larger amount of information would require greater effort to be integrated at the lexico-semantic level. Accordingly, we predicted worse performance (lower accuracy and/or slower response times) and higher N400 amplitudes when processing the gender-incongruent words compared with the gender-congruent words in the Italian native speakers (Paolieri et al., 2020). We did not expect any differences between the gender-incongruent and gender-congruent words in the Spanish native speakers since they do not have any knowledge of Italian.

Additionally, we hypothesized that the mechanism underlying the gender congruency effect would be like the one underlying the ambiguity effect: the activation of multiple information at the lexico-semantic level that requires more effort to be activated. Accordingly, we predicted greater N400 amplitudes for the ambiguous words than for the unambiguous words (Haro et al., 2017; Rabovsky et al., 2012; for contradictory results with bilinguals see Taler et al., 2016). Moreover, we expected the effects of ambiguity to be stronger for the Spanish native speakers compared with the Italian native speakers, because the latter performed the task in their second language (L2). According to the Sense Model proposed by Finkbeiner et al. (2004), bilingual speakers (especially those unbalanced) may not be aware of all the senses of polysemous words in L2 compared with L1 words, even in the case in which the translation-equivalents share the same number of meanings between languages (Chen et al., 2014; Crossley & Skalicky, 2019).

2. Methods

2.1. Participants

2.1.1. Italian native speakers

Twenty-three Italian native speakers with Spanish as their second language (mean age: 23.3; SD age: 2.11; 16 females) participated in this study. All participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007) to subjectively measure their linguistic history with their mother tongue (Italian) and second language (Spanish). All participants were living in Spain by the time of testing (around 3 years). They were late learners of Spanish (around 14 y.o.) but still they had good proficiency in Spanish (76 out of 100). Besides Italian and Spanish, they also reported using a different language (i.e., English) but only a small amount of the time (less than 10%) (see Appendix A). They also completed a verbal fluency task in both Italian and Spanish and in a mixed condition, each in an independent block, in which they were asked to produce as many examples as possible belonging to different semantic categories (e.g., fruits, animals, etc.) in 1 min (see Appendix A). As inclusion criteria, none of the participants had any kind of hearing impairment, uncorrected visual impairment, or language or neurological impairment. All participants took part in the experiment voluntarily and gave their informed consent in accordance with the protocol of the Ethics Committee of the University of Granada. One participant was removed after EEG preprocessing due to a large number of artifacts produced by movements and eye blinks.

2.1.2. Spanish native speakers

Twenty-six Spanish native speakers (mean age: 22.44; SD age: 3.92; 18 females) participated in this study, following the same inclusion criteria as bilinguals. Six participants were removed after EEG preprocessing due to a large number of artifacts produced by movements and eye blinks.

2.2. Materials

A set of 90 Spanish nouns was selected from a larger pool of 186 (see Appendix B for the complete list of stimuli and their characteristics). The selected nouns referred to the same inanimate entities in both Spanish and Italian and had arbitrary gender. A group of 58 native Spanish speakers and a group of 72 native Italian speakers completed the selection questionnaire (these were different speakers from those who participated in the main experiment). To do so, they were asked to score the ambiguity of the words in their native language by deciding whether each word had more than one meaning (ambiguous words) or only one meaning (unambiguous words). The 90 stimuli selected comprised a total of 30 ambiguous words (those with more than one meaning) and 60 unambiguous words (those with only one meaning). The 30 ambiguous words had the same grammatical gender in Spanish and Italian. Thirty of the 60 unambiguous words had the same grammatical gender in Spanish and Italian (e.g., bottle – *botella/bottiglia* – in Spanish and Italian is feminine), while the other 30 unambiguous words had the opposite grammatical gender in Spanish and Italian (e.g., the Spanish word for candy –*caramel*– is masculine, while the Italian word – *caramella*– is feminine).[1] In each set of 30 words (ambiguous-gender congruent words, unambiguous-gender congruent words, unambiguous-gender incongruent words), one-third had opaque gender

[1] Due to practical limitations, only 60 unambiguous and 30 ambiguous words were selected for the ambiguity condition, and 60 congruent and 30 incongruent words for the gender congruency condition. However, our statistical approach (mixed-effects models with random slopes by participants and items) account for the variability in the data and to model the effects of both the ambiguity condition and the gender congruency condition on our dependent variables. The interactions between the two conditions in the model was excluded to avoid inflating the explained variance with the decompensation.

desinences (e.g., -e, or consonants) and two thirds had transparent gender desinences (i.e., -a for feminine/-o for masculine). In addition, all three sets of words were matched in terms of the Levenshtein distance between Spanish and Italian (i.e., orthographic overlap or cognate status), their number of letters, and their frequency of use in Spanish (EsPal database, Duchon et al., 2013) or Italian (PhonItalia database, Goslin et al., 2014). From the 90 selected words, we used the Wuggy software (Keuleers & Brysbaert, 2010) to create 90 pseudowords to be used as fillers following the Spanish phonological constraints. We made sure that none of the pseudo words had meanings in Italian.

2.3. Procedure

The stimuli were presented on a computer using E-Prime version 2.0 (Schneider et al., 2002). They were presented in bold 20-point Geneva font placed in the center of the screen. The participants were tested individually. They were asked to decide whether the string of letters presented formed a real Spanish word, or a pseudoword. Each trial began with an image of an eye displayed for 2000 ms, which indicated to participants that they were allowed to blink. The image was followed by a fixation point (i.e., "+"), which appeared in the center of the screen for 500 ms. Immediately after this, a string of letters (a word or a pseudoword) replaced the fixation point and the participants had to decide whether the string was a Spanish word or not. They were instructed to press the "yes" labeled key of a keyboard with their right hand if the string of letters was a word and to press the "no" labeled key with their left hand if it was not. The string of letters remained on the screen until the participant responded or until time ran out (after 2000 ms). After responding, a feedback message (i.e., "ERROR" or "CORRECT", and the time in milliseconds) was displayed for 750 ms. The stimuli were grouped in an experimental block of 180 trials (90 words and 90 pseudowords per block). The block was repeated so that each stimulus would be presented twice and more observations could be taken. The experimental session lasted approximately 35 min. The dependent variables derived from the participants' performance were reaction time (RT) from the onset of the stimulus until the beginning of the response, and accuracy (ACC). We also recorded brain responses with EEG.

2.4. Electrophysiological recording

The electrophysiological recording was conducted using a SynAmps2 64-channels Quik-Cap plugged into a Neuroscan SynAmps RT amplifier. The electrical signal was amplified with a 30 Hz low-pass filter and a continuous sample rate of 500 Hz. Impedances were kept below 5 kΩ. The electrodes were referenced to a vertex electrode online (REF) and the grounding electrode (GND) was mounted on the participant's forehead. The offset values were the voltage difference between each electrode and the reference. Vertical and horizontal eye movements were also recorded from bipolar pairs of vertical (VEOG) and lateral (HEOG) electrodes to enable corrections of blink artifacts. Individual averages were re-referenced off-line to the average of all electrodes. Before analyzing the data, artifacts (such as eye movements) were removed using independent component analyses (ICA). Additionally, other artifacts were defined as events in which there was a difference of ±100 μV in amplitude within less than 50 ms, or when the absolute amplitude exceeded ±100 μV. Trials with artifacts (3.12%) were rejected and recordings from electrodes with a high level of artifacts (>1%) were interpolated using the average value of the group of nearest electrodes. Epochs from −200 to 800 ms with respect to the presentation of the target picture were averaged and analyzed. We applied baseline correction using average EEG activity in the 200 ms prior to target onset as reference. The EEG-accepted epochs were averaged for each participant and each electrode across the three experimental conditions (ambiguous-gender congruent, unambiguous-gender congruent, and unambiguous-gender incongruent). Only correct response trials were included in the averages. ERP extraction, averaging and cleaning were conducted with EEGlab (Delorme & Makeig, 2004) and ERPlab (Lopez-Calderon & Luck, 2014) MATLAB software toolboxes. We analyzed the N400 component and, based on Haro et al. (2017), selected the 350–450ms time windows after word-onset on a cluster of central electrodes: FC1, FCZ, FC2, C1, CZ, C2, CP1, CPZ, CP2.

2.5. Statistical analyses

Before conducting the analyses, we filtered the data. Words which had an accuracy rating of 50% or lower were eliminated. In total, three items were removed ("as", "bistec" and "dique".[2] We calculated the mean RT and sd for each participant to filter those RT below or upper 2.5 sd from the mean for the RT analysis and also for the ERP analysis. All analyses (behavioral and ERP) were performed using linear mixed-effects models, as implemented in the lme4 package (version 1.1.21; Bates et al., 2015) in R using participants and items as crossed random effects. The models included as fixed effects Group (bilingual, monolingual), Gender congruency (congruent, incongruent), Ambiguity (ambiguous words, unambiguous words), and Repetition (the first or second time the word was presented, new or repeated). Additionally, we included as covariates some of the words' characteristics: Logarithmic lexical frequency, Normalized Levenshtein distance, Number of letters, Grammatical gender of the target (masculine, feminine) and the Desinence (transparent, opaque). Besides, in the ERP model we included the RT as a predictor.

We transformed the categorical predictors applying a sum contrast (Schad et al., 2020): Group (bilingual = 1, monolingual = −1), Gender congruency (congruent = 1, incongruent = −1), Ambiguity (ambiguous = 1, unambiguous = −1), Repetition (new = 1; repeated = −1), Grammatical gender of the target (feminine = −1, masculine = 1), Desinence (transparent = −1, opaque = 1). We

[2] We believe that participants were not sure that "as" or "bistec" were real Spanish words because they can also be considered loans from English or French. In the case of "dique", this is not a high frequency word, and many speakers may not have identified it as a Spanish real word.

transformed the naming latencies to the logarithmic RT to correct for normality of the distribution. The rest of the continuous variables were scaled. We fitted the maximal model first (Barr et al., 2013) and simplified it for non-convergence or singularities. We considered as significant any factor whose t-statistic was above two. The planned comparisons were analyzed using the emmeans package (version 1.8.1-1; Lenth et al., 2018).

Brysbaert and Stevens (2018) recommended a minimum of 1600 observations in repeated measures designs to have enough power to detect significant effects for RT. In our case we ended up with a total of 85 items and 47 participants (7238 observations for RT).

3. Results

3.1. Behavioral results

3.1.1. Accuracy

The results revealed a main effect of the Group, showing that the Italian native speakers participants produced more errors (M = 0.90) compared with the Spanish native speakers (M = 94). There was also a main effect of Gender congruency modulated by an interaction with Group. The results showed that the Italian native speakers were less accurate in the gender-incongruent condition (M = 0.84) compared with the gender-congruent condition (M = 0.95), z = 3.32, p > .001. However, no differences were found for the Spanish native speakers (z = 1.57, p > .05). See Fig. 1. Additionally, there was an effect of the Logarithmic lexical frequency, such as the higher the lexical frequency, the higher the accuracy. See Table 1.

3.1.2. Response times

The results showed a main effect of Repetition, such that participants took longer to process new words (M = 602 ms) than repeated words (M = 584 ms). There was also a main effect of Ambiguity, showing that the ambiguous words took longer to process (M = 608 ms) than the unambiguous words (M = 584 ms). Additionally, there was an interaction between Group and Ambiguity that was modulated by Repetition. In particular, Spanish native speakers were more sensitive to the ambiguity compared with the Italian native speakers, especially during the first presentation (see Fig. 1). Additionally, there was a main effect of Logarithmic lexical frequency, such as the higher the lexical frequency, the faster the RT. See Table 1.

3.2. ERP results

3.2.1. N400: 350–450 ms time window

In the first place, there was a main effect of the RT, revealing that the effects on the N400 component correlated with the effects of the manipulations on the lexical decision task. In particular, we observed that the slower responses evoked higher N400 amplitudes.

Secondly, the results showed a main effect of Repetition, such as higher N400 amplitudes were evoked by new words (M = −0.50 mV) compared with repeated words (M = −0.10 mV).

Interestingly, there was an interaction between Group and Gender congruency, showing that the Italian native speakers evoked higher N400 amplitudes when processing the gender-incongruent words (M = −0.40 mV) compared with the gender-congruent words (M = -0.08 mV), z = 2.57; p < .01. The Spanish native speakers did not differ (z = −0.25; p < .05). See Fig. 2.

Additionally, there were two-way interactions between Group and Repetition, Group and Ambiguity, and Ambiguity and Repetition that converged into a three-way interaction between the three factors. In particular, we could observe that the ambiguous words evoked higher N400 amplitudes compared with the unambiguous words, but only for the Spanish native speakers during the first presentation (see Fig. 2).[3]

3.2.2. Further analysis N400: 350–450 ms time window visual inspection of the scalp distribution

In order to explore whether the gender-congruency effect had the same underlying mechanism as the ambiguity effect, we compared the scalp distribution of the N400 amplitudes of the two effects. In particular, we subtracted the ERP evoked by the gender congruent words from that of the gender incongruent words in the time window 350–450 ms of the Italian native speakers to calculate a gender_congruency index. Similarly, we subtracted the ERP evoked by the unambiguous words from that of the ambiguous words during the first presentation (new items) in the time window 350–450 ms of the Spanish native speakers to calculate an ambiguity index. Thus, we created a dataset in which we had the voltage for each of the selected electrodes (FC1, FC2, FCZ, C1, C2, CZ, CP1, CP2, CPZ) of the gender-congruency index for each participant in the Italian native group, and the ambiguity index for each participant in the Spanish native group. We created a model in which as the dependent variable we included the voltage of the indices. As fixed effects we included the group (Italian native [gender-congruency effect], Spanish native [ambiguity effect]), the laterality (right, left, central), and the horizontality (frontocentral, centroparietal, central). We selected sum contrasts for all the fixed effects (selecting "central" as baseline for both laterality and horizontality). Following the visual inspection of Fig. 2, we found that more negative voltages were evoked in frontal and left areas. Moreover, the analysis revealed a tendency for an interaction between the indices and the laterality (see Table 2). We performed paired-wise comparisons to doublecheck the interaction, and we found that the Italian natives evoked more negative amplitudes at the left side compared with Spanish native. This means that in contrast to the gender

[3] In order to make sure that there were no baseline differences for each of the conditions we performed further analyses presented in Appendix C.

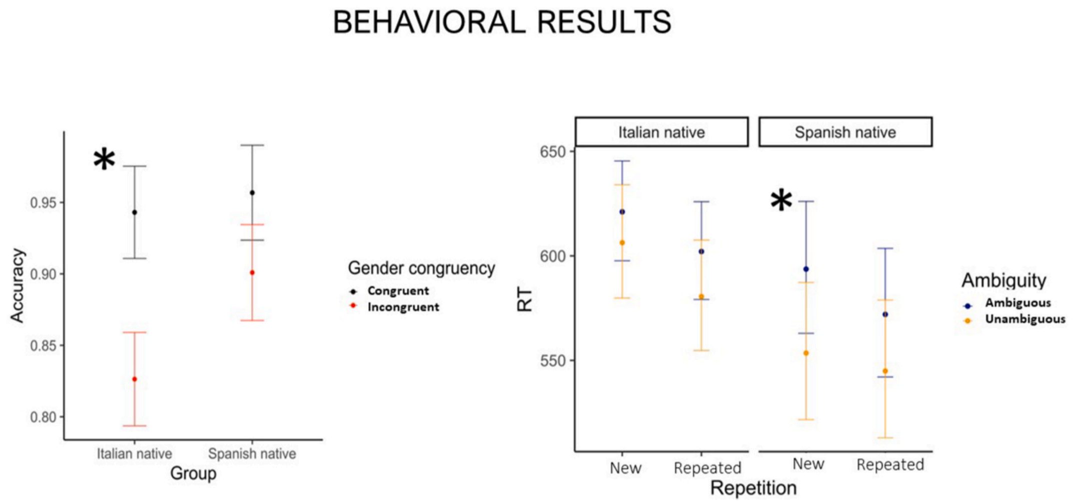

Fig. 1. At the left, model-based predictions for the gender congruency effect; behavioral results of the accuracy score for Italian and Spanish natives. At the right, model-based predictions for the interaction between the ambiguity effect and the repetition; behavioral results of the response time for Italian and Spanish natives.

congruency effect, the ambiguity effect was stronger in the left areas. Thus, despite we have to take into account that the analysis is between groups, it seems that different areas are involved in the two effects, being the ambiguity effect more lateralized at the left side than the gender congruency effect (see Table 3).

4. Discussion

The present study was designed to explore whether access to L2 words presented in isolation during comprehension was modulated by gender congruency with their translation equivalents in L1. We also wished to shed light on the mechanisms behind this modulation. To pursue these aims, a group of Spanish native speakers and a group of Italian native speakers with Spanish as their L2 performed a lexical decision task in Spanish. These speakers evaluated the lexicality of words that could have the same grammatical gender assignment in Italian and Spanish (gender congruent) or the opposite gender between the languages (gender incongruent). All the words were concrete nouns with an arbitrary gender assignment. Additionally, the words could be either ambiguous (i.e., words with more than one meaning) or unambiguous (i.e., words with only one meaning).

We first predicted that the gender-incongruent nouns would activate more information than the gender-congruent nouns at the lexico-semantic level for the Italian native speakers (Paolieri et al., 2020); this would be reflected by worst performance (lower accuracy and/or slower RT) and higher N400 amplitudes. In line with our prediction, the results showed that while there were no differences in the Spanish native speakers between the gender-congruent and gender-incongruent nouns, the Italian native speakers had more difficulties processing gender-incongruent nouns than processing gender-congruent nouns. This was reflected behaviorally as lower accuracy for the gender-incongruent nouns than for the gender-congruent nouns. Moreover, at the electrophysiological level we observed higher amplitudes of the N400 component for gender-incongruent nouns than for gender-congruent nouns. This harder processing of the gender-incongruent words in the second or weaker language is most likely caused by the automatic activation of the grammatical gender in the native language, which is mismatching the automatic gender activation in the second or weaker language (Paolieri et al., 2020).

Altogether, our results indicate that grammatical gender is an intrinsic part of the lexical representation that is automatically activated during lexical access (Cubelli et al., 2005; Cubelli & Paolieri, 2008; Paolieri et al., 2010, 2020). As a reminder, previous studies that observed the gender congruency effect used tasks where the determiner was presented (e.g., *la botella* — the bottle), which reinforces the chances to observe the effect (e.g., La Heij et a., 1998). Importantly, only one previous study showed that the grammatical gender feature is activated automatically in language comprehension when accessing bare nouns, that is, when no agreement context is required (Paolieri et al., 2020). We gather further evidence of this automatic gender activation when exclusively using L2. Furthermore, against the previous design in which the participants had to activate their L1 to perform the task (Catalan-Spanish bilinguals had to decide whether a Spanish word was the correct translation of a Catalan word), in the present study access to the participants' L1 (Italian) was not required, and not even mentioned during the experimental task or instructions. Therefore, we consider that our results constitute solid evidence to support the notion that grammatical gender is part of the lexical representation and its activation is automatic, not only because we used bare nouns (without the determiner) but also because no reference to L1 was present during the task.

The automatic activation of gender information at the lexical level triggers the gender-congruency effect (for a review, see Sá-Leite et al., 2019). That is, a facilitated access of L2 words when the grammatical gender is the same between languages (gender congruent) because the L2 lexical representation receives higher activation compared to when the grammatical gender is the opposite in their L1

Table 1
Fixed and random effects from the LME model of accuracy, response times, and voltage of the N400 component.

ACCURACY							
Fixed Effects	Estimate	SE	t	p	sd part	sd item	effect size
(Intercept)	0.754	0.061	12.341	0.000	0.020	0.117	
group1	−0.019	0.004	−4.785	0.000			−0.194
repeated1	−0.007	0.004	−1.718	0.086		0.025	−0.0743
status1	−0.022	0.016	−1.340	0.180	0.005		−0.224
gen_con1	0.041	0.016	2.477	0.013			0.412
genGRAM1	0.019	0.011	1.701	0.089			0.189
log_frq	0.049	0.019	2.578	0.010			−0.5
NLD	0.056	0.056	1.013	0.311			−0.285
terminacion1	0.003	0.010	0.245	0.807			0.026
num_letters	0.012	0.007	1.601	0.109			−0.419
group1:repeated1	0.002	0.003	0.554	0.579			Bil: 0.0588; Mon: 0.0898
group1:status1	−0.002	0.003	−0.838	0.402			Bil: 0.248; Mon: 0.200
group1:gen_con1	0.015	0.003	5.445	0.000			Bil: 0.559; Mon: 0.265
repeated1:status1	−0.003	0.004	−0.599	0.549			R1: 0.250; R2: 0.198
repeated1:gen_con1	0.003	0.004	0.603	0.547			R1: 0.437; R2: 0.386
group1:repeated1:status1	−0.002	0.003	−0.767	0.443			BilR1: 0.295; BilR2: 0.200; MonR1: 0.20; MonR2: 0.195
group1:repeated1:gen_con1	0.001	0.003	0.263	0.792			BilR1: 0.592; BilR2: 0.526; MonR1: 0.283; MonR2: 0.246
RESPONSE TIME							
Fixed Effects	Estimate	SE	t	p	sd part	sd item	effect size
(Intercept)	6.465	0.042	152.573	0.000	0.122	0.056	
group1	0.033	0.018	1.840	0.066		0.021	0.446
repeated1	0.015	0.004	3.763	0.000	0.023		0.205
status1	0.022	0.008	2.727	0.006			0.299
gen_con1	−0.015	0.008	−1.787	0.074	0.007		−0.201
genGRAM1	−0.008	0.007	−1.214	0.225			−0.113
log_frq	−0.050	0.012	−4.077	0.000			0.667
NLD	−0.054	0.035	−1.528	0.127			0.358
terminacion1	0.005	0.007	0.719	0.472			0.0637
num_letters	0.003	0.005	0.604	0.546			−0.132
group1:repeated1	0.003	0.004	0.845	0.398			Bil: 0.251; Mon: 0.159
group1:status1	−0.007	0.004	−2.013	0.044			Bil: 0.202; Mon: 0.396
group1:gen_con1	−0.002	0.004	−0.557	0.577			Bil: 0.559; Mon: 0.265
repeated1:status1	0.001	0.002	0.533	0.594			R1: 0.314; R2: 0.284
repeated1:gen_con1	0.001	0.002	0.271	0.786			R1: 0.437; R2: 0.386
group1:repeated1:status1	−0.004	0.002	−1.970	0.049			BilR1: 0.161; BilR2: 0.243; MonR1: 0.468; MonR2: 0.324
group1:repeated1:gen_con1	−0.001	0.002	−0.346	0.730			BilR1: 0.231; BilR2: 0.226; MonR1: 0.156; MonR2: 0.191
N400							
Fixed Effects	Estimate	SE	t	p	sd part	sd item	effect size
(Intercept)	1.950	0.865	2.254	0.024	1.09940218586205	0.406295590352828	
group1	0.078	0.177	0.442	0.659			0.0392
repeated1	−0.226	0.044	−5.078	0.000			−0.113
status1	−0.116	0.072	−1.612	0.107			−0.0583
gen_con1	0.038	0.065	0.580	0.562			0.019
RT.log	−0.353	0.122	−2.898	0.004			0.177
genGRAM1	0.044	0.054	0.816	0.414			0.0221
log_frq	−0.081	0.095	−0.855	0.393			0.0407
NLD	−0.067	0.276	−0.242	0.809			0.0167
terminacion1	−0.036	0.052	−0.700	0.484			−0.0183
num_letters	0.020	0.037	0.554	0.580			−0.0359
group1:repeated1	0.173	0.044	3.895	0.000			Bil: 0.0264; Mon: 0.2000
group1:status1	0.094	0.042	2.237	0.025			Bil: 0.011; Mon: 0.106
group1:gen_con1	0.117	0.026	4.591	0.000			Bil: 0.0778; Mon: 0.0398
repeated1:status1	−0.184	0.042	−4.368	0.000			R1: 0.1508; R2: 0.0341
repeated1:gen_con1	0.016	0.025	0.618	0.536			R1: 0.0269; R2: 0.0111
group1:repeated1:status1	0.130	0.042	3.079	0.002			BilR1: 0.0382; BilR2: 0.0162; MonR1: 0.2633; MonR2: 0.0520
group1:repeated1:gen_con1	0.017	0.025	0.689	0.491			BilR1: 0.0944; BilR2: 0.0612; MonR1: 0.0407; MonR2: 0.0389

(gender incongruent). Indeed, we found reduced N400 amplitudes associated with the processing of gender-congruent nouns, which are typically related with easier lexical integration. Moreover, we observed that the results are independent of the transparency of the gender marking. Recall that we selected nouns with transparent and opaque desinences to mark the grammatical gender of the words. When including the desinence in the models, this factor did not predict any of our results (accuracy, RT or N400). Hence, our results

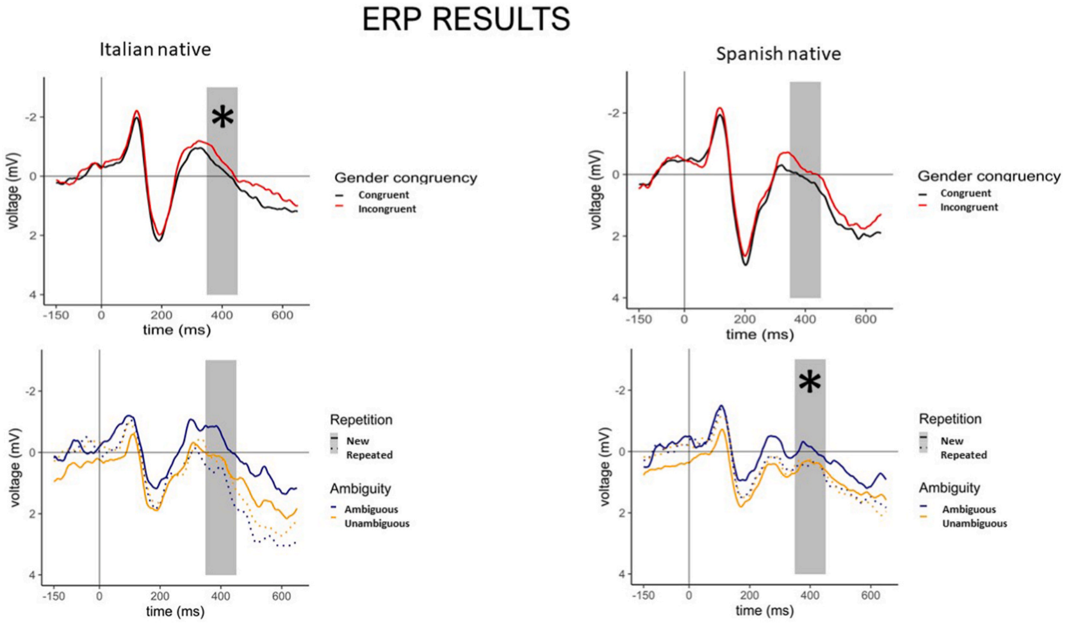

Fig. 2. On the top: model-based predictions of the N400 amplitudes for the gender congruency effect; at the left the Italian native speakers; at the right the Spanish native speakers. On the bottom: model-based predictions of the N400 amplitudes for the interaction between the ambiguity effect and the repetition; at the left the Italian native speakers; at the right the Spanish native speakers.

Table 2
Further analysis N400: 350–450 ms time window comparison of the scalp distribution between the gender_congruency index calculated for the Spanish natives and the ambiguity index calculated for the Spanish natives.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	−0.787496255713796	0.334910739202083	38.9999979991655	−2.3513616123209	0.0238519043172857
group1	−0.536986763052271	0.334910739202083	38.9999979999163	−1.6033727802567	0.116921594407491
lateral1	−0.146227106722514	0.13556403074276	312.000000761263	−1.07865711812589	0.281573811735008
lateral2	0.262581943824054	0.13556403074276	312.000000761263	1.93695881116369	0.0536529235063203
horizontal1	−0.100609610780258	0.13556403074276	312.000000761263	−0.742155645778715	0.458551373420901
horizontal2	0.291661316649419	0.13556403074276	312.000000761263	2.15146536327813	0.0322068155328788
group1:lateral1	−0.169175540525262	0.13556403074276	312.000000761263	−1.24793825912629	0.212989421907515
group1:lateral2	0.240356250995784	0.13556403074276	312.00000076128	1.77300903254989	0.0772027959177389
group1:horizontal1	−0.0837669371298944	0.13556403074276	312.000000761263	−0.61791418174078	0.537082676268755
group1:horizontal2	0.216737902191726	0.13556403074276	312.00000076128	1.59878620460171	0.110880278465373
lateral1:horizontal1	0.00514391606504219	0.191716490846374	312.000000761263	0.0268308482089008	0.978611801804124
lateral2:horizontal1	−0.154572290320404	0.191716490846374	312.00000076128	−0.806254535736656	0.420710204604986
lateral1:horizontal2	0.0104457747565031	0.191716490846374	312.00000076128	0.054485530745884	0.956583208053652
lateral2:horizontal2	0.198947120695933	0.191716490846374	312.000000761263	1.03771522114575	0.300206214485958
group1:lateral1:horizontal1	0.00568978740825565	0.191716490846374	312.000000761263	0.0296781324503534	0.976342726863319
group1:lateral2:horizontal1	−0.12993074311931	0.191716490846374	312.00000076128	−0.677723353612943	0.498449377123397
group1:lateral1:horizontal2	0.0361679118187669	0.191716490846374	312.00000076128	0.188653107821324	0.850487356710988
group1:lateral2:horizontal2	0.114519287361785	0.191716490846374	312.00000076128	0.597336655058802	0.550715894041688

seem to indicate that gender-incongruent nouns are harder to integrate, most likely because a greater amount of information is activated at the lexico-semantic level when gender is incongruent between languages than when it is congruent.

Secondly, we predicted that this greater effort to integrate the extra information evoked by gender-incongruent nouns in bilinguals would be carried out by the same mechanisms that are responsible for integrating the extra information activated during the processing of ambiguous words (as opposed to unambiguous words) in native speakers. In particular, we envisaged that the cost of integrating the greater amount of information activated by gender-incongruent nouns than by gender-congruent nouns would be similar to the cost of integrating the greater amount of information activated by ambiguous nouns than by unambiguous nouns. In short, we expected to find greater N400 amplitudes for the ambiguous nouns than for the unambiguous nouns, especially in the Spanish native speakers. Accordingly, the results showed that higher N400 amplitudes were evoked by Spanish native speakers when processing the ambiguous words compared with the unambiguous words. Importantly, the N400 amplitudes were predicted by the RT, such that both measurements were capturing the same effect. However, this ambiguity effect was modulated by the repetition of the items. In particular, the higher N400 amplitudes for the ambiguous vs. unambiguous words were only evoked during the first presentation of the words;

Table 3
Pairwised comparisons of the further analysis N400 350–450 ms time window of the scalp distribution between the gender_congruency index calculated for the Spanish natives and the ambiguity index calculated for the Spanish natives.

group	lateral	emmean
Spanish native	left	−1.639885666
Italian native	left	−0.227561059
Spanish native	right	−0.821544824
Italian native	right	−0.2282838
Spanish native	central	−1.512018566
Italian native	central	−0.295683619

contrast	lateral	estimate	SE	df	t.ratio	p.value
Spanish native - Italian native	left	−1.412324607	0.72261417	52.65011688	−1.954465697	0.05596593
Spanish native - Italian native	right	−0.593261024	0.72261417	52.65011688	−0.820992791	0.415352336
Spanish native - Italian native	central	−1.216334947	0.72261417	52.65011688	−1.68324259	0.098249493

during the second presentation, the N400 amplitudes were reduced to the level of unambiguous words. This may be indicating that the first time they process the word, all the available meanings are active. However, it seems that during the second presentation the participants spontaneously select one meaning only (probably the most prevalent) while inhibiting the rest. This is a novel finding because to our knowledge, no previous research showed this spontaneous meaning selection during the second presentation of an ambiguous word in a lexical decision task, and its correspondent decrease in the N400 amplitudes.

Interestingly, despite the fact that selected Spanish words and their translation equivalents in Italian were ambiguous in both languages (as rated by independent native speakers from a similar population as the participants of the study), we did not observe differences in the N400 amplitudes evoked by the Italian native speakers between ambiguous and unambiguous words. This pattern contrasts with the results by Taler et al. (2016), who found lower N400 amplitudes evoked by bilinguals when processing ambiguous words relative to unambiguous words especially in L2. Nevertheless, these findings are in line with the Sense Model (Finkbeiner et al., 2004), which postulates that L1 words are likely associated with many more semantic senses than their L2 translation equivalent. That is, even though the translation equivalent may have as many semantic meanings associated with the same word, the L2 learners may not be familiar with all those senses (Chen et al., 2014; Crossley & Skalicky, 2019). Accordingly, it may be the case that our Italian native speakers were not proficient enough in Spanish to identify all the senses the Spanish words have in common with the Italian translation equivalents.

In order to explore whether the mechanism underlying gender congruency effect is similar to that of the ambiguity effect, we compared the N400 amplitudes of both effects between groups. We found similar N400 effects between groups. However, the scalp distribution seems slightly different (Fig. 3): the ambiguity effect (Spanish native speakers) is more left-lateralized than the gender congruency effect (Italian native speakers). Moreover, while the gender congruency effect wave seems to start around the 250 ms, we can observe that the ambiguity effect starts around the 350 ms (see Fig. 1). Therefore, it seems that the nature of the gender congruency effect may not be exactly the same as the ambiguity effect. If the gender (in)congruency effect was triggered only by the activation of different semantic features at the semantic level (Athanasopoulos & Kasai, 2008; Boutonnet et al., 2012; Konishi, 1993; Saalbach et al., 2012), we should observe a similar pattern as in the case of the ambiguity effect. However, our data does not confirm this prediction. In contrast, if we consider that the gender congruency is generated at the lexical level (Paolieri et al., 2020), what we observe at the N400 time window may be the difficulty to integrate extra lexical information that got activated around the 250 ms (consistent with the prominent models of lexical access, e.g., Indefrey & Levelt, 2004) and extends to the semantic level. Still, we should be extremely cautious with this statement such as the calculation were performed between groups. Moreover, the current data and statistics are not strong enough to reach convincing conclusions about this comparison and thus the gender by ambiguity comparison should be considered an unresolved issue. All in all, our data suggest that the information about grammatical gender is located at the lexical level, gets automatically activated during lexical selection, and has a close relationship with the semantic information in bilingual language comprehension.

5. Conclusion

In conclusion, grammatical gender is a rich attribute that is activated during lexical selection. This experiment has shown that even when processing L2 words presented in isolation (inanimate nouns with an arbitrary gender assignment), bilinguals activate the grammatical gender of the corresponding L1 nouns, which influences access to L2 representations. Specifically, when the gender is incongruent between languages, a greater amount of information is active at the lexical level, which is more costly to integrate at the semantic level compared to when the gender is congruent. These results provide strong evidence of the automatic activation of grammatical gender since the experiment that was conducted entirely in the L2, with no need to activate the L1 to perform the task. Moreover, we explored whether the mechanism responsible for managing the greater effort required to process the extra information activated by incongruent gender in bilinguals would be similar to the mechanism responsible for managing the extra information activated by ambiguous words in the native language of the speakers. The results do not seem to support this assumption, pointing to the nature of the gender congruency effect as lexical. Still, further research should confirm it. It will be interesting in future research to explore how the two dimensions of gender congruency and ambiguity interact in a fully factorial design that also includes ambiguous gender-incongruent words.

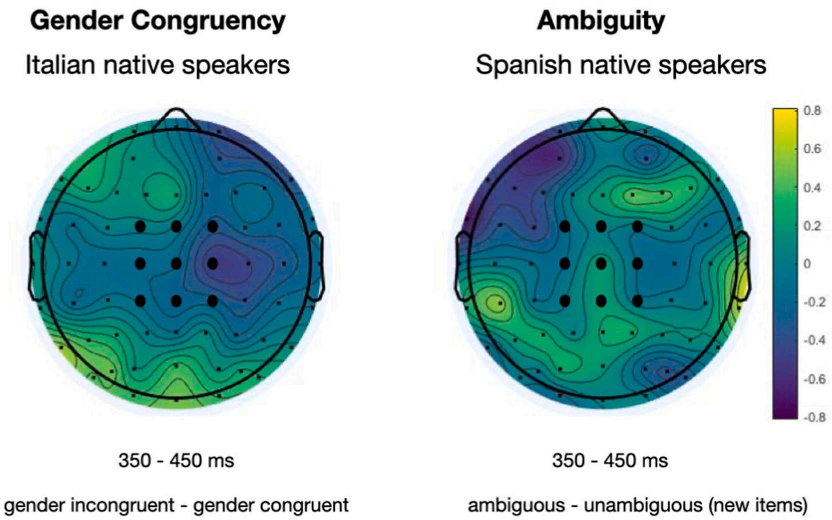

Fig. 3. At the left, scalp distribution of the gender congruency of Italian native speakers represented by the subtraction of gender incongruent and gender congruent conditions. At the right, scalp distribution of the ambiguity effect of Spanish native speakers represented by the subtraction of ambiguous and unambiguous conditions only for new items. The brighter colors indicate positive amplitudes; the darker colors indicate negative amplitudes.

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.

Data availability

Data will be made available on request.

Acknowledgments

This work has been supported by Spanish Ministry of Science, Innovation and Universities (RED2018-102615-T), by Feder-Andalucía (A-SEJ-416-UGR20) to D.P., by Spanish Ministry of Science and Innovation (PID2019-111359 GB-I00/AEI/10.13039/501100011033) and by the Universitat Rovira i Virgili to P. F. (2019PFR-URV-B2-32). Funding for open access charge: Universidad de Granada / CBUA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jneuroling.2023.101156.

References

Alario, F., Ayora, P., Costa, A., & Melinger, A. (2008). Grammatical and nongrammatical contributions to closed-class word selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 960.
Amsel, B. D. (2011). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia, 49*(5), 970–983.
Athanasopoulos, P., & Kasai, C. (2008). Language and thought in bilinguals: The case of grammatical number and nonverbal classification preferences. *Applied PsychoLinguistics, 29*(1), 105–123.
Barber, H. A., Otten, L. J., Kousta, S. T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language, 125*(1), 47–53.
Bates, E., Devescovi, A., Pizzamiglio, L., D'amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics, 57*(6), 847–862.
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., Mäechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., … Bolker, M. B. (2015). Package 'lme4'. *Convergence, 12*(1), 2.
Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240–250.
Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language & Cognitive Processes, 22*(5), 633–660.
Bobb, S. C., Von Holzen, K., Mayor, J., Mani, N., & Carreiras, M. (2020). Co-activation of the L2 during L1 auditory processing: An ERP cross-modal priming study. *Brain and Language, 203*, Article 104739.

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition, 10*(3), 299–314.
Bordag, D., & Pechmann, T. (2008). Grammatical gender in speech production: Evidence from Czech. *Journal of Psycholinguistic Research, 37*(2), 69–85.
Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research, 1479*, 72–79.
Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of cognition, 1*(1).
Casado, A., Palma, A., & Paolieri, D. (2018). The influence of sex information on gender word processing. *Journal of Psycholinguistic Research, 47*(3), 557–583.
Chen, B., Zhou, H., Gao, Y., & Dunlap, S. (2014). Cross-language translation priming asymmetry with Chinese-English bilinguals: A test of the sense model. *Journal of Psycholinguistic Research, 43*(3), 225–240.
Cole, P., & Segui, J. (1994). Grammatical incongruency and vocabulary types. *Memory & Cognition, 22*(4), 387–394.
Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45*(4), 721–736.
Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283.
Crossley, S. A., & Skalicky, S. (2019). Making sense of polysemy relations in first and second language speakers of English. *International Journal of Bilingualism, 23*(2), 400–416.
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.
Cubelli, R., & Paolieri, D. (2008). The processing of grammatical gender as abstract lexical feature. *Language and Cognition: A long story. Festschrift in honour of Ino Flores D'Arcais*, 73Á86 (Padova, Italy: Cleup).
De Groot, A. M. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(5), 824.
De Martino, M., Bracco, G., & Laudanna, A. (2011). The activation of grammatical gender information in processing Italian nouns. *Language & Cognitive Processes, 26*(4–6), 745–776.
Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21.
Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. *Handbook of bilingualism: Psycholinguistic approaches*, 179–201.
Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45*(4), 1246–1258.
Duyck, W., & Brysbaert, M. (2004). Forward and backward number translation requires conceptual mediation in both balanced and unbalanced bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 889.
Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51*, 1–22.
Goslin, J., Galluzzi, C., & Romani, C. (2014). PhonItalia: A phonological lexicon for Italian. *Behavior Research Methods, 46*(3), 872–886.
Guasch, M., Ferre, P., & Haro, J. (2017). Pupil dilation is sensitive to the cognate status of words: Further evidence for non-selectivity in bilingual lexical access. *Bilingualism: Language and Cognition, 20*(1), 49–54.
Halberstadt, L., Kroff, J. R. V., & Dussias, P. E. (2018). Grammatical gender processing in L2 speakers of Spanish. *Journal of Second Language Studies, 1*(1), 5–30.
Haro, J., Demestre, J., Boada, R., & Ferré, P. (2017). ERP and behavioral effects of semantic ambiguity in a lexical decision task. *Journal of Neurolinguistics, 44*, 190–202.
Holocomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effect in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 721–742.
Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: Electrophysiological evidence for cross-language competition. *Brain Research, 1371*, 100–109.
Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1–2), 101–144.
Iniesta, A., Paolieri, D., Serrano, F., & Bajo, M. T. (2021). Bilingual writing coactivation: Lexical and sublexical processing in a word dictation task. *Bilingualism: Language and Cognition, 24*(5), 902–917.
Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science, 15*(5), 314–318.
Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627–633.
Klassen, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus, 28*(1), 9–28.
Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research, 22*(5), 519–534.
Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*(5), 497–514.
Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science, 23*(3), 159–163.
Kroll, J. F., & De Groot, A. M. (Eds.). (2009). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.
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.
Kroll, J. F., Sumutka, B. M., & Schwartz, A. I. (2005). A cognitive view of the bilingual lexicon: Reading and speaking words in two languages. *International Journal of Bilingualism, 9*(1), 27–48.
Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470.
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.
La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture-word tasks. *Psychological Research, 61*(3), 209–219.
Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language, 59*(3), 312–330. Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. Frontiers in human neuroscience, 8, 213.
Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version, 1*(1), 3.
Lopez-Calderon, J., & Luck, S. J. (2014). Erplab: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, 213.
Macizo, P., Bajo, T., & Martin, M. C. (2010). Inhibitory processes in bilingual language comprehension: Evidence from Spanish-English interlexical homographs. *Journal of Memory and Language, 63, 232al of Memory and Language, 63terlexical homograp*.
Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*.
Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition, 6*(2), 97–115.
Morales, L., Paolieri, D., & Bajo, T. (2011). Grammatical gender inhibition in bilinguals. *Frontiers in Psychology, 2*, 284.
Morales, L., Paolieri, D., Dussias, P. E., Kroff, J. R. V., Gerfen, C., & Bajo, M. T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition, 19*(2), 294–310.
Palma Duràn, C., & Pillon, A. (2011). The role of grammatical category information in spoken word retrieval. *Frontiers in Psychology, 2*, 338.
Paolieri, D., Cubelli, R., Macizo, P., Bajo, T., Lotto, L., & Job, R. (2010). Grammatical gender processing in Italian and Spanish bilinguals. *The Quarterly Journal of Experimental Psychology, 63*(8), 1631–1645.
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*, 1–11.

Paolieri, D., Lotto, L., Leoncini, D., Cubelli, R., & Job, R. (2011). Differential effects of grammatical gender and gender inflection in bare noun production. *British Journal of Psychology, 102*(1), 19–36.
Paolieri, D., Padilla, F., Koreneva, O., Morales, L., & Macizo, P. (2018). Gender congruency effects in Russian–Spanish and Italian–Spanish bilinguals: The role of language proximity and concreteness of words. *Bilingualism: Language and Cognition*, 1–18.
Plaut, D. C., & Shallice, T. (1993). Perseverative and semantic influences on visual object naming errors in optic aphasia: A connectionist account. *Journal of Cognitive Neuroscience, 5*(1), 89–117.
Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6*, 11.
Sá-Leite, A. R., Fraga, I., & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review, 26*(4), 1148–1173.
Sá-Leite, A. R., Haro, J., Comesaña, M., & Fraga, I. (2021). Of beavers and Tables: The role of animacy in the processing of grammatical gender within a picture-word interference task. *Frontiers in Psychology, 12*, Article 661175. https://doi.org/10.3389/fpsyg.2021.661175
Saalbach, H., Imai, M., & Schalk, L. (2012). Grammatical gender and inferences about biological properties in German-speaking children. *Cognitive Science, 36*(7), 1251–1267.
Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition, 10*(3), 257–275.
Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review, 26*(6), 1767–1786.
Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language, 110*, Article 104038.
Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime*. Pittsburgh, PA: Psychology Software Tools Version 2.0. 8.90. .
Slobin, D. I. (1997). *Mind, code, and text. Essays on language function and language type*. Dedicated to T. Givón.
Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition, 28*(3), 387–422.
Taler, V., López Zunini, R., & Kousaie, S. (2016). Effects of semantic richness on lexical processing in monolinguals and bilinguals. *Frontiers in Human Neuroscience, 10*, 382.
Van Hell, J. G., & De Groot, A. M. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*(3), 193–211.
Wang, M., & Schiller, N. O. (2019). A review on grammatical gender agreement in speech production. *Frontiers in Psychology, 9*, 2754.
Wang, M., Shao, Z., Chen, Y., & Schiller, N. O. (2018). Neural correlates of spoken word production in semantic and phonological blocked cyclic naming. *Language, Cognition and Neuroscience, 33*(5), 575–586.
Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1–25.
West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*(6), 1024–1037.
Zappa, A., Mestre, D., Pergandi, J. M., Bolger, D., & Frenck-Mestre, C. (2022). Cross-linguistic gender congruency effects during lexical access in novice L2 learners: Evidence from ERPs. *Language, Cognition and Neuroscience, 37*(9), 1073–1098.
Zu Wolfsthurn, S. V. G., Pablos, L., & Schiller, N. O. (2021). Noun-phrase production as a window to language selection: An ERP study. *Neuropsychologia, 162*, Article 108055 (c).