

**When the End Matters: Influence of Gender Cues during Agreement Computation in
Bilinguals**

Sendy Caffarra^{a*}, Horacio Barber^{a,b}, Nicola Molinaro^{a,c}, Manuel Carreiras^{a,c}

^a BCBL, Basque Center on Cognition, Brain and Language, Mikeletegi 69, 20009, Donostia, Spain; s.caffarra@bcbl.eu, n.molinaro@bcbl.eu, m.carreiras@bcbl.eu; ^b University of La Laguna, Institute for Biomedical Technologies (ITB), Campus de Guajara, 38205, Tenerife, Spain; hbarber@ull.es; ^c Ikerbasque, Basque Foundation for Science, Maria Diaz de Haro 3, 48013, Bilbao, Spain

*Corresponding author. Email: s.caffarra@bcbl.eu

Sendy Caffarra, BCBL, Paseo Mikeletegi 69, 20009 Donostia-San Sebastian, Spain.

tel.: +34 943 309 300

This work was supported by the Spanish Ministry [PSI 2014-54500-P, PSI2015-65694-P]; the Basque Government [PI_2015_1_25]; the Severo Ochoa [SEV-2015-0490]; and the AThEME project [FP7-SSH-2013-1-GA613465].

Abstract

The present ERP study was aimed at testing whether form-function mappings can differently affect sentence comprehension in early bilinguals with a range of linguistic profiles. Basque-Spanish and Spanish-Basque early bilinguals were presented with Spanish sentences with article-noun gender agreement violations. The gender of the target noun could be retrieved based on the word-form (i.e., transparent nouns) or only on a lexical representation (i.e. opaque nouns). While Basque-dominant bilinguals showed an impact of gender-to-ending consistency on agreement computation, Spanish-dominant bilinguals' agreement processing was not affected by form-function mappings. A multiple regression analysis on early ERP responses from all participants showed that the more Spanish was produced on a daily basis, the easier the detection of gender violation for opaque nouns. The present results suggest that the strength of the lexical representation of gender is not fixed and can change depending on the linguistic habits of early bilinguals.

Keywords: grammatical gender; gender-to-ending consistency; agreement; bilingualism

Introduction

Learning and effectively computing the grammatical gender system of a new language frequently represents a challenge for bilinguals (Franceschina, 2005; Dewaele & Véronique, 2001; Holmes & Dejean de la Bâtie, 1999; Montrul & Potowski, 2007; Unsworth, 2008). One of the factors that can facilitate gender learning and processing is the availability of distributional regularities (Caselli, Leonard, Volterra, & Campagnoli, 1993; Henzl, 1975; Holmes & Dejean de la Bâtie, 1999; Mills, 1986; Matthews, 2010; Pérez-Pereira, 1991). Some authors have also highlighted the potential influence of already-known languages on the way new gender assignment rules are acquired (Dechert & Raupach, 1989; Franceschina, 2005; Gass, 1984; Gass & Selinker, 1992; MacWhinney, 2005, 2012; Odlin, 1989; White, 2003). However, it is still unclear whether knowledge and usage of other languages can influence the way in which people treat the distributional features of a given grammatical gender system. The present ERP study investigates whether bilinguals differently compute distributional regularities during gender agreement computation, depending on their linguistic knowledge and habits.

Probabilistic patterns in language use have been considered particularly useful for language acquisition and comprehension (MacWhinney, 1987; Romberg & Saffran, 2011). Extracting a systematic mapping between the formal level and the functional level in language use facilitates grammar learning (Clair, Monaghan, & Christiansen, 2010; Gerken, Wilson, & Lewis, 2005; Pérez-Pereira, 1991) and helps detection of role assignment, co-reference and argument attachment (MacWhinney, 1987; MacWhinney, Bates, & Kliegl, 1984). The stronger the probabilistic form-function mapping, the more easily can people detect and rely on this (Mills, 1986). According to this view, the strength of a form-function mapping would depend on

its availability (i.e., number of linguistic items providing a consistent form-function correspondence in a given language) and on its reliability (i.e., how consistent is the relation between the formal and the functional level in a given language). When a probabilistic pattern covers a high number of items (i.e., high availability) and has few or no exceptions (i.e., high reliability), it provides a strong distributional cue, which is considered to be influential in language acquisition and comprehension (MacWhinney, 1987, 2004).

However, the strength of a distributional cue does not seem to be fixed and strictly determined on the basis of probabilistic patterns within a specific language. Some evidence would suggest that the validity of the same cue can change across languages (MacWhinney et al., 1984) and can be influenced by the characteristics of other known languages (Scheutz & Eberhard, 2004). Within the second language processing literature, some authors have already emphasized the effects of potential interactions between different known languages (Gass, 1984; MacWhinney, 2005, 2012; White, 2003). Although a few studies have confirmed the potential influence of previously-learned languages on grammar acquisition and computation, it is still unclear whether people treat the distributional regularities of a given language differently depending on their language profile. We will address this question by examining gender-to-ending consistency in Spanish and its impact on agreement computation. Specifically, we will test whether early bilinguals change their sensitivity to Spanish distributional gender cues depending on their linguistic habits.

Grammatical gender cues in monolinguals and bilinguals

The grammatical gender of a noun is considered as a morphosyntactic feature, abstractly stored in the lexicon (Levelt, Roelofs, & Meyer, 1999). Across different languages, gender can entail consistent relations with the meaning and/or with the word-form of a noun (Corbett, 1991).

In Romance languages, it is possible to observe a consistent correspondence between a specific noun ending and a gender class. In the case of Spanish, for example, (where there are two gender classes: masculine and feminine) the noun endings “-a” and “-o” represent a highly available and reliable distributional cue to gender (66% of Spanish singular nouns show these endings; almost 100% of singular Spanish nouns ending in “-o” are masculine and 96% of singular nouns ending in “-a” are feminine, Harris, 1991). Transparent nouns have an ending that represents a valid gender cue (e.g., *casa*, house_F), while opaque nouns have an ending that is not informative of grammatical gender (e.g., *verniz*, paint_F; Bates, Devescovi, Pizzamiglio, D’Amico & Hernandez, 1995).

The impact of gender cues has been widely investigated in behavioural studies presenting words or word pairs outside a sentence context (Italian: Bates et al., 1995; Bates, Devescovi, Hernandez, & Pizzamiglio, 1996; De Martino, Bracco, & Laudanna, 2010; French: Spalek, Franck, Schriefers, & Frauenfelder, 2008; Taft & Meunier, 1998; Hebrew: Gollan & Frost, 2001; Spanish: Afonso, Domínguez, Álvarez, & Morales, 2014; Hernandez, Kotz, Hofmann, Valentin, Dapretto, & Bookheimer, 2004; German: Schiller, Münte, Horemans, & Jansma, 2003). Results from languages with a strong gender-to-ending consistency have shown that formal gender cues can influence gender access. For instance, when monolinguals have to perform an explicit task on gender (i.e., gender decision task, determiner retrieval, grammatically judgment on word pairs) they are usually faster and/or more accurate with isolated transparent nouns compared to opaque nouns (auditory modality: Bates et al., 1995; Bates et al., 1996; Spalek et al., 2008; visual modality: Afonso et al., 2014; Gollan & Frost, 2001; Hernandez et al. 2004; Schiller et al. 2003; Taft & Meunier, 1998).

However, the weight of these formal gender cues can change depending on the type of task employed and on the type of processes investigated. When the target nouns are presented in isolation and the task is not focused on gender, the effect of gender-to-ending consistency is less consistently reported (studies on monolinguals: Bates et al., 1995, 1996; Padovani & Cacciari, 2003, but see De Martino et al., 2010). In addition, when both gender-to-ending consistency and gender agreement are manipulated, behavioural studies have shown that available formal gender cues do not influence the processing of agreement dependencies in isolated word pairs (Bates et al., 1996). The authors interpreted these data as suggesting that although the distributional information conveyed by noun endings can facilitate gender retrieval, it represents redundant (and sometime ambiguous or misleading) information, which might not be essential for gender access and agreement computation (Delfitto & Zamparelli, 2009; Vigliocco & Franck, 1999). According to this view, adult monolinguals mainly use an abstract representation of the gender feature stored in the lexicon, since it is more reliable than the information conveyed by the word form level (Delfitto & Zamparelli, 2009).

If we look at the case of monolingual children and bilinguals, the weight of formal cues to gender seems to be much more relevant as compared to the case of monolingual adults. In languages where there is a strong gender-to-ending consistency, developmental studies have shown that children are highly sensitive to formal gender cues and they usually know the grammatical gender of transparent nouns better than those of opaque nouns (Caselli et al., 1993; Chini, 1995; Mills, 1986; Matthews, 2010; Pérez-Pereira, 1991). In addition, behavioural studies comparing highly proficient English-German bilinguals and German monolinguals on the same tasks (i.e., picture naming, grammaticality judgment) have shown that formal gender cues usually have a stronger impact on bilingual performance as compare to that of monolinguals

(Bordag, Opitz, & Pechmann, 2006; here German monolinguals did not show a clear effect of gender-to-ending consistency, in contrast to Schiller et al., 2003). The authors interpreted these results as suggesting that bilinguals (whose dominant language lacks a grammatical gender system) might not have stored gender as a fixed feature of the noun but rather compute gender each time it is needed on the basis of any available cues (Bordag et al., 2006; see also Bordag & Pechmann, 2007, 2008). In contrast to monolinguals, bilinguals' lexical representation of grammatical gender can be weak and unstable, with concomitant delays in gender access (for similar conclusions see also Hopp, 2013; Lemhofer, Schriefers, & Hanique, 2010) and an over-reliance on formal gender cues (Bordag et al., 2006).

We conducted an ERP study on Spanish gender agreement processing in order to test whether the strength of gender lexical representation and the weight of formal gender cues would change depending on the linguistic habits of highly proficient bilinguals.

The time course of gender agreement processing

The ERP (Event-Related Potentials) technique enables collection of fine-grained temporal information about linguistic mechanisms. This technique has been widely employed in order to investigate the time course of gender agreement processing both in monolinguals and bilinguals. However, less is known about the processing of distributional regularities of grammatical gender. ERP studies focussing on agreement processing in sentence comprehension have typically shown that gender violations elicit a greater posterior positivity around 500 ms after stimulus onset (i.e., P600, monolinguals: Barber & Carreiras, 2005; Barber, Salillas, & Carreiras, 2004; Deutsch & Bentin, 2001; Gunter, Friederici, & Schriefers, 2000; Hagoort, 2003; Hagoort & Brown, 1999; Molinaro, Vespignani, & Job, 2008; Wicha, Moreno & Kutas, 2004; highly proficient bilinguals: Bañon, Fiorentino, & Gabriele, 2014; Dowens, Guo, Guo, Barber, &

Carreiras, 2011; Dowens, Vergara, Barber, & Carreiras, 2010; Rossi, Kroll & Dussias, 2014), as compared to the corresponding correct sentence. This positive effect has been interpreted as reflecting processes of syntactic re-analysis and repair (Friederici, 2002; Molinaro, Kim, Vespignani, & Job, 2008; Molinaro, Vespignani, Zamparelli, & Job, 2011), monitoring (Van de Meeredonk, Kolk, Chwilla, & Vissers, 2009), or higher-level integration (Brouwer, Fitz, & Hoeks, 2012). With local agreement violations that do not involve deeper lexical-semantic processes, the P600 can frequently be preceded by a greater left anterior negativity (monolinguals: Barber & Carreiras, 2005; Barber et al., 2004; Gunter et al., 2000; Molinaro et al., 2008; highly proficient bilinguals: Dowens et al., 2010). The LAN effect has been interpreted as reflecting processes of initial detection of morphosyntactic violations (Friederici, 2002), initial difficulties in integrating mismatching information (Gunter et al., 2000), or increased working memory demands (Coulson, King & Kutas, 1998). When languages with a strong gender-to-ending consistency have been studied (e.g., Romance languages), most of the ERP studies focussing on gender agreement tested transparent nouns (Barber & Carreiras, 2005; Barber et al., 2004; Dowens et al., 2010, 2011; but see Molinaro et al., 2008). Few ERP studies have manipulated both gender agreement and gender-to-ending consistency of the target nouns (Caffarra & Barber, 2015; Caffarra, Siyanova-Chanturia, Pesciarelli, Vespignani, & Cacciari, 2015). Some data from Spanish and Italian monolinguals showed that article-noun gender violations elicited a LAN-P600 response which was not modulated by the transparency of the target noun (for similar behavioural results see Bates et al., 1996). There was a main effect of noun transparency, with a greater sustained negativity (around 200 ms after the stimulus presentation) for transparent nouns as compared to opaque and irregular nouns. This effect suggests that during sentence reading monolinguals can detect distributional properties of the

noun ending, but this information does not influence agreement computation. The authors interpreted this lack of interaction as suggesting that native speakers of Romance languages compute gender agreement dependencies mainly relying on the lexical representation of gender, regardless of available cues at the word-form level.

The present study

The present ERP study will investigate the impact of formal gender cues on the time course of agreement computation during bilingual sentence reading. Spanish article-noun pairs, embedded in sentences, could agree or not in gender. Also, the target noun could provide a reliable cue to gender (i.e., transparent noun) or not (i.e., opaque noun). Basque-Spanish highly proficient bilinguals had to judge the grammaticality of the sentences. As in Bordag et al. (2006), the bilinguals' dominant language lacked a grammatical gender system (in Basque, as in the case of English, nouns are not categorized based on grammatical gender classes). Following previous studies (Bordag et al., 2006; Hopp, 2013; Lemhofer et al., 2010), we hypothesized that highly proficient Basque-Spanish bilinguals might have an unstable lexical representation of Spanish gender and should be sensitive to gender-to-ending consistency during agreement processing. In contrast to what has been observed in Spanish monolinguals (Caffarra & Barber, 2005), we predicted that noun transparency would influence bilingual agreement processing, with different ERP responses for gender violations of transparent and opaque nouns.

In the case of transparent nouns, gender should be easily retrieved on the basis of probabilistic form-function mappings, and gender violations should elicit LAN-P600 responses, similar to those already reported for article-noun gender mismatches in monolinguals (Barber & Carreiras, 2005; Gunter et al., 2000; Molinaro et al., 2008) and in highly proficient English-Spanish bilinguals (Dowens et al., 2010). In the case of opaque nouns, detecting and computing

agreement violation should be more difficult, since it requires an entrenched abstract representation of gender stored in the lexicon. If Basque-Spanish bilinguals have a weak lexical representation of grammatical gender, the ERP responses to gender violations of opaque nouns should be reduced or at least delayed as compared to transparent nouns (Bordag et al., 2006; Hopp, 2013; Lemhofer et al., 2010).

Experiment 1: material and methods

Participants

Twenty-four Basque-Spanish bilinguals took part in the experiment (15 women) and were paid for their participation (20 €). Ages ranged from 18 to 40 years with a mean of 24.71 years (SD: 6.33). All participants were born and lived in the Basque Country, an area which provides multiple social settings where Basque and Spanish are highly available and used. The participants reported having been exposed to both Spanish and Basque since birth and that they started to learn both languages early in life (before 12 years old). Language assessment was performed before the experimental session and consisted of a semi-structured interview with a Basque native speaker, a semi-structured interview with a Spanish native speaker, a Basque picture naming test and a Spanish picture naming test. During the interview, participants had to answer questions about three main topics: introduce yourself (i.e., who, where are they from, what did they study, why), hobbies (i.e., sports, music, art, dance, etc.), spatial localization (i.e., how did you get here?). Based on their responses, a score from one to five was assigned: one for a total lack of knowledge (i.e., the participant knows some words, but he/she is unable to create a sentence); two for a basic level (i.e., the participant can produce simple sentences but with a lot of difficulties and mistakes); three for an intermediate level (i.e., the participant is able to have

simple conversations but with some mistakes; he/she shows difficulties and low fluency with long sentences); four for an advanced level (i.e., the participant is able to have conversations on a wide range of topics; he/she can produce long sentences with high fluency but there are still some mistakes); five for an extremely high proficiency level (i.e., the participant can talk easily about everything without errors). During the naming test, participants were asked to name 77 colored pictures (including animals, body parts and man-made objects selected from Snodgrass & Vanderwart, 1980) in Basque and Spanish. The number of correctly-named pictures was the final naming score. This language assessment showed that participants were highly proficient in both languages (Spanish mean scores at the interview: 5/5 and at the naming test: 76/77; Basque mean scores at the interview: 5/5 and at the naming test: 75/77, see Table 1). All participants reported being dominant in Basque, since it was the language they started to speak with their parents (AoA of Basque: 0; see Table 1) and the most frequently used language on a daily basis (self-reported amounts of Basque produced and heard on a daily basis: 63% and 55%). In addition, they reported using Basque with most of their relatives and friends. They started to use Spanish only after being enrolled in kindergarten or school (mean AoA of Spanish: 5) and they reported using it in their adult life, especially during formal occasions and with non-Basque speaking people. All participants were right-handed (based on the Edinburgh Handedness Inventory; Oldfield, 1971) and reported normal or corrected-to-normal vision. None of the participants had a history of neurological disorder.

(Table 1 about here)

Materials

Two hundred and forty Spanish nouns were used (length range: 4–8 letters; see Caffarra & Barber, 2015). Half of them were transparent (60 masculine nouns; e.g. *lago*, “lake_m”, and 60

feminine nouns; e.g., *playa*, “beach_f”) and half were opaque (60 masculine nouns; e.g., *reloj*, “watch_m”, and 60 feminine nouns; e.g., *flor*, “flower_f”). The ending of transparent nouns was highly predictive of a specific gender class (i.e., “-o” for masculine and “-a” for feminine), while the ending of opaque nouns was uninformative of grammatical gender (i.e., “-e”, “-n”, “-l”, “-s”, “-j”, “-r”, “-d”, “-z”; final groups of letters strongly associated with a specific gender class were excluded, e.g., “-sis”, “-tis”, “-í”, “-briz”, “-triz”, “-driz”, “-ie”, “-umbre”; according to Bull, 1965; Clegg, 2010). The cue availability (i.e., how often a specific cue is available in the lexicon, MacWhinney et al., 1984) and the cue reliability (i.e., how strong the gender-to-ending correspondence is, MacWhinney et al., 1984) of each opaque ending were always lower than those of transparent endings (cue availability: 70% of feminine Spanish nouns end in –a; 62% of masculine Spanish nouns end in –o; cue reliability: 96% of the Spanish nouns ending in –a are feminine; and 99% of the Spanish nouns ending in –o are masculine; Harris, 1991). Irregular nouns (where the endings “-a” and “-o” represent misleading cues to grammatical gender; e.g., *mano*, “hand_f”, *problemaa*, “problem_m”) were not used. In addition, nouns with biological gender (i.e., referring to human beings or animals where the sex of the referent is clear; e.g., *abuela*, “grandmother_f”) were not used in order to avoid any interaction between the grammatical gender and the conceptual information of the sex of the referent (Vigliocco & Frank, 1999). Transparent and opaque nouns were matched for number of letters, AoA, frequency, familiarity, imageability, concreteness and number of Basque-Spanish cognates (all $ps > .05$, see Table 2).

(Table 2 about here)

These 240 nouns were paired with a definite article that could either agree or not in grammatical gender. Each word pair was embedded in a sentence context (7 words on average, SD:1, range: 4-11). The target noun position varied across trials but it always appeared at least

two words before the end of the sentence and at least one word after the beginning of the sentence (50% of the time it appeared in the second position, 36% in the fourth and 14% in the fifth). Sixteen Spanish-Basque early bilinguals who did not participate in the EEG experiment rated the cloze probability of the target nouns (12 women; mean age: 23.63 years; SD: 6.22). Overall, participants were not able to guess the selected target noun, since it frequently appeared in initial positions of the sentence (86% of the time it appeared between the second and the fourth position). As a result, transparent and opaque nouns showed equally low cloze probability ratings (transparent nouns: CP: 0 %, SD: 0; opaque nouns CP= .01%, SD: .04; $p>.05$).

Each experimental sentence could belong to one of the following four conditions: Transparent Agreement, Transparent Disagreement, Opaque Agreement, Opaque Disagreement (60 trials per condition, see Table 3). Each sentence was assigned to one of the four conditions in a counterbalanced order, so that no participant saw the same version of the sentence in different experimental conditions.

One hundred and twenty filler sentences were also presented (7 words on average, SD: 2, range: 4-12), half of them were grammatically correct and the other half included different types of grammatical errors (i.e., number disagreement, substitution or omission of function words, word category violations) in order not to make participants exclusively focus on article-noun gender agreement violations. The target word presented in the fillers frequently appeared between the second and the fourth position of the sentence (74% of the times) and it was never at the end. Overall, participants were presented with 360 sentences during the EEG session.

(Table 3 about here)

Procedure

Participants were facing the computer screen in a darkened, silent room. The experimental session was divided into two phases: an online sentence comprehension study and an offline gender decision task.

During the first phase, the EEG was recorded and participants were required to perform a grammatically judgment task on a series of sentences. The sentences were visually presented and the presentation order was random. At the beginning of each trial, a fixation cross was displayed in the centre of the screen for 2000 ms and this was followed by a blank screen for 300 ms. The sentences were presented word by word. Each word was in white letters on a black background and appeared for 300 ms, followed by a blank screen for 300 ms. The final word was displayed with a full stop and after the sentence offset a question appeared on the screen asking whether the sentence was correct. Participants had 2000 ms to give their response by pressing one of the two response buttons. The button position was counterbalanced across subjects. The next trial was presented after an interval that could vary randomly between 300 ms and 700 ms. At the beginning of the recording session, subjects were advised to blink during the presentation of the fixation cross in order to reduce the probability of eye movements during the critical epochs. A practice session preceded the real experiment in order to familiarize participants with the task. The EEG session lasted almost an hour and it was divided in four blocks of 90 trials each.

The second phase always followed the EEG recording and was aimed at testing the participants' knowledge of Spanish grammatical gender. All target nouns previously embedded in sentences (n : 240) were now presented in isolation. The participants had to perform an offline gender decision task on single nouns. At the beginning of each trial a fixation cross appeared in the center of the screen for 800 ms, followed by a blank of 300 ms. Then a single noun appeared

on the screen until participants' response or for a maximum of 3000 ms. The participants were asked to decide as quickly as possible whether the word was feminine or masculine by pressing one of the two response buttons. Accuracy rates and reactions times (RTs, measured from the onset of the target noun) were calculated.

Electrophysiological Recordings and Analyses

The electroencephalogram (EEG) was recorded from 27 electrodes placed in an elastic cap: Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P3, P4, P7, P8, O1, O2, Fz, Cz, Pz. Two external electrodes were placed on mastoids and four were placed around the eyes (two on the ocular canthi, one above and one below the right eye). All sites were referenced online to the left mastoid. Data were recorded and amplified at a sampling rate of 500 Hz. Impedance was kept below 5 K Ω for the electrodes on the scalp, and below 10 K Ω for the external channels. EEG recordings were re-referenced offline to the average activity of the two mastoids. Then, the data were filtered offline with a bandpass of 0.01–30 Hz (24 dB/oct). Artefacts exceeding ± 100 μ V in amplitude were rejected. For each target word, an epoch of 1200 ms was obtained including a 200-ms prestimulus baseline. For each condition, average ERP waveforms, time locked to the onset of the target noun, were computed only on the trials followed by a correct response. On average, 13.9 % of trials were excluded (due to artefacts or incorrect responses) and the number of rejections did not differ across conditions ($F(3,69) = 1.38, p = 0.26$).

Statistical analyses were carried out on different time windows, which were selected in order to check for the presence of LAN and P600 effects (LAN: 350–500; P600: 500–700; 700–900 according to Barber & Carreiras, 2005; Caffarra & Barber, 2015; Schacht et al., 2010). Nine cluster electrodes were taken into account, representing the mean amplitude of three sites in

adjacent positions: left anterior (F3, F7, FC5), left central (C3, T7, CP5), left posterior (P3, P7, O1), medial anterior (Fp1, Fp2, Fz), medial central (FC1, FC2, Cz), medial posterior (CP1, CP2, Pz), right anterior (F4, F8, FC6), right central (C4, T8, CP6), right posterior (P4, P8, O2). These clusters represented nine levels of a topographic factor named “Cluster”. A repeated measures analysis of variance (ANOVA) was performed for each time window, including the two experimental factors (Transparency and Agreement) and the additional topographic factor (Cluster) as within-subject factors. The Greenhouse-Geisser procedure was applied where the sphericity assumption was violated. The p values derived from multiple post-hoc tests were adjusted by using Hochberg’s procedure (1988). Effects of topographic factors are reported when they interact with the experimental factors.

Results

Behavioural results: offline gender decision task

Participants knew the grammatical gender of most of the target nouns (average accuracy: 96% , SD: 3%, range: 93%–99%) and they showed mean RTs of 790 ms (SD: 140 ms).

Accuracies and RTs from accurate trials were analysed using a one-way ANOVA, including Transparency as a two-level factor (Transparent, Opaque). These analyses revealed that transparent nouns showed higher accuracy and faster responses than opaque nouns (Accuracy: $F(1,23)=23.99, p<.001$; RTs: $F(1,23)=40.22, p<.001$, see Table 4).

(Table 4 about here)

Behavioural results: online grammaticality judgment

Participants performed the grammaticality judgment task with a mean accuracy of 96% (SD: 4). This high score suggested that participants could easily detect grammatical errors in

almost all the sentences (the four experimental conditions showed average accuracy above 94%). A two-way ANOVA including Transparency and Agreement as two within-subject factors showed a main effect of Transparency ($F(1,23)=6.13, p<.05$), suggesting that the judgment of transparent sentences was more accurate as compared to the judgment of opaque sentences (96% for transparent nouns vs. 95% for opaque nouns).

EEG results

Figure 1 shows ERPs in response to the agreement manipulation for transparent and opaque nouns. The topographic distribution of average potentials is shown in Figure 2. Visual inspection of ERPs for transparent nouns revealed a greater negativity peaking around 400 ms for the disagreement condition, followed by a greater posterior positivity 500 ms after the stimulus onset. In the case of opaque nouns, the disagreement condition seems to elicit only a greater positivity 500 ms after the stimulus onset.

350–500 ms¹

Statistical analyses revealed a main effect of Transparency ($F(1,23)=4.93, p<.05$). Specifically, transparent nouns elicited more negative responses compared to those elicited by opaque nouns. In addition, the interaction between Agreement and Cluster was significant ($F(8,184)=7.40, p<.001$). Post-hoc tests showed that the disagreement condition elicited a greater negativity compared to the agreement condition and this effect was localized in the left anterior ($t(23)=3.21, p<.05$) and left central clusters ($t(23)=3.28, p<.05$). Finally, the three-way interaction Transparency x Agreement x Cluster was significant ($F(8,184)=3.27, p<.05$). Post-hoc comparisons showed that, for transparent nouns, the disagreement condition was more negative than the agreement condition in the left anterior cluster ($t(23)=3.11, p<.05$; in all the

other clusters: $ps > .05$). Opaque nouns did not show any significant agreement effect (all $ps > .05$).

(Figure 1 about here)

500–700 ms

A main effect of Agreement was found ($F(1,23)=45.23, p<.001$) with more positive amplitudes for the disagreement condition than the agreement condition. In addition, the Agreement factor significantly interacted with the Cluster factor ($F(8,184)=9.00, p<.001$). Post-hoc tests revealed a significant agreement effect in all the clusters, but was greater at the central-posterior sites (left anterior: $t(23)=2.89, p<.01$; medial anterior: $t(23)=4.37, p<.001$; right anterior: $t(23)=5.15, p<.001$; left central: $t(23)=4.42, p<.001$; medial central: $t(23)=6.44, p<.0001$; right central: $t(23)=6.43, p<.0001$; left posterior: $t(23)=5.49, p<.0001$; medial posterior: $t(23)=7.18, p<.0001$; right posterior: $t(23)=5.46, p<.0001$).

700–900 ms

There was a main effect of Agreement ($F(1,23)=64.93, p<.001$) with a greater positivity for the disagreement condition compare to the agreement one. The two-way interaction Agreement x Cluster was also significant ($F(8,184)=52.64, p<.001$) and post-hoc comparisons suggested that the agreement effect was distributed over central-posterior sites (left central: $t(23)=5.31, p<.0001$; medial central: $t(23)=6.44, p<.0001$; right central: $t(23)=7.21, p<.0001$; left posterior: $t(23)=11.61, p<.0001$; medial posterior: $t(23)=10.55, p<.0001$; right posterior: $t(23)=10.52, p<.0001$).

(Figure 2 about here)

Discussion

When Basque-Spanish early bilinguals were reading Spanish sentences, the ERP responses were modulated by Agreement and Transparency. Between 350 and 500 ms, transparent nouns elicited a greater negativity than opaque nouns. In line with previous ERP studies, agreement violations elicited a LAN-P600 pattern. Interestingly, this biphasic response was modulated by noun transparency. In contrast to Spanish monolinguals (Caffarra & Barber, 2015), Basque-dominant bilinguals show an impact of noun transparency on agreement processing, with delayed ERP responses for gender violations of the opaque nouns compared to the transparent ones. We think that these persistent difficulties in computing agreement dependencies with opaque nouns reflect an unstable representation of grammatical gender (Bordag et al., 2006) and provide evidence that the strength of gender lexical representation can change depending on an individual's linguistic background.

However, the impact of gender-to-ending consistency on agreement computation can be related to two different characteristics of the present participant sample: 1) the simple knowledge of another language that lacks a grammatical gender system; 2) the dominance of this gender-free language. In order to tease apart these two possible explanations, we ran the same experiment with Spanish-Basque early bilinguals who were dominant for Spanish (i.e., they learned Spanish from birth and were used to speaking Spanish on daily basis). If there were again an interaction between Transparency and Agreement, we should conclude that the presence of Basque as a second language is enough to explain the present pattern of results. In the case of a lack of interaction, we should conclude that participants' dominance for Basque accounts for our findings.

Experiment 2: material and methods

Participants

Twenty-four Spanish-Basque bilinguals participated in the experiment (14 women; mean age: 23.88, SD: 4.62; range: 19–40; they did not differ in age from the Basque-dominant group: $t(23)=1.43, p>.05$) in exchange for a small sum (20 €). They were born and lived in the Basque Country. They reported having been in contact with both Spanish and Basque since birth and that they started to learn both languages early in life (before 7 years old). At the language assessment performed before the experimental session, participants were highly proficient in both languages, although higher performances were reached in Spanish (Spanish mean scores at the interview: 5/5 and in the naming test: 76/77; Basque mean scores at the interview: 4.5/5 and at the naming test: 67/77, see Table 5). All participants reported being dominant in Spanish, since it was the language initially used with their parents (AoA of Spanish: 0; see Table 5) and the most frequently-used language on a daily basis (self-reported amounts of Spanish produced and heard in daily life: 65% and 60%, respectively). In addition, they reported using Spanish with their family and most of their friends. They started to use Basque only after being enrolled in kindergarten or school (mean AoA of Basque: 4) and they reported using it with non-Spanish-speaking people. All participants were right-handed (based on the Edinburgh Handedness Inventory; Oldfield, 1971) and reported normal or corrected-to-normal vision. No participant had a history of neurological disorder.

(Table 5 about here)

Materials

The experimental materials were the same as those used in Experiment 1.

Procedure

Procedure and EEG analyses were the same as those employed in Experiment 1. On average, 14.6 % of trials were excluded due to incorrect responses or artefacts. The number of rejections did not differ across conditions [$F(3,69)= 0.58, p = 0.63$].

Results

Behavioural results: offline gender decision task

Participants provided the correct response for the majority of the target nouns (average accuracy: 97%, SD: 3%, range: 90%–100%) and they showed mean RTs of 714 ms (SD: 107 ms). Accuracies and RTs from accurate trials were analysed using a one-way ANOVA, including Transparency as a two-level factor (Transparent, Opaque). The results showed that responses to transparent nouns were more accurate and faster than those to opaque nouns (Accuracy: $F(1,23)=28.42, p<.001$; RTs: $F(1,23)=65.25, p<.001$, see Table 6). In addition, the overall ANOVA including the factor Group showed a main effect of Group for RTs, suggesting that Basque-dominant group was slightly slower than Spanish-dominant ($F(1,46)=4.41, p<.05$; Basque-dominant: 790 ms, SD: 140 ms; Spanish-dominant: 714 ms, SD: 107 ms).

(Table 6 about here)

Behavioural results: online grammaticality judgment

Participants performed the grammaticality judgment task with a mean accuracy of 96% (SD: 4), suggesting that they could easily detect grammatical errors in most of the sentences. A two-way ANOVA including Transparency and Agreement as two within-subject factors did not show any significant effects or interaction (all $ps >.05$), suggesting that the four experimental

conditions did not significantly differ in accuracy rate (all averages were above 95%). The overall ANOVA did not show any significant effects involving the factor Group (all $p > .05$).

EEG results

Figure 3 shows the average waveforms in response to the agreement manipulation for transparent and opaque nouns. Figure 4 shows the topographic distribution of average potentials. Visual inspection of ERPs for both types of nouns revealed that the disagreement condition elicited a greater left negativity at around 400 ms, followed by a greater posterior positivity. A repeated measures ANOVA was performed including Transparency, Agreement and Cluster as within-subjects factors. Additional analyses including data from both groups are reported at the end of this section.

350–500 ms²

A main effect of Transparency was reported ($F(1,23)=5.71, p<.05$) with more negative responses for transparent nouns, compared to opaque nouns. A significant interaction between Agreement and Cluster was also found ($F(8,184)=7.05, p<.001$). Post-hoc comparisons showed that there was a greater negativity for the disagreement condition, compared to the agreement one in the left anterior ($t(23)=3.76, p<.01$) and the left central clusters ($t(23)=4.71, p<.0001$). No interaction between Transparency and Agreement was significant (Transparency x Agreement: $F(1,23)<1, p=.75$; Transparency x Agreement x Cluster: $F(8,184)=1.14, p=.33$).

(Figure 3 about here)

500–700 ms

The statistical analyses revealed a main effect of Agreement ($F(1,23)=22.76, p<.001$), with an increased positivity for the disagreement condition compared to the agreement one. There was also a significant interaction Agreement x Cluster ($F(8,184)=22.98, p<.001$),

suggesting that the agreement effect was central-posteriorly distributed (left central: $t(23)=4.16$, $p<.01$; medial central: $t(23)=4.27$, $p<.01$; right central: $t(23)=4.38$, $p<.01$; left posterior: $t(23)=5.77$, $p<.0001$; medial posterior: $t(23)=6.43$, $p<.0001$; right posterior: $t(23)=6.61$, $p<.0001$).

700–900 ms

There was a main effect of Agreement ($F(1,23)=22.13$, $p<.001$), with the disagreement condition eliciting a greater positivity compared to the agreement condition. The Agreement factor significantly interacted with the Cluster factor ($F(8,184)=63.98$, $p<.001$) and post-hoc tests revealed that the agreement effect was distributed over central-posterior sites (left central: $t(23)=4.96$, $p<.001$; medial central: $t(23)=4.10$, $p<.01$; right central: $t(1,23)=5.13$, $p<.001$; left posterior: $t(23)=7.79$, $p<.0001$; medial posterior: $t(23)=8.09$, $p<.0001$; right posterior: $t(23)=8.74$, $p<.0001$)³.

(Figure 4 about here)

Discussion

Spanish-dominant bilinguals showed a pattern of ERP responses similar to Spanish monolinguals (Caffarra & Barber, 2015). Transparent nouns elicited more negative waveforms than opaque nouns around 400 ms after the target noun onset. Agreement violations elicited a LAN effect followed by a P600 effect. No interaction was observed between these two experimental factors. Thus, it seems that the bilinguals' language dominance can account for the different patterns of results reported in Experiment 1.

Overall analysis on the LAN effect size for opaque nouns

A general ANOVA including the factor Group did not lead to significant interactions between Group, Agreement and Transparency (all $ps >.05$) in the three time windows taken into

account. The absence of interaction with Group in the LAN time window suggests that the averaged brain responses of the two bilingual groups did not significantly differ. However, language dominance can still have a progressive impact on the LAN effect size observed for each participant. In order to better understand which factors might gradually influence the LAN effect with opaque nouns, further analyses were performed on this subset of data.

A linear multiple regression analysis was performed on the data collected from all the participants (48) in order to see what factors would account for LAN effect modulations in response to agreement violations with opaque nouns. The analysis included the LAN effect size in the 350–500 ms time window as a dependent variable (i.e., amplitude difference between the average brain responses to the disagreement and the agreement conditions recorded in a group of left central-anterior electrodes: F3, F7, FC5, C3, T7, CP5). The following participant characteristics were treated as independent demeaned variables: AoA of Spanish, Spanish proficiency level measured at the interview, Spanish proficiency level measured at the naming test, self-reported amount of Spanish production, self-reported amount of Spanish comprehension. The intercept value was equal to $-0.96 \mu\text{V}$ ($\text{SE}=0.26$). The amount of Spanish production had the strongest effect and reached significance ($t(46)=2.20$, $p<.05$, see Figure 5). Specifically, the estimated increase of LAN effect size was $0.06 \mu\text{V}$ ($\text{SE}=0.03$) per each unit change in the self-reported amount of Spanish production. No other variable had a strong effect (all $ps>.05$)⁴.

(Figure 5 about here)

Finally, to make sure that the large age range did not influence our pattern of results additional analyses were performed after excluding those participants who were above 35 years of age (three from the Basque-dominant group and one from the Spanish-dominant group). Similar

patterns of results were obtained for both groups and for the overall linear regression analysis. The only difference regarded the main effect of Transparency between 350 and 500 ms, which did not reach significance in the Basque-dominant group ($F(1,20)= 2.50, p =.13$). This was possibly due to the smaller sample size and the reduced statistical power.

General Discussion

The present ERP studies were aimed at investigating whether the strength of gender lexical representation and the weight of formal gender cues could change depending on the presence of (or the dominance for) a second language. Basque-Spanish and Spanish-Basque bilinguals were presented with Spanish sentences that could contain gender agreement errors between a determiner and a noun. Spanish target nouns could provide a reliable or an ambiguous cue to gender. Previous ERP data on Spanish monolinguals has shown an effect of Transparency and Agreement, without any interaction between the two factors (Caffarra & Barber, 2015; Caffarra, Siyanova-Chanturia, et al., 2015). Specifically, transparent nouns elicited a greater negativity than opaque nouns around 400 ms after the target noun, suggesting that monolinguals can detect distributional information conveyed by the noun ending during sentence reading. A LAN-P600 pattern was similarly reported for gender violations of both transparent and opaque nouns, suggesting that gender-to-ending consistency does not have a strong impact on monolingual agreement computation.

On the basis of previous studies on gender processing in highly proficient bilinguals (Bordag et al., 2006; Bordag & Pechmann, 2007, 2008; Hopp, 2013) and gender acquisition (Lemhofer et al., 2010), we hypothesized that bilingual lexical representation of gender could be more unstable. In this case, formal gender cues could be less redundant and more helpful in

recovering gender and computing agreement dependencies. Thus, in contrast to what has been seen in Spanish monolinguals, we predicted that noun transparency should influence agreement processing in the case of bilinguals.

The present ERP results from Basque-dominant bilinguals confirmed these predictions. A significant interaction between Transparency and Agreement was reported between 350 ms and 500 ms. While gender errors in transparent nouns elicited a greater LAN-P600 pattern (similar to Barber & Carreiras, 2005; Barber et al., 2004; Dowens et al., 2010), gender violations of opaque nouns only showed a P600 effect. However, results from Spanish-dominant bilinguals were similar to those of Spanish monolinguals (Caffarra & Barber, 2015), showing LAN-P600 responses for gender disagreement with both transparent and opaque nouns. Both bilingual groups also showed a main effect of transparency between 350 and 500 ms, indicating that they were sensitive to the presence of available gender cues. Moreover, both groups showed high accuracy rates (above 95% of correct responses) on the grammaticality judgment task performed during the EEG recording and on the gender decision task performed after the EEG session. Basque-dominant bilinguals were slower at deciding the gender of target nouns than Spanish-dominant bilinguals.

The ERP results from the Basque-dominant group suggest that while agreement violations of transparent nouns can be detected as early as 350 ms, in the case of opaque nouns the processing of agreement violations is delayed, although successfully carried out (i.e., participants showed high accuracy rates at both grammaticality judgments and gender decision tasks). The fact that ERP responses to opaque and transparent nouns differ in the LAN time window suggests that Basque-dominant bilinguals encounter problems with opaque nouns at an

early stage of agreement processing. During this first stage, the parser should be able to detect a mismatch between the different grammatical features extracted from distinct sentential constituents (Mancini, Molinaro, & Carreiras, 2013; Molinaro, Barber, & Carreiras, 2011). If the lexical representation of gender is not fully entrenched, however, feature extraction could be particularly difficult in the case of opaque nouns, whose grammatical gender can only be recovered through a lexical route (Delfitto, & Zamparelli, 2009; Gollan & Frost, 2001). Thus, the lack of a strong LAN effect for gender violations in opaque nouns can be interpreted as the result of problematic feature extraction from a noun that does not provide any reliable formal cues to gender.

Nonetheless, this difficulty in gender retrieval does not prevent Basque-dominant bilingual from correctly identifying agreement violations of opaque nouns and successfully performing the gender decision task. In the absence of transparent endings, other potential cues, such as determiners, could allow successful performances at explicit tasks on gender (Afonso et al., 2014; Taft & Meunier, 1998). Spanish definite articles are never ambiguous or misleading, and, thus, they represent a strong valid cue that can almost always predict the gender of the following noun (Franck, Vigliocco, Anton-Mendez, & Collina, 2008; Wicha et al., 2004). It could therefore be that most of Basque-dominant bilinguals were relying on the entire determiner-noun combination in order to successfully perform the task, instead of extracting the relevant feature directly from the opaque noun (which implies relying on an abstract representation of gender stored in the lexicon). Basque-dominant bilinguals would have been able to compute agreement violations of opaque nouns mainly by monitoring the correctness of the entire determiner-noun chunk at a later stage of agreement processing. In line with this potential explanation, some of the Basque-dominant bilinguals (8 out of 24) explicitly reported

silently generating the corresponding definite article for each noun (i.e., *el* for masculine and *la* for feminine) presented during the gender decision task⁵.

We are not claiming that difficult feature extraction automatically results in a lack of LAN effect, but rather a less consistent early response across trials and across participants. Figure 5 shows that, although the ERP grand average does not exactly reflect what happens at the individual level (Tanner, Inoue, & Osterhout, 2014; Tanner, Mclaughlin, Herschensohn, & Osterhout, 2013), part of the individual variability is explained by linguistic habits. Based on this, we can conclude that the more the participants used Spanish on a daily basis, the more likely it was to observe a LAN effect in response to the opaque nouns⁶. We interpret this trend as reflecting the fact that lexical representation of gender is more likely to be unstable in the Basque-dominant group compared to the Spanish-dominant one.

In addition, it should be noted that the higher variability in the LAN results could also be related to a stronger cross-linguistic transfer in the case of Basque-dominant bilinguals. In contrast to Spanish, Basque is an agglutinative language, where words are formed by joining affixes to the stem. Each affix is a morpheme, which has one function (e.g., expressing the grammatical feature of number, Hualde & de Urbina, 2003) and cannot be changed or influenced by other affixes. This results in a one-to-one correspondence between a specific formal cue and a function (with few exceptions, Jacobsen, 1972). In the case of Basque, the parser can easily extract each grammatical feature based on sub-lexical cues at the word-form level (Hualde & de Urbina, 2003). Basque-dominant bilinguals frequently have to rely on form-function correspondences in order to retrieve grammatical features. These bilinguals could have transferred their Basque syntactic routines to Spanish (MacWhinney, 2005, 2012), overestimating any available formal cues that can potentially lead to a grammatical feature (e.g.,

gender). The transfer of those syntactic preferences that were widely employed in the dominant language could contribute to explaining why Basque-dominant bilinguals showed a different pattern of results for transparent (where they can retrieve gender based on formal cues) and opaque Spanish nouns (where it is not possible to take advantage of any form-function mapping, Caffarra, Zimnukhova, & Mancini, 2016).

The ERP results from the Basque and Spanish-dominant groups suggest that language dominance plays a role in determining the weight of the lexical and the form-based route in agreement computation: early learners of Spanish who are used to speaking Spanish on a daily basis might rely on a fully-entrenched lexical representation of Spanish gender, with no need to compute gender based on formal cues (similarly to the case of Spanish monolinguals). Early learners of Basque who used Basque on a daily basis, however, would have a less stable representation of Spanish gender and would rely on any available formal cues during agreement computation (e.g., noun ending, determiner).

This interpretation is in line with previous behavioural results showing asymmetric performances between transparent and opaque nouns in highly proficient bilinguals (Bordag et al., 2006; Caffarra et al. 2016). The present pattern of results further suggests that this different sensitivity is not due to the simple presence of a second gender-free language, but mainly depends on language dominance, and more specifically on language use.

Results from a multiple regression analysis performed on the LAN effect size showed that the stronger predictor of reduction of the LAN effect with opaque nouns was the self-reported amount of Spanish speech produced on a daily basis. This result seems to be at odds with previous studies suggesting that left negativities modulation is mainly related to factors of

AoA (Weber-Fox & Neville, 1996) and proficiency level (Rossi, Gugler, Friederici, & Hahne, 2006; Steinhauer, White, & Drury 2009). However, it should be noted that most previous research investigating the role of influential factors on second language processing have examined the case of late learners (for an overview see Caffarra, Molinaro, Davidson, & Carreiras, 2015). Here, we considered the case of early bilinguals who were born in the Basque country and were exposed to Spanish and Basque since birth. All the ERP studies so far conducted on early bilingual syntactic processing have focused on a direct comparison between monolingual and early bilingual parsing (see Grosjean, 1989 for a critical analysis of this approach). To our knowledge, this is the first ERP study that considers different groups of early bilinguals, trying to shed new light on what factors can influence the time course of their syntactic analysis. On the basis of the present results, linguistic dominance, and more specifically language use, seems to be an influential factor in the way people treat distributional information in sentence reading. Additional studies on early bilinguals are needed in order to better define the role of other factors (e.g., AoA, proficiency) and to generalize the present result to other language pairs.

Conclusions

The present ERP study investigated whether the role of formal gender cues during agreement computation can change, depending on the presence or the dominance of a second language. Spanish-dominant bilinguals (similarly to Spanish monolinguals) have a stable lexical representation of gender, and distributional cues to gender do not play a differential role in agreement computation. However, when bilinguals are used to speaking a gender-free

L2 syntactic processing

agglutinative language (i.e., Basque) the lexical representation of gender is less stable and formal cues are more influential in agreement processing.

Funding

This work was supported by the Spanish Ministry [PSI 2014-54500-P, PSI2015-65694-P]; the Basque Government [PI_2015_1_25]; the Severo Ochoa [SEV-2015-0490]; and the AThEME project [FP7-SSH-2013-1-GA613465], which has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 613465.

Acknowledgements

The authors wish to thank Margaret Gillon-Dowens for reviewing the manuscript and giving insightful comments on this work.

Notes

1. No significant ERP effects involving the experimental factors were found before 350 ms (all $ps > .05$).
2. No significant ERP effects involving the experimental factors were found before 350 ms (all $ps > .05$).
3. Additional analyses using two topographic factors (Longitude, anterior: F3, F7, FC5, F4, F8, C6, Fp1, Fp2, Fz, central: C3, T7, CP5, FC1, FC2, Cz, C4, T8, CP6, posterior: P3, P7, O1, CP1, CP2, Pz, P4, P8, O2; and Latitude, right: F4, F8, FC6, C4, T8, CP6, P4, P8, O2, medial: Fp1, Fp2, Fz, FC1, FC2, Cz, CP1, CP2, Pz, left: F3, F7, FC5, C3, T7, CP5, P3, P7, O1) also showed a similar pattern of results.
4. The same multiple regression performed on the LAN effect size for agreement violations with transparent nouns did not lead to any large effects (all $ps > .05$; intercept: -1.10, SE: .24). Also, no significant effects were found in the P600 time windows (500-700: all $ps > .05$; intercept: 2.23, SE: .26; 700-900: all $ps > .05$; intercept: 3.24, SE: .30).
5. This strategy could explain why Basque-dominant bilinguals were slightly slower than Spanish-dominant bilinguals at the gender decision task.
6. The lack of a clear-cut distinction between the two bilingual groups can probably account for the lack of significant interactions between Group, Agreement and Transparency in the overall ANOVA.

References

- Afonso, O., Domínguez A., Álvarez, C. J., & Morales, D. (2014). Sublexical and Lexico-Syntactic Factors in Gender Access in Spanish. *Journal of Psycholinguistic Research*, 43(1), 13–25. doi: 10.1007/s10936-012-9236-0
- Bañon, J. A., Fiorentino, R., & Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1–L2 similarity and structural distance. *Second Language Research*, 30(3), 275–306. doi: 10.1177/0267658313515671
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, 17(1): 137–153. doi: 10.1162/0898929052880101
- Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement? In M. Carreiras & C. Clifton (Eds.), *On-line study of sentence comprehension; eye-tracking, ERP and beyond* (pp. 309–328). Brighton: Psychology Press.
- Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, 58 (7), 992–1004. doi:10.3758/BF03206827
- Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, 57, 847–862. doi: 10.3758/BF03206800
- Bordag, D., Opitz, A., & Pechmann, T. (2006). Gender processing in first and second languages. The role of noun termination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1090–1101. doi: 10.1037/0278-7393.32.5.1090

- Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition*, 10, 299–314. doi: 10.1017/S1366728907003082
- Bordag, D., & Pechmann, T. (2008). Grammatical gender in translation. *Second Language Research*, 24(2), 139–166. doi: 10.1177/0267658307086299
- Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, 1446, 127–143. doi: 10.1016/j.brainres.2012.01.055
- Bull, W. (1965). *Spanish for teachers: Applied Linguistics*. New York: Ronald Press Co.
- Caffarra, S., & Barber, H. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, 1605, 83–92. doi: 10.1016/j.brainres.2015.02.018
- Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: An empirical review. *Neuroscience & Biobehavioral reviews*, 51, 31–47. doi: 10.1016/j.neubiorev.2015.01.010
- Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, 52, 1019–1030. doi: 10.1111/psyp.12429
- Caffarra, S., Zimnukhova, S., & Mancini, S. (2016). What usage can do: The effect of language dominance on simultaneous bilinguals' morphosyntactic processing. *Linguistics Vanguard*, 2, 43–53. doi: 10.1515/lingvan-2016-0020
- Caselli, M. C., Leonard, L., Volterra, V., & Campagnoli, M. G. (1993). Toward mastery of Italian morphology: A cross-sectional study. *Journal of Child Language*, 20(22), 377–393. doi: 10.1017/S0305000900008333

- Chini, M. (1995). *Genere grammaticale e acquisizione. Aspetti della morfologia nominale in italiano L2*. Milano: Franco Angeli.
- Clair, M. C., Monaghan, P., & Christiansen, M. H. (2010). Learning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition*, 116, 341–360. doi: 10.1016/j.cognition.2010.05.012
- Clegg, J. H. (2010). Native Spanish Speaker Intuition in Noun Gender Assignment. *Language Design*, 12, 5-18.
- Corbett, G. (1991). *Gender*. Cambridge, England: Cambridge University Press.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13(1), 21–58. doi: 10.1080/016909698386582
- De Martino, M., Bracco, G., & Laudanna, A. (2010). The activation of grammatical gender information in processing Italian nouns. *Language and Cognitive Processes*, 26, 745–776. doi: 10.1080/01690965.2010.491977
- Dechert, H.W., & Raupach, M. (1989). *Transfer in language production*. Norwood: Ablex.
- Delfitto, D., & Zamparelli, R. (2009). *Le strutture del significato*. Bologna: Il Mulino.
- Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, 45, 200–224. doi: 10.1006/jmla.2000.2768
- Dewaele, J. M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, 4, 275–297. doi: 10.1017/S136672890100044X

- Dowens, M.G., Guo, T., Guo, J., Barber, H., & Carreiras, M., (2011). Gender and number processing in Chinese learners of Spanish – Evidence from Event Related Potentials. *Neuropsychologia*, 49, 1651–1659. doi: 10.1016/j.neuropsychologia.2011.02.034
- Dowens, M.G., Vergara, M., Barber, H.A., & Carreiras, M. (2010). Morphosyntactic Processing in Late Second-Language Learners. *Journal of Cognitive Neuroscience*, 22(8), 1870–1887. doi: 10.1162/jocn.2009.21304
- Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam: John Benjamins.
- Franck, J., Vigliocco, G., Anton-Mendez, I., & Collina, S. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, 23, 329–374. doi: 10.1080/01690960701467993
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trend in Cognitive Sciences*, 6 (2), 78–84. doi: 10.1016/S1364-6613(00)01839-8
- Gass, S. (1984). A review of interlanguage syntax: language transfer and language universals. *Language Learning*, 34, 115–32. doi: 10.1111/j.1467-1770.1984.tb01007.x
- Gass, S., & Selinker, L. (1992). *Language transfer in language learning: Revised edition*. Amsterdam: John Benjamins.
- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249–268. doi: 10.1017/S030500090400678
- Gollan, T., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, 30, 627–651. doi: 10.1023/A:1014235223566
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain Language*, 36, 3–15. doi: 10.1016/0093-934X(89)90048-5

- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556–568. doi: 10.1162/089892900562336
- Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, 15(6), 883–899. doi: 10.1162/089892903322370807
- Hagoort, P., & Brown, C. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, 28, 6, 715–728. doi: 10.1023/A:1023277213129
- Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22, 27–62.
- Henzl, V. M. (1975). Acquisition of grammatical gender in Czech. *Papers and Reports on Child Language Development*, 10, 188–198.
- Hernandez, A. E., Kotz, S. A., Hofmann, J., Valentin, V. V., Dapretto, M., & Bookheimer, S. Y. (2004). The neural correlates of grammatical gender decisions in Spanish. *Neuroreport*, 15(5), 863–866. doi: 10.1097/00001756-200404090-00026
- Hochberg, B. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4), 800–802. doi: 10.1093/biomet/75.4.800
- Holmes, V. M., & Dejean de la Bâtie, B. (1999). Assignment of grammatical gender by native speakers and foreign learners of French. *Applied Psycholinguistics*, 20, 479–506.
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. doi: 10.1177/0267658312461803
- Hualde, J. I., & de Orbina, J. O. (2003). *A grammar of Basque*. Berlin: Mouton de Gruyter.

- Jacobsen, W. H. (1972). Nominative-ergative syncretism in Basque. *International Journal of Basque Linguistics and Philology*, 6(1), 67–109.
- Lemhofer, F., Schriefers, H., & Hanique, I. (2010). Native language effects in learning second-language grammatical gender: A training study. *Acta Psychologica*, 135, 150–158. doi: 10.1016/j.actpsy.2010.06.001
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1–38.
- MacWhinney, B. (1987). The Competition Model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum.
- MacWhinney, B. (2004). Parameters or cues? *Bilingualism: Language and Cognition*, 7, 35–36.
- MacWhinney, B. (2005). A Unified Model of Language Acquisition. In J. F. Kroll & A. M. B. de Groot, (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.
- MacWhinney, B. (2012). The Logic of the Unified Model. In S. Gass & A. Mackey (Eds.), *Handbook of Second Language Acquisition* (pp. 211–227). New York: Routledge.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23, 127–150. doi: 10.1016/S0022-5371(84)90093-8
- Mancini, S. Molinaro, N., & Carreiras, M. (2013). Anchoring Agreement in Comprehension. *Language & Linguistics Compass*, 7(1), 1–21. doi: 10.1111/lnc3.12008
- Matthews, C. (2010). On the nature of phonological cues in the acquisition of French gender categories: Evidence from instance-based learning models. *Lingua*, 120 (4), 879–900. doi: 10.1016/j.lingua.2009.06.007

- Mills, A. E. (1986). *The acquisition of gender. A study of English and German*. Berlin: Springer-Verlag.
- Molinaro, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 47, 908–930. doi: 10.1016/j.cortex.2011.02.019
- Molinaro, N., Kim, A., Vespignani, F., & Job, R. (2008). Anaphoric agreement violation: An ERP analysis of its integration. *Cognition*, 106, 963–974. doi: 10.1016/j.cognition.2007.03.006
- Molinaro, N., Vespignani, F., & Job, R. A. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research*, 1228, 161–176. doi: 10.1016/j.brainres.2008.06.064
- Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011). Why brother and sister are not just siblings: Repair processes in agreement computation. *Journal of Memory and Language*, 64, 211–232. doi: 10.1016/j.jml.2010.12.002
- Montrul, S., & Potowski, K. (2007). Command of gender agreement in school-age Spanish–English bilingual children. *International Journal of Bilingualism*, 11, 301–328. doi: 10.1177/13670069070110030301
- Oldfield, R.C. (1971). “The assessment and analysis of handedness: the Edinburgh inventory.” *Neuropsychologia*, 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Odlin, T. (1989). *Language transfer: cross-linguistic influence in language learning*. New York: Cambridge University Press.
- Padovani, R., & Cacciari, C. (2003). Il ruolo della trasparenza morfologica nel riconoscimento di parole in Italiano. *Giornale Italiano di Psicologia*, 30, 749–771.

- Pérez-Pereira, M. (1991). The acquisition of gender: what Spanish children tell us. *Journal of Child Language*, 18, 571–590. doi: 10.1017/S0305000900011259
- Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews. Cognitive Science*, 1(6), 906–914. doi: 10.1002/wcs.78
- Rossi, S., Gugler, M.F., Friederici, A.D., & Hahne, A. (2006). The Impact of Proficiency on Syntactic Second-language Processing of German and Italian: Evidence from Event-related Potentials. *Journal of Cognitive Neuroscience*, 18(12), 2030–2048. doi: 10.1162/jocn.2006.18.12.2030
- Rossi, E., Kroll, J. F., & Dussias, P. E. (2014). Clitic pronouns reveal the time course of processing gender and number in a second language. *Neuropsychologia*, 62, 11–25. doi: 10.1016/j.neuropsychologia.2014.07.002
- Schacht, A., Martín-Loeches, M., Casado, P., Rahman, R. A., Sel, A., & Sommer, W. (2010). How Is Sentence Processing Affected by External Semantic and Syntactic Information? Evidence from Event-Related Potentials. *PLoS ONE*, 5, e9742. doi: 10.1371/journal.pone.0009742
- Scheutz, M. J., & Eberhard, K. M. (2004). Effects of morphosyntactic gender features in bilingual language processing. *Cognitive Science*, 28, 559–588. doi: 10.1016/j.cogsci.2004.03.001
- Schiller, N. O., Münte, T. F., Horemans, I., & Jansma, B. M. (2003). The influence of semantic and phonological factors on syntactic decisions: An event-related brain potential study. *Psychophysiology*, 869–877. doi: 10.1111/1469-8986.00105

- Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174-215. doi: 10.1037/0278-7393.6.2.174
- Spalek, K., Franck, J., Schriefers, H., & Frauenfelder, U. (2008). Phonological regularities and grammatical gender retrieval in spoken word recognition and word production. *Journal of Psycholinguistic Research*, 37 (6), 419-442. doi: 10.1007/s10936-008-9074-2
- Steinhauer, K., White, E.J., & Drury, J.E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Language Research*. 25 (1), 13-41. doi: 10.1177/0267658308098995
- Taft, M., & Meunier, F. (1998). Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research*, 27, 23-45. doi: 10.1023/A:1023270723066
- Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based Individual Differences in On-line L2 Grammatical Comprehension. *Bilingualism: Language and Cognition*, 17(2), 277-293. doi: 10.1017/S1366728913000370
- Tanner, D., Mclaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP Evidence. *Bilingualism: Language and Cognition*, 16(2), 367-382. doi: 10.1017/S1366728912000302
- Unsworth, S. (2008). Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research*, 24(3), 365-395. doi: 10.1177/0267658308090185
- Van de Meeredonk, N., Kolk, H.H.J., Chwilla, D.J., & Vissers C.T.W.M. (2009). Monitoring in Language Perception. *Language and Linguistics Compass*, 3(5), 1211-1224. doi: 10.1111/j.1749-818x.2009.00163.x

- Vigliocco, G., & Franck, J. (1999). When Sex and Syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, 40, 455–478. doi: 10.1006/jmla.1998.2624
- Weber-Fox, C. M., & Neville, H.J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256. doi: 10.1162/jocn.1996.8.3.231
- White, L. (2003). The initial state. In L. White (Ed.), *Second language acquisition and universal grammar* (pp. 58–99). Cambridge: Cambridge University Press.
- Wicha, N. Y. Y., Moreno, E., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16, 1272–1288. doi: 10.1162/0898929041920487

Table 1. Summary of the characteristics of Basque-dominant participants (mean values and SD are reported for each language and measure). AoE (age of exposure): age at which a specific language was available in the surrounding environment (Basque country). AoA (age of acquisition): age at which people started to speak the language. Proficiency was rated by a trained native speaker, by means of a semi-structured interview and a picture naming test. Participants were required to estimate how much they were used to speaking and listening to a specific language on a daily basis.

		Spanish	Basque
Starting point	AoE (years)	Birth	Birth
	AoA (years)	4.79 (2.60)	Birth
Proficiency	Interview (tot: 5)	4.88 (0.34)	5 (0)
	Naming test (tot: 77)	75.92 (1.61)	75.33 (2.32)
Daily use	Language production (%)	31.04 (15.11)	62.50 (16.35)
	Language exposure (%)	35.83 (16.40)	55.42 (18.41)

Table 2. Lexical features of the target nouns and statistical comparisons.

	Opaque		Transparent		Comparisons	
	<i>Mean (SD)</i>	<i>Range</i>	<i>Mean (SD)</i>	<i>Range</i>	<i>t</i>	<i>p</i>
N° letters	5.4 (1.0)	4.0-8.0	5.4 (1.0)	4.0-8.0	0	1
AoA (years)	2.1 (0.5)	1.0-3.0	2.4 (0.6)	2.0-4.0	.14	.89
Log Frequency	1.4 (0.6)	0.1-2.6	1.4 (0.5)	0.3-2.9	.04	.97
Familiarity	5.7 (0.9)	2.6-6.9	5.7 (0.8)	3.1-6.9	.63	.53
Imageability	5.0 (1.3)	1.9-6.7	5.2 (1.1)	2.0 -6.6	1.08	.28
Concreteness	5.1 (1.1)	2.5-6.8	5.0 (1.0)	2.4-6.7	.24	.81
	<i>n</i>		<i>n</i>		χ^2	<i>p</i>
N° cognates	50		40		1.78	.18

Table 3. Examples of experimental sentences for each condition.

Transparent nouns

Agreement	<i>El zapato le quedaba estrecho.</i>	The _M shoe _M was tight for him/her.
Disagreement	<i>La zapato le quedaba estrecho.</i>	The _F shoe _M was tight for him/her.

Opaque nouns

Agreement	<i>El bosque está lleno de gente.</i>	The _M forest _M is full of people.
Disagreement	<i>La bosque está lleno de gente.</i>	The _F forest _M is full of people.

Table 4. Behavioural results of the gender decision task for Basque-dominant bilinguals. Mean RTs and percentage of correct responses are provided for transparent and opaque nouns. Standard deviations are given in parentheses.

	Accuracy (%)	RTs (ms)
Transparent nouns	98 (2)	776 (147)
Opaque nouns	93 (3)	817 (134)

Table 5. Summary of the characteristics of Spanish-dominant participants (mean values and SD are reported for each language and measure). AoE (age of exposure): age at which a specific language was available in the surrounding environment (Basque country). AoA (age of acquisition): age at which people started to acquire the language. Proficiency was rated by a trained native speaker, by means of a semi-structured interview and a picture naming test. Participants were required to estimate how much they were used to speaking and listening to a specific language.

		Spanish	Basque
Starting point	AoE (years)	Birth	Birth
	AoA (years)	Birth	3.75 (1.40)
Proficiency	Interview (tot: 5)	5 (0)	4.54 (0.51)
	Naming test (tot: 77)	76.42 (1.25)	66.92 (4.65)
Daily use	Language production (%)	65.21 (8.91)	25.00 (9.67)
	Language exposure (%)	59.58 (11.22)	30.00 (11.80)

Table 6. Behavioural results of the gender decision task for Spanish-dominant bilinguals. Mean RTs and percentage of correct responses are provided for transparent and opaque nouns. Standard deviations are given in parentheses.

	Accuracy (%)	RTs (ms)
Transparent nouns	98 (2)	696 (110)
Opaque nouns	95 (3)	732 (105)

Figure 1. ERP grand average waveforms for Basque-dominant bilinguals. Negativity is plotted upwards.

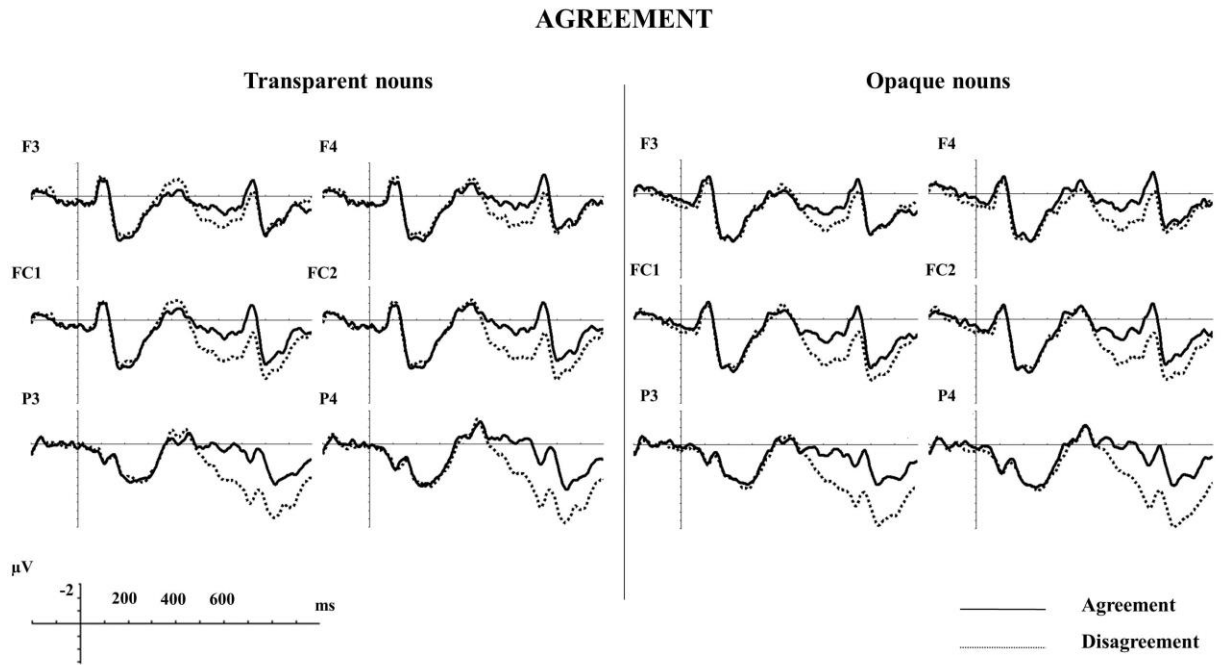


Figure 2. Topographic maps of the agreement effect for Basque-dominant bilinguals. These were computed from the subtraction between the disagreement and the agreement conditions.

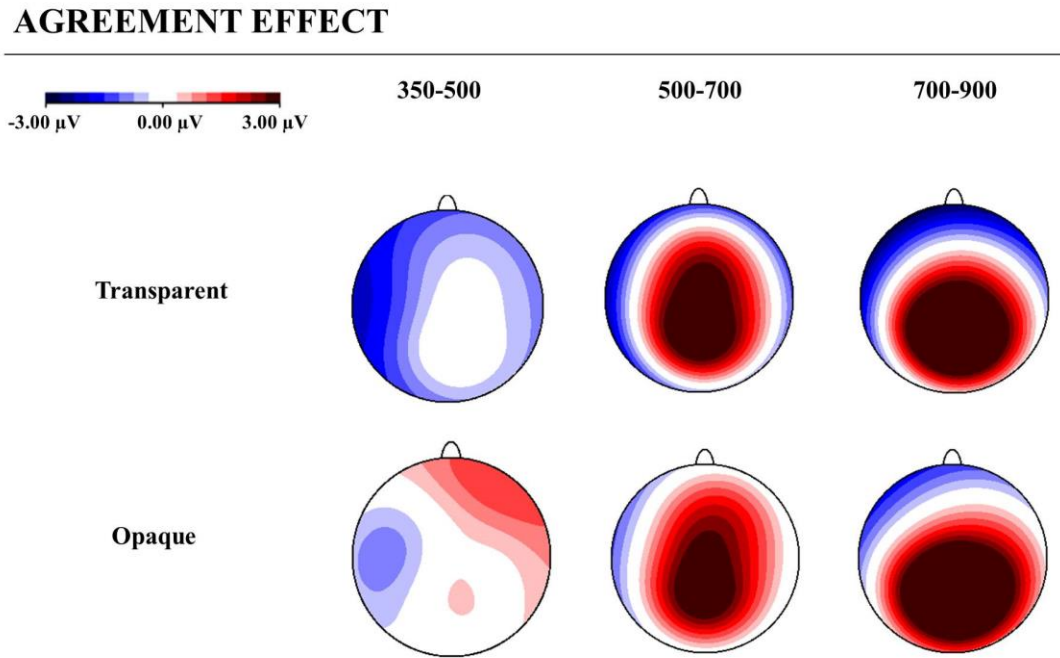


Figure 3. ERP grand average waveforms for Spanish-dominant bilinguals. Negativity is plotted upwards.

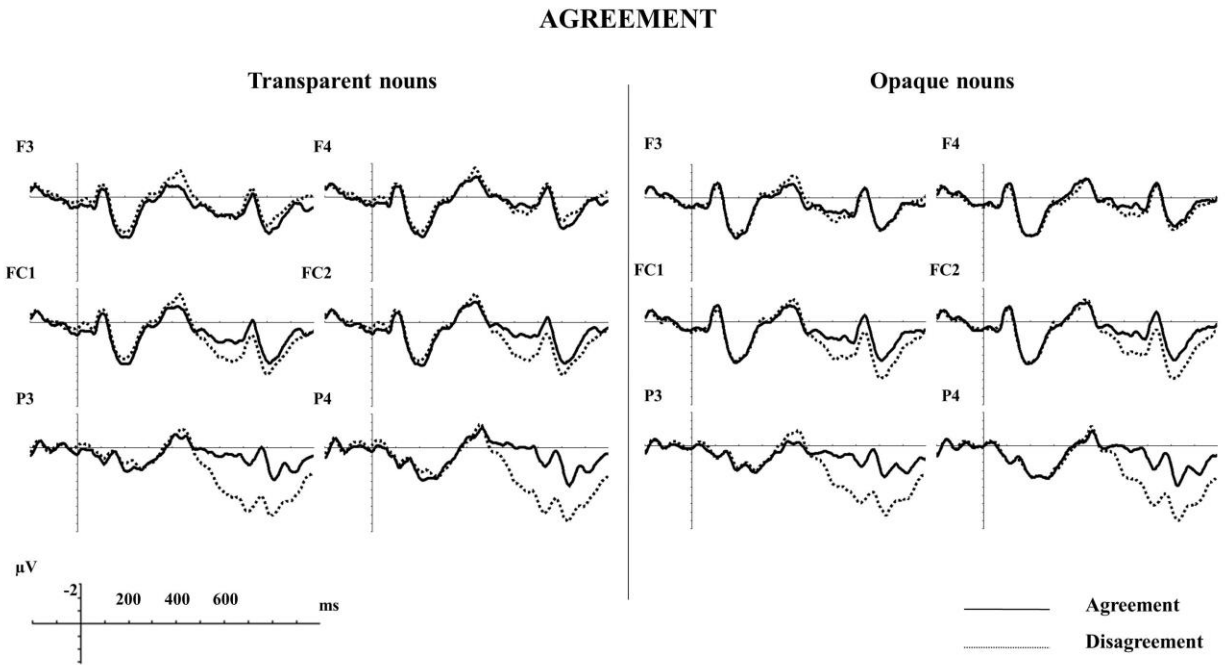


Figure 4. Topographic maps of the agreement effect for Spanish-dominant bilinguals. These were computed from the subtraction between the disagreement and the agreement conditions.

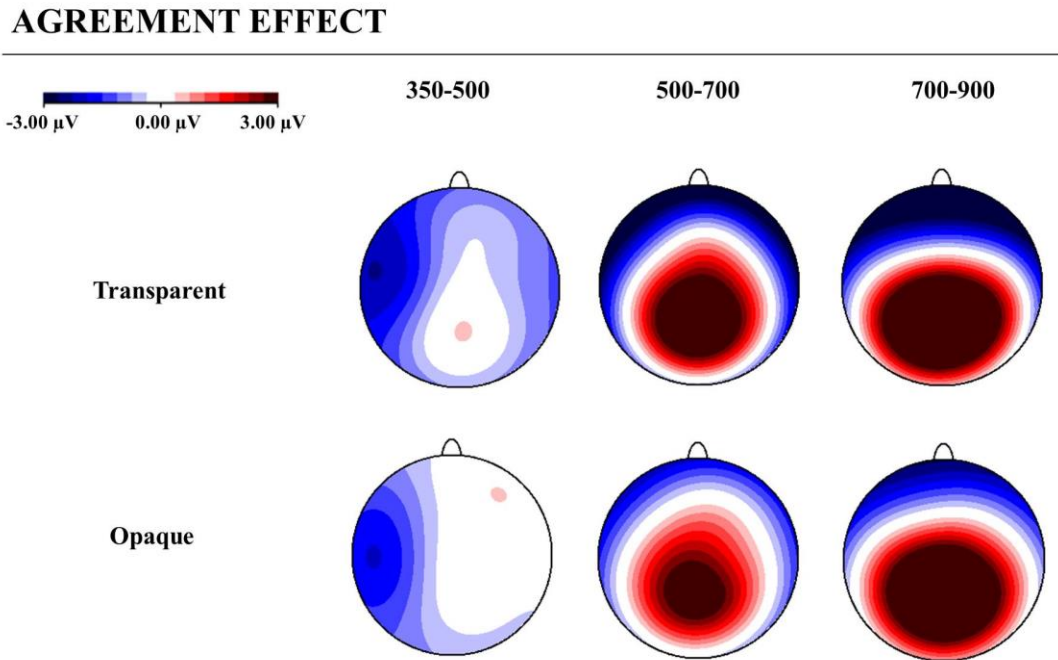


Figure 5. The impact of Spanish usage on the LAN effect size. The x-axis shows the self-reported measures of Spanish production while the y-axis shows the magnitude of the LAN effect in response to agreement violations with opaque nouns.

