Received: 17 July 2020 | Revised: 14 October 2020 | Accepted: 12 November 2020

DOI: 10.1111/psyp.13737

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers

Anne L. Beatty-Martínez[1,2] | **Michelle R. Bruni**[3] | **María Teresa Bajo**[4,5] | **Paola E. Dussias**[6,7]

[1]Department of Psychology, McGill University, Montreal, QC, Canada

[2]Centre for Research on Brain, Language, and Music, Montreal, QC, Canada

[3]Department of Psychology, University of California-Riverside, Riverside, CA, USA

[4]Department of Experimental Psychology, University of Granada, Granada, Spain

[5]Mind, Brain and Behavior Research Center, University of Granada, Granada, Spain

[6]Department of Spanish, Italian and Portuguese, Pennsylvania State University, State College, PA, USA

[7]Center for Language Science, Pennsylvania State University, State College, PA, USA

Correspondence

Anne L. Beatty-Martínez, Department of Psychology, Room 740, McGill University, 2001 Avenue McGill College, Montréal, QC, CA, H3A 1G1.

Email: anne.beatty-martinez@mail.mcgill.ca

Funding information

Ministerio de Economía y Competitividad, Grant/Award Number: PCIN-2015-132; National Institute on Aging, Grant/Award Number: F32-AG064810; National Science Foundation, Grant/Award Number: BCS-1535124, OISE-1545900 and OISE-0968369; Eunice Kennedy Shriver National Institute of Child Health & Human Development, Grant/Award Number: 5R21HD071758; Ministerio de Ciencia, Innovación y Universidades, Grant/Award Number: PGC2018-093786-B-I00

Abstract

Studies of Spanish grammatical gender have shown that native speakers exploit gender cues in determiners to facilitate speech processing and are sensitive to gender mismatches. However, past research has not considered attested distributional asymmetries between masculine and feminine gender, collapsing performance on trials with one or the other gender into a single analysis. We use event-related potentials to investigate whether masculine and feminine grammatical gender elicit qualitatively different brain responses. Forty monolingual Spanish speakers read sentences that were well-formed or contained determiner-noun gender violations. Half of the nouns were masculine and the other half were feminine. Consistent with previous research, brain responses varied along a continuum between LAN- and P600-dominant effects for both gender categories. However, results showed that individuals' ERP response dominance (LAN/P600) systematically differed across the two genders: participants who showed a LAN-dominant response to masculine-noun violations were more likely to show a P600 effect in response to feminine-noun violations. Correlations with individual difference measures further revealed that responses to masculine-noun violations were modulated by performance on the AX-CPT, a measure of cognitive control, whereas responses to feminine-noun violations were modulated by lexical knowledge, as indexed by verbal fluency. Together, the results demonstrate that even when processing features of language that belong to the same "natural class," native speakers can exhibit patterns of brain activity attuned to distributional patterns of language use. The inherent variability in native speaker processing is, therefore, an important factor when explaining purported deviations from the "native norm" reported in other types of populations.

KEYWORDS

event-related potentials, LAN, language variation, morphosyntax, P600, Spanish

1 | INTRODUCTION

Spanish nouns are classified into *masculine* and *feminine*, and Spanish agreement rules require that other grammatical elements (e.g., determiners, adjectives, etc.) match the gender of the noun they modify. Seminal work on grammatical gender processing indicates that native Spanish speakers exploit grammatical gender cues encoded in determiners to facilitate online sentence processing (Lew-Williams & Fernald, 2007) and are sensitive to grammatical gender assignment violations, as measured by event-related potentials (ERPs; Molinaro et al., 2011). However, this line of research has generally assumed that grammatical gender is a uniform construct, collapsing the two gender categories into a single analysis. Here, we propose that the distributional asymmetry between masculine and feminine gender categories—due to varying degrees of reliability and schematicity among their members—makes this assumption an oversimplification (Beatty-Martínez & Dussias, 2019). Before exploring our proposal more fully, we draw on the extensive literature on grammatical gender processing using ERPs, focusing specifically on the interplay between predictive and integrative mechanisms and how they may relate to dissociable ERP components. Our focus will be on the ERP technique because, in contrast to behavioral measures, ERPs provides continuous high temporal resolution indices at different neural stages of processing reflected by distinct components. Psycholinguistic research investigating the ERP correlates of grammatical gender processing in Spanish has revealed evidence of a biphasic pattern consisting of a left anterior negativity (LAN) around 300 ms after stimulus onset and a P600 after 500 ms (Barber & Carreiras, 2005; Caffarra & Barber, 2015; Molinaro et al., 2011). This pattern has been argued to reflect initial detection of a morphosyntactic expectancy violation (reflected in the LAN), followed by processes of repair or reanalysis and structural integration (reflected in the P600; Molinaro et al., 2011; Osterhout et al., 2004).

Despite the ubiquity of studies on grammatical gender processing, most previous research has collapsed across masculine and feminine gender categories and has not provided detailed distributional information about their stimuli. A notable exception is Caffarra and Barber (2015) who examined whether brain responses to gender violations would be differentially modulated depending on the consistency between noun endings and a specific gender class. The rationale was that the gender information for transparent nouns (these are nouns whose endings have a strong consistency associated with a specific gender; e.g., *-a* for feminine and *-o* for masculine) would be retrieved more rapidly compared to opaque nouns (i.e., nouns with ambiguous endings; e.g., *-e*). While transparent nouns elicited a greater sustained negativity around 200 ms after the stimulus onset—suggesting that participants were sensitive to noun endings—a LAN-P600 biphasic pattern was similarly reported for gender violations overall, regardless of the gender-to-ending consistency manipulation. A potential issue of this and other similar approaches is that they do not evaluate the impact of other gender cues beyond the word-final phoneme. Some evidence has shown that the correspondence between the gender of a noun and its phonological shape is not limited to the final phoneme, but may be based on additional factors such as the noun's penultimate and final rhyme or its final syllable (Eddington, 2002). In the current study, we take this hypothesis a step further by postulating that cue reliability is an attribute intrinsic to the gender category (i.e., whether the noun is masculine or feminine). We propose that the weighting of masculine versus feminine gender cues have come to differ systematically because masculine nouns are more diverse and heterogeneous as a category, while feminine nouns tend to cluster closer together and to exhibit similar characteristics, a point we elaborate below. A prediction that follows from this assumption is that brain responses to gender violations should systematically vary as a function of category membership; that is, whether a noun is masculine or feminine.

To our knowledge, only a few studies have investigated electrophysiological sensitivity to masculine and feminine gender violations separately. Alemán Bañón and Rothman (2016) examined native Spanish speakers' sensitivity to noun–adjective agreement violations during sentence comprehension. ERPs were time-locked to predicate adjectives in relative clauses (e.g., un uniforme que parecía sucia, "a$_{\text{MASC}}$ uniform$_{\text{MASC}}$ that looked dirty$_{\text{FEM}}$"), where half of the nouns were masculine and the other half, feminine. They found that while both types of gender agreement violations yielded a modulation of the P600, this effect emerged earlier for feminine-marked adjectives. The difference in P600 onset was taken to suggest that violations in feminine-marked adjectives were detected and revised more quickly, consistent with previous reports linking the timing of the P600 to the detection of a structural deviance (Sassenhagen & Bornkessel-Schlesewsky, 2015). Notwithstanding, there should be a clear distinction drawn between nouns, whose gender is inherent and fixed[1], and other grammatical elements such as determiners and adjectives, which are assigned gender via agreement with the noun. Indeed, noun–adjective agreement processing has been shown to differ from the processing of violations

[1]Exceptions include nouns that refer to humans but do not immediately indicate biological gender (e.g., *el/la testigo*; "the$_{\text{M/F}}$ witness"), although studies have found that the generic use of masculine forms may lead to biased representations of gender during language processing on the basis of social stereotypes (e.g., Carreiras et al., 1996; Gygax et al., 2020).

related to noun gender assignment (Barber & Carreiras, 2003; Dewaele & Véronique, 2001; Kupisch et al., 2013; Molinaro et al., 2011), and while adjectives and nouns have overlapping cues to gender, there are differences in marking consistency between them.

At this point, it is important to note that the reliable identification of the LAN effect has been called into question (Caffarra et al., 2019; Molinaro et al., 2015; Tanner, 2015), as there have been a number of studies on grammatical gender processing in which this effect is not evident[2] (e.g., Alemán Bañón et al., 2012; Alemán Bañón & Rothman, 2016; Barber & Carreiras, 2005; Guajardo & Wicha, 2014; Martín-Loeches et al., 2006; Wicha et al., 2004). This is especially true within the second language processing literature, where the biphasic LAN-P600 pattern is upheld as the benchmark of native-like linguistic representation, and any deviations from the target are treated as reflecting constraints associated with incomplete acquisition or deficient processing in bilingual speakers (Alemán Bañón et al., 2014, 2017; Gillon Dowens et al., 2010, 2011; see DeLuca et al., 2019; Dussias et al., 2019; McLaughlin et al., 2010, for further discussion).

In the morphosyntactic literature more generally, there is increasing evidence for individual differences with respect to processing strategies. Of relevance here, intersubject variability in ERP responses has previously been linked to not only experiential factors (see Caffarra et al., 2015, for a review), but also individual difference measures in both linguistic and cognitive domains (Batterink & Neville, 2013; Kim et al., 2018; Morgan-Short et al., 2012; Pakulak & Neville, 2010; Pélissier, 2020; Qi et al., 2017; Zirnstein et al., 2018). This work has shown that brain responses elicited by grammatical violations may not always be related to fixed and clearly identifiable components, but rather vary on a continuum ranging from preponderant-negative to preponderant-positive responses (Kim et al., 2018; Qi et al., 2017; Tanner, 2019; Tanner et al., 2013, 2014). An important implication is that systematic individual differences may give rise to nonrepresentative grand average ERPs. In particular, some studies have claimed that biphasic responses such as the LAN-P600 pattern are a spurious result of grand averaging individuals showing negativity- and positivity-dominant brain responses (Osterhout, 1997; Tanner & van Hell, 2014; c.f., Caffarra et al., 2019).

In sum, although much research has sought to understand the functional significance of brain responses to gender violations, the conditions that elicit them remain poorly understood. We suggest that these discrepant findings are, to some degree, attributable to representational differences between the two gender categories, individual differences in participants' weighting of gender cues, and an interaction between the two. Because we hypothesize that masculine and feminine gender violations would elicit qualitatively different brain responses, it will be important to examine whether systematic individual differences in participants' responses exist, and to the extent that they do, whether they are consistent across the two gender conditions in terms of distribution and variance.

1.1 | On the differential behavior of masculine and feminine gender in Spanish

Prior research has shown that Spanish speakers are sensitive to the systematic correspondences between a noun's phonemic make up and its gender, but one important difference between masculine and feminine gender categories is that nouns belonging to the masculine category have overall less phonological characteristics in common, which leads to biases in gender assignment (Eddington, 2002; see Beatty-Martínez & Dussias, 2019, for a review). For example, Spanish-speaking children are more likely to assign masculine gender to nouns with irregular (i.e., ambiguous) phonological cues (Pérez-Pereira, 1991), a tendency that has been linked to the higher rates of irregular masculine nouns in child-directed speech (Smith et al., 2003). More generally, native speakers have been shown to favor masculine over feminine when assigning gender to determiners of unknown nouns and loanwords. One recent example that illustrates this strong bias toward masculine gender assignment comes from the widespread criticism on Twitter following the Royal Spanish Academy's (Spanish: La Real Academia Española; RAE) decision in favor of the feminine form as the appropriate gender for COVID-19 (RAE, 2020). An online search comparing "el covid" to "la covid" shows how the masculine form predominates throughout most of the Spanish-speaking world (Data source: Google Trends (https://www.google.com/trends). This default status of masculine gender is further exemplified in nouns of Greek origin ending in -a (e.g., *el problema* "the problem," *el planeta* "the planet") despite that the vast majority of nouns with this phonological characteristic are feminine (e.g., *la pluma* "the feather," *la planta* "the plant"; Eddington, 2002). Perhaps more importantly, masculine determiners are also used in nominalizations of non-gender marked words (e.g., el fumar mata "the$_{\text{MASC}}$ smoking kills"; reemplaza

[2]Some studies on gender violations have reported an N400 rather than a LAN (e.g., Wicha et al., 2004). The degree to which the two components are functionally dissociable has been widely debated given that they occur in the same time window and have been elicited in similar morphosyntactic violation paradigms (see Molinaro et al., 2015, for a discussion). For example, Guajardo and Wicha (2014) have argued that the topographical distribution of the LAN may actually be an artifact of the spatiotemporal overlap between a broadly distributed N400 and an emerging right-posterior P600 (c.f., Mancini et al., 2011; Ojima et al., 2005, who reported LANs in the absence of a P600). Following Bornkessel-Schlesewsky and Schlesewsky (2019), we consider the LAN and N400 as part of the same family of language-related negativities, where topographical differences can be attributed to the nature of the linguistic representations that are involved.

este *aunque* por un *sin embargo*, “replace this$_{MASC}$ *still* for a$_{MASC}$ *nevertheless*”; see Harris, 1991, for further examples).

While it may be challenging to probe gender assignment criteria for nouns whose gender speakers know, converging evidence for this asymmetry comes from studies on reading comprehension and naturally occurring speech errors. For example, native Spanish speakers show faster reading times for feminine relative to masculine nouns in the processing of long-distance agreement dependencies (López Prego, 2015). They are also less accurate at detecting gender violations in masculine nouns, as compared to feminine nouns when performing under increased processing demands (López Prego & Gabriele, 2014). Furthermore, in an interesting experiment on hermaphroditic feminine nouns[3] Historically, *el* and *la* derived from Latin demonstrative articles ILLE and ILLA, respectively (e.g., ILLE REX > *el rey*; ILLA CASA > (e) la casa > *la casa*), an exception being feminine nouns beginning in a stressed /a/ (e.g., ILLA AQUA > el(a) agua > *el agua*). For these exceptions, definite determiners and determiners ending in /-una/ (*una* “a$_{FEM}$”; *alguna* “some$_{FEM}$”; *ninguna* “none$_{FEM}$”) must carry masculine gender if they immediately precede the noun to avoid a phonetic infelicity involving a word-final /a/ followed by a word-initial /a/., Eddington and Hualde (2008) found that native speakers tend to “incorrectly” overextend masculine agreement as shown in the example below:

(1)	(a)	Echa	tod**o**	el	agua	fría	en	el	barreño
		pour	all$_{MASC}$	the$_{MASC}$	water$_{FEM}$	cold$_{FEM}$	in	the	basin
instead of									
	(b)	Echa	tod**a**	el	agua	fría	en	el	barreño
		pour	all$_{FEM}$	the$_{MASC}$	water$_{FEM}$	cold$_{FEM}$	in	the	basin

(Eddington & Hualde, 2008, p. 4)

Finally, evidence from bilingual speakers also sheds light on the asymmetrical relationship between masculine and feminine. That is, a number of studies have shown that Spanish–English bilinguals tend to use masculine-marked determiners when code-switching (Beatty-Martínez & Dussias, 2019; Montes-Alcalá and Lapidus Shin, 2011; Valdés Kroff, 2016), leading to the production of mixed noun phrases with a masculine-marked determiner and an English noun (e.g., *el purse*, “the$_{MASC}$ purse $_{FEM}$”), even in instances when English nouns have a clear Spanish translation equivalent that would encourage the use of a Spanish feminine determiner (e.g., *la cartera*, “the$_{FEM}$ purse$_{FEM}$”). Against this backdrop, we argue that the distributional asymmetry between masculine and feminine gender reflects underlying differences in the representation of the two genders, with implications for gender expectancy and processing more generally.

1.2 | The present study

Following the proposal by Beatty-Martínez and Dussias (2019), we aim to test whether the distributional asymmetry between masculine and feminine gender categories differentially impact the qualitative use of gender cues during online sentence comprehension. To this end, we collected electrophysiological data from 40 monolingual speakers of Spanish while they read sentences that were either well-formed or contained a grammatical gender violation between a determiner and a target noun. Half of the target nouns were masculine (e.g., *vestido* “dress”) and half were feminine (e.g., *cartera* “purse”) in gender. The logic is that (1) if the attested distributional asymmetry between masculine and feminine gender reflects differences intrinsic to the characteristics of nouns (Beatty-Martínez & Dussias, 2019; Eddington, 2002), and (2) if readers and listeners make use of gender cues encoded in determiners to anticipate the gender of an upcoming noun (i.e., the gender anticipatory effect; Lew-Williams & Fernald, 2007; Wicha et al., 2004), masculine and feminine gender should differentially affect the neurocognitive processing of gender violations.

The first point appeals to the idea that there are key differences in the relative *schematicity* between the two gender categories (see Figure 1), where feminine nouns (yellow dots) hold a tighter range of variance compared to masculine nouns (blue dots; Beatty-Martínez & Dussias, 2019) or in the words of Eddington (2002, p. 66): “a random throw of the dart onto a map of nouns organized according to phonological similarities, has a much higher probability of landing in a neighborhood of masculine nouns, even if they do not dominate feminine nouns numerically.”

We follow recent accounts in assuming that the P600 is sensitive to processes of categorization and recognition (Sassenhagen et al., 2014), signaling in this case the degree to which nouns are classified as belonging to a particular gender category. This approach thus predicts an interaction between determiner-noun congruency and category schematicity, such that the P600 response to gender violations will be more robust for feminine nouns (see Alemán-Bañón & Rothman, 2016; Mehravari et al., 2015, for similar arguments). Concerning the second point, we draw on the

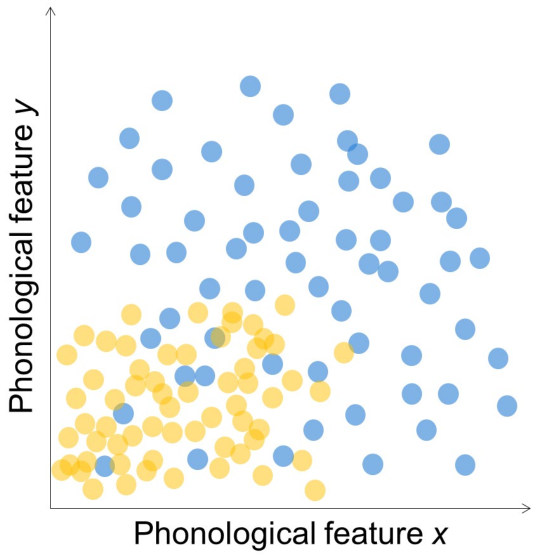

FIGURE 1 Hypothetical illustration of feminine (yellow) and masculine (blue) category structure. Exemplars closer to the x- and y-intercept represent members whose phonemic makeup is more prototypical of the category, and those departing from it represent members sharing less phonological characteristics in common. While masculine and feminine gender categories are very similar numerically (masculine and feminine gender nouns are distributed approximately equally in Spanish; Bull, 1965); they nonetheless show varying degrees of schematicity: the masculine gender category is more schematic, allowing exemplars that are not closely similar to each other. Conversely, the feminine gender category is less schematic, showing higher degrees of similarity among its members

premise of the predictive encoding framework developed by Bornkessel-Schlesewsky and Schlesewsky (2019), in that we expect gender cues encoded in determiners to elicit larger prediction errors in the form of a LAN in contexts in which expectations are not met (i.e., for violation conditions, where the noun of the unexpected gender is encountered). Based on evidence just reviewed, as well as previous research showing that comprehenders build up stronger noun gender expectations when presented with a feminine determiner than with a masculine determiner (see Footnote [4]Simply put, while pre-target feminine gender cues categorically exclude masculine referents, there is greater uncertainty about upcoming referents following masculine gender cues (e.g., nominalizations: *el nadar es buen ejercicio* "swimming is good exercise," hermaphroditic feminine nouns: *el agua* "the$_{\text{MASC}}$ water$_{\text{FEM}}$," and foreign or code-switched nouns: *el firewall* "the firewall"; Beatty-Martínez & Dussias, 2019; Butt & Benjamin, 2012; Eddington & Hualde, 2008).; Beatty-Martínez, 2019; Beatty-Martínez & Dussias, 2017; Dussias et al., 2013; Valdés Kroff et al., 2017; see also Molinaro et al., 2011), we predict that masculine-noun violations (e.g., **la vestido*; "*the$_{\text{FEM}}$ dress$_{\text{MASC}}$") will show a more reliable LAN effect than feminine-noun violations. By using ERPs to examine brain potentials to masculine and feminine gender violations separately, we can provide, to our knowledge, the first direct evidence at the neural level for an underlying difference in the processing of masculine and feminine gender in Spanish nouns. Moreover, to the extent that qualitative differences in brain responses exist between the two genders, we aim to identify whether they are related to individual differences in lexical knowledge and in the ability to efficiently adapt when outcomes contradict expectations. To this end, we will correlate two individual difference measures—category verbal fluency and the AX-continuous performance task. For verbal fluency, the rationale is that individuals with larger vocabularies (i.e., those who name more noun exemplars in the task) have greater experience with gender-to-noun-form correspondences, further defining the differences in the relative schematicity between the two gender categories. The AX-CPT is a measure of cognitive control engagement that pits effects of reactive response inhibition against the effects of proactive goal maintenance and conflict monitoring, and that has been previously shown to modulate individuals' brain responses during sentence comprehension (Zirnstein et al., 2018). We present an analogy between processing determiner-noun grammatical gender violations and AY response times (as elaborated in below in Section 3.2.4), an index of processing difficulty involved in inhibiting an incorrect response when predictions are disconfirmed.

2 | METHOD

2.1 | Participants

We tested 40 (30 female) right-handed native speakers of Spanish from the University of Granada, Spain. All participants gave informed consent and the procedures had the approval of the Pennsylvania State University Institutional Review Board. Participants were paid 10 euros per hour for their participation. Participants provided proficiency self-ratings for Spanish (the native language) and any additional languages that they had learned inside and outside of the home. While participants reported having studied some English in school, all reported that they were functionally monolingual, which is a common characteristic in the Andalusian region of Spain (Perrotti, 2012). Participant characteristics are summarized in Table 1.

2.2 | Experimental design and procedure

Two hundred and forty semantically low-constraint sentences were constructed using the same masculine–feminine noun pairs from a previous electrophysiological study (Beatty-Martínez & Dussias, 2017). Each member of an experimental masculine–feminine noun pair was combined with both a masculine and a feminine determiner resulting in four experimental conditions: congruent masculine trials (e.g., La mujer compró *el*$_{\mathrm{MASC}}$ *vestido*$_{\mathrm{MASC}}$ en la tienda), incongruent masculine trials (e.g., La mujer compró **la*$_{\mathrm{FEM}}$ *vestido*$_{\mathrm{MASC}}$ en la tienda; "*" indicates incorrect gender), congruent feminine trials (e.g., La mujer compró *la*$_{\mathrm{FEM}}$ *cartera*$_{\mathrm{FEM}}$ en la tienda), and incongruent feminine trials (e.g., La mujer compró **el*$_{\mathrm{MASC}}$ *cartera*$_{\mathrm{FEM}}$ en la tienda). Determiners were singular either definite or indefinite articles (e.g., *el/la*, "the$_{\mathrm{MASC/FEM}}$"; *un/una*, "a(n)$_{\mathrm{MASC/FEM}}$") and were kept the same across masculine and feminine noun pairs. For reasons described in the introduction (see also Footnote [3]), we did not include animate nouns or hermaphroditic feminine nouns as experimental material. Word frequency and orthographic length for all target nouns were obtained using the CLEARPOND database (Marian et al., 2012). Target masculine–feminine word pairs were matched in orthographic word length (masculine nouns: $M = 6.86$; feminine nouns: $M = 6.73$; $t(457.41) = 0.71$, $p = .477$) and word frequency (masculine nouns: $M = 1.19$; feminine nouns: $M = 1.24$; $t(460.90) = -0.95$, $p = .342$). See Table 2 for a sample experimental quartet.

Each participant saw 60 trials per condition resulting in a total of 240 unique trial quartets (see Supplementary Materials for complete list of experimental stimuli). The 240 unique trial quartets were distributed across four experimental lists in a Latin square design, such that each list contained only one version of each sentence, and no participant saw more than one version of each quartet. Each list contained 60 trials per condition, and all target nouns were counterbalanced across conditions. Experimental sentences were pseudo-randomized among 80 fillers. Filler trials were either syntactically correct or contained other types of agreement violations (e.g., between subjects and verbs: El perro *escondió/*escondieron* el hueso en el jardín. "The dog hid$_{\mathrm{SG/*PL}}$ the bone in the garden."). The total of 320 items was split into two experimental blocks with a short break between blocks to avoid fatigue, and the order of the blocks was counterbalanced between participants.

During ERP recording, participants were seated in a comfortable armchair in front of a computer screen in a sound-attenuated room with dimmed lighting. Participants were instructed to relax, minimize eye blinks and movements, and to read each sentence as normally as possible. Before the experiment proper began, participants were presented with a six-trial practice block to familiarize them with the trial sequence. Each trial started with a fixation cross "+" presented at the center of the screen (duration: 1,000 ms) after which a sentence was presented word by word using rapid serial visual presentation (RSVP), at a rate of 300 ms per word with a 350 ms interstimulus interval. Sentences were presented in lower case letters except for the first letter of the first word; the last word of each sentence was signaled with a period. The target noun was always in sentence-medial position. See Figure 2 for an example of the trial sequence.

Participants were asked to read for comprehension. Each sentence was followed by a picture asking for a sentence-picture relatedness judgment (e.g., a picture of a dress for the sample trial sentence in Figure 2). After the onset of the picture, participants were asked to respond whether the picture was semantically related or unrelated to the preceding sentence by, respectively, pressing "yes" or "no" on a button box. Sentence-picture association accuracy was 96% ($SD = 1.81\%$; ranging from 85 to 100%) and did not differ across experimental and filler trials ($F\text{s} < 1$), revealing that participants were paying attention to the task. Pictures were used in lieu of comprehension questions to keep the duration of the experiment relatively short, avoiding participant fatigue.

TABLE 1 Participant self-reported characteristics

Measure	*M*	*SD*	95% *CI*
Age, years	22	2.47	[21.2, 22.8]
Self-ratings: Spanish (/10)	9.4	0.9	[9.1, 9.7]
Self-ratings: English (/10)	4.0	1.5	[3.5, 4.5]

Note: Means, standard deviations, and 95% confidence intervals for age and proficiency self-ratings. Proficiency self-ratings were made on a 10-point scale ranging from 1 (not proficient) to 10 (highly proficient). All values represent raw, non-standardized scores.

TABLE 2 Experimental design and stimulus examples

Determiner	Noun	NP congruency	Sample trial
Masculine	Masculine	Congruent	La señora compró **el vestido** en la tienda.
Feminine	Masculine	Incongruent	La señora compró **la vestido** en la tienda.
Feminine	Feminine	Congruent	La señora compró **la cartera** en la tienda.
Masculine	Feminine	Incongruent	La señora compró **el cartera** en la tienda.

Note: NP = noun phrase; English translation: The woman bought **the dress/purse** at the store.

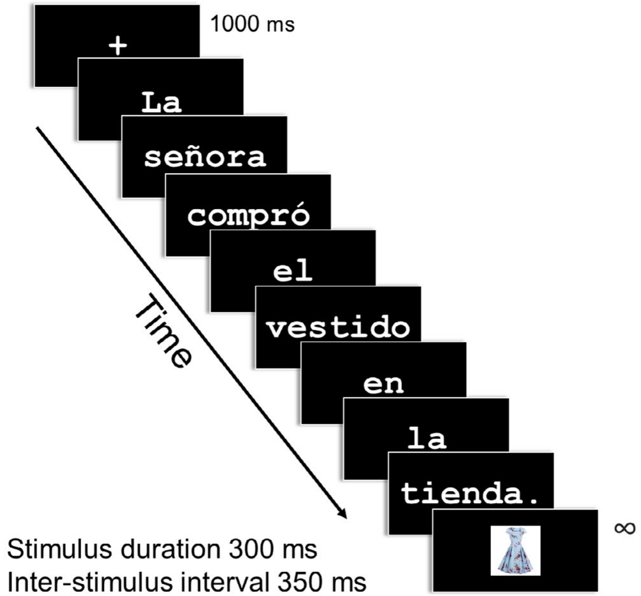

FIGURE 2 Trial sequence for the ERP sentence reading task. Each trial started with a fixation cross "+" presented at the center of the screen after which a sentence was presented word by word. Each word appeared in the middle of the screen for 300 ms with a 350 ms interstimulus interval. The target noun was always in sentence-medial position. Each sentence was followed by a picture, and participants were asked to respond whether the picture was semantically related or unrelated to the preceding sentence by, respectively, pressing "yes" or "no" on a button box. Pictures stayed on the screen until a response was recorded

2.3 | Individual difference measures

2.3.1 | Verbal category fluency

Participants performed a verbal category fluency task, in which they were asked to generate as many exemplars as possible that belong to a semantic category within a 30-second time limit. The task included four categories (out of eight: animals, body parts, clothing, colors, fruits, furniture, musical instruments, and vegetables) that were counterbalanced across participants. Participants were asked to avoid producing repetitions and names of people or places. Responses were recorded on a digital recorder. Performance was analyzed by calculating the total number of unique exemplars produced across categories.

2.3.2 | AX-continuous performance task

The distractor version of the AX-continuous performance task (AX-CPT; Ophir et al., 2009) was used to measure participants' ability to suppress irrelevant contextual information. As depicted in Figure 3, participants were presented five letters at the center of the screen, starting with a red cue, followed by three white distractors, and ending with a red probe. Each letter was presented for 300 ms, with a 1,000 ms interval between letters, and no extra pause between sequence trials. Participants were required to respond "no" to all cue and distractor letters. For the probe, they were required to respond "yes" only when they detected an AX sequence (i.e., an X-probe preceded by an A-cue). AX trials made up 70% of the total number of trials and each of the other three cue-probe combinations (AY, BX, and BY) made up one-third of the remaining 30%. As a result of this distribution of trials, cue information has been shown to serve a predictive function, with A-cues luring participants to generate strong predictions about the upcoming probe. However, the prepotent tendency to make a target "yes" response following A-cues creates an inappropriate expectancy bias for AY trials. As such, participants who greatly rely on cue information are likely to demonstrate increased error rates and slower response times (RTs) in this condition. Participants completed 10 practice trials including all four experimental conditions, and they were provided with feedback on accuracy and RT after each practice trial. The experimental block consisted of 100 trials with a small break halfway through. Error rates and RTs were recorded for each condition. RTs were computed from correct responses that were either above 100 ms or below 1,200 ms.

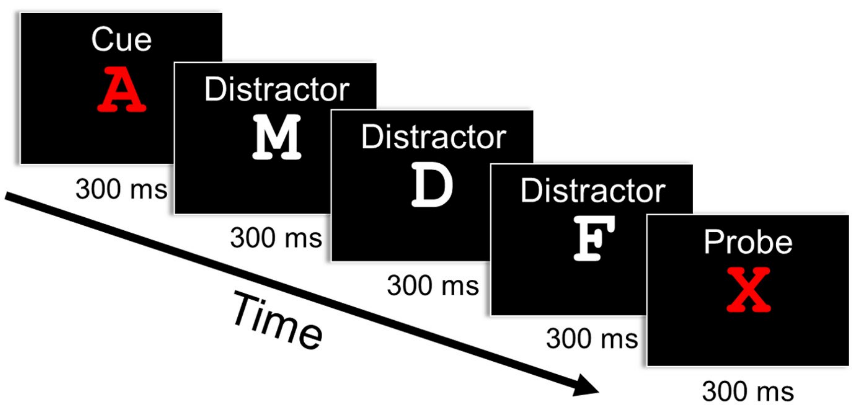

FIGURE 3 AX-CPT trial sequence. A-cues occurred in 80% of sequence trials. Of these, only 10 trial sequences had non-target (i.e., non-X) probes

2.4 | ERP recording

Electrophysiological data were recorded from 32 electrodes mounted in an elastic cap according to the international 10–20 system (Quick-Cap, Neuroscan Inc.) and a Synamps2 amplifier (Neuroscan, Inc.) with a 24-bit analog to digital conversion (online sampling rate: 500 Hz; 0.05–100 Hz band-pass filter). Impedances were kept below 5 KΩ. During recording, electrodes were referenced to a vertex electrode (REF), and the grounding electrode (GND) was mounted on the forehead. Blinks and eye movements were measured by placing bipolar pairs of lateral (HEOG) electrodes at the outer canthi of both eyes and vertical (VEOG) electrodes above and beneath the left eye. Using ERPLAB (López Calderón & Luck, 2014), ERPs were computed offline in each condition for each participant, time-locked to the onset of the target noun in each sentence, relative to a 200 ms pre-stimulus baseline. Trials contaminated by ocular or muscle artifacts were rejected using an amplitude threshold of ±100 μV. On average, 15.5% of trials were rejected due to artifacts, which did not differ across condition (*F*s < 2). Data were re-referenced to the average of the left and right mastoids, and they were refiltered with a Gaussian low-pass filter (25 Hz half-amplitude cutoff; slope 12 db/Oct). Individual averages were computed from artifact-free trials. Grand average amplitudes were then computed for each condition using the filtered data.

2.5 | ERP analysis

Amplitude measurement windows and electrode sites of interest were determined using a collapsed localizer (Luck & Gaspelin, 2017; see also Luck, 2014). Specifically, we averaged across incongruent and congruent conditions for both masculine and feminine genders. Based on the timing and scalp distribution of the collapsed waveform, the LAN time window was defined between 300 and 500 ms and was measured at F3, FC3, and C3. Accordingly, the P600 time window was defined between 500 and 800 ms and was measured at Pz, P3, and P4, where components reached their respective maxima. This approach is useful in reducing not only the problem of multiple comparisons inherent in ERP research, but also the probability of component overlap by measuring effects only at sites where the component of interest was large and the other relatively small, thus, preserving maximal discriminatory power (Kappenman & Luck, 2012; Luck, 2014; Luck & Gaspelin, 2017). Importantly, electrode regions and time windows are consistent with previous reports on grammatical gender processing (see Molinaro et al., 2011, for similar time windows).

Grand average ERPs were entered into a repeated measures ANOVA with noun gender (two levels: masculine and feminine), congruency (two levels: congruent gender and incongruent gender), and electrode site (three levels: F3, FC3, and C3 for the LAN; Pz, P3, and P4 for the P600) as within-subject factors. The Greenhouse–Geisser correction with more than 1 degree of freedom was used in all relevant analyses to avoid Type I errors for violations of sphericity. In such cases, the corrected *p* value is reported. Partial eta squared was calculated to determine the effect sizes for both significant and nonsignificant comparisons.

3 | RESULTS

3.1 | Grand mean ERP results

3.1.1 | Grand average ERP time-locked to target nouns

Grand average ERP results for congruent and incongruent conditions are depicted in Figure 4 for masculine nouns (Figure 4a,c) and feminine nouns (Figure 4b,d). Visual inspection of the waveforms and scalp topography plots showed that masculine gender violations elicited a biphasic waveform characterized by an anterior negativity with a left lateralization between approximately 300 ms and 500 ms poststimulus onset (LAN), followed by a posteriorly distributed positivity (P600). Feminine gender violations did not show an earlier negativity, but instead a large broadly distributed positivity (P600). Statistical analysis confirmed these observations. In the 300–500 ms time window, there was a significant Noun Gender x Congruency interaction ($F(1, 39) = 4.38$, $p = .043$, $\eta_p^2 = .101$). A breakdown of this interaction showed a significant congruency effect for masculine nouns ($F(1, 39) = 11.22$, $p = .002$, $\eta_p^2 = .223$), but not for feminine nouns ($p = .739$), indicating that the LAN effect was larger for masculine versus feminine nouns. In the 500–800 ms time window, the Noun Gender × Congruency × Electrode interaction was significant ($F(1.56, 60.75) = .534$, $p = .012$, $\eta_p^2 = .120$). Subsequent analyses of this interaction revealed congruency main effects at Pz ($F(1, 39) = 13.10$, $p = .001$, $\eta_p^2 = .251$) and P4 ($F(1, 39) = 14.08$, $p = .001$, $\eta_p^2 = .265$) electrodes sites, and a Noun Gender x Congruency interaction at P3 ($F(1, 39) = 4.39$, $p = .043$, $\eta_p^2 = .101$). Follow-up paired sample *t* tests for this interaction showed significant mean differences between congruent and incongruent conditions for feminine nouns ($t(39) = 3.56$, $p = .001$), but not for masculine nouns ($p = .680$), indicating that the while both genders showed a P600 effect to congruency manipulations, this effect was more broadly distributed for feminine than masculine nouns.

3.1.2 | Control analysis

Because our experimental design inherently involved systematic pretarget word differences between masculine and

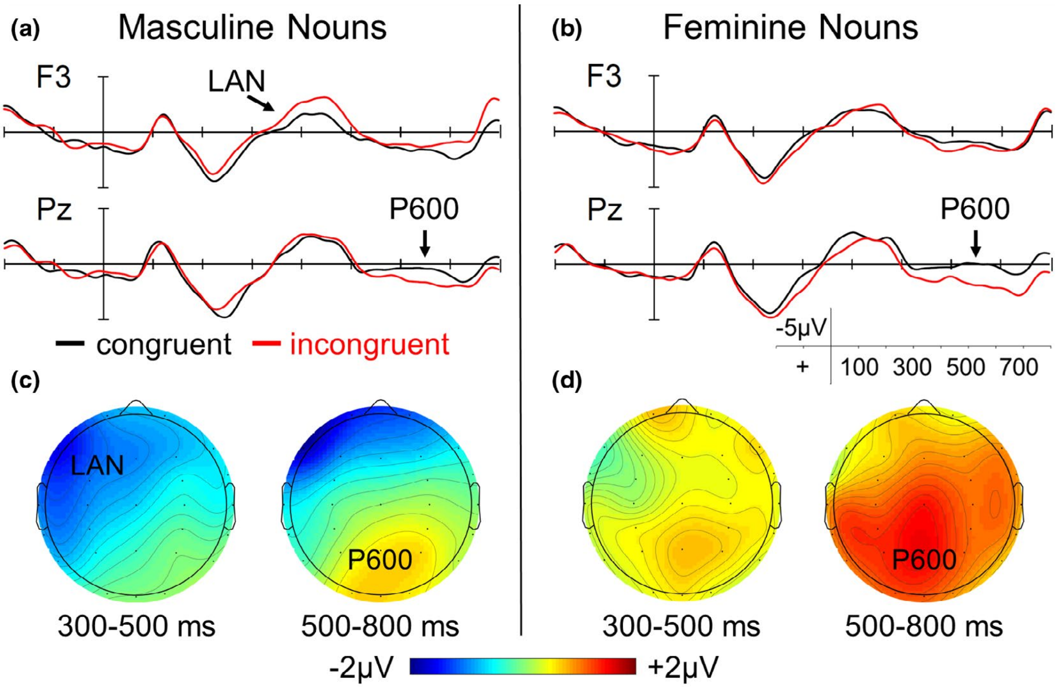

FIGURE 4 Grand-averaged ERPs time-locked to the onset of masculine (a) and feminine (b) target nouns at F3 and Pz electrodes for congruent (black) and incongruent (red) conditions. Negative is plotted up. Scalp topographic maps of the congruency effect in masculine (c) and feminine (d) target nouns. Differences were calculated by subtracting mean congruent noun amplitudes from mean incongruent noun amplitudes. Scale is from −2 (blue) to +2 (red) microvolts

feminine conditions (i.e., masculine-noun violations were always preceded by feminine determiners la/una; feminine-noun violations were always preceded by masculine determiners el/un), we additionally analyzed ERPs time-locked to the onset of pretarget determiners to rule out the possibility of different spillover or offset artifacts on the target nouns (e.g., Steinhauer & Drury, 2012). As shown in Figure 5, masculine- and feminine-gendered determiners did not yield differential ERP responses, validating the comparison of target nouns across congruent and incongruent conditions[5].

3.2 | Individual difference analyses

3.2.1 | The negativity-to-positivity continuum

Because the argument made here is that masculine and feminine gender violations should elicit qualitatively different brain responses, and the biphasic response was only observed for masculine nouns, we employed an individual differences approach to determine if there were differences in the amount of variance between the two gender conditions. For the purposes of these analyses, we selected the most representative electrode for each of the two components: F3 and Pz for the LAN and P600, respectively.

Figure 6 shows the distribution of participants' LAN and P600 effect amplitudes for masculine and feminine nouns. Consistent with previous findings (Qi et al., 2017; Tanner et al., 2013, 2014; Tanner & van Hell, 2014), participants' ERP responses varied along a continuum ranging between negativity- and positivity-dominance, with a subset of participants (i.e., those in the upper right quadrants) showing signs of a biphasic LAN-P600 pattern. Furthermore, there was a strong inverse relationship between LAN and P600 effects, indicating that participants who exhibited a predominant P600 tended to show a less robust LAN, and vice versa (masculine nouns: $r = -.44$, $p = .004$; feminine nouns: $r = -.60$, $p < .001$). While these observations hint at a possible LAN-P600 trade-off (Kim et al., 2018), comparisons of LAN and P600 effects across masculine and feminine-noun conditions showed the opposite association. First, mean LAN amplitudes were inversely correlated across the two gender conditions ($r = -.44$, $p = .005$). Second, and perhaps more importantly, the LAN effect for masculine nouns showed a positive correlation with the P600 effect for feminine nouns ($r = .55$, $p < .001$). Overall, this suggests that participants

[5]We also considered the possibility that participants may have generated different expectations for the gender of the determiner based on prior sentence context. If this were so, we would expect different patterns of ERP responses between determiners with the contextually expected gender and those with the contextually unexpected gender (e.g., Foucart et al., 2014; Molinaro et al., 2017; Wicha, et al., 2004). The finding that masculine- and feminine-gendered determiners did not yield differential ERP responses confirms that the sentences were overall low constraint and had comparable cloze probability.

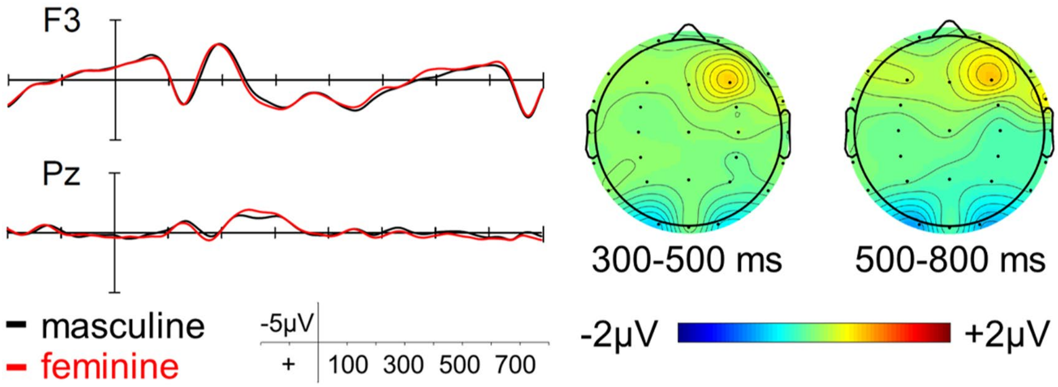

FIGURE 5 Left: Grand-averaged ERPs time-locked to the onset of masculine (black) and feminine (red) determiners at F3 and Pz electrodes. Negative is plotted up. Right: Scalp topographic maps showing grand mean effects between 300–500 ms and 500–800 ms. Differences were calculated by subtracting mean feminine determiner amplitudes from mean masculine determiner amplitudes. Scale is from −2 (blue) to +2 (red) microvolts

who showed highly predominant LANs to masculine violations, showed highly predominant P600s to feminine-noun violations. Finally, there were no reliable associations between the P600 effect elicited by masculine gender violations and brain responses for feminine nouns at either time window (feminine LAN: $r = .15$, $p = .351$; feminine P600: $r = -.07$, $p = .691$). In all, these findings are relevant as they demonstrate that individual differences reflect not only differences in processing strategies across participants, but also their interaction with intrinsic differences between masculine and feminine gender.

To further assess whether there were differences in the range and type of variability elicited by masculine and feminine gender conditions, measures of response magnitude and dominance were calculated for each participant, following Tanner et al. (Grey et al., 2017; Tanner et al., 2013, 2014; Tanner & van Hell, 2014). The Response Magnitude Index (RMI) is a measure of participants' absolute magnitude of congruency effects, where larger RMI values indicate relatively greater ERP responses to the gender violations across both time windows, regardless of polarity. The Response Dominance Index (RDI) is a measure of participant's relative preponderance of a negativity over a positivity, regardless of magnitude. The RDI was calculated by fitting each participant's least squares distance from the equal effect sizes lines (the dashed line in Figure 6; see Footnote [6]) with perpendicular offsets for both masculine and feminine congruency effects (see formulas below, derived from Tanner et al., 2014, where LAN and P600 refer to the mean amplitude at F3 and Pz electrode sites, respectively).

$$\mathrm{RMI} = \sqrt{\left(\mathrm{LAN}_{\mathrm{Congruent}} - \mathrm{LAN}_{\mathrm{Incongruent}}\right)^2 + \left(\mathrm{P600}_{\mathrm{Incongruent}} - \mathrm{P600}_{\mathrm{Congruent}}\right)^2}$$

$$\mathrm{RDI} = \frac{\left(\mathrm{P600}_{\mathrm{Incongruent}} - \mathrm{P600}_{\mathrm{Congruent}}\right) - \left(\mathrm{LAN}_{\mathrm{Congruent}} - \mathrm{LAN}_{\mathrm{Incongruent}}\right)}{\sqrt{2}}$$

[6]We note that dashed line implicitly assumes that LAN and P600 effects are commensurate which is unlikely the case (i.e., if one ERP has a maximum of 5μV and the other has a maximum of 2μV, then, a 1μV modulation is 20% for the former but 50% for the latter). However, we wish to point out that the theoretical takeaway from the individual differences analyses is not whether one ERP effect is larger/smaller than the other, rather if participants' response dominance measures differ between the two genders.

3.2.2 | Response magnitude

As depicted in Figure 7, congruency effect magnitudes across masculine and feminine nouns were positively correlated ($r = .68$, $p < .001$) and showed a highly similar distribution of variation (F-test for equality of two variances: $F(39, 39) = 0.69$, $p = .243$). Importantly, participants were not differentially sensitive to gender violations between masculine ($M = 3.26$, $SD = 2.01$) and feminine ($M = 3.21$, $SD = 2.53$) nouns ($t(39) = 0.18$, $p = .860$).

3.2.3 | Response dominance

Distributions of RDI values in masculine and feminine nouns are depicted in Figure 8. More negative values indicate a preponderance of LAN effects, whereas more positive values indicate a preponderance of P600 effects. Values closer to zero reflect biphasic responses (i.e., relatively equal-sized LAN and P600 effects). Several features of the distributions in Figure 8 are worth noting. First, both distributions are closely centered around zero, but the tails of the distributions are asymmetric, with RDI values for masculine nouns extending in the negative direction and RDI values for feminine nouns extending in the positive direction.

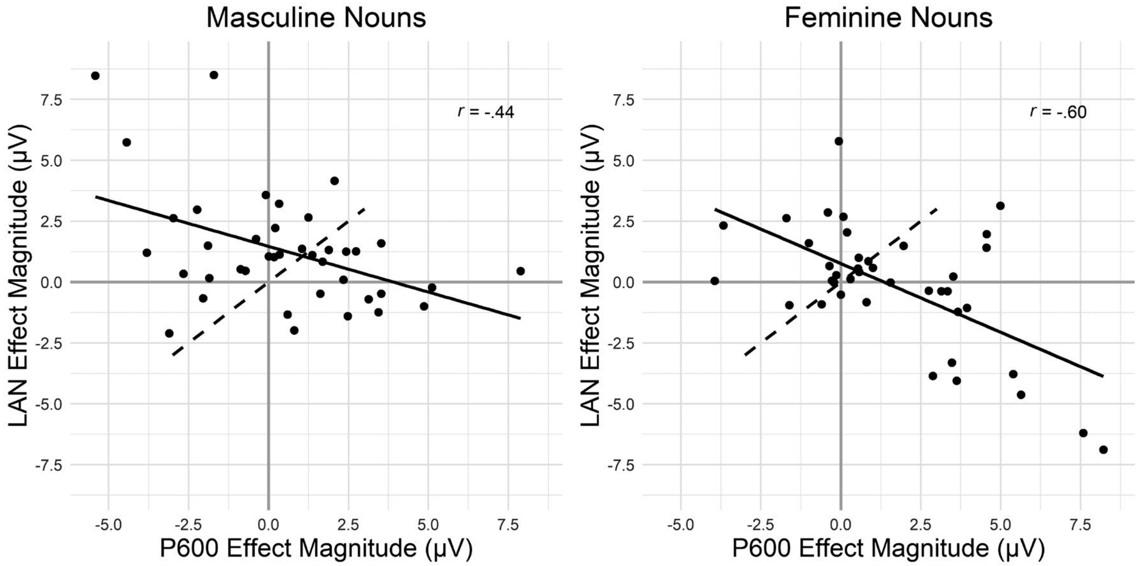

FIGURE 6 Scatterplots showing the distribution of LAN (*y* axis) and P600 (*x* axis) effect amplitudes for masculine (left panel) and feminine (right panel) noun conditions. Each dot represents a data point from a single participant. Mean amplitudes were extracted from F3 and Pz electrode sites within the LAN and P600 windows, respectively. The solid lines indicate the best-fit regression line relating participants' LAN and P600 effect magnitudes. The dashed lines represent equal LAN and P600 effect amplitudes. Participants above and left of the dashed line showed LAN-dominant responses to gender violations, while participants below and right of the dashed line showed primarily P600-dominant responses

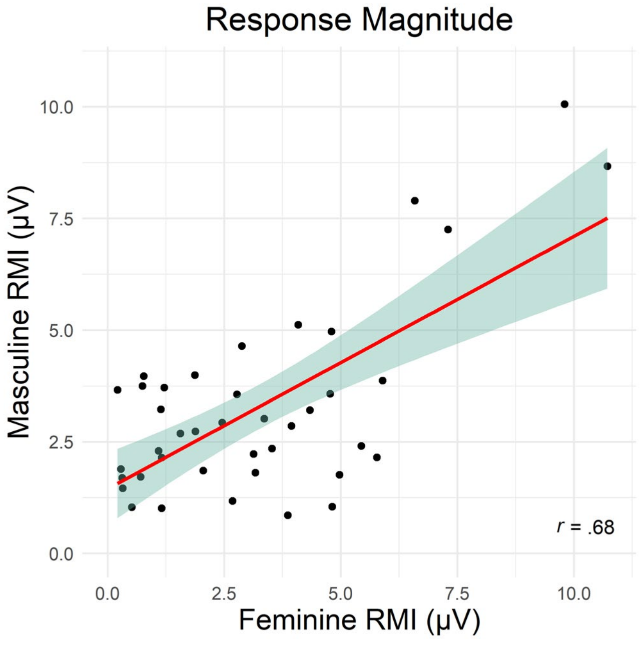

FIGURE 7 Correlation between participants' response magnitude indices for masculine and feminine nouns. Larger values indicate relatively greater ERP responses to the gender violations across both time windows, regardless of polarity

FIGURE 8 Violin plots visualizing the distribution of RDI values as a function of noun gender. More negative values reflect LAN-dominant responses to gender violations, whereas more positive values reflect P600-dominant responses. Values closer to zero reflect biphasic responses (i.e., relatively equal-sized LAN and P600 effects). The black dots represent the mean RDI value across participants and the black lines represent the 95% confidence intervals

Second, in contrast to other studies showing a positive association with participants' RDI scores across different conditions or tasks (e.g., Tanner & van Hell, 2014), RDI values for masculine and feminine conditions were inversely correlated ($r = -.33$, $p = .041$). This indicates that participants tended to show opposite polarity responses (i.e., LAN or P600) in masculine and feminine gender conditions. This is depicted in Figure 9. Furthermore, there were no variance differences

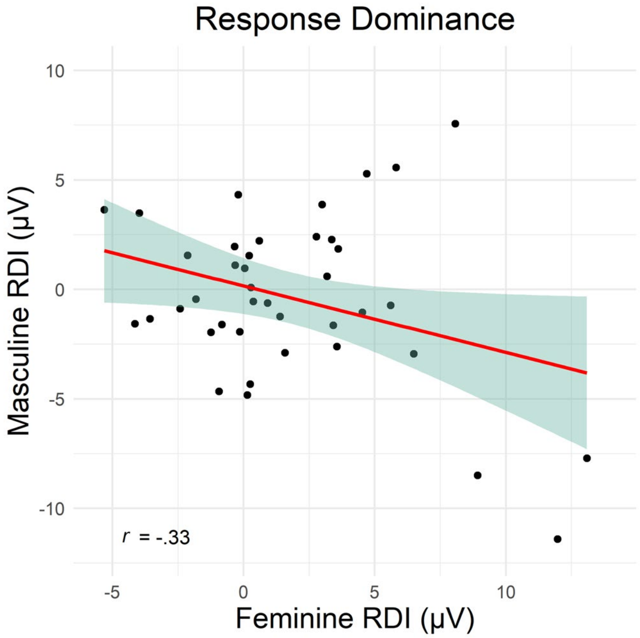

FIGURE 9 Correlation between participants' response dominance indices for masculine and feminine-noun conditions. More positive values indicate P600-dominant responses, whereas more negative values indicate LAN-dominant responses to the gender violations across both time windows, regardless of magnitude

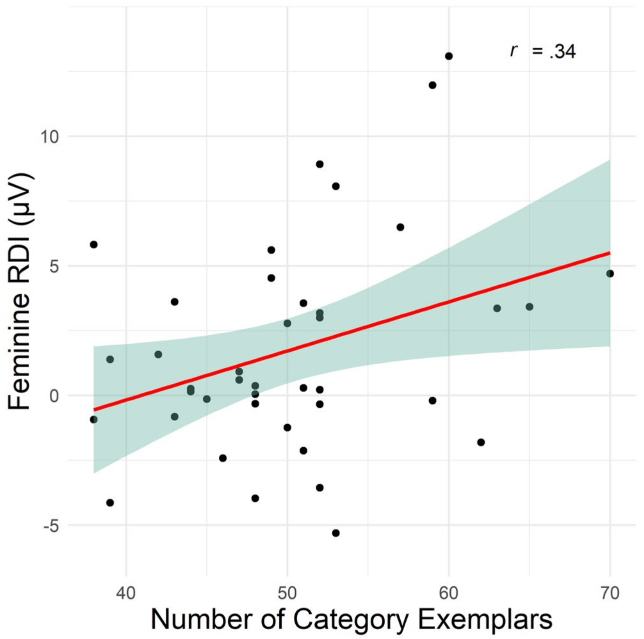

FIGURE 10 Correlation between participants' verbal fluency performance (total number of exemplars; *x* axis) and their RDI values (in microvolts) for feminine nouns (*y* axis). For verbal fluency, greater number of exemplars indicate better performance, while lower number of exemplars indicate poorer performance in the task. For RDI, more positive values indicate relatively P600-dominant responses, whereas more negative values indicate LAN-dominant responses to the gender violations across both time windows, regardless of magnitude

between masculine and feminine RDI values (F-test for variance: $F(39, 39) = 0.87$, $p = .664$). This finding is particularly important as it reveals that the variation in participants' brain responses is systematic rather than a function of chance due to sampling or measurement errors. Finally, RDI values were significantly more positive for feminine nouns ($M = 1.77$, $SD = 4.11$) than for masculine nouns ($M = -0.38$, $SD = 3.88$) as confirmed by a paired t test assuming equal variance ($t(39) = -2.10$, $p = .042$). This result further supports the relationship between asymmetries in the morphological encoding of gender information (where feminine nouns share higher homology than masculine nouns) and P600 effects related to morphological decomposition and reanalysis (for a similar rationale, see Mehravari et al., 2015).

3.2.4 | Relationships between RDI values and verbal fluency

Verbal fluency performance was based on participants' total number of unique category exemplars ($M = 50.3$ exemplars, $SD = 7.4$, 95% $CI = 47.9$–52.6). For the individual differences analyses, we used this measure as an index of lexical knowledge and correlated it with RDI values for both masculine and feminine-noun conditions separately. For masculine nouns, there was no relation between participants' RDI and verbal fluency scores ($r = .07$, $p = .683$). Notwithstanding, a more positive RDI in response to feminine gender violations was associated with better verbal fluency performance ($r = .34$, $p = .031$; see Figure 10). Thus, participants with greater lexical knowledge were more likely to show P600-dominant effects.

3.2.5 | Relationships between RDI values and the AX-CPT

Table 3 displays the means, standard deviations, and 95% confidence intervals for AX-CPT error rates and RTs across conditions. Linear and generalized mixed-model analyses were performed using the lme4 package (Bates et al., 2015). Models included dummy coded fixed effects of condition (AY, BX, and BY) with random intercepts for subjects and items. We focused specifically on comparing AY versus BY to measure the degree to which the cue bias negatively impacted probe responses, using BY as the reference level. For the individual differences analyses, average RTs in AY trials were calculated for each participant and taken as an index of processing difficulty involved in inhibiting an incorrect response when predictions are disconfirmed (e.g., Zirnstein et al., 2018).

Consistent with previous analyses on young adults, the AY condition revealed higher error rates ($\beta = .83$, $SE = 0.18$,

TABLE 3 Means, standard deviations, and 95% confidence intervals for AX-CPT error rates and RTs

Measure	*M*	*SD*	*95% CI*
AX error rate	0.12	0.11	[0.08, 0.15]
AY error rate	0.34	0.20	[0.27, 0.40]
BX error rate	0.23	0.19	[0.17, 0.30]
BY error rate	0.18	0.17	[0.13, 0.23]
AX RT (ms)	291.11	38.36	[278.04, 304.18]
AY RT (ms)	402.13	89.59	[371.55, 432.70]
BX RT (ms)	232.74	88.41	[202.66, 262.81]
BY RT (ms)	246.57	96.80	[213.94, 279.21]

Note: AX-CPT (the AX variant of the continuous performance task): A and B represent cues; X and Y represent probes. Two participants had missing data due to computer malfunction.

$z = 4.73$, $p < .001$) and longer response times ($\beta = .24$, $SE = 0.01$, $z = 72.2$, $p < .001$) relative to the BY control condition[7]. This suggests that on AY sequences, participants relied on contextual information (i.e., A-cues) to anticipate upcoming probe responses and, as a result, had greater difficulty inhibiting incorrect responses.

For masculine nouns, a correlational analysis (see Figure 11) indicated a significant inverse relationship between participants' reactive response inhibition ability and their response dominance to gender violations ($r = -.33$, $p = .038$). Longer response times on correct trials in the AY condition, indicating greater reliance on contextual information, were associated with negativity-dominant responses. In other words, participants with better inhibitory control ability also experienced less difficulty associated with encountering a prediction error. Inhibitory control ability and RDI values for feminine nouns were not significantly correlated ($r = -.27$, $p = .099$). This finding is consistent with previous work showing that cognitive control ability mediates prediction error costs in comprehension (Zirnstein et al., 2018).

4 | DISCUSSION

Applying ecological approaches that emphasize natural language use, we examined the extent to which distributional asymmetries between masculine and feminine grammatical gender affect the processing of gender violations in native monolingual Spanish speakers. Consistent with previous research, native Spanish speakers showed electrophysiological sensitivity to gender congruency violations between

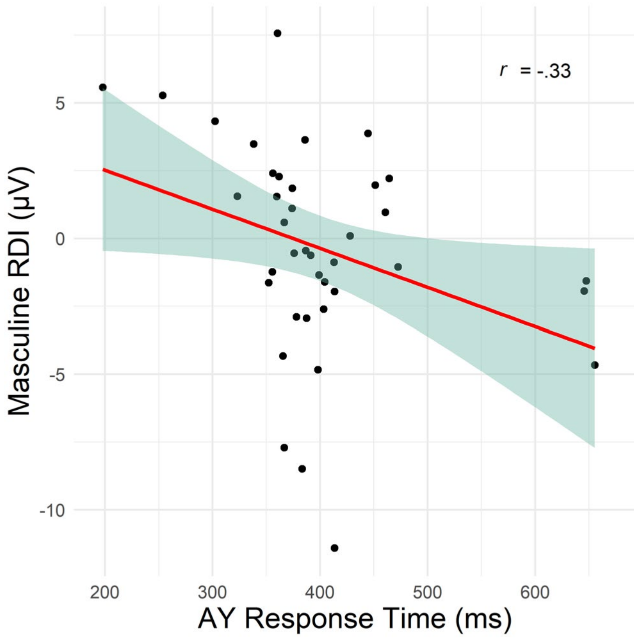

FIGURE 11 Correlation between participants' average response time (ms) on AY trials (*x* axis) and their RDI values (in microvolts) for masculine nouns (*y* axis). On the *x* axis, larger values indicate slower RTs on Y-probes in AY trials. On the *y* axis, more positive values indicate relatively P600-dominant responses, whereas more negative values indicate LAN-dominant responses to the masculine-noun gender violations across both time windows, regardless of magnitude

determiners and nouns (Caffarra et al., 2015). Notwithstanding, a striking feature that emerges from the experiment described here is that there are underlying systematic differences in the processing of the two gender categories. Although grand mean analyses showed a reliable biphasic LAN-P600 pattern for masculine nouns only, individual brain responses varied along a continuum between LAN-dominant, biphasic[8], and P600-dominant effects for both gender conditions. What is most striking about this result is that individuals' ERP response dominance (LAN or P600) systematically differed across the two genders: participants who showed a LAN-dominant response to masculine gender violations were more likely to show a P600 effect in response to violations of feminine nouns. These findings run counter to prior suggestions that an individual's ERP response dominance is stable,

[7]Unlike in the AY condition, BX and BY had similar error rates ($\beta = .30$, $SE = 0.18$, $z = 1.65$, $p = .099$), and BX yielded faster response times relative to BY ($\beta = -.03$, $SE = 0.01$, $z = 68.04$, $p = .047$), suggesting that participants were anticipating X-probes upon detection of the B-cue (see Beatty-Martínez et al., 2020 for a similar pattern).

[8]Only a subset of the participants showed some degree of biphasic response across conditions (33% for masculine nouns and approximately 23% for feminine nouns; i.e., those in the upper right quadrants of the scatterplots; in Figure 6). Notwithstanding, this suggests that both responses (LAN and P600) can co-occur within individuals for whom the two genders are represented more similarly. This finding is also consistent with mixed evidence on the gender anticipatory effect (Lew-Williams & Fernald, 2007), with some studies showing similar effects for both masculine and feminine gender cues (Beatty-Martínez, 2019; Dussias et al., 2013), and others only for feminine cues (Dussias et al., 2013; Valdés Kroff et al., 2017).

reflecting a consistent processing strategy (e.g., Tanner, 2019). Instead, they demonstrate that there are intrinsic differences between masculine and feminine gender interacting with intersubject variability. What is important to consider is that, by most past accounts, masculine and feminine gender conditions would be assumed to be derived from the same construct. The results obtained here suggest that they are not.

4.1 | On the functional significance of the LAN and P600 components

The findings reported here deepen our understanding of the functional significance underlying language-related ERP effects. First, we discuss the pattern of the LAN, specifically its prevalence in response to masculine but not feminine-noun violations. One possibility, alluded to in the introduction, is that the LAN may reflect morpheme-based expectations (Bornkessel-Schlesewsky & Schlesewsky, 2009) as influenced by cue validity, where a cue is highly valid when it is high in both applicability (i.e., available when needed) and reliability (i.e., neither misleading nor ambiguous; MacWhinney et al., 1984; see also, Li et al., 1993). This interpretation is congenial with the proposal that prediction errors manifested in the form of language-related negativities are weighted by the interplay between form-based bottom-up factors and top-down processes (e.g., contextually induced pre-activation of the target gender; Bornkessel-Schlesewsky & Schlesewsky, 2009, 2019; Lotze et al., 2011). Recall that, in Spanish, masculine gender has a default status that distinguishes it from feminine gender, as evidenced by its use wider range of applications[4]. Thus, a possible explanation for the LAN (or lack thereof) is that feminine gender cues preceding masculine-noun violations may have induced greater prediction error than masculine gender cues preceding feminine-noun violations (see Figure 12). Most telling is the association between the LAN effect and the AX-CPT, where the LAN effect for masculine nouns was attenuated by reactive response inhibition ability. The implication here is that stronger expectations may incur greater prediction error but engaging cognitive control may be able mediate the conflict between the predicted and the encountered gender. This result is consistent with previous work on semantic prediction during online reading (e.g., Gernsbacher et al., 1990; Zirnstein et al., 2018), and we expect processes that generate predictions (e.g., other types of agreement processes; c.f., Tanner, 2019) to show similar associations. Thus, we hope this result opens an exciting range of possibilities of studying the role of cognitive control in mediating error costs when comprehenders' predictions are disconfirmed. More speculatively, if the LAN effect is related to prediction error costs based on the associational strength of gender cues, then, we expect that the reliability of this effect should depend on how important

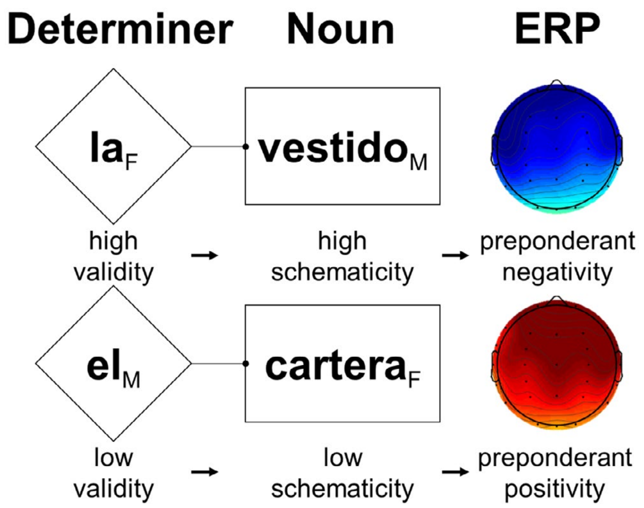

FIGURE 12 A scheme of the purported systematic differences in determiner cue validity, category schematicity, and brain responses to grammatical gender violations for masculine and feminine nouns. Because feminine gender cues encoded in determiners categorically exclude masculine referents (but not vice versa), individuals' predictions regarding the gender of an upcoming noun will be more reliable, and hence, stronger than those based on masculine cues. Violation of these stronger predictions thus results in the preponderance of LAN-dominant effects. Similarly, invalidly cued feminine nouns will elicit P600-dominant effects because feminine nouns have higher category coherence with relatively lower degrees of phonological schematicity relative to masculine nouns

gender cues are for interpretation in a given language (see Molinaro et al., 2011, for a review of studies reporting LAN effects across different languages) and even within experiments, given past findings showing that individuals adapt and readjust their expectations during comprehension to reduce future prediction error (Beatty-Martínez, 2019; Dell & Chang, 2014; Fine & Jaeger, 2013; Kleinman et al., 2015).

Whereas the LAN was predominantly observed for masculine-noun violations, the P600 effect was observed for both types of violations. Notwithstanding, it is noteworthy that the P600 effect was more broadly distributed for feminine-noun violations than that for masculine-noun violations. This is perhaps unsurprising given previous evidence that suggests that the P600 is particularly sensitive to grammatical violations with visibly more robust and reliable cues (e.g., morphological complexity: the sheep were *graze; the sheep should *grazing; "*" indicates ungrammaticality; Mehravari et al., 2015). We also attribute this finding to differences in the relative schematicity between the two gender categories, where feminine nouns hold a tighter range of variance compared to masculine nouns (Beatty-Martínez & Dussias, 2019, see Figures 1 and 12). Furthermore, our results revealed that category verbal fluency performance modulated participants' response dominance. Specifically, participants with greater lexical knowledge had a higher likelihood of showing a

P600-dominant response to feminine- (but not masculine-) noun violations. We interpret this asymmetrical relation as reflecting differences in the representational strength of the two gender categories. Under this proposal, the P600 can be seen as a reliable indicator of the degree to which nouns are classified as belonging to a particular gender category. Thus, our findings point to a novel interpretation of the P600 as an index of category exemplar strength, with categories of higher coherence (here, the feminine gender category), yielding a more salient neurophysiological response to gender violations. Altogether, we extend previous findings showing similar associations (Batterink & Neville, 2013; Pakulak & Neville, 2010) by considering the usage-based view that tokens of linguistic experience build up cognitive representations and affect language form and language processing (Beatty-Martínez & Dussias, 2018; Bybee, 2010; Dell & Chang, 2014). Future research may seek to determine the extent to which other experiential and/or language background factors such as reading skill and education will additionally show similar associations (e.g., Dąbrowska, 2012; Hoff et al., 2018; Montag & MacDonald, 2015).

4.2 | Final considerations

Indeed, there is much to be gained by considering subject- and item-related variability in grammatical gender processing, and fortunately there is a growing trend in this direction (e.g., Caffarra et al., 2019; Tanner, 2019). The current study provides an important contribution to this work: Although the use of individual difference analyses offers an opportunity for a more nuanced view of brain potentials, they may nonetheless involve the same risks as those observed with grand averages in the absence of well-defined envelope of language variation and a well-honed understanding of how the implications of such variation affect language processing. In this article, we have argued that accounting for distinct sources of linguistic variation in natural language use is key to understanding variability in language processing. The present work aligns nicely with a growing body of research implicating both patterns of natural language use and subject variability as key determinants of sources of variation in language processing more generally (Boland et al., 2016; Dussias et al., 2019; Fricke et al., 2019; Wigdorowitz et al., 2020). What we have yet to address is how this type of usage-based approach can be applied more widely to questions about the dynamics of the language system.

From a theoretical perspective, the ramification of these findings extends beyond the functional significance of these components. Of particular interest, our findings cut across traditional component interpretations that have been assumed in the electrophysiological literature on second language processing, namely that the absence of the biphasic LAN-P600 response can be taken as evidence for constraints associated with morphosyntactic representations and native-like attainment more generally (Batterink & Neville, 2013; Hahne & Friederici, 2001). Our experiment revealed that masculine and feminine gender violations elicit qualitatively different ERP response patterns in monolingual Spanish speakers, owing to underlying differences in the representation of the two gender categories. Especially noteworthy is that the systematic variability in dominance of the LAN and the P600 would have been overlooked entirely had we not carefully evaluated both stimulus- and participant-related characteristics. Thus, these findings provide further support to the argument that native-likeness is not, in and of itself, a reliable yardstick for identifying processing capabilities, since participants displayed a negativity-to-positivity continuum of responses that were additionally modulated by linguistic and domain general individual difference measures (Kim et al., 2018; Qi et al., 2017; Tanner, 2019; Tanner et al., 2013, 2014). Instead, speakers' ability to establish correspondences between a noun's phonological shape and a particular gender category (e.g., Eddington, 2002) may be a more suitable benchmark. In previous studies, both L2 learners and heritage speakers of Spanish have been found to overgeneralize masculine gender in gender assignment errors, predominantly substituting feminine for masculine forms (e.g., **el* pared, "the$_{\mathrm{MASC}}$ wall$_{\mathrm{FEM}}$"), in production (Montrul et al., 2008) and comprehension (McCarthy, 2007). These observations suggest that feminine gender cues may not have as high predictive validity as for (some) native speakers. With increased knowledge of lexical gender and proficiency (e.g., Hopp, 2016; Lemhöfer et al., 2014; Morgan-Short et al., 2012), however, gender-to-noun-form correspondences are expected to be identified and cluster within categories, gradually developing in a manner potentially analogous to native speakers. In short, variability in the language experiences of individuals (L1 and L2 speakers alike) and in the ability of individuals to learn from these experiences, may better explain variability of language processing itself (e.g., Dussias et al., 2019; Fricke et al., 2019). Moving forward, language processing research must reexamine assumptions regarding the status of gender, recognizing masculine and feminine gender categories as heterogeneous and gradient, with implications for morphosyntactic expectancy and processing.

ACKNOWLEDGMENTS

We are grateful to Christian A. Navarro Torres and Rena Torres Cacoullos for their helpful comments and discussions during the preparation of this manuscript. This work was supported in part by National Institutes of Health grants F32-AG064810 to A. B. M and 5R21HD071758 to P. E. D., National Science Foundation grants BCS-1535124, OISE-0968369, OISE-1545900, to P. E. D., Ministerio de Ciencia, Innovación y Universidades grant PGC2018-093786-B-I00 to M. T. B. and Ministerio de Economía y Competitividad grant PCIN-2015-132 to M. T. B. The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

AUTHOR CONTRIBUTION

Anne L. Beatty-Martínez: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Visualization; Writing-original draft; Writing-review & editing. **Michelle R. Bruni:** Conceptualization; Data curation. **María Teresa Bajo:** Funding acquisition; Resources; Software; Supervision. **Paola E. Dussias:** Conceptualization; Funding acquisition; Project administration; Resources; Software; Supervision; Validation; Writing-review & editing.

ORCID

Anne L. Beatty-Martínez https://orcid.org/0000-0003-4306-2052

REFERENCES

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2012). The processing of number and gender agreement in Spanish: An event-related potential investigation of the effects of structural distance. *Brain Research*, *1456*, 49–63. https://doi.org/10.1016/j.brainres.2012.03.057

Alemán Bañón, J., 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*, 275–306. https://doi.org/10.1177/0267658313515671

Alemán Bañón, J., Miller, D., & Rothman, J. (2017). Morphological variability in second language learners: An examination of electrophysiological and production data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1509–1536. https://doi.org/10.1037/xlm0000394

Alemán Bañón, J., & Rothman, J. (2016). The role of morphological markedness in the processing of number and gender agreement in Spanish: An event-related potential investigation. *Language, Cognition and Neuroscience*, *31*, 1273–1298. https://doi.org/10.1080/23273798.2016.1218032

Barber, H. A., & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex*, *39*, 465–482. https://doi.org/10.1016/S0010-9452(08)70259-4

Barber, H. A., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17*, 137–153. https://doi.org/10.1162/0898929052880101

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience*, *25*, 936–951. https://doi.org/10.1162/jocn

Beatty-Martínez, A. L. (2019). *Revisiting Spanish grammatical gender in monolingual and bilingual speakers: Evidence from event-related potentials and eye-movements*. The Pennsylvania State University.

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, *95*, 173–189. https://doi.org/10.1016/j.jml.2017.04.002

Beatty-Martínez, A. L., & Dussias, P. E. (2018). Tuning to languages: experience-based approaches to the language science of bilingualism. *Linguistics Vanguard*, *4*, 9. https://doi.org/10.1515/lingvan-2017-0034

Beatty-Martínez, A. L., & Dussias, P. E. (2019). Revisiting masculine and feminine grammatical gender in Spanish: Linguistic, psycholinguistic, and neurolinguistic evidence. *Frontiers in Psychology*, *10*, 1–11. https://doi.org/10.3389/fpsyg.2019.00751

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 1022–1047. https://doi.org/10.1037/xlm0000770

Boland, J. E., Kaan, E., Valdés Kroff, J. R., & Wulff, S. (2016). Psycholinguistics and variation in language processing. *Linguistics Vanguard*, *2*, 1–8. https://doi.org/10.1515/lingvan-2016-0064

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). Processing syntax and morphology: A neurocognitive perspective. In R. D. Van Valin (Ed.). *Oxford surveys in syntax and morphology*, (pp. 1–360). Oxford, UK: Oxford University Press.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology*, *10*(298), 1–17. https://doi.org/10.3389/fpsyg.2019.00298

Bull, W. (1965). *Spanish for teachers: Applied Linguistics*. Ronald Press.

Butt, J., & Benjamin, C. (2012). *A new reference grammar of modern Spanish*. Routledge.

Bybee, J. L. (2010). *Language, usage and cognition*. University Press. https://doi.org/10.1017/CBO9781107415324.004

Caffarra, S., & Barber, H. A. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, *1605*, 83–92. https://doi.org/10.1016/j.brainres.2015.02.018

Caffarra, S., Mendoza, M., & Davidson, D. (2019). Is the LAN effect in morphosyntactic processing an ERP artifact? *Brain and Language*, *191*, 9–16. https://doi.org/10.1016/j.bandl.2019.01.003

Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: An empirical review. *Neuroscience and Biobehavioral Reviews*, *51*, 31–47. https://doi.org/10.1016/j.neubiorev.2015.01.010

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender nformation in constructing a mental model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, *49A*, 639–663. https://doi.org/10.1080/713755647

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, *2*, 219–253. https://doi.org/10.1075/lab.2.3.01dab

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological*

Sciences, *369*(20120394), 1–9. https://doi.org/10.1098/rstb.2012.0394

DeLuca, V., Miller, D., Pliatsikas, C., Rothman, J., Schwieter, J., & Paradis, M. (2019). Brain adaptations and neurological indices of processing in adult second language acquisition. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the neuroscience of multilingualism* (pp. 170–196). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch8

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. https://doi.org/10.1017/S136672890100044X

Dussias, P. E., Valdés Kroff, J. R., Beatty-Martínez, A. L., & Johns, M. A. (2019). What language experience tells us about cognition: Variable input and interactional contexts affect bilingual sentence processing. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the neuroscience of multilingualism* (pp. 467–484). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch23

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition*, *35*, 353–387. https://doi.org/10.1017/S0272263112000915

Eddington, D. (2002). Spanish gender assignment in an analogical framework. *Journal of Quantitative Linguistics*, *9*, 49–75. https://doi.org/10.1076/jqul.9.1.49.8482

Eddington, D., & Hualde, J. I. (2008). El abundante agua fría: Hermaphroditic Spanish nouns. *Studies in Hispanic and Lusophone Linguistics*, *1*, 5–32. https://doi.org/10.1515/shll-2008-1004

Fine, A. B., & Florian Jaeger, T. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*, 578–591. https://doi.org/10.1111/cogs.12022

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1461–1469. https://doi.org/10.1037/a0036756

Fricke, M., Zirnstein, M., Navarro-Torres, C. A., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition*, *22*, 200–207. https://doi.org/10.1017/S1366728918000482

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 430–445.

Gillon Dowens, M., Guo, T., Guo, J., Barber, H. A., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish: Evidence from event related potentials. *Neuropsychologia*, *49*, 1651–1659. https://doi.org/10.1016/j.neuropsychologia.2011.02.034

Gillon Dowens, M., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, *22*, 1870–1887. https://doi.org/10.1162/jocn.2009.21304

Grey, S., Tanner, D., & van Hell, J. G. (2017). How right is left? Handedness modulates neural responses during morphosyntactic processing. *Brain Research*, *1669*, 27–43. https://doi.org/10.1016/j.brainres.2017.05.024

Guajardo, L. F., & Wicha, N. Y. Y. (2014). Morphosyntax can modulate the N400 component: Event related potentials to gender-marked post-nominal adjectives. *Neuroimage*, *91*, 262–272. https://doi.org/10.1016/j.neuroimage.2013.09.077

Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2020). The masculine form in grammatically gendered languages and its multiple interpretations: A challenge for our cognitive system. *Language Sciences*, *83*, 101328. https://doi.org/10.1016/j.langsci.2020.101328

Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141. https://doi.org/10.1017/s1366728901000232

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22*, 27–62. www.jstor.org/stable/4178707

Hoff, E., Quinn, J. M., & Giguere, D. (2018). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental Science*, *21*, 1–16. https://doi.org/10.1111/desc.12536

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, *32*, 277–307. https://doi.org/10.1177/0267658315624960

Kappenman, E. S., & Luck, S. J. (2012). ERP Components: The ups and downs. In *The Oxford Handbook of event-related potential components* (pp. 3–30). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0014

Kim, A. E., Oines, L., & Miyake, A. (2018). Individual differences in verbal working memory underlie a tradeoff between semantic and structural processing difficulty during language comprehension: An ERP investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 406–420. https://doi.org/10.1037/xlm0000457

Kleinman, D., Runnqvist, E., & Ferreira, V. S. (2015). Single-word predictions of upcoming language during comprehension: Evidence from the cumulative semantic interference task. *Cognitive Psychology*, *79*, 68–101. https://doi.org/10.1016/j.cogpsych.2015.04.001

Kupisch, T., Akpınar, D., & Stöhr, A. (2013). Gender assignment and gender agreement in adult bilinguals and second language learners of French. *Linguistic Approaches to Bilingualism*, *3*, 150–179. https://doi.org/10.1075/lab.3.2.02kup

Lemhöfer, K., Schriefers, H., & Indefrey, P. (2014). Idiosyncratic grammars: Syntactic processing in second language comprehension uses subjective feature representation. *Journal of Cognitive Neuroscience*, *26*, 1428–1444. https://doi.org/10.1162/jocn_a_00609

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. https://doi.org/10.1111/j.1467-9280.2007.01871.x

Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language*, *32*, 169–192.

Lotze, N., Tune, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Meaningful physical changes mediate lexical-semantic integration: Top-down and form-based bottom-up information sources interact in the N400. *Neuropsychologia*, *49*, 3573–3582. https://doi.org/10.1016/j.neuropsychologia.2011.09.009

López-Calderón, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

López Prego, B. (2015). *The online use of markedness information in L1 and L2 Spanish gender agreement*. University of Kansas.

López Prego, B., & Gabriele, A. (2014). Examining the impact of task demands on morphological variability in native and non-native

Spanish. *Linguistic Approaches to Bilingualism*, *4*, 192–221. https://doi.org/10.1075/lab.4.2.03lop

Luck, S. J. (2014). *An introduction to the event-related potential technique*, Cambridge, MA: MIT University Press. https://doi.org/10.1017/CBO9781107415324.004

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*, 146–157. https://doi.org/10.1016/j.gde.2016.03.011

MacWhinney, B., Bates, E. A., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *2*, 127–150. https://doi.org/10.1016/S0022-5371(84)90093-8

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

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. https://doi.org/10.1016/j.brainres.2006.03.094

Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011). When persons disagree: An ERP study of unagreement in Spanish. *Psychophysiology*, *48*, 1361–1371. https://doi.org/10.1111/j.1469-8986.2011.01212.x

McCarthy, C. (2007). *Morphological variability in second language Spanish*. McGill University.

McLaughlin, J., Tanner, D., Pit, I., Frenck-Mestre, C., Inoue, K., Valentine, G. D., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, *60*, 123–150. https://doi.org/10.1111/j.1467-9922.2010.00604.x

Mehravari, A. S., Tanner, D., Wampler, E. K., Valentine, G. D., & Osterhout, L. (2015). Effects of grammaticality and morphological complexity on the P600 event-related potential component. *PLoS ONE*, *10*(10), 1–16. https://doi.org/10.1371/journal.pone.0140850

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, *66*, 156–159. https://doi.org/10.1016/j.cortex.2014.06.009

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*, 908–930. https://doi.org/10.1016/j.cortex.2011.02.019

Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition*, *164*, 61–73. https://doi.org/10.1016/j.cognition.2017.03.012

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in eight and twelve-year-old children and adults. *Journal of Experimental Psychology: General*, *144*, 447–468. https://doi.org/10.1037/xge0000054

Montes-Alcalá, C., & Lapidus Shin, N. (2011). Las keys versus el key: Feminine gender assignment in mixed-language texts. *Spanish in Context*, *8*, 119–143. https://doi.org/10.1075/sic.8.1.06mon

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*, 933–947. https://doi.org/10.1162/jocn_a_00119

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, *58*, 503–553. https://doi.org/10.1111/j.1467-9922.2008.00449.x

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience*, *17*, 1212–1228. https://doi.org/10.1162/0898929055002436

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 15583–15587. https://doi.org/10.1073/pnas.0903620106

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*, 494–522. https://doi.org/10.1006/brln.1997.1793

Osterhout, L., McLaughlin, J., Kim, A. E., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 309–327). https://doi.org/10.4324/9780203509050

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *12*, 2728–2744. https://doi.org/10.1162/jocn.2009.21393

Pélissier, M. (2020). Comparing ERPs between native speakers and second language learners: Dealing with individual variability. In A. Edmonds, P. Leclercq, & A. Gudmestad (Eds.), *Interpreting language-learning data* (pp. 39–69). Language Science Press. https://doi.org/10.5281/zenodo.4032282

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18*, 571–590. https://doi.org/10.1017/S0305000900011259

Perrotti, L. (2012). *Grammatical gender processing in L2 speakers of Spanish: Does cognate status help?* The Pennsylvania State University.

RAEinforma. (2020). *#RAEconsultas Si se sobrentiende el sustantivo tácito «enfermedad», lo más adecuado sería el uso en femenino: «la COVID-19» [Tweet]*. https://twitter.com/RAEinforma/status/1242355343465820160

Qi, Z., Beach, S. D., Finn, A. S., Minas, J., Goetz, C., Chan, B., & Gabrieli, J. D. E. (2017). Native-language N400 and P600 predict dissociable language-learning abilities in adults. *Neuropsychologia*, *98*, 177–191. https://doi.org/10.1016/j.neuropsychologia.2016.10.005

Sassenhagen, J., & Bornkessel-Schlesewsky, I. (2015). The P600 as a correlate of ventral attention network reorientation. *Cortex*, *66*, A3–A20. https://doi.org/10.1016/j.cortex.2014.12.019

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, *137*, 29–39. https://doi.org/10.1016/j.bandl.2014.07.010

Smith, P., Nix, A., Davey, N., López Ornat, S., & Messer, D. (2003). A connectionist account of Spanish determiner production. *Journal of Child Language*, *30*, 305–331. https://doi.org/10.1017/S0305000903005622

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*, 135–162. https://doi.org/10.1016/j.bandl.2011.07.001

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A commentary on “Grammatical agreement processing in reading: ERP findings and future directions” by Molinaro et al., 2014. *Cortex*, *66*, 149–155. https://doi.org/10.1016/j.cortex.2014.04.007

Tanner, D. (2019). Robust neurocognitive individual differences in grammatical agreement processing: A latent variable approach. *Cortex*, *111*, 210–237. https://doi.org/10.1016/j.cortex.2018.10.011

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*(2), 277–293. https://doi.org/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*, 367–382. https://doi.org/10.1017/S1366728912000302

Tanner, D., & van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289–301. https://doi.org/10.1016/j.neuropsychologia.2014.02.002

Valdés Kroff, J. R. (2016). Mixed NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing. In R. E. Guzzardo Tamargo, C. M. Mazak, & M. C. Parafita Cuoto (Eds.), *Spanish-English Codeswitching in the Caribbean and the US* (pp. 281–300). John Benjamins.

Valdés Kroff, J. R., Dussias, P. E., Gerfen, C., Perrotti, L., & Bajo, M. T. (2017). Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism*, *7*, 163–198.

Wicha, N. Y. Y., Moreno, E. M., & 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. https://doi.org/10.1162/089892904192048

Wigdorowitz, M., Pérez, A. I., & Tsimpli, I. M. (2020). A holistic measure of contextual and individual linguistic diversity. *International Journal of Multilingualism*. Advance online publication. https://doi.org/10.1080/14790718.2020.1835921

Zirnstein, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition*, *176*, 87–106. https://doi.org/10.1016/j.cognition.2018.03.001

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

Experimental stimuli

How to cite this article: Beatty-Martínez AL, Bruni MR, Teresa Bajo M, Dussias PE. Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers. *Psychophysiology*. 2021;58:e13737. https://doi.org/10.1111/psyp.13737