Gender Congruency Behavioral and Electrophysiological evidence

The Gender Congruency Effect in Catalan–Spanish Bilinguals: Behavioral and Electrophysiological evidence

Daniela Paolieri[1], Josep Demestre[2], Marc Guasch[2], Teresa Bajo[1], Pilar Ferré[2]

[1]Mind, Brain and Behavior Research Center; Department of Experimental Psychology; University of Granada, Granada, Spain
[2]Universitat Rovira i Virgili; Department of Psychology and Research Center for Behavior Assessment, Tarragona, Spain.

*Acknowledgements
This research was supported by Grants PCIN-2015-165-C02-01, PCIN-2015-165-C02-02, PCIN-2015-132 and PGC2018-093786-B-I00 from the Spanish Ministry of Economy and Competitiveness.

Address for correspondence:

Daniela Paolieri
Mind, Brain and Behavior Research Center, Department of Experimental Psychology, C/ Profesor Clavera, s/n, University of Granada, 18071; Granada, Spain.
E-mail: dpaolieri@ugr.es

Pilar Ferré
Research Center for Behavior Assessment; Department of Psychology;
Carretera de Valls, s/n, Rovira i Virgili University, 43007; Tarragona, Spain.
E-mail: mariadelpilar.ferre@urv.cat

Abstract

The present study examines whether processing a word in one language is affected by the grammatical gender of its translation equivalent in another language. To this end, a group of Catalan–Spanish bilinguals performed a translation–recognition task while event-related potentials (ERPs) were recorded. Participants were presented with Catalan and Spanish pairs of words and had to decide if they were translation equivalents. Correct translations included words that were gender congruent (*estiu* MAS/*verano* MAS–summer) or gender incongruent (*tardor* FEM/*otoño* MAS–autumn). The behavioral results showed that participants were faster and more accurate in the gender-congruent condition than in the incongruent condition. The ERP data showed a reduced N400 for the congruent condition. The facilitative effect of gender congruency observed in this study constitutes evidence of the obligatory access to grammatical gender information during bare noun processing and suggests that the bilinguals' gender systems interact, even in highly proficient early bilinguals.

Keywords: Gender congruency, gender processing in bilinguals, translation–recognition task, ERPs, N400.

Introduction

Grammatical gender is one of the most interesting grammatical categories (Corbett, 1991), which plays a central role in the processing of languages with two (e.g., Italian, Spanish and Catalan), three (e.g., German, Czech and Russian) or more grammatical classes (e.g., Bantu languages have between 12 and 20 classes; Corbett, 2011; Nurse & Philippson, 2003). A fluent speaker of those languages has to completely control gender information to be able to make the correct agreement between the elements of the sentence in the appropriate way.

From a psycholinguistic point of view, some models of language production propose that grammatical gender information is represented at a different level than those that represent semantic and phonological information (Caramazza & Miozzo, 1997; Levelt, Roelofs, & Meyer, 1999). According to these models, grammatical gender information would only be retrieved when nouns are integrated in syntactic contexts in which an agreement relation has to be established between two (or more) linguistic elements. Thus, they propose that gender information would not be retrieved when agreement is not required. In contrast, some recent evidence suggests that this grammatical property also seems to be automatically activated during lexical access outside of a sentence context, that is, in the production of bare nouns (Alario, Ayora, Costa, & Melinger, 2008; Cubelli, Lotto, Paolieri, Girelli, & Job, 2005; Duràn & Pillon, 2011; Paolieri, Lotto, Leoncini, Cubelli, & Job, 2011; see also Wang, Shao, Chen, & Schiller, 2018, for a comparable lexico-syntactic classifier feature). Several theoretical proposals have been made to account for such findings, in terms of different dynamics of activation of grammatical gender information. The model proposed by Duràn and Pillon (2011) suggests two mechanisms of gender activation, one involving

bidirectional links between lexical and grammatical category nodes and the other one involving links between grammatical category nodes and gender morpheme nodes. In a similar way, but giving a more central role to grammatical gender during lexical access as an inherent grammatical property of nouns, the model proposed by Cubelli et al. (2005) considers grammatical gender as obligatory for accessing the complete morphology of languages such as Italian and Spanish.

The activation of grammatical gender during language processing has been reported not only in production, but also in comprehension (e.g., Colé, Pynte, & Andriamamonjy, 2003; Colé & Seguí, 1994; De Martino, Bracco, & Laudanna, 2011; De Martino, Bracco, Postiglione, & Laudanna, 2017) where the interest is centered mainly in the easiness for lexical access due to the morphological transparency of nouns (the statistical-phonological regularities in a particular language between words endings and their gender). In this context, Golland and Frost (2001) have provided the main explanatory attempt with the proposal of a "double route" model of grammatical gender selection. The model considers two different forms of activation of gender information: a first path correlates with the morphological ending, and a second one incorporates an abstract representation of grammatical gender, located at the lexical level. According to this model, when the two sources of information coincide, nouns would be accessed faster and more accurately. Otherwise, the presence of a conflict between the two sources of information would make the processing more difficult.

The above reviewed research, which has focused on the native language of the speakers, suggests that grammatical gender is activated during language production and comprehension, both when nouns are presented in sentence contexts and when they are presented in isolation. Several studies examining grammatical gender effects in bilinguals have reached similar conclusions. In many gendered languages, grammatical

gender is completely independent from conceptual information (Aronoff, 1994). For example, the word for *car* is masculine in Spanish (*coche*) and in Catalan (*cotxe*), but feminine in Italian (*macchina*) and neuter in German (*Auto*); even if the words refer to the same concept, these nouns have a different grammatical gender in each language. Bilingual studies have taken advantage of the existence of this type of incongruences between languages to study how the grammatical gender systems of bilinguals interact during language processing. The usual approach has been to compare nouns that have one gender value in one language and a distinct gender value in the other language (i.e., incongruent or heterogeneric nouns) with nouns that have the same gender value in both languages (i.e., congruent or homogeneric nouns). The so-called gender congruency effect, that is, the advantage in processing (i.e., shorter reaction times and/or greater accuracy) for congruent nouns in comparison to incongruent nouns is taken as evidence of the interaction between the gender systems of the two languages.

A large number of studies have obtained evidence of such an interaction in bilingual speech production by means of different paradigms and in distinct bilingual populations. Indeed, gender congruency effects have been observed both with symmetrical gender systems (i.e., languages that share the number and types of gender values; e.g., Czech–German, Bordag & Pechmann, 2007, 2008; Greek–German, Salamoura & Williams, 2007; German–Dutch, Lemhöfer, Spalek, & Schriefers, 2008; Italian–Spanish; Paolieri et al., 2010) as well as with asymmetrical gender systems (i.e., languages with different number and types of gender values; e.g., Spanish–German, Klassen, 2016; Russian–Spanish, Paolieri, Padilla, Koreneva, Morales, & Macizo, 2018). Furthermore, these effects have been reported both between highly similar languages, such as Spanish and Italian, and between dissimilar languages, such as Russian and Spanish (Paolieri et al., 2018). Importantly, the effects seem to be obtained

not only when gender agreement is required but also with bare nouns (see Sá-Leite, Fraga, & Comesaña, 2019, for a review of this line of research and the main variables investigated). These findings provide support to the idea that grammatical gender is an intrinsic lexical feature that is available even when agreement is not required (Paolieri et al., 2018). Furthermore, they constitute evidence in favor of an integrated gender system in bilinguals, in which the representation of grammatical gender values would be shared across languages (Klassen, 2016). In relation to that, gender congruency effects also support the notion of language non-selective lexical access in bilinguals (see Bialystok, Craik, & Luk, 2012; Kroll & Bialystok, 2013; Kroll, Bobb, & Hoshino, 2014 for reviews). Although most influential models of language processing in bilinguals do not speak explicitly about gender, they propose an integrated lexicon (e.g., BIA+ model; Dijkstra & Van Heuven, 2002 and Multilink; Dijkstra et al., 2019) or a direct influence between the two languages at the level of lexical representations (RHM; Kroll & Stewart, 1994). For that reason, gender effects can be explained in the context of cross-language interactions at every level of lexical access (see also Sá-Leite et al., 2019 for a review).

Gender congruency effects in bilinguals have been mostly observed in production tasks (Bordag & Pechmann, 2007, 2008; Klassen, 2016; Lemhöfer et al., 2008; Paolieri et al., 2010; Paolieri et al., 2018; Salamoura & Williams, 2007). In language comprehension, the role played by grammatical gender has focused mainly on its syntactic function in noun phrases during visual (Lemhöfer et al., 2008) or auditory word recognition tasks (Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Dussias, Valdés Kroff, Guzzardo, & Gerfen, 2013; Lew-Williams & Fernald, 2007; Morales et al., 2016; van Heugten & Johnson, 2011; van Heugten & Shi, 2009), but its importance as an intrinsic lexical feature has scarcely been considered. Two of these studies are

worth mentioning here. In the first one, Lemhöfer et al. (2008) investigated the effects of gender congruency in visual word recognition. In particular, they used a lexical decision task in which second language (L2) words were primed by gender-marked determiners. These authors observed that performance in L2 (Dutch) was influenced by the gender in the first language (L1, German): the German translation equivalents which were congruent in gender between languages were recognized faster than gender-incongruent words. These results showed that the grammatical gender systems of the bilinguals' two languages interact during visual word recognition, at least through preactivation of grammatical gender information by means of the presentation of gender-marked determiners. Morales et al. (2016) reached similar conclusions in a study in which the effects of the L1 gender on L2 auditory word recognition were investigated with a visual world paradigm. In this study, Italian-Spanish bilinguals were presented with pairs of objects on a computer screen. One of these objects was the target, whose name was presented auditorily preceded by a definite article. The grammatical gender of the L1 (Italian) translation could be either the same or different from the L2 (Spanish) noun. The results of this study showed that participants looked at the target pictures significantly less when the translation equivalents were incongruent in gender with the Spanish nouns than when they were congruent. This finding constituted evidence that even when performing a task in the L2, the native language influences the pattern and duration of visual fixations in bilinguals.

It should be noted that in the two above mentioned studies (Lemhöfer et al., 2008; Morales et al., 2016), L2 nouns were preceded by determiners, that is, they were presented in sentence contexts. Gender congruency effects in bare noun comprehension have not been studied yet. On the other hand, participants in most of the reviewed comprehension and production studies, although being highly proficient in both

languages, did not have a proficiency level in L2 comparable to the one they had in their native language (L1). There is robust evidence that cross-language activation is stronger for low proficient bilinguals than for highly proficient bilinguals (e.g., Sundermann & Kroll, 2006). Hence, proficiency and other variables related to the characteristics of the bilinguals might also influence grammatical gender processing. Indeed, language production studies have shown that the gender congruency effect is less reliable in bilinguals with a similar level of proficiency in the two languages and with an early age of acquisition of the L2 than in unbalanced and late bilinguals (Bordag & Pechmann, 2007; Costa, Kovacic, Franck & Caramazza, 2003).

All the above research on gender congruency in bilinguals has relied on behavioral measures (i.e., reaction times and accuracy), and not on measures that are more sensitive to the time course of processing, such as the recording of event-related brain potentials (ERPs). ERPs can offer a fine-grained measure of the cognitive processes stimulated by the experimental manipulation (Luck, 2005), providing valuable information about the time course of gender activation. In the monolingual domain, ERP studies have focused on the N400 component, which has been considered as an index of anomaly detection in situations in which the gender of the predicted target does not match the gender of the determiner (Wicha et al., 2005). This ERP component has also been related to the activation of gender information encoded at the lexicon (Friederici & Jacobsen, 1999). In this line, Barber and Carreiras (2005) and Caffarra, Janssen, and Barber (2014) found that gender disagreement in word pairs formed by a noun and an adjective elicited an "N400-type effect": disagreeing words produced larger negative deflections as compaed to agreeing words. Wang et al. (2018) reported similar findings in a bare noun production task in Mandarin Chinese, in which they manipulated the congruency of a lexico-syntactic classifier feature (comparable to

grammatical gender). The results showed that the classifier incongruent condition elicited larger N400 effects in comparison to the classifier congruent condition. These findings converge with the behavioral evidence discussed earlier that shows the automatic activation of lexico-syntactic features (including grammatical gender) in noun production in the native language. However, it is still unclear if there is also a neural signature of gender congruency effects in word comprehension in bilinguals.

The Present Study

In the present study, we aimed to examine whether lexical access in highly proficient early Catalan–Spanish bilinguals is modulated by the grammatical gender of the words during a translation recognition task, focusing on gender-congruency effects. To date, no evidence of gender congruency effects has been firmly established in Catalan-Spanish bilinguals either in production or in comprehension. Catalan and Spanish are both Romance languages, and have similar gender systems in terms of gender values and gender agreement. In both languages, grammatical gender is organized into two different categories, namely masculine and feminine (Lloret &Viaplana, 1997).

Participants had to decide whether a Spanish word was the correct translation of a Catalan word while ERP were recorded. Half of the correct translations had the same gender value in Catalan and Spanish (the congruent condition, e.g., handkerchief, *mocador*$_{\text{MAS}}$ in Catalan and *pañuelo*$_{\text{MAS}}$ in Spanish) and the other half had a different gender value (the incongruent condition, e.g., fork, *forquilla*$_{\text{FEM}}$ in Catalan and *tenedor*$_{\text{MAS}}$ in Spanish). It remains to be established if the grammatical gender of the words is automatically accessed when this kind of bilinguals (i.e., highly proficient bilinguals who acquired both languages in early childhood and who are exposed to both

languages on a daily basis) perform a task involving the recognition of bare nouns. Furthermore, to our knowledge, it is the first time that ERPs are recorded in order to examine the neural signature of grammatical gender congruency effects in bilinguals. If Catalan-Spanish bilinguals activate the grammatical gender of the non-target language during lexical access, we would expect comparable effects in behavioral and electrophysiological measures: A faster, more accurate and easier (as indexed by a reduction in the amplitude of the N400) lexical access for gender-congruent stimuli in comparison with incongruent ones.

Method

Participants

A sample of 31 participants (28 females, mean age = 20.69, SD = 4.41), took part in the study. They were undergraduate students at the Universitat Rovira i Virgili (Tarragona, Spain) who were paid to participate in the experiment. All of them signed an informed consent form before starting the experiment. Participants were Catalan-Spanish bilinguals. To assess their proficiency in Catalan and Spanish, they completed a questionnaire in which they rated their ability in reading, writing, speaking, and listening on a 7-point Likert scale (1 = 'very poor'; 7 = 'native-like'). They were also asked for the age of exposure to each language, as well as for their language preference and use. Concretely, they assessed their preference and use on a 1 (exclusively in Catalan) to 7 (exclusively in Spanish) scale, where the middle point (i.e., 4) indicated no preference at all for either of the two languages and the same frequency of use for both of them. The questionnaire revealed that all the participants were exposed to Catalan from birth, and their average age of exposure to Spanish was 1.79 years (SD = 2.23). Particiapnts' ratings showed that they were all highly proficient in both languages (see

Table 1). However, planned paired comparison between the average reported proficiency in Catalan (M = 6.97, SD = 0.18) and in Spanish (M = 6.82, SD = 0.44) showed an advantage for Catalan in, $t(30) = 3.58$, $p = .001$. The questionnaires also showed that Catalan was the language preferred in most cases (M = 3.40, SD = 0.98) and that it was also the more frequently used by these participants overall (M = 3.34, SD = 1.22). Thus, participants were early sequential bilinguals of Catalan and Spanish, being highly proficient in both languages, but slightly more dominant in Catalan than in Spanish.

<Insert Table 1 about here>

Materials

A set of 76 Catalan-Spanish translation pairs were selected from a larger pool of 224 pairs. All the words selected were nouns. The selection criteria were that: 1) they were inanimate nouns, 2) they were not identical cognates between Catalan and Spanish, and 3) Catalan-Spanish bilinguals knew the correct translations, as well as their gender. Concerning the last criterion, a group of 62 participants (52 females, mean age = 22.32, SD = 3.34), different from those participating in the main experiment, but coming from the same population, completed a written production test. They were presented with the word in Catalan or in Spanish together with the article (e.g., *el carrer*, the street) and were asked to produce the translation in the other language (*la calle*). To conduct this production test, the 448 words (224 Catalan words and their 224 Spanish translations) were distributed in eight questionnaires. In each questionnaire, half of the items were Catalan words and the other half were Spanish words (of note, the same questionnaire did not include the Catalan word and its corresponding Spanish translation). Furthermore, half of the words had the same gender across languages and

the other half had a different gender. Participants first produced the translations for words in one language (Catalan or Spanish) and then the translations for words in the other language. The order of the production language was counterbalanced across questionnaires. Each word was presented to at least 15 participants. The percentage of participants who knew the translation in the other language (included the article) was computed.

We selected translation pairs that were known by at least 80% of the participants. Among the final 76 selected pairs, 38 had the same grammatical gender in Catalan and in Spanish (gender-congruent condition), while the other 38 had a different gender in the two languages (gender-incongruent condition). In each set, half of the nouns were masculine and the other half were feminine. Gender congruent and gender incongruent pairs were matched in number of letters, lexical frequency (Catalan data was obtained from NIM, Guasch, Boada, Ferré, & Sánchez-Casas, 2013, and Spanish data was obtained from ESPAL, Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013), formal similarity between the Catalan word and its Spanish translation (Normalized Levensthein Distance, NLD, data obtained from Guasch et al., 2013), percentage of participants who knew the Spanish translation (data obtained from the above mentioned production test), as well as in several subjective variables regarding the Spanish words (there were no normative data for those variables in Catalan), such as concreteness, imageability, familiarity (data obtained from Duchon et al., 2013), valence and arousal (data obtained from Stadhagen-Gonzalez, Imbault, Pérez Sánchez, & Brysbaert, 2017, and Guasch, Ferré, & Fraga, 2016) (all *p*s > .112) (see Table 2). In addition, the number of nouns with transparent and opaque gender endings[1] was matched as closely as possible across the two congruency conditions both in Spanish

[1] Spanish and Catalan nominal categorization systems have formal regularities related to the distribution of noun endings, with gender transparent (-a feminine; for Spanish and Catalan; -o masculine for Spanish) and gender opaque nouns (with different endings) (see El-Yousseph, 2010; Harris, 1991).

(Congruent pairs= 29 Transparent/9 Opaque; Incongruent pairs= 26 Transparent /12 Opaque; χ^2 (1) =.263, p=.608), and in Catalan (Congruent pairs= 13 Transparent/25 Opaque; Incongruent pairs= 16 Transparent/22 Opaque; χ^2 (1) =.223, p=.637). The critical stimuli are reported in Appendix A.

Finally, due to the characteristics of the task (i.e., a translation recognition task), a set of non-translation pairs was required. To this end, we selected 76 Catalan words (38 masculine and 38 feminine) that were randomly paired with 76 Spanish words to create the non-translation pairs. Furthermore, mimicking the experimental conditions, in half of the pairs the gender of the two words was the same, while in the other half the gender was different. Words in the critical (i.e., translation) and non-critical (i.e., non-translation) pairs were matched in length, lexical frequency, and NLD between the Catalan and the Spanish word in each pair (all *p*s > .103).

<Insert Table 2 about here>

Procedure

After participants provided informed consent, the EEG cap was placed and they were seated in a comfortable chair about 60 cm from the stimulus monitor. The entire experimental session was carried out in a sound- attenuated room. Participants performed a translation recognition task in which they were asked to decide whether the second word of the pair was a correct translation of the first one. All the stimuli were presented as bare nouns. Participants had to answer by pushing the 'YES' or the 'NO' button of a keypad. The EEG was monitored while participants performed the task. The stimuli were presented one at a time at the center of a screen in black font on a white background, using the DMDX program (Forster & Forster, 2003). The computer

generated a pseudo-random order of presentation for each participant. Each trial started with an image of an eye displayed for 2,000 ms, which indicated participants that they were allowed to blink, followed by a 500 ms fixation point ('#'). Just after the fixation point, a Catalan word was presented for 250 ms, followed by a Spanish word. The Spanish word remained on the screen until the participant responded or 2,500 ms had elapsed. There was a 1,000 ms ISI between the trials. The experiment began with eight practice trials, followed by the 152 experimental trials.

EEG recording

The electroencephalogram was recorded from 30 Ag/AgCl active electrodes attached on the ActiCap system (Brain Products GmbH, Gilching, Germany) placed on the scalp according to the extended 10–20 system (Pivik et al., 1993). Vertical and horizontal electrooculogram (EOG) was recorded through two additional electrodes attached below the left eye and the corner of the right eye. All electrodes were referenced online to the left mastoid. An additional electrode was attached to the right mastoid for offline referencing. Electrode impedances were kept below 5 kΩ. The EEG and EOG signals were recorded and digitized with PyCorder and amplified by an antiCHamp amplifier (Brain Vision LLC). All channels were amplified with a band pass of 0.01–100 Hz and at a sampling rate of 500 Hz.

ERPs were time-locked to the second word of each pair and averaged over a 900 ms epoch, including a 100 ms pre-stimulus-onset baseline. The EEG was refiltered offline with a 25-Hz, low-pass, zerophase shift digital filter. Automatic and manual rejections were carried out to exclude periods containing movement or technical artifacts (the automatic EOG rejection criterion was ± 70 mV). After these rejection procedures, additional visual inspection of the resulting signal was carried out for each

subject individually. Using these criteria, a mean of 5.14 % of trials were rejected (SD = 6.62). The paired-samples t-test revealed no differences between the two experimental conditions in the rate of artifact rejection, t (30) = .333, p =.744. Nine regions of interest (ROIs) were created. The left frontal (LF) ROI included sites F7, F3, FC5; the left central (LC) ROI included sites CP5, T7, C3; the left posterior (LP) ROI included sites P7, P3, O1; the midline frontal (MF) ROI included sites FP1, FP2, Fz; the midline central (MC) ROI included sites FC1, FC2, Cz; the midline posterior (MP) ROI included sites CP1, CP2, Pz; the right frontal (RF) ROI included sites F8, F4, FC6; the right central (RC) ROI included sites CP6, T8, C4; the right posterior (RP) ROI included sites P8, P4, O2.

Mean N400 amplitude was calculated for each subject at each ROI between 300 and 500 ms separately for pairs of words with congruent and incongruent gender. The time window for the N400 was consistent with previous studies (e.g., Guo, Misra, Tam, & Kroll, 2012; Ma, Chen, Guo, & Kroll, 2017; Moldovan, Demestre, Ferré, & Sánchez-Casas, 2016).

Results

Behavioral results

The data corresponding to incorrect responses were discarded from the analysis. Reaction times with values below 200 ms and above 2,000 ms were removed from the analysis and were treated as missing points. Reaction times that were more than two SDs above or below the mean for a given participant in all conditions were also removed. This led to the exclusion of 5 % of the data. No participant had an error rate above 15% and thus all participants were included in the analysis. Mean RTs and error rates (and standard deviations) are summarized in Table 3.

Separate ANOVAs were conducted based on the participants and items response latencies and error rates. The analyses were based on a one-way ANOVA with the factor gender congruency (gender-congruent vs. gender-incongruent). Gender congruency was manipulated within-subjects and between-items.

<Insert Table 3 about here>

The ANOVA on the RTs revealed a main effect of the gender congruency factor, F_1 (1, 30) = 13.52, $p < .001$, $\eta_p^2 = .311$, F_2 (1, 74) = 4.61, $p = .035$, $\eta_p^2 = .059$), showing that RTs to gender-incongruent pairs (655 ms) were significantly longer than RTs to gender-congruent pairs (631 ms). The ANOVA on the error rates also revealed a main effect of the gender congruency factor, F_1 (1, 30) = 5.55, $p = .025$, $\eta_p^2 = .156$, F_2 (1, 74) = 4.35, $p = .041$, $\eta_p^2 = .055$), showing that participants committed more errors with gender-incongruent pairs (4.6 %) than with gender-congruent pairs (2.6 %).

ERP results

An ANOVA with the factors gender congruency (gender-congruent vs. gender incongruent), Hemisphere (Left, Midline, Right), and Latitude (Anterior, Central, Posterior) was used to analyze the N400 data. The Greenhouse and Geisser (1959) correction was applied to all repeated measures having more than one degree of freedom in the numerator. In such cases, the corrected *p*-value is reported. Bonferroni adjustment was used for post hoc pairwise comparisons. In addition, only the main effect of the gender congruency factor and the interactions between this factor and the two topographical factors (Hemisphere & Latitude) are reported when significant.

The ANOVA revealed a significant main effect of gender congruency, F (1, 30) = 5.80, p = .022, η_p^2 = .162, at the N400 time window (i.e., from 300 ms to 500 ms). As can be seen in Figures 1 and 2, the gender-incongruent condition exhibits a clear modulation of the N400 component, starting around 300 ms post-stimulus. Indeed, the mean amplitude for the gender-incongruent condition (M = 2.46 μV) was significantly more negative than that for the gender-congruent condition (M = 3.07 μV) during the N400 epoch.

The two-way interaction between gender congruency and latitude did not reach significance, F (2, 60) = 3.48, p = .065, η_p^2 = .104. The three-way interaction between gender congruency, latitude and hemisphere was significant, F (4, 120) = 3.19, p = .038, η_p^2 = .096. Post-hoc pairwise comparisons revealed that at the three frontal ROIs (i.e., LF, MF, RF) there were no significant differences between the two conditions (all ps > .15). At central ROIs, the N400 was significant at the midline-central ROI (p = .011) and the right-central ROI (p = .005), and was not significant at the left-central ROI (p = .077). Moreover, the comparisons revealed that at posterior ROIs, the N400 effects were larger at the posterior-midline ROI (p = .007) than at the left-posterior ROI (p = .009), and that such effects were even larger at the right posterior ROI (p = .003).

<Insert Figure 1 about here>

<Insert Figure 2 about here>

Discussion

The aims of this study were to examine the grammatical gender congruency effect in highly proficient balanced bilinguals during a bare noun translation recognition

task, and to explore the neural basis of such an effect. In order to pursue these aims, highly proficient balanced Catalan–Spanish bilinguals, who live immersed in both languages, and who acquired them in early childhood, performed a translation recognition task in which gender congruent and gender incongruent word translations were included as stimuli. Both the behavioral and the ERP results clearly show a gender congruency effect: Participants' responses to gender-congruent pairs were faster and more accurate than to gender-incongruent pairs. In addition, the mean amplitude of the N400 component was larger for the later than for the former. These results are in accordance with the predictions of faster lexical access for gender congruent stimuli with respect to gender incongruent ones, supporting the hyphotesis that the two gender systems of highly proficient bilinguals interact even in a bare noun recognition task.

The gender congruency effect in bilinguals' word production is a robust effect observed in different bilingual populations and different tasks, especially in sentence contexts (e.g., Bordag & Pechmann, 2007; Klassen, 2016; Lemhöfer et al., 2008; Paolieri et al., 2018; Salamoura & Williams, 2007). This effect has been studied to a much lesser extent in language comprehension despite the fact that it can shed light on the nature of gender representation. The limited research conducted in this area has focused mainly on the syntactic function of gender, with studies on the recognition of words presented visually or auditorily (Dahan et al., 2000; Dussias et al., 2013; Lemhöfer et al., 2008; Lew-Williams & Fernald, 2007; Morales et al., 2016; van Heugten & Johnson, 2011; van Heugten & Shi, 2009).

To the best of our knowledge, the present study is the first one to examine the gender congruency effect and its electrophysiological correlates in a bilingual population by using bare nouns in a language comprehension task. The present behavioral findings lend support to the assumption that the grammatical gender systems

of the two languages of a bilingual are integrated at the lexical level: the L1 representations activate the corresponding representations in the L2; when the two representations have the same gender, then the activation is higher than when their genders are different. In this way, word selection and word recognition are facilitated as reflected in a decrease in naming times and recognition latencies (see also Cubelli et al., 2005; Paolieri et al., 2010). The ERP results obtained in the present study clearly support this hyphotesis too. Concretely, gender-incongruent translation equivalents elicited a larger negative deflection in the 300-500 ms time window in comparison to gender-congruent pairs. This negative deflection was widely distributed across the scalp, being more prominent at central and posterior sites. Importantly, the three-way interaction between gender congruency, latitude and hemisphere reflects the typical central-posterior distribution of the N400, as can be seen in the voltage maps in Figure 2.

It is worth mentioning here that comprehension does not require lexical selection, since gender features are already assigned in the verbal utterance or written text. Therefore, the presence of gender congruency behavioral and electrophysioloical effects in a situation where the access to gender information is not required strongly supports the proposal 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; Wang et al., 2018). Indeed, a reduced N400 is typically interpreted as revealing easier lexical integration (see Kutas & Federmeier, 2011 for an overview). In accordance with this, monolingual ERP studies have shown that words disagreeing in gender elicit larger negative-going waves that those agreeing in gender when presented in minimal contexts (i.e., determiner-noun and adjective-noun pairs; Barber & Carreiras, 2005). Similarly, gender related N400 effects

have been found during the production of bare nouns, where an incongruent feature (a lexico-syntactic classifier similar to grammatical gender), elicited larger N400 effects in comparison to a congruent condition (Wang et al., 2018). Hence, the present findings seem to indicate that gender-incongruent pairs across languages are harder to integrate than congruent ones. On the other hand, regarding the mechanism underlying the effect, it should be noted that purely (morpho)syntactic violations have been more commonly associated in the literature with ERP components such as the P600 or the (early)left anterior negativity (E/LAN). For instance, studies examining gender disagreement in Spanish and Italian (Barber & Carreiras, 2005; Martín-Loeches, Nigbur, Casado, Hohlfeld, & Sommer, 2006; Molinaro, Vespignani, & Job, 2008) reported a biphasic ERP response with a LAN followed by a posteriorly distributed P600 effect. In contrast, when the incongruence is lexico-semantic, rather than syntactic, the ERP modulations are observed in the N400 time window (Osterhout & Mobley, 1995; see also Kutas & Federmeier, 2011 for an overview). Thus, the fact that gender congruency pairs produce a reduced N400 effect is consistent with the idea that the gender systems of Catalan and Spanish are integrated at a lexical level.

To conclude, the present study provides new and important evidence of cross-language grammatical gender activation, by contributing to the limited data available on comprehension, being the first study that has investigated the neural signature of grammatical gender congruency effects in bilinguals. Behavioral and ERPs results showed a clear gender congruency facilitating effect in highly proficient early Catalan-Spanish bilinguals through a bare noun translation-recognition task. These findings constitute strong evidence of the obligatory access to grammatical gender information during bare noun processing in highly proficient bilinguals, and suggest that gender information is a feature stored at a noun's lexical representation.

References

Alario, F. X., 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*, 960-981.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes.* Cambridge, MA: MIT Press.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17,* 137-153.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, *16*, 240-250.

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition, 10,* 299- 314.

Bordag, D., & Pechmann, T. (2008). Grammatical gender in speech production: Evidence from Czech. *Journal of Psycholinguistic Research, 37,* 69-85.

Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia*, *63*, 124– 134.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from the ‘tip-of-the-tongue' phenomenon. *Cognition*, *64*, 309-343.

Colé, P., Pynte, J., & Andriamamonjy, P. (2003). Effect of grammatical gender on visual word recognition: Evidence from lexical decision and eye movement experiments. *Perception & psychophysics*, *65*, 407-419.

Colé, P., & Seguí, J. (1994). Grammatical incongruency and vocabulary types. *Memory & Cognition, 22,* 387-394.

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

Corbett, G. (2011). Number of Genders. In M. Dryer & M. Haspelmath (Eds.), *The World atlas of language structures online*. Munich, Germany: Max Planck Digital Library.

Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition, 6,* 181-200.

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,* 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 (pp. 73-86). Padova, Italy: Cleup*.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42*, 465-480.

De Martino, M., Bracco, G., & Laudanna, A. (2011). The activation of grammatical gender information in processing Italian nouns. *Language and Cognitive Processes*, *26*, 745-776.

De Martino, M., Bracco, G., Postiglione, F., & Laudanna, A. (2017). The influence of grammatical gender and suffix transparency in processing Italian written nouns. *The Mental Lexicon*, *12*, 107-128.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175-197.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*, 657-679.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*, 1246-1258.

Duràn P. C., & Pillon, A., (2011). The Role of Grammatical Category Information in Spoken Word Retrieval. *Frontiers in Psychology, 2,* 16 November 2011.

Dussias, P. E., Valdés Kroff, J. R, Guzzardo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition, 35*, 353-387.

El-Yousseph, N. (2010). Gender assignment to new words in Catalan. Masteroppgave. Universitat Autònoma de Barcelona.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116-124.

Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, *28*, 467-484.

Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research, 30*, 627-651.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

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

Guasch, M., Ferré, P., & Fraga, I. (2016). Spanish norms for affective and lexico-semantic variables for 1,400 words. *Behavior Research Methods*, *48*, 1358-1369.

Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. (2012). On the time course of accessing meaning in a second language: an electrophysiological and behavioral investigation of translation recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1165–1186.

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22*, 27-62.

Klassen, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus*, *28*, 9-28.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*, 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*, 159-163.

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*, 149-174

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.

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,* 312-330.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*, 193-198.

Lloret, M., & Viaplana, J. (1997). On the morphological category of gender in Catalan and in Spanish. *Advances in Morphology, 97,*171.

Luck, S. J. (2005). Ten simple rules for designing ERP experiments. *Event-Related Potentials: A Methods Handbook*, *262083337*.

Ma, F., Chen, P., Guo, T., & Kroll, J. F. (2017). When late second language learners access the meaning of L2 words : Using ERPs to investigate the role of the L1 translation equivalent. *Journal of Neurolinguistics*, *41*, 50–69.

Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., & Sommer, W. (2006). Semantics prevalence over syntax during sentence processing: A brain potential study of noun–adjective agreement in Spanish. *Brain Research, 1093,* 178-189.

Moldovan, C. D., Demestre, J., Ferré, P., & Sánchez-Casas, R. (2016). The role of meaning and form similarity in translation recognition in highly proficient balanced bilinguals: A behavioral and ERP study. *Journal of Neurolinguistics*, *37*, 1–11.

Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research*, *1228*, 161-176.

Morales, L., Paolieri, D., Dussias, P., Valdés Kroff, J., Gerfen, C., & Bajo, T., (2016). The Gender Congruency Effect During Bilingual Spoken-Word Recognition. *Bilingualism: Language and Cognition, 19,* 294-310.

Nurse, D. & Philippson, G. (2003). Towards a historical classification of the Bantu languages. The Bantu Languages ed. by Derek Nurse & Gérard Philippson, 164–181. London: Routledge

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*, 739-773.

Paolieri, D., Cubelli, R., Macizo, P., Bajo, T., Lotto, L., & Job, R. (2010). Grammatical Gender Processing in Italian and Spanish Bilinguals . Q*uarterly Journal of Experimental Psychology, 63*, 1631-45.

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,* 19-36.

Paolieri, D., Padilla, F., Koreneva, O., Morales L., & Macizo, P. (2018). Gender congruency effect in second language production: Evidence from Russian-Spanish bilinguals. *Bilingualism: Language and Cognition*. 1-18.

Pivik, R. T., Broughton, R. J., Coppola, R., Davidson, R. J., Fox, N., & Nuwer, M. R. (1993). Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology*, *30*, 547–558.

Sá-Leite, A. R., Fraga, I., & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review*, 1-26

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,* 257-275.

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A. P., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, *49*, 111-123.

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*, 387-422.

van Heugten, M., & Johnson, E. (2011). Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. *Journal of Child Language, 38*, 87-100.

van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science, 12*, 419-425.

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*, 575-586.

Wicha, N. Y., Orozco-Figueroa, A., Reyes, I., Hernandez, A., Gavaldón de Barreto, L., & Bates, E. A. (2005). When zebras become painted donkeys: Grammatical gender and semantic priming interact during picture integration in a spoken Spanish sentence. *Language and Cognitive Processes, 20,* 553-587.

Appendix A.

Critical stimuli (Catalan and Spanish) and their translation equivalents in English.

Condition	Gender	Catalan	Spanish	*English*
InCon.	m/f	ajut	ayuda	*help*
InCon.	m/f	núvol	nube	*cloud*
InCon.	m/f	dubte	duda	*doubt*
InCon.	m/f	queixal	muela	*tooth*
InCon.	m/f	sopar	cena	*dinner*
InCon.	m/f	berenar	merienda	*snack*
InCon.	m/f	greix	grasa	*grease*
InCon.	m/f	cim	cima	*top*
InCon.	m/f	llit	cama	*bed*
InCon.	m/f	nas	nariz	*nose*
InCon.	m/f	front	frente	*front*
InCon.	m/f	bolet	seta	*mushroom*
InCon.	m/f	carrer	calle	*street*
InCon.	m/f	cap	cabeza	*head*
InCon.	m/f	raïm	uva	*grape*
InCon.	m/f	avantatge	ventaja	*advantage*
InCon.	m/f	enciam	lechuga	*lettuce*
InCon.	m/f	somriure	sonrisa	*smile*
InCon.	m/f	genoll	rodilla	*knee*
InCon.	f/m	gana	hambre	*hungry*
InCon.	f/m	llimona	limón	*lemon*
InCon.	f/m	tardor	otoño	*autumn*
InCon.	f/m	por	miedo	*fear*
InCon.	f/m	abraçada	abrazo	*hug*
InCon.	f/m	salutació	saludo	*greeting*
InCon.	f/m	dada	dato	*fact*
InCon.	f/m	estona	rato	*little while*
InCon.	f/m	sabata	zapato	*shoe*
InCon.	f/m	vall	valle	*valley*
InCon.	f/m	disfressa	disfraz	*costume*
InCon.	f/m	pols	polvo	*powder*
InCon.	f/m	forquilla	tenedor	*fork*
InCon.	f/m	dent	diente	*tooth*
InCon.	f/m	xocolata	chocolate	*chocolate*
InCon.	f/m	figa	higo	*fig*
InCon.	f/m	joguina	juguete	*toy*
InCon.	f/m	butxaca	bolsillo	*pocket*
InCon.	f/m	trobada	encuentro	*meeting*
Con.	m/m	ferro	hierro	*iron*
Con.	m/m	ou	huevo	*egg*
Con.	m/m	colze	codo	*elbow*

Con.	m/m	soroll	ruido	*noise*
Con.	m/m	amagatall	escondite	*hiding place*
Con.	m/m	patiment	sufrimiento	*suffering*
Con.	m/m	parc	parque	*park*
Con.	m/m	abric	abrigo	*coat*
Con.	m/m	arbre	árbol	*tree*
Con.	m/m	turmell	tobillo	*ankle*
Con.	m/m	nus	nudo	*knot*
Con.	m/m	assaig	ensayo	*test*
Con.	m/m	aparell	aparato	*apparatus*
Con.	m/m	menjador	comedor	*dinning room*
Con.	m/m	cop	golpe	*knock*
Con.	m/m	pit	pecho	*chest*
Con.	m/m	molí	molino	*windmill*
Con.	m/m	estiu	verano	*summer*
Con.	m/m	pantà	pantano	*swamp*
Con.	f/f	cadira	silla	*chair*
Con.	f/f	fulla	hoja	*sheet*
Con.	f/f	llet	leche	*milk*
Con.	f/f	poma	manzana	*apple*
Con.	f/f	neu	nieve	*snow*
Con.	f/f	pluja	lluvia	*rain*
Con.	f/f	cantonada	esquina	*corner*
Con.	f/f	força	fuerza	*force*
Con.	f/f	taula	mesa	*table*
Con.	f/f	amistat	amistad	*friendship*
Con.	f/f	samarreta	camiseta	*t-shirt*
Con.	f/f	cendra	ceniza	*ash*
Con.	f/f	pèrdua	pérdida	*lost*
Con.	f/f	roda	rueda	*wheel*
Con.	f/f	botiga	tienda	*store*
Con.	f/f	xemeneia	chimenea	*chimney*
Con.	f/f	emoció	emoción	*emotion*
Con.	f/f	dreta	derecha	*right*
Con.	f/f	ampolla	botella	*bottle*

Note. InCon. = Incongruent; Con = Congruent; m = masculine ; f = feminine

Table 1. *Mean and standard deviation (in parentheses) of participants' self-rated proficiency in Catalan and Spanish*

Skills	**Catalan**	**Spanish**
Listening	7.00 (0.00)	7.00 (0.00)
Speaking	6.97 (0.18)	6.62 (0.54)
Reading	7.00 (0.00)	6.97 (0.18)
Writing	6.90 (0.30)	6.69 (0.58)
Average	**6.97 (0.18)**	**6.82 (0.44)**

Note. The anchor points for the 7-point Likert scale were 1 = 'very poor', and 7 = 'native-like'.

Table 2. *Characteristics of the critical stimuli (means and standard deviation in parentheses)*

Variable	Gender-congruent translations	Gender-incongruent translations
Translation Agreement	96.45% (6.13%)	94.72% (5.79%)
NLD	0.41 (0.21)	0.38 (0.22)
Catalan Frequency	59.57 (76.54)	98.18 (317.53)
Spanish Frequency	36.27 (31.16)	62.04 (93.19)
Catalan Length	5.53 (1.78)	5.68 (1.86)
Spanish Length	6.16 (1.44)	5.63 (1.50)
Valence	5.54 (1.52)	6.00 (1.54)
Arousal	5.03 (1.12)	4.87 (1.00)
Familiarity	6.01 (0.54)	6.09 (0.53)
Imageability	5.74 (0.98)	5.56 (1.12)
Concreteness	5.34 (1.02)	5.27 (1.06)

Note. All *p*s > .112. NLD = Normalized Levenshtein Distance between translations; Valence and Arousal rated in a 9-point Likert scale; Familiarity, Imageability, and Concreteness rated in a 7-point Likert scale.

Table 3. *Mean reaction times (in ms) and error rates (in %) in the two experimental conditions (SDs in parentheses)*

Condition	Mean RTs	Error Rates
Gender-congruent translations	631 (126)	2.6 (3.3)
Gender-incongruent translations	655 (131)	4.6 (4.3)

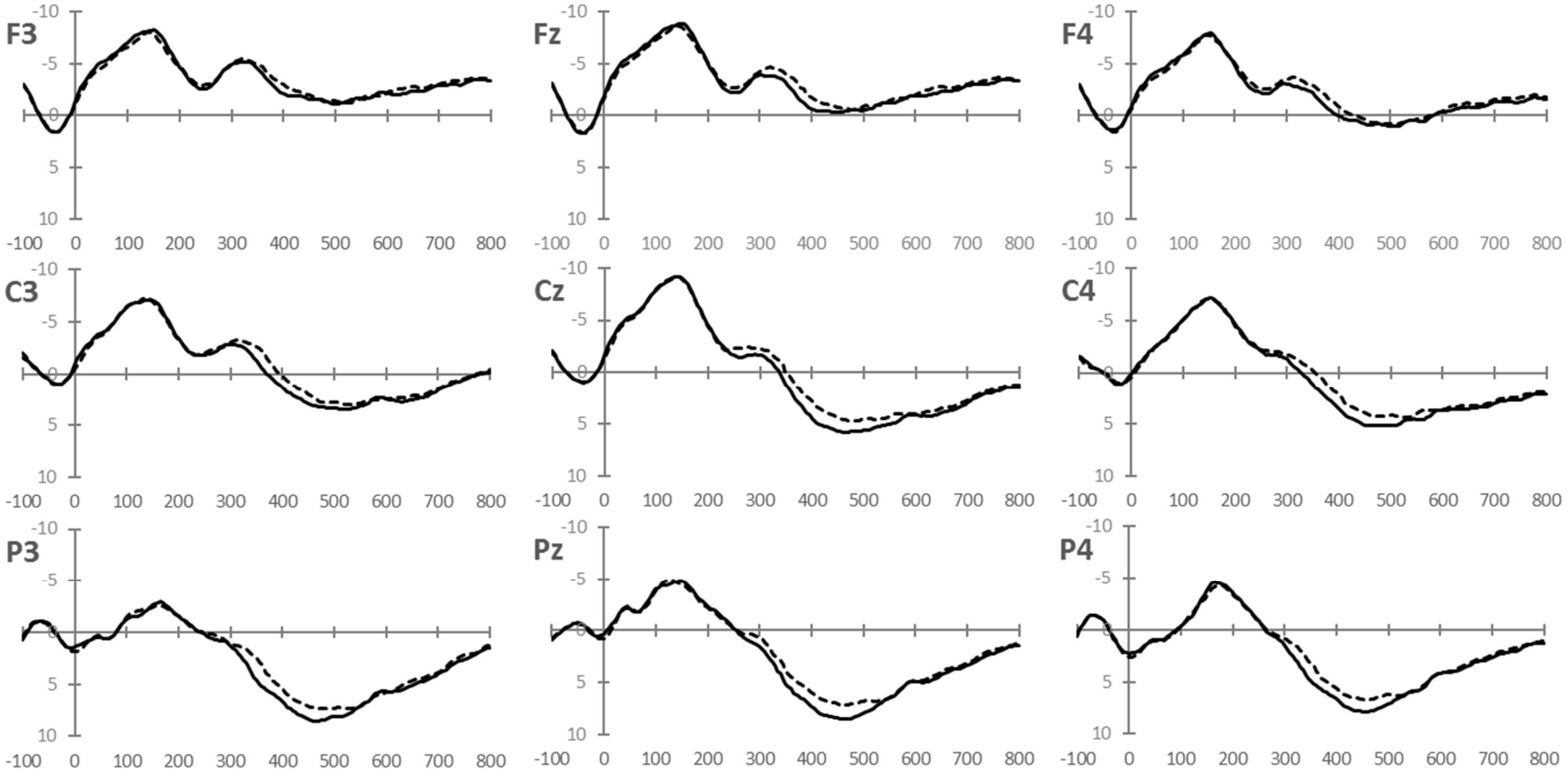

Figure 1. Grand average ERPs in microvolts (N=31) from nine representative electrodes for the two conditions of the gender congruency factor. ERPs are time-locked to the onset of the critical word with a -100ms to 0ms prestimulus interval. The black solid line refers to gender-congruent words and the black dotted line refers to gender-incongruent words. Negative values are plotted up.

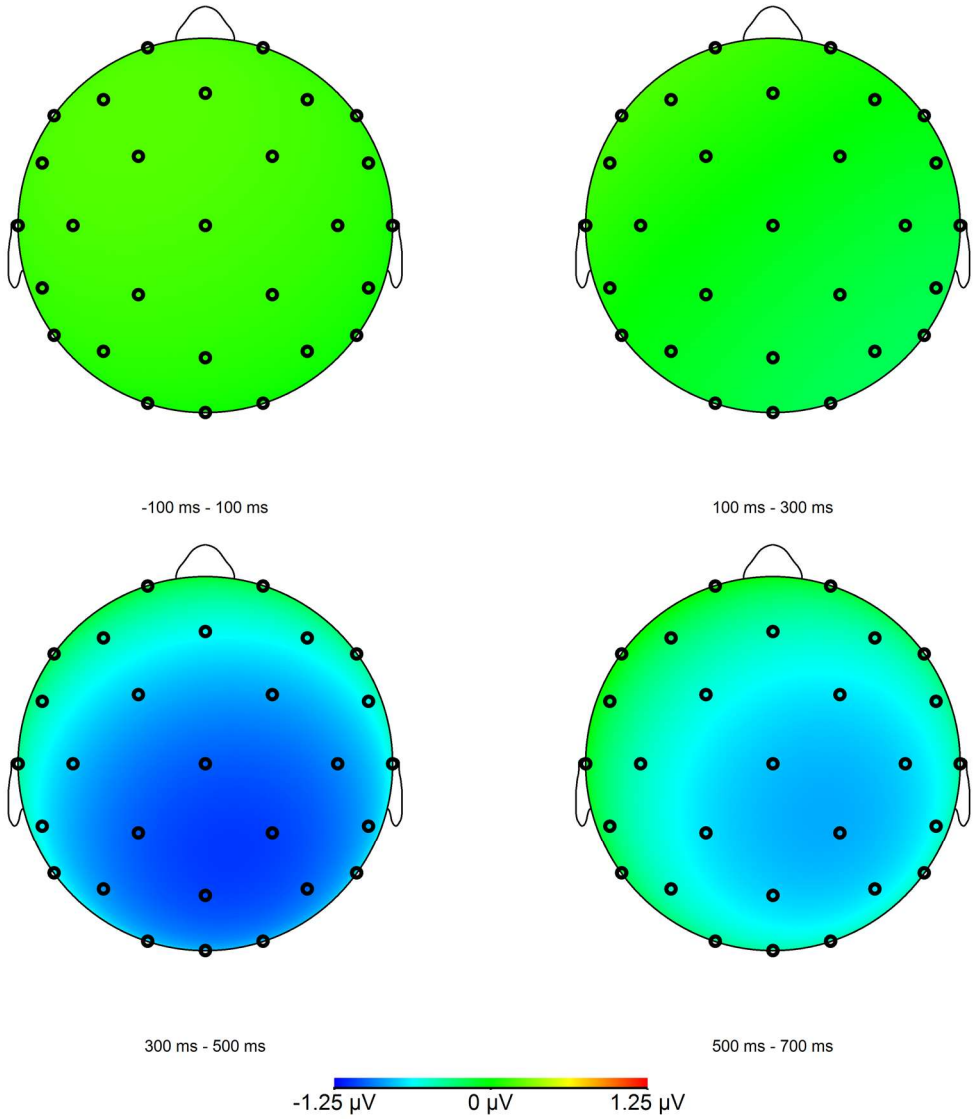

Figure 2. Topographical distribution of the gender congruency effect (gender-incongruent minus gender-congruent) in four time intervals (-100-100 ms, 100-300 ms, 300–500 ms and 500-700 ms).