

Research report

Clitic pronouns reveal the time course of processing gender and number in a second language



Eleonora Rossi^{a,b,*}, Judith F. Kroll^{a,b}, Paola E. Dussias^{b,c}

^a Department of Psychology, Pennsylvania State University, Moore Building, University Park, PA 16802-3104, USA

^b Center for Language Science, Pennsylvania State University, USA

^c Department of Spanish, Italian, & Portuguese, Pennsylvania State University, USA

ARTICLE INFO

Article history:

Received 3 February 2014

Received in revised form

3 July 2014

Accepted 7 July 2014

Available online 15 July 2014

Keywords:

ERPs

Second-language processing

Grammatical gender

Clitic pronouns

Bilingualism

ABSTRACT

This study investigates grammatical gender and number processing marked on clitic pronouns in native Spanish speakers and in late English–Spanish bilinguals using ERPs. Spanish clitic pronouns were chosen as a critical grammatical structure which is absent in English, and which encodes both grammatical gender and number. Number, but not grammatical gender, is present in English, making this structure a prime one to investigate second language processing. Results reveal a P600 effect in native speakers for violations of both gender and number. Late but relatively proficient English–Spanish bilinguals show a P600 effect only for number violations occurring at the clitic pronoun, but not for gender violations. However a post-hoc analysis reveals that a subset of highly proficient late bilinguals does reveal sensitivity to violations of grammatical gender marked on clitic pronouns. Taken together these results suggest that native-like processing is possible for highly proficient late second language learners for grammatical features that are not present in the speakers' native language, even when those features are encoded on a grammatical morpheme which itself is absent in the speakers' native language.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

A longstanding question is whether late bilinguals who acquire a second language (L2) after the hypothesized critical period for language acquisition are able to process grammatical structures in a native-like way (Birdsong, 1999; Lenneberg, 1967). Early behavioral studies suggest that late L2 speakers are less sensitive than native speakers in processing grammatical violations (Johnson & Newport, 1989; Weber-Fox & Neville, 1999), while other findings demonstrate that native-like attainment is possible when L2 speakers achieve a high level of proficiency (Birdsong & Molis, 2001; McDonald, 2000). The fact that it is possible to document the presence of proficient late L2 speakers exhibiting native-like performance has suggested that L2 achievement cannot be solely explained by a maturational account. Although age of acquisition (AoA) is an important factor for successful L2 learning, it has often been confounded with proficiency, making it unclear to which extent they interplay in L2 processing, and which of the two account the most for successful L2 processing (Hahne, 2001; Kotz, 2009; Steinhauer, White, & Drury, 2009). Early acquirers tend to be

more proficient in the L2 but these two factors have not always been assessed independently. Additionally, many studies have not always examined on-line measures of processing, relying more directly on off-line measures such as grammaticality judgments. Despite behavioral evidence that highly proficient late L2 speakers may be able to attain native-like performance, the question remains open as to whether they process linguistic information in the same way as native speakers.

Within a Universal Grammar (UG) framework, two views have been put forward to explain L2 language processing performance. One focuses on representational/processing deficit (Clahsen & Felser, 2006; Hawkins & Chan, 1997), proposing that adult L2 learners cannot attain native proficiency for grammatical features that are not instantiated in their first language (L1), especially for functional categories such as grammatical gender and number. The Shallow Structure Hypothesis (Clahsen & Felser, 2006), posits that late L2 learners are constrained in accessing full syntactic information in the L2. On this account, late L2 learners compensate for the lack of hierarchical syntactic processing by relying on lexical, semantic and pragmatic information to process and/or resolve syntactic ambiguities in real time (Clahsen & Felser, 2006; Papadopoulou & Clahsen, 2003). An alternative account focuses on representational accessibility. For example, the Full Transfer Full Access hypothesis (Schwartz & Sprouse, 1996) claims that L2 attainment is possible as a function of increased proficiency, even for grammatical structures and features that are not present in the L1.

* Corresponding author. Tel.: +1 814 867 2441; fax: +1 814 863 7002.

E-mail addresses: exr22@psu.edu (E. Rossi), jfk7@psu.edu (J.F. Kroll), pdussias@psu.edu (P.E. Dussias).

On the other hand, cognitive and computational accounts of L2 processing hypothesize that L2 native-like attainment is possible even for grammatical features that are not represented in the L1. Critically, these accounts do not attribute L2 performance to the inability to access certain morpho-syntactic structures, but to overall diminished cognitive resources when processing the L2 in real time (Hopp, 2007; McDonald, 2000). Support for this view has been provided by studies showing that individual differences in memory resources constrain sentence processing (Just & Carpenter, 1992). Furthermore, imposing an additional processing load on native speakers appears to simulate L2 performance (Blackwell & Bates, 1995; Dick et al., 2001; Dussias & Piñar, 2010; Roberts, Marinis, Felser, & Clahsen, 2004). Critically, grammatical constructions differ in the demands that they make on cognitive resources. For example, more complex syntactic structures have been found to require greater processing capacity than simpler structures (Johnson, Shenkman, Newport, & Medin, 1996; Miyake, Carpenter, & Just, 1994), suggesting that some grammatical structures may be particularly taxing for L2 speakers. A number of neuroimaging studies (Hasegawa, Carpenter, & Just, 2002) have supported this view by showing that the L2 elicits greater activation than the L1 in areas dedicated to negotiating cognitive load, and also that alternative sentential structures engage computational resources differentially both in the L1 and L2 (Just, Carpenter, Keller, Eddy, & Thulborn, 1996). Although greater neural activation has been associated with greater cognitive demands, the question of which specific cognitive components are engaged by language processing in the bilingual's two languages is still open. Fiebach, Schleewsky, and Friederici (2001) proposed that a working memory component (supported by Broca's area) is responsible for maintaining active syntactic information during on-line sentence processing. Support for this model comes from studies that have utilized specific tasks to correlate working memory abilities with processing performance (Andrews, Birney, & Halford, 2006), demonstrating a central role for working memory in maintaining active linguistic information while information is progressively integrated both in the L1 and the L2. For example McDonald (2006) showed that native speakers who scored at ceiling in an auditory grammaticality judgment task perform like L2 speakers when required to remember (and subvocalize) seven-digit numbers. A recent meta-analysis examining 79 studies on the role of working memory in L2 language processing (Linck, 2013) confirmed a positive robust relationship of working memory with L2.

The advent of behavioral and neuroimaging techniques that enable an analysis of language processing at high temporal and spatial resolutions has reshaped the debate about the factors and the neural mechanisms that drive similarities and differences between native and L2 processing. Critically, several Event Related Potential (ERP) and functional magnetic resonance imaging (fMRI) studies have clarified the relative roles of age of acquisition (AoA) and proficiency in L2 processing, highlighting how AoA might have been confounded with proficiency and language dominance (Steinhauer et al., 2009; Steinhauer, 2014). For example, a number of studies have demonstrated that when AoA is kept constant, proficiency and exposure are the driving factors that determine similarity in the neural representations and in the electrophysiological correlates of language processing in native and late L2 speakers (fMRI evidence: Abutalebi, Cappa, & Perani, 2001; Perani et al., 1998; ERP evidence: Osterhout et al., 2008).

The goal of the present paper is to investigate morpho-syntactic processing in native Spanish speakers and in native English-speaking adults who acquired Spanish as the L2 after childhood. We use ERPs to investigate pronominal reference marked on a particular grammatical morpheme, the Spanish clitic pronoun. Clitics have been shown to be particularly difficult to acquire for adult second language learners (Montrul, 2011), but

also for native speakers acquiring their L1 (Antelmi, 1997; Guasti, 1994; Pérez-Leroux, Cuza, & Thomas, 2011) and also highly vulnerable in aphasic speakers (Rossi, 2007; Rossi, 2013). Specifically, we investigate gender and number agreement marked at the clitic pronoun in Spanish using ERPs in native speakers of Spanish and in English L2 speakers of Spanish. In what follows we first introduce the logic of language-related ERPs, and then discuss the relevant background on gender and number agreement.

The multidimensional nature of ERPs, with components characterized in terms of temporal resolution, amplitude, and scalp distribution, has been central in understanding native and bilingual language processing (Kotz, 2009). The ERP record shows that distinct linguistic processes lead to different neural responses. Traditionally, lexico-semantic processes have been shown to be indexed primarily by a negative ongoing wave between 200 and 600 ms with maximum amplitude at around 400 ms over centroparietal sites, the so-called N400 effect (Kutas & Hillyard, 1980)¹. The N400 can be modulated by factors such as cloze probability (Kutas & Van Petten, 1994) with its amplitude being affected by expectancy (see Kutas and Federmeier (2011), for a recent review on the N400). Most relevant for the present study is that the N400 has been examined in relation to a number of different aspects of L2 processing. For example, analyses of the N400 in recent studies of very early stages of adult language learning have shown that the brain outpaces behavior in demonstrating sensitivity to the new L2 (McLaughlin, Osterhout, & Kim, 2004; McLaughlin et al., 2010; Osterhout, McLaughlin, Pitkanen, Frenck-Mestre, & Molinaro, 2006).

Whereas the N400 has been interpreted as signaling lexico-semantic processing, the P600 component is a positive ongoing wave between 400 and 900 ms with maximum amplitude around 600 ms observed predominantly at posterior sites (Osterhout & Holcomb, 1992; Osterhout, McKinnon, Bersick, & Corey, 1995; Osterhout & Nicol, 1999). The P600 is elicited by a number of typical morpho-syntactic violations, including violations of phrase-structure and subcategorization (Ainsworth-Darnell, Shulman, & Boland, 1998; Hagoort, Brown, & Groothusen, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992) and number, gender, and case violations elicited in a variety of grammatical structures (Foucart & Frenck-Mestre, 2011, 2012; Frenck-Mestre, Foucart, Carrasco, & Herschensohn, 2009; Frenck-Mestre, Osterhout, McLaughlin, & Foucart, 2008; Osterhout & Mobley, 1995). The P600 has generally been interpreted as signaling morpho-syntactic reanalysis and repair, suggesting a relatively late, non-automatic process. A P600 effect has also been linked to increasing integration difficulties and increased processing demands (Kaan & Swaab, 2003). Recent evidence suggests that a P600 can also be elicited in contexts in which syntax and semantics are in competition, as in the case of sentences which are grammatically correct but which contain apparent semantic anomalies (Kim & Osterhout, 2005), as for example when the subject of an active sentence makes a plausible theme of the verb (e.g., *The hearty meal was devouring...*). Such data indicate that even among native speakers, neural responses can be modulated by different types of linguistic constraints.

Finally, an earlier language-related component, the Left Anterior Negativity (LAN), has been related to rule-based automatic parsing, resulting in a left-anterior scalp distribution with a maximum amplitude in a temporal window ranging from 150 to 500 ms. The LAN has been implicated in the processing of violations of syntactic constraints and phrase-structure violations

¹ The negativity of the N400 does not have to be interpreted in absolute terms, but in relative terms, i.e., an N400 effect is observed in contrast to a baseline that is proportionally less negative.

(Neville et al., 1991; Weber-Fox & Neville, 1996). An even earlier component, the ELAN (early left anterior negativity), which peaks between 100 and 300 ms, has been interpreted as evidence of an initial syntactic structure building phase. The ELAN has been integrated in psycholinguistic models, which propose a syntax-first serial view of language processing (Friederici, 1995), in which the presence of the ELAN would signal an extremely early phrase-structure building process, which represents a first highly automatic first pass parse. Even though the ELAN is not one of the ERP components that are to be expected for the manipulations reported here, it is important to mention that the absence of an ELAN in young children and late L2 learners has been interpreted to reflect inaccessibility of early and automatic syntactic parsing mechanisms (Hahne, 2001; Hahne, Eckstein, & Friederici, 2004) which could reveal a maturational process on the one hand, and a lack of syntactic processing automaticity for late L2. For a recent review on the ELAN and related neurolinguistics models see Steinhauer & Drury (2011).

Gender and number concord has been a fruitful field of research in the second language acquisition literature and it has also been extensively explored in native language processing. A number of seminal studies have utilized off-line and on-line tasks to investigate whether adult L2 learners whose L1 does not encode grammatical gender can process it in a native-like manner (Friederici, Garrett, & Jacobsen, 1999) in a variety of syntactic contexts and languages (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008; Sagarra & Herschensohn, 2010). Overall, results from the behavioral literature suggest that L2 learners are able to process features that are also present in their L1 (e.g., number concord for English–Spanish bilinguals). However, for unique L2 structures (e.g., grammatical gender for native English speakers), a more mixed pattern of results has been reported. Importantly, some studies reveal that at least a subset of advanced L2 learners achieve native-like processing for gender concord, even when it is not encoded in the L1. For example, Sagarra & Herschensohn (2010) tested native speakers of Spanish and L1 English beginning and intermediate learners of L2 Spanish on a noun–adjective gender and number agreement task using an off-line and on-line sentence processing task. They demonstrated that only intermediate, but not beginning, Spanish learners were sensitive to the manipulations of gender and number. In line with computational accessibility approaches of L2 grammatical processing (Hopp, 2010) the results were supportive of the Full Transfer Full Access hypothesis (Herschenson, 2000, 2001; Schwartz & Sprouse, 1996), showing that even late L2 learners are able to access L2-unique new grammatical features which are required by the L2. Importantly, Sagarra & Herschensohn (2010) showed a positive correlation between working memory and gender agreement violation in sentences for the intermediate learner group, indicating that speakers with higher working memory were more sensitive to gender agreement violations. Recently, syntactic and cognitive accounts of gender and number processing have proposed that grammatical gender concord might be cognitively more complex to process than number agreement (Antón-Méndez, 1999; Antón-Méndez, Nicol, & Garrett, 2002). Gender agreement has been proposed to be a feature at the lemma level, while number is a morphological feature that marks quantity. As such gender has been proposed to involve first a step in which gender has to be attributed at the lemma level and also a subsequent stem in which gender needs to be computed. (Antón-Méndez, 1999; Antón-Méndez et al., 2002). This hypothesis has been supported by results showing that English–Spanish learners with higher working memory were more sensitive to grammatical gender agreement violations than speakers with lower-memory span (Sagarra, 2007).

In recent years, the ERP literature has also added important data on gender and number agreement processes. In native speakers, violations of gender and number agreement (e.g., “*der Milch” instead of the correct feminine “die Milch” the milk) have

been shown to elicit primarily a P600 and/or a LAN/P600 (Gunter, Friederici, & Schriefers, 2000; Barber, Salillas, & Carreiras, 2004). In addition, the presence of the biphasic component (LAN/P600) has been related to violations presented in isolated noun phrases (eliciting a LAN) or in sentence contexts, eliciting a LAN/P600 (see Barber and Carreiras (2005) for similar evidence for noun–adjective gender violations. For recent discussions of these three major language-related ERP components see Swaab, Ledoux, Camblin, and Boudewyn (2011); see Molinaro, Barber, and Carreiras (2011) for a review of ERP studies on grammatical gender processing).

For late L2 learners, previous research has suggested that there may be constraints in complete processing of grammatical gender, especially for speakers whose L1 does not encode it (Sabourin & Stowe, 2008; Sabourin, Stowe, & de Haan, 2006). More generally, grammatical gender processing at lower levels of L2 proficiency shows qualitative differences, such as the absence of the LAN (Hahne & Friederici, 2001), although LANs have been reported in recent studies for more proficient L2 speakers (Friederici, Steinhauer, & Pfeifer, 2002; Rossi, Gugler, Friederici, & Hahne, 2006). In addition, less proficient L2 speakers have been reported to show an N400 or a posterior N400 as a response to grammatical violations, or a delayed P600 (Rossi et al., 2006) and most reliably no P600 (Weber-Fox & Neville, 1996). The absence of the LAN in L2 speakers has been interpreted as greater variability in L2 processing (Osterhout et al., 2006). However, it is important to note that the presence of a LAN component is highly variable even within native speakers (Hagoort et al., 1993; Kim & Osterhout, 2005), making it difficult to interpret its absence in L2.

Despite the difficulty of acquiring grammatical gender in the L2, a number of recent ERP studies have revealed that L2 speakers who have extensive immersion experience (Gillon Dowens, Vergara, Barber, & Carreiras, 2010) or who are highly proficient (Foucart & Frenck-Mestre, 2011, 2012; Tokowicz & MacWhinney, 2005) show a similar neural signature to native speakers (for evidence using an artificial language see Morgan-Short, Sanz, Steinhauer, and Ullman (2010)), even for late bilinguals whose L1 does not encode this feature (Foucart & Frenck-Mestre, 2011, 2012) and whose two languages differ in script (Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011). The majority of ERP studies that examined grammatical gender processing have however investigated violations in open class words (i.e., nouns or adjectives) and most of the times in a local or adjacent relationship (e.g., determiner–noun; noun–adjective). Very few ERP studies have investigated gender and number agreement in a coreferential relationship between a pronoun and its antecedent, such as reflexives or subject and object pronouns (Molinaro, Kim, Vespignani, & Job, 2008; Silva-Pereyra, Gutierrez-Sigut, & Carreiras, 2012; Osterhout & Mobley, 1995). Osterhout and Mobley (1995) elicited gender and number agreement violations on reflexive and subject pronouns in native speakers of English. The results showed a P600 effect for both types of violations. In this study however, gender and number agreement was tested in English, a language that marks biological gender but does not encode grammatical gender. Other studies have broadened this question, attempting to tease apart the role of biological and grammatical gender by examining languages that encode both types of gender, like German, Dutch, and Spanish (Hammer, Jansma, Lamers, & Münte, 2005; Lamers, Jansma, Hammer, & Münte, 2006, 2008; Schmitt, Lamers, & Münte, 2002; Silva-Pereyra et al., 2012). The majority of the results in these studies show that a referential pronoun violation for both biological and grammatical gender elicits a P600 component. To summarize, the majority of ERPs studies which examined gender and number agreement across a variety of grammatical structures (subject–verb,

determiner–noun, noun–adjective, phrase structure violations², reflexive and subject pronouns) in monolingual speakers, and in a number of different languages, converge in showing that these violations elicit a P600 or LAN/P600 biphasic components (De Vincenzi et al., 2003). Importantly, the observed ERP effects have been found to be modulated by a number of variables, e.g., the type of grammatical structure and whether the structure of interest is tested in isolation or embedded in a sentence context (Barber & Carreiras, 2005; Osterhout & Mobley, 1995). In addition, variability in the presence, magnitude, temporal and/or scalp distribution of these grammatical-related components is influenced by subject-dependent variables as working memory capacity (King & Kutas, 1995; Vos, Gunter, Kolk, & Mulder, 2001) and proficiency, even in native speakers (Pakulak & Neville, 2010).

The goal of the present study is to investigate whether native Spanish speakers and native English speakers who learned Spanish as the L2 later in life show sensitivity to gender and number agreement violations that are marked on a specific pronominal morpheme (the clitic pronoun). This study will contribute to the current literature on gender and number agreement and more specifically to the ERP literature on the neural correlates of referential pronoun resolution by adding novel data on how late L2 learners process gender and number marked on clitic pronouns, a specific structure that has not been previously investigated using ERPs. Moreover, the present study will also contribute to the recent discussion on potential differences between grammatical gender and number processing by adding new real-time neural data. In what follows, a short introduction to clitic pronouns will be provided, and the rationale and hypotheses for the two reported studies will be outlined.

From a linguistic perspective, clitic pronouns are grammatical morphemes that do not bear phonological stress. Clitics are introduced in postverbal positions, as well as in the form of lexically specified complements (Kayne, 1991; Rizzi, 1982), and they are “cliticized” together with the verb to their final position through a Clitic Placement Rule by which clitics move from their base-generated position to their final position (Rizzi, 1982). Most importantly for the purposes of this paper clitics appear in different sentential positions relative to English. In finite Spanish constructions, clitics are placed obligatorily before the finite verb (or before the finite auxiliary), as for example in: “*Maria la come*” (Maria eats it), while in English the pronoun always appears after the verb (Mary eats it). Moreover Spanish clitic pronouns encode grammatical gender (and number), which must agree with the previously introduced referent (as in “*Maria pela la manzana_{fem.sing} y la_{fem.sing} come*” (Maria peels the apple and eats it). Clitic pronouns have been shown to be a particularly difficult grammatical structure to acquire and process. Clitic pronouns are produced relatively later in L1 acquisition, i.e., around 22 months (Antelmi, 1997; Guasti, 1994), with initial omissions or substitutions of the full noun instead of the clitic counterpart till approximately the age of 27 months. Clitics are also highly vulnerable in aphasic speakers (Lonzi & Luzzatti, 1993; Miceli et al., 1989; Miceli & Mazzucchi, 1990; Rossi, 2007; Rossi, 2013 for Italian; Reznik, Dubrovsky, & Maldonado, 1995 for Spanish; Stavrakaki & Kouvava, 2003 for Greek). Specifically for the present study, for L1 English learners of Spanish as the L2, clitic pronouns are particularly taxing for a number of reasons. First they represent a linguistic structure which is absent in English (Grüter, 2006; Santoro, 2007) and the word order in finite sentences, as mentioned before, does not overlap with the English one. Finally, and most importantly for this study, Spanish

clitic encodes grammatical gender (and number), a feature that, like the clitic itself, is absent in English.

While this is the first study, to our knowledge, that specifically investigates gender and number processing marked on clitic pronouns using ERPs in English late L2 learners of Spanish, some predictions can be made based on previous ERP studies of gender and number agreement. Native speakers of languages that encode grammatical gender and number show either a P600 component or a biphasic LAN/P600 (Barber & Carreiras, 2005; Barber et al., 2004; Hagoort & Brown, 1999; Hammer et al., 2005; Lamers et al., 2006, 2008; Schmitt et al., 2002; Silva-Pereyra et al., 2012). A similar ERP signature has been reported for English, which does not encode grammatical gender, but only marks biological gender (Molinero et al., 2008; Osterhout & Mobley, 1995) demonstrating overall that native speakers are sensitive to violations of gender and number across languages. Importantly, the present study diverges from the previous ERP literature in that ERPs were locked to the clitic pronoun itself, a closed-class grammatical particle that stands in an agreement relationship with the sentential verb. Also, clitic pronouns are of particular interest as they stand in an anaphoric dependency with their antecedents (i.e. the noun phrase that precedes them), and have been suggested to produce an immediate reactivation of the lexical and grammatical features of their antecedents (De Vincenzi, 1999).

Here, we report two ERP experiments. Experiment 1 investigates the electrophysiological correlates of gender and number processing marked on direct-object clitic pronouns in a group of native Spanish speakers. Specifically, we ask whether native speakers of Spanish are sensitive to violations of gender and number marked on clitic pronouns, and whether the expected neurophysiological response (i.e., P600 or LAN/P600) will be similar for gender and number violations. As noted earlier, there are few, if any, prior ERP studies that have examined the neurophysiological correlates for this particular grammatical structure. The results from the first study will therefore and foremost provide (first) evidence on the ERP signature for the processing of gender and number marked on clitics, and will highlight potential differences in the neural signature between the two, informing accounts that predict processing differences between gender and number (Antón-Méndez, 1999; Antón-Méndez et al., 2002).

Experiment 2 investigates whether late, but relatively proficient, English–Spanish bilinguals show sensitivity for grammatical gender and number marked on clitics. Critically for the purposes of this study, neither clitic pronouns (as a grammatical structure) nor grammatical gender is represented in the L1 of the selected speakers, while number agreement is present in English, and is marked on pronouns (though not clitic pronouns). Specifically, Experiment 2 will allow to test the following predictions. First, under the hypothesis that late L2 learners cannot access and process grammatical structures that are unique to the L2, (MacWhinney, 1997, 2005; Clahsen & Felser, 2006), late English–Spanish bilinguals should not reveal sensitivity to the clitic structure. Specifically, English–Spanish bilinguals are not expected to show any significant ERP component that has been related to the access (i.e., LAN) and/or reanalysis (P600) of clitic processing. However, if late bilinguals are sensitive to features of the L2 that are shared with the L1 (as number agreement which is present both in English and Spanish), then the prediction is that they should show sensitivity to number, but not to gender. Under this scenario, English–Spanish bilinguals should reveal either a LAN and/or P600 for number violations but not for gender violations. Finally, if late bilinguals are able to fully acquire sensitivity to the L2, even when it differs from the L1, then we would expect to find sensitivity for both gender and number violations.

² The ERP literature on phrase structure violations will not be reviewed, as phrase structure was not examined here. However, a recent literature (Steinhauer & Drury, 2011) has challenged Friederici’s proposal (Friederici, 2002) that phrase structure violations elicit an early left anterior negativity (ELAN).

2. Experiment 1: the processing of gender and number marked on clitic pronouns in native Spanish speakers

2.1. Method

2.1.1. Participants

Twenty-six native speakers of Spanish (11 females, 15 males; mean age: 29.5 yrs.; SD=6.3) were recruited and paid for participation from the student population at Pennsylvania State University in accordance to IRB approval. They were all right-handed (as assessed by a handedness questionnaire), had normal or corrected-to-normal vision, and no history of neurological disorders. All participants completed a language history questionnaire to assess their language history and skills in both Spanish and English. They rated their language knowledge using a scale from 1 to 10 (1 being the lowest and 10 being the highest score) for oral comprehension, oral production, reading and writing. As the participants were native Spanish speakers immersed in a non-native language environment (English), only speakers who were assessed to have maintained their dominance in Spanish were included. Heritage speakers (born in the US and who grew up in a bilingual environment) or native Spanish speakers who had become dominant in English (self-reporting a higher proficiency in English than Spanish, or an equal score in both languages) were excluded. According to these criteria, five participants who had switched language dominance were excluded. One participant was eliminated due to equipment failure. Also, due to excessive EEG artifact, two additional participants had to be excluded from the analysis (see details for the EEG analysis below). To sum up, data from 18 native speakers of Spanish were analyzed. Participants' demographics and language characteristics are summarized in Table 1.

2.1.2. Materials and design

Three hundred and eighty four experimental sentences were created (192 experimental stimuli and 192 fillers). Each experimental item began with a preamble, which contained an antecedent (a determiner and a noun) followed by a clitic pronoun. For example: "Después de lavar *los cuchillos*, Andrea *los* colocó en la mesa del comedor" (After washing the knives, Andrea placed them on the dinner table). There were four experimental conditions. In one condition gender and number marked on the clitic correctly matched with the antecedent. We will refer to this correct condition as +Gender+Number. In a second condition, gender was uniquely violated, i.e., the gender marked on the clitic mismatched the gender of the antecedent while number was kept correct (−Gender+Number). In the third condition, only the number marked on the clitic was violated while gender was kept correct (+Gender−Number). Finally, in the fourth condition, both gender and number on the clitic were violated, yielding to a double violation (−Gender−Number). Examples for each condition are provided in Table 2. There were a total of 48 sentences per condition. Gender (feminine, masculine) and number (singular and plural) marked on the antecedent were equally represented within each condition. For example, out of the 48 sentences for each condition, 12 experimental items contained a masculine plural antecedent (i.e., "los mangos", the mango), 12 a feminine plural antecedent (i.e., "las manzanas", the apple), 12 a singular masculine antecedent (i.e., "el libro", the book), and finally 12 singular feminine antecedent (i.e., "la pera", the pear).

The length of the sentences varied from 10 to 14 words. To avoid end-of sentence effects, at least three words followed the critical region of interest (i.e., the clitic). Moreover, gender (feminine and masculine) and number (singular and plural) were counterbalanced equally across experimental items. Importantly,

Table 1

Demographics and language characteristics of native Spanish participants. Self-assessed ratings were on a scale from 1 to 10 where 10 was most proficient.

| Gender | Age (yrs) | Rating Spanish | Rating English |
|-----------|------------|----------------|----------------|
| M | 27 | 10 | 6.8 |
| M | 25 | 10 | 9 |
| F | 30 | 10 | 6.6 |
| M | 37 | 10 | 7.4 |
| F | 31 | 10 | 6.2 |
| M | 30 | 10 | 7.8 |
| M | 22 | 9 | 10 |
| M | 34 | 10 | 8.4 |
| M | 21 | 10 | 10 |
| M | 26 | 10 | 9 |
| M | 31 | 9.2 | 7.6 |
| F | 30 | 10 | 9 |
| F | 27 | 10 | 7.8 |
| M | 29 | 9.75 | 7 |
| M | 24 | 10 | 9.5 |
| M | 35 | 10 | 9.75 |
| F | 28 | 9.8 | 9.6 |
| F | 32 | 10 | 8.2 |
| Mean (SD) | 28.8 (4.3) | 9.9 (0.3) | 8.3 (1.2) |

pre-critical words were of the same type across conditions. Overall, as far as the critical experimental material goes, two-thirds of the experimental sentences (134) contained a violation, and one-third (48) did not contain a violation. Materials used for the experimental sentences were matched for frequency and word length. Specifically, word frequency (log) and word length (in letters) values were checked for four words preceding and for two words following the critical word (the clitic) using word values available in the NIM database (Guasch, Boada, Ferré, & Sánchez-Casas, 2013). Relevant statistics showed that there were no significant differences neither for frequency nor for length for none of the words. A complete list of all the experimental materials and the relevant frequency and length values is available at request from the authors.

In order to counterbalance the overall ratio of correct and incorrect trials across the experimental paradigm, two-thirds of the fillers (134) was correct, while one-third (48) was incorrect. In this way, the ratio of sentences with and without violations was counterbalanced. Filler sentences included additional grammatical structures such as subject-verb agreement (e.g., *El regalo para los bebés estaba/*estaban en el suelo*; the present for the babies was/*were on the floor), tense agreement (e.g., *Carolina le ha ordenado a María que limpie/*limpia su escritorio con un trapo*; Carolina ordered Maria to clean her desk with a cloth), and sentences with high and low attachment preferences (e.g., *Amelia fotografió a la prometida del príncipe que fue jugador de fútbol en Italia*; Amelia photographed the fiancée of the prince who was a soccer player in Italy). In addition to functioning as a balance between correct and incorrect trials, a diverse range of fillers sentences were used to reduce the likelihood that participants would adopt a strategy that was focused on the clitic constructions themselves. All sentences were counterbalanced across four experimental lists using a Latin square design. The presentation of the materials was pseudo-randomized to avoid having consecutive trials drawn from the same condition and also to ensure that no more than two sentences would contain a feminine or a masculine (singular or plural) clitic. Additionally, four lists were created in which the order of items was opposite to the original one. This manipulation diminished general-tiredness effects that would be more prominent at the end of each experimental list. As such, there were a total of eight experimental lists. A given participant received only one version. Twelve practice items (6 correct and 6 incorrect) were presented at the beginning of the experiment to familiarize

Table 2

Examples of the experimental sentences for each condition.

| | Correct clitic position |
|-------------------|--|
| + Gender + Number | Antes de leer el libro, Ana <u>lo</u> sacó de la envoltura de plástico |
| – Gender + Number | ^a Antes de comer <u>la manzana</u> , Ana <u>lo</u> peló con un cuchillo |
| + Gender – Number | ^a Antes de leer <u>el libro</u> , Ana <u>los</u> sacó de la envoltura de plástico |
| – Gender – Number | Antes de comer <u>la manzana</u> , Ana <u>los</u> peló con un cuchillo |

^a : violation.

participants with the task. Practice sentences contained an equal number of experimental items and fillers. Finally, comprehension questions ($n=20$) were created to ensure participants' attention. Comprehension questions were interspersed at regular intervals but unbeknownst to the participants, only referred to the filler sentences.

2.1.3. Procedure

Participants sat in a sound-proof chamber optimized for EEG recordings. Experimental stimuli were presented through a Faraday-caged mirror monitor connected with a stimulus presentation computer located in the EEG control room. Stimuli were presented using E-Prime 2.0 (Psychological Software Tools, Inc., Pittsburgh, PA). Each trial started with a fixation cross for 2500 ms to allow participants to blink freely. Sentences were then presented word by word at a fixed rate. Each word was presented for 300 ms, followed by an interstimulus interval (ISI) of 350 ms. Participants were instructed to read each sentence and to perform an acceptability judgment at the end of each sentence as quickly and as accurately as possible. When prompted with a fixation cross, they pressed a red response key to signal that they considered the sentence not acceptable, while they pressed a green key to signal that they considered the sentence acceptable. They were explicitly instructed to try to minimize movements while pressing the response keys. Participants were also instructed to read each sentence carefully, as they would be asked to answer periodic comprehension questions. Responses to the comprehension questions could be given using the same response keys. Response Times (RTs) and accuracy for the acceptability judgments were collected for analysis.

2.1.4. EEG data acquisition and analysis

EEG activity was recorded from 28 Ag/Ag/Cl-sintered electrodes mounted in a Quik cap (Neuroscan Inc.). EEG activity was recorded at the following International 10–20 locations: O1/OZ/O2, P3/PZ/P4, P7/P8, CP3/CP4, TP7/TP8, C3/CZ/C4, T7/T8, FC3/FC4, FT7/FT8, F3/FZ/F4, F7/F8, FP1/FP2. All electrodes were referenced to the right mastoid during recording and re-referenced offline to linked mastoids. Bipolar horizontal and vertical electro-oculographic (EOG) activity was recorded for artifact rejection purposes. Vertical EOG was recorded from two electrodes placed above and below the left eye (VEOU, VEOL) and one lateral to the right eye (HEOR), and one to the left eye (HEOL). All electrode impedances were kept below 5 kilo-Ohms ($k\Omega$) throughout recording. The EEG signals were continuously recorded with a band-pass filter from 0.05 to 100 Hz (Hz) and a Notch filter, with a sampling rate of 500 Hz. ERPs were digitally filtered off-line at a low-pass of 30 Hz (24 dB setting, and zero-phase-shift).

ERPs were time-locked to the onset of the critical stimulus (i.e., the clitic) and averaged off-line for each participant. Epochs were extracted with a baseline of –150 ms and an epoch duration of 950 ms post-stimulus onset. There were no bad channels. All trials that contained vertical and horizontal eye movements, excessive muscle artifact (exceeding 100 microvolts) were not included in the analysis. Only correct trials were included in the ERP analysis.

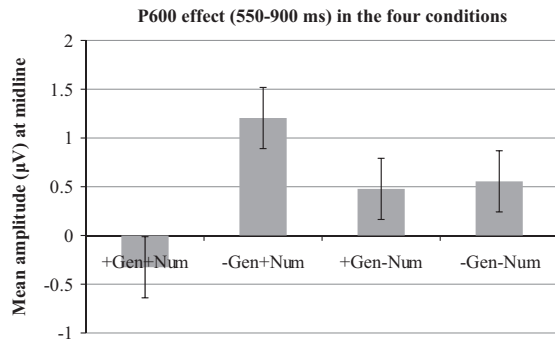
To allow for adequate statistical power, and for an optimal signal-to-noise ratio, the minimum percentage of trials per participant per condition to enter the analysis was 50% (24 trials). Following these criteria, two native Spanish participants were excluded from the analysis, leaving data from 18 participants. For the final group of participants, a total of 22% of trials were eliminated due to artifacts and incorrect behavioral responses (20% for the +Gen+Num condition, 22% for the –Gen+Num condition, 25% for the +Gen–Num condition and 20% for the –Gen–Num condition). Grand averages were then computed for each condition. Based on visual inspection and previous literature, two time windows were chosen to capture the major components of interest. A first window from 300–500 ms was chosen to capture an N400 or LAN components. A second window from 550–900 ms was chosen to roughly correspond to the P600. Grand mean amplitudes values were calculated over these windows for each participant for each condition. For each time window, a series of repeated-measures ANOVAs were performed on mean amplitude values with Gender (Correct; Incorrect) and Number (Correct; Incorrect) as factors. Also, based on the topographical distribution of electrodes in the montage, and in line with previous studies performed in our labs in which the same ERP protocol was used (Guo, Misra, Tam, & Kroll, 2012; Misra, Guo, Bobb, & Kroll, 2012), separate repeated-measures analyses of variance were performed for the midline electrode sites and two lateral sites. For the midline sites, four levels of electrode site (Oz, Pz, Cz, Fz) were included as a variable. For the first lateral site, the ANOVA included hemisphere (2: left/right) and electrode site (7 levels: O1/O2, P3/P4, CP3/CP4, C3/C4, FC3/FC4, F3/F4, FP1/FP2). For the second lateral site, hemisphere (2: left/right) and electrode site (5 levels: P7/P8, TP7/TP8, T7/T8, FT7/FT8, F7/F8) were included in the analysis. Finally, in the analysis of ERP data, the Greenhouse–Geisser correction was applied to account for non-sphericity of the data (Greenhouse & Geisser, 1959).

2.2. Results: behavioral analysis

A series of 2×2 repeated measures ANOVAs were performed for accuracies and RTs on the acceptability judgment task (factors: Gender [correct, incorrect] \times Number [correct, incorrect]). For the RTs analysis, only correct trials were analyzed. Additionally, RTs below 300 ms and above 5000 ms were excluded as outliers. Any data point below or above 2.5 SD for a given participants was also eliminated. In the analysis on accuracy, there was a main effect of Gender, $F(1, 17)=5.4$, $p < 0.05$. Native Spanish speakers were more accurate for the conditions in which gender was correct, and less accurate when gender was correct but number was violated (i.e., +Gender–Number). The analysis on the RTs, revealed a gender by number interaction $F(1, 17)=6.06$, $p < 0.05$. Native Spanish speakers were faster to respond in conditions with pure gender violations (–Gender+Number), but were slower in the double violation condition (–Gender–Number), and in the condition in which only number was violated (+Gender–Number). Behavioral data are reported in Table 3.

Table 3Summary of the behavioral results values (Mean Accuracy and RTs) for the native Spanish speakers' group ($n=18$) for the end of the sentence acceptability judgment.

| | +Gender+Number Mean (SE) | –Gender+Number Mean (SE) | +Gender – Number Mean (SE) | – Gender – Number Mean (SE) |
|----------|-----------------------------|-----------------------------|-------------------------------|--------------------------------|
| Accuracy | 0.91 (0.07) | 0.91 (0.09) | 0.85 (0.14) | 0.91 (0.11) |
| RTs | 973.9 (108.6) | 825.6 (71.2) | 925.8 (100.6) | 941.6 (95.0) |

**Fig. 1.** Magnitude of the P600 effect in the four conditions for native Spanish speakers (bars represent Standard Error).

2.3. Results: ERP analysis

Only results that showed a main effect of Gender or Number and an interaction between the two, or any interaction between the experimental conditions and electrode site or hemisphere will be reported. Significant electrode site or hemisphere results not interacting with the experimental conditions will not be reported, as general hemisphere and electrode site differences are to be expected.

At the first window of interest (300–500), there was a main effect of number and gender at the midline (number: $F(1, 17)=4.996$, $p < 0.05$); gender ($F(1, 17)=12.542$, $p < 0.05$) with a larger negativity for conditions in which gender and number were correct. An interaction with electrode site for both main effects revealed that the largest negativity was observed at CZ (number: $F(3, 51)=7.236$, $p < 0.05$; gender: $F(3, 51)=5.701$, $p < 0.05$).

At the 550–900 window chosen to target the P600 component there was a main effect of gender which was strongest at midline sites $F(1, 17)=5.383$, $p < 0.05$, and continued as a trend towards significance at the first lateral site $F(1, 17)=4.157$, $p=0.057$. Also there was a significant gender by number interaction (midline $F(1, 17)=6.487$, $p < 0.05$; first-lateral $F(1, 17)=4.342$, $p=0.053$, second lateral $F(1, 17)=4.950$, $p < 0.05$). Follow up tests showed that the effect was strongest for the condition in which gender was purely violated ($F(1, 17)=7.760$, $p < 0.05$) relative to when number was purely violated ($F(1, 17)=4.513$, $p=0.049$) or for the double violation condition. Also, the P600 effect was overall right-lateralized (first-lateral: Gender \times Hemisphere $F(1, 17)=9.663$, $p < 0.05$; Gender \times Number \times Hemisphere $F(1, 17)=5.794$, $p < 0.05$; second lateral: Gender \times Hemisphere $F(1, 17)=9.663$; Gender \times Number \times Hemisphere $F(1, 17)=6.343$, $p < 0.05$. Fig. 1 exemplifies the magnitude of the P600 effect in the four conditions. Fig. 2 shows an overview of grand mean waveforms comparing all conditions in a selection of electrodes (midline and first lateral line electrodes), and Fig. 3 illustrates the most representative electrode for the P600 component (PZ) highlighting the difference between the correct condition (i.e., +Gen+Num) and the three violation conditions (i.e., –Gen+Num, +Gen–Num, –Gen–Num). For graphing purposes only, the grand mean waveforms were further low-pass filtered at 15 Hz (12 dB).

2.4. Discussion Experiment 1

The aim of Experiment 1 was to investigate the processing of grammatical gender and number agreement in the context of pronominal references in native speakers of Spanish. The results show that native Spanish speakers are sensitive to violations caused by lack of gender and number correspondence between the antecedent and the subsequent clitic pronoun. The major ERP component that emerged in response to the three violations was a P600. These results are in line with previous ERP studies that investigated the processing of pronominal references in a variety of languages (Barber & Carreiras, 2005; Barber et al., 2004; Hagoort & Brown, 1999; Hammer et al., 2005; Lamers et al., 2006, 2008; Schmitt et al., 2002; Silva-Pereyra et al., 2012). A novel aspect of our results is that number and gender agreement was analyzed at the clitic pronoun, which has not been investigated in the literature before. Also, our results allow to generalize previous behavioral and neurophysiological findings to a novel grammatical structure. The results also show that the P600 effect was largest for the condition in which grammatical gender was uniquely violated, while in the other conditions in which number agreement was violated (i.e., the condition in which number alone was violated and in the condition in which number and gender were both violated) the P600 effect was smaller. Importantly, these results are in line with theories suggesting that gender processing is more taxing than number processing, even in native speakers (Antón-Méndez, 1999; Antón-Méndez et al., 2002; Sagarra, 2007). However, further research is needed to further clarify what specifically makes gender processing more taxing than number processing.

3. Experiment 2: the processing of gender and number marked on clitic pronouns in English–Spanish bilinguals

3.1. Method

3.1.1. Participants

In Experiment 2, 28 native speakers of English who were adult learners of Spanish as the L2 were tested (21 females, 7 males; mean age: 22.8 yrs.; $SD=4.47$). All recruiting criteria (other than language pairing) were identical to Experiment 1. As in Experiment 1, all participants completed a language history questionnaire to assess their language history and skills in both Spanish and English. They rated their language knowledge using a scale from 1 to 10 (1 being the lowest and 10 being the highest score) for oral comprehension, oral production, reading and writing. All participants started learning Spanish after age 14. For technical reasons, seven participants' data had to be excluded from the analysis. Therefore 21 native English L2 learners of Spanish (17 females, 4 males; mean age: 22.8 yrs.; $SD=4.76$) were included in the final analysis. Participants' demographics and language characteristics are summarized in Table 4.

3.1.2. Materials and design

Materials and design were identical to the ones described in Experiment 1.

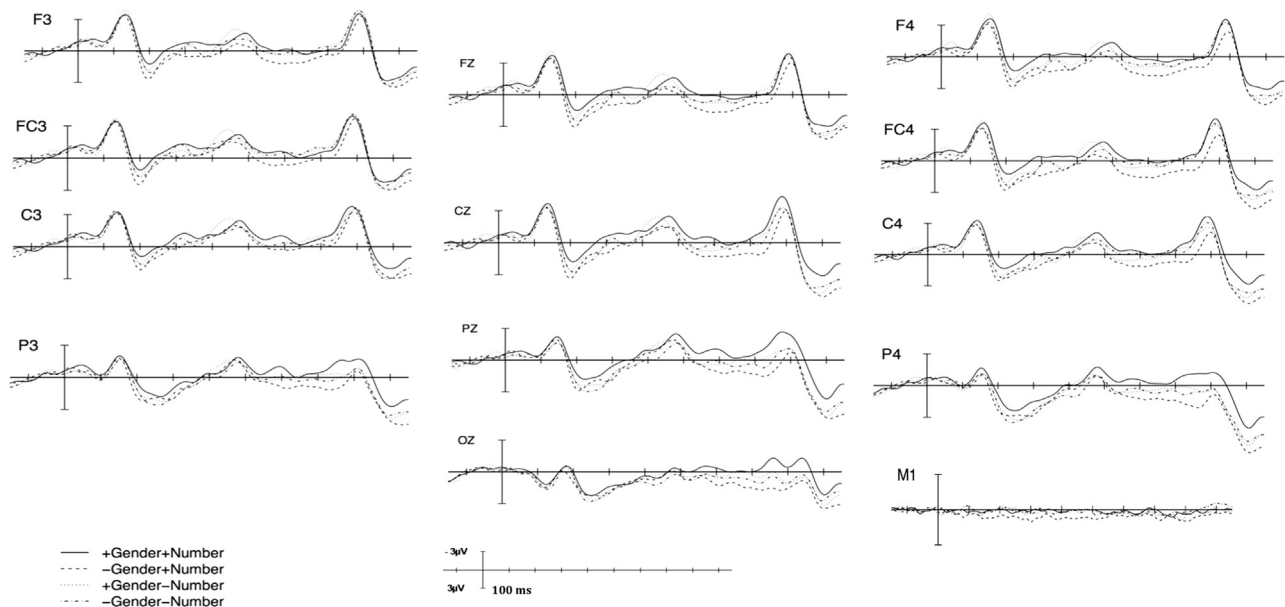


Fig. 2. Grand mean waveforms comparing all conditions for native Spanish speakers.

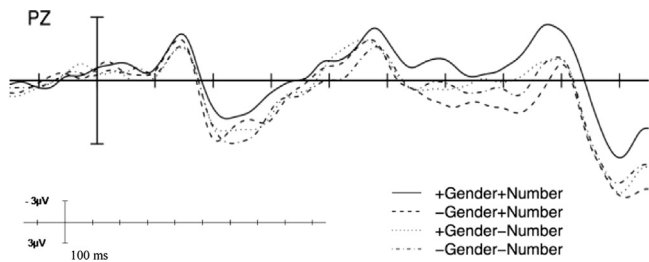


Fig. 3. The P600 effect at electrode PZ highlighting the difference between the four conditions.

Table 4

Demographics and language characteristics of L2 Spanish participants. Self-assessed ratings were on a scale from 1 to 10 where 10 was most proficient.

| Gender | Age (yrs) | Rating Spanish | Rating English |
|-----------|------------|----------------|----------------|
| F | 23 | 9 | 10 |
| F | 30 | 8.8 | 9.8 |
| F | 25 | 8.3 | 9.6 |
| F | 20 | 7 | 10 |
| F | 21 | 7 | 9.2 |
| F | 39 | 8.4 | 10 |
| M | 20 | 6.6 | 9.4 |
| F | 30 | 8 | 9.6 |
| F | 21 | 9.6 | 9.4 |
| F | 21 | 8.3 | 10 |
| F | 21 | 8.3 | 10 |
| F | 19 | 9 | 10 |
| M | 20 | 9 | 9.9 |
| M | 20 | 8.3 | 9.8 |
| M | 22 | 8.3 | 9.8 |
| F | 27 | 8.3 | 9.8 |
| Mean (SD) | 23.7 (5.4) | 8.3 (0.8) | 9.7 (0.3) |

3.1.3. Procedure

The procedure was identical to the one described in Experiment 1.

3.1.4. EEG data acquisition and analysis

EEG data acquisition was identical to Experiment 1. As in Experiment 1, ERP data analysis was performed on correct responses

and artifact-free trials only. To allow for adequate statistical power and for an optimal signal-to-noise ratio, the minimum percentage of trials per participant per condition to enter the analysis was 50% (24 trials). By these criteria, five participants had to be excluded from the analysis. Therefore, data from 16 English–Spanish bilinguals were analyzed. In the final sample, a total of 29% of trials were eliminated due to artifact and incorrect behavioral responses (22% for the +Gen+Num condition, 35% for the –Gen+Num condition, 20% for the +Gen–Num condition and 30% for the –Gen–Num condition). Grand averages were then computed for each condition. Two windows were chosen to capture the major components of interest. Based on the visual inspection of the grand mean waveforms, a first window from 250–400 ms was chosen to capture a LAN/N400 components. A second window from 550–900 ms was chosen to roughly correspond to the P600. Mean amplitudes values were calculated over these windows for each participant. Grand mean analyses were calculated for each time window.

3.2. Results: behavioral analysis

The statistical parameters used in the data analysis were identical to those used in Experiment 1³. Overall, for accuracy on the grammaticality judgment task, a significant interaction of gender \times number $F(1, 20)=13.028$, $p < 0.05$ and a main gender effect $F(1, 20)=10.258$, $p < 0.05$ show that English–Spanish bilinguals are less accurate to determine pure gender violations, while they are more accurate for the conditions which include a number violation. The RTs analysis did not reveal any significant effects. A summary of the behavioral results is shown in Table 5.

3.3. Results: ERP analysis

As in Experiment 1, only those results that show a main effect of Gender or Number and an interaction between the two, or an interaction between the experimental conditions and electrode site or hemisphere, will be reported. Significant electrode site or

³ For the behavioral results the analyses were performed on the original 21 participants, before the five participants were removed due to excessive EEG artifacts. Given the relatively low number of analyzable dataset, this permitted a stronger analysis at the behavioral level. Note, however, that the results do not change even when the analyses are performed on the smaller n .

Table 5

Summary of the behavioral results values (Mean Accuracy and RTs) for the end of the sentence acceptability judgment for the native English speakers, late L2 Spanish learners' group ($n=21$).

| | +Gender+Number Mean (SE) | –Gender+Number Mean (SE) | +Gender–Number Mean (SE) | –Gender–Number Mean (SE) |
|----------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Accuracy | 0.88 (0.11) | 0.66 (0.31) | 0.71 (0.30) | 0.72 (0.31) |
| RTs | 805.2 (74.2) | 904.9 (116.1) | 858.3 (100.4) | 793.6 (73.4) |

hemisphere results not interacting with the experimental conditions will not be reported, as general hemisphere and electrode site differences are to be expected.

At the first window of interest (250–400), a significant gender by number interaction was found at the second lateral sites ($F(1, 15)=4.944$, $p<0.05$) showing that pure gender violation (–Gen+Num) elicited the largest negativity. The effect was largest in the left hemisphere at frontal sites (hemisphere \times electrode $F(6, 96)=3.540$, $p<0.05$), in line with the presence of a LAN component. At the 550–900 window there was a main effect of number distributed at midline sites $F(1, 15)=4.566$, $p<0.05$, and first lateral sites $F(1, 15)=5.909$, $p<0.05$ showing a reliable larger positivity for conditions in which number was violated, (i.e., +Gen–Num condition and –Gen–Num condition). No other effects were observed relative to gender violations. Fig. 4 exemplifies the magnitude of the P600 effect in the four conditions. Fig. 5 shows an overview of grand mean waveforms comparing all conditions in a selection of electrodes (midline and first lateral line electrodes), and the most representative electrode for the P600 component (CZ) highlighting the difference between the correct condition (i.e., +Gen+Num) and the three violation conditions (i.e., –Gen+Num, +Gen–Num, –Gen–Num). For graphing purposes only, the grand mean waveforms were further low-pass filtered at 15 Hz (12 dB).

3.4. Interim discussion Experiment 2

The aim of Experiment 2 was to investigate whether English proficient L2 learners of Spanish would show sensitivity to gender and number using a structure (clitic pronouns) that also instantiates a co-referential grammatical structure. Clitic pronouns were chosen because they are not represented in English and are therefore a unique L2 structure. An interesting feature of clitics is that they encode grammatical gender, which is absent in English, and also number, which is instead present in English. Results show that overall, adult English–Spanish bilinguals show an early left distributed negativity in response to all violation conditions which included a violation of number. Critically, a P600 response to number violations demonstrates that English–Spanish bilinguals are sensitive to number violations marked on object clitics. However, the data do not show any P600 effect related to gender processing marked at the clitic, suggesting no sensitivity to violations of grammatical gender marked on clitic pronouns.

4. Post-hoc analysis

In order to investigate whether relative proficiency would modulate sensitivity to gender agreement, a post-hoc analysis was performed. For this exploratory analysis, participants were further separated in two proficiency groups (very-high proficiency and high proficiency) using the overall accuracy score on the online grammaticality judgment, as a proxy of overall proficiency. Participants with an overall score of 90% or above were identified as being very-high proficiency, while those who scored from 88% below were labeled as high proficiency. According to these criteria,

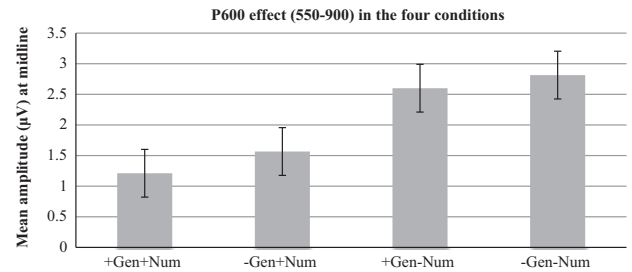


Fig. 4. Magnitude of the P600 effect in the four conditions for English L2 Spanish speakers (bars represent Standard Error).

9 participants were identified as very-high proficient ($n=9$; average accuracy: 0.94; SD=0.02; range: 0.90–0.97), and 7 as high proficient ($n=7$; average accuracy: 0.75; SD=0.13; range: 0.45–0.88). New repeated measure analyses were performed for the midline and at the two lateral lines for the 550–900 ms window, this time however with proficiency-group (high and very-high) as a between subject factor. Results show a main effect of number (midline: $F(1, 14)=7.680$, $p<0.05$; first lateral: $F(1, 14)=6.221$, $p<0.05$; second lateral=n.s.) showing that conditions in which number was violated elicited a greater positivity, with no difference between the two groups. No main effect of gender was observed; however a Gender \times Number \times Proficiency-group interaction (midline: $F(1, 14)=4.375$, $p=0.055$; first lateral: $F(1, 14)=8.232$, $p<0.05$; second lateral=n.s.) revealed that very-high proficient speakers showed a larger positivity for the condition in which gender was purely violated. Subsequent post-hoc analyses showed that differences for the pure gender violation between the two groups were most obvious at the first lateral sites ($F(1, 14)=3.618$, $p=0.07$). To seek additional convergence with the new above reported results, an additional exploratory analysis was performed in which the average mean amplitude (calculated at midline) for the condition in which gender was purely violated was correlated with the overall accuracy behavioral score (again as a proxy for proficiency). Results show a significant positive correlation between the mean P600 amplitude and the accuracy scores ($r=0.439$, $p<0.05$) suggesting again that the P600 effect was largest for more proficient L2 speakers.

Fig. 6 shows the magnitude of the P600 effect in the four conditions for both sub-groups. Fig. 7 plots the grand mean waveforms in all conditions for a larger subset of electrodes and the most representative electrode for the P600 component (PZ) for the high proficiency group. Fig. 8 plots the same for the very-high proficiency group. For plotting purposes only, the grand mean waveforms were further low-pass filtered at 15 Hz (12 dB). Finally Fig. 9 shows the correlation plot.

This exploratory analysis reveals that the observed sensitivity depends on the relative proficiency of the L2 speakers. Highly-proficient L2 learners show a similar ERP signature to native speakers, whereas the high proficiency L2 speakers show sensitivity to number agreement (also encoded in the L1) but not to gender (absent in the L1). Overall, the results are in line with previous behavioral and ERP data showing that at least a subset of highly proficient L2 speakers can process L2-specific grammatical

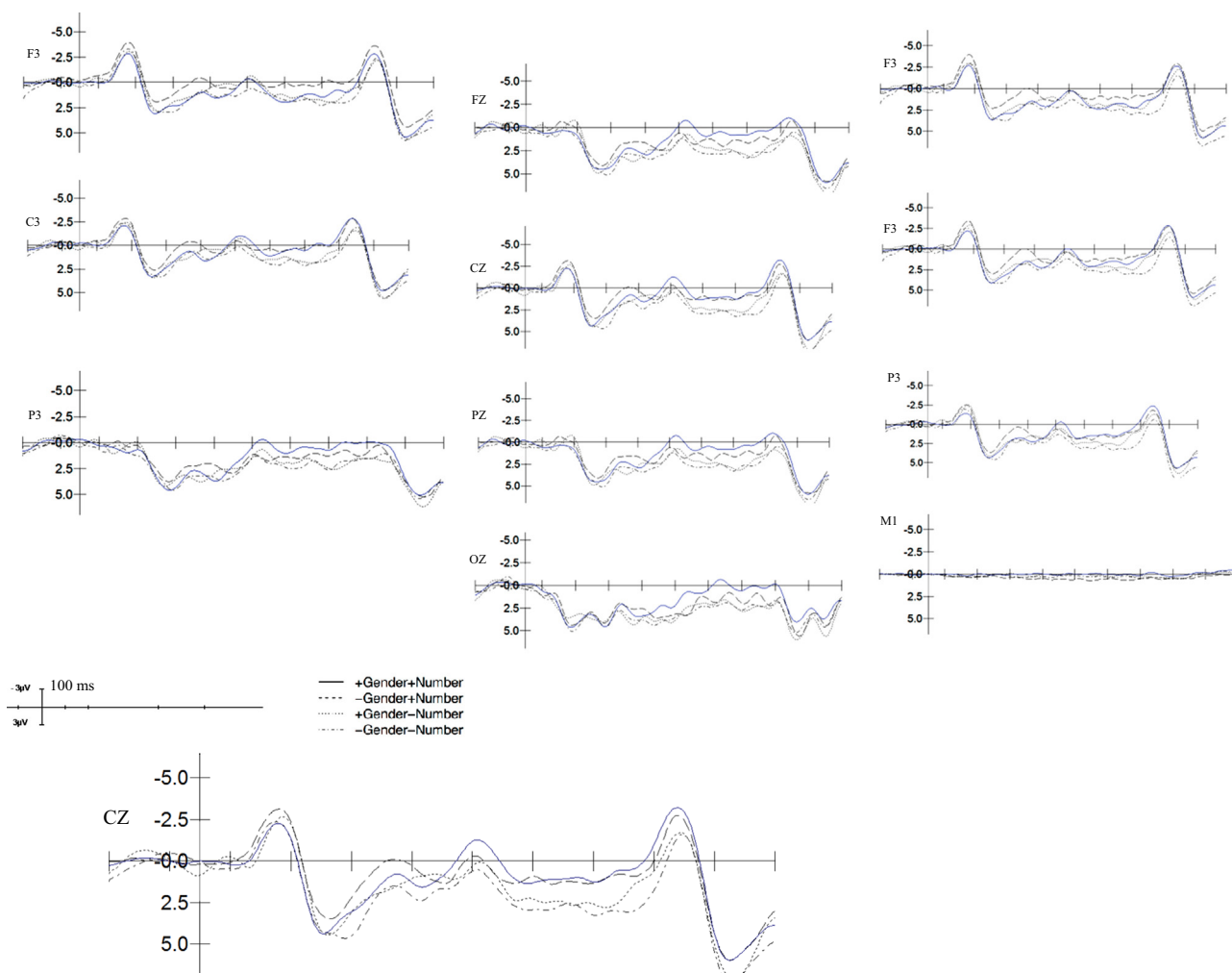


Fig. 5. Grand mean waveforms comparing all conditions for English L2 Spanish speakers for a subset of electrodes, and for the electrode CZ.

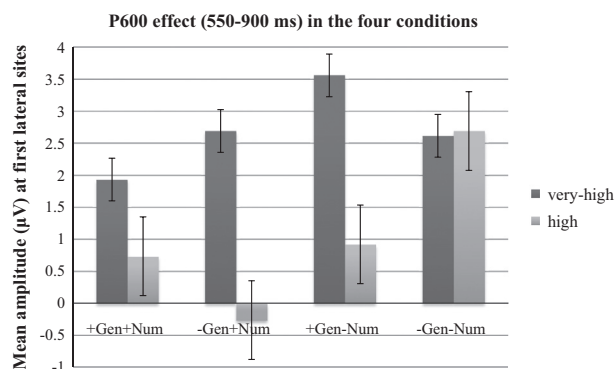


Fig. 6. Magnitude of the P600 effect in the four conditions for both sub-groups (bars represent Standard Error).

structures under the conditions imposed by on-line processing (Foucart & Frenck-Mestre, 2011, 2012).

5. General discussion

The goal of this study was to investigate the neurophysiological correlates of grammatical gender and number processing when native Spanish speakers (Experiment 1) and late English–Spanish bilinguals (Experiment 2) processed sentences that included a

pronominal reference between an antecedent and a clitic pronoun. Because of their temporal sensitivity, ERPs were chosen as the method that best reveals the earliest on-line neural responses to different grammatical violations. Clitic pronouns were used to further examine how gender and number marking on closed-class structures affects sentence processing (Molinero et al., 2008; Osterhout & Mobley, 1995). Given the novelty of examining the clitic structure, the first goal was to determine what kind of neurophysiological signature would be observed in response to gender and number violations marked on clitic pronouns in native speakers of Spanish. Previous behavioral research (Rossi, Diaz, Kroll, & Dussias, under review) shows that native Spanish speakers are sensitive to violations of clitic position in different sentential contexts (therefore tapping into word order processes), and are able to utilize gender and number information marked on clitics to predict upcoming linguistic information. However, in that particular set of studies, gender and number were correct, while clitic position was violated. As such, the question of whether native Spanish speakers are sensitive to gender and number violations marked on clitics while processing sentences online still remain unanswered.

The results of Experiment 1 demonstrate that native speakers of Spanish are sensitive to pure violations of gender, number, and also combined violations of gender and number marked on clitics. This sensitivity is primarily revealed by the emergence of the P600 component largest for pure gender violations, followed by number and double violations, suggesting a syntactic reanalysis and repair

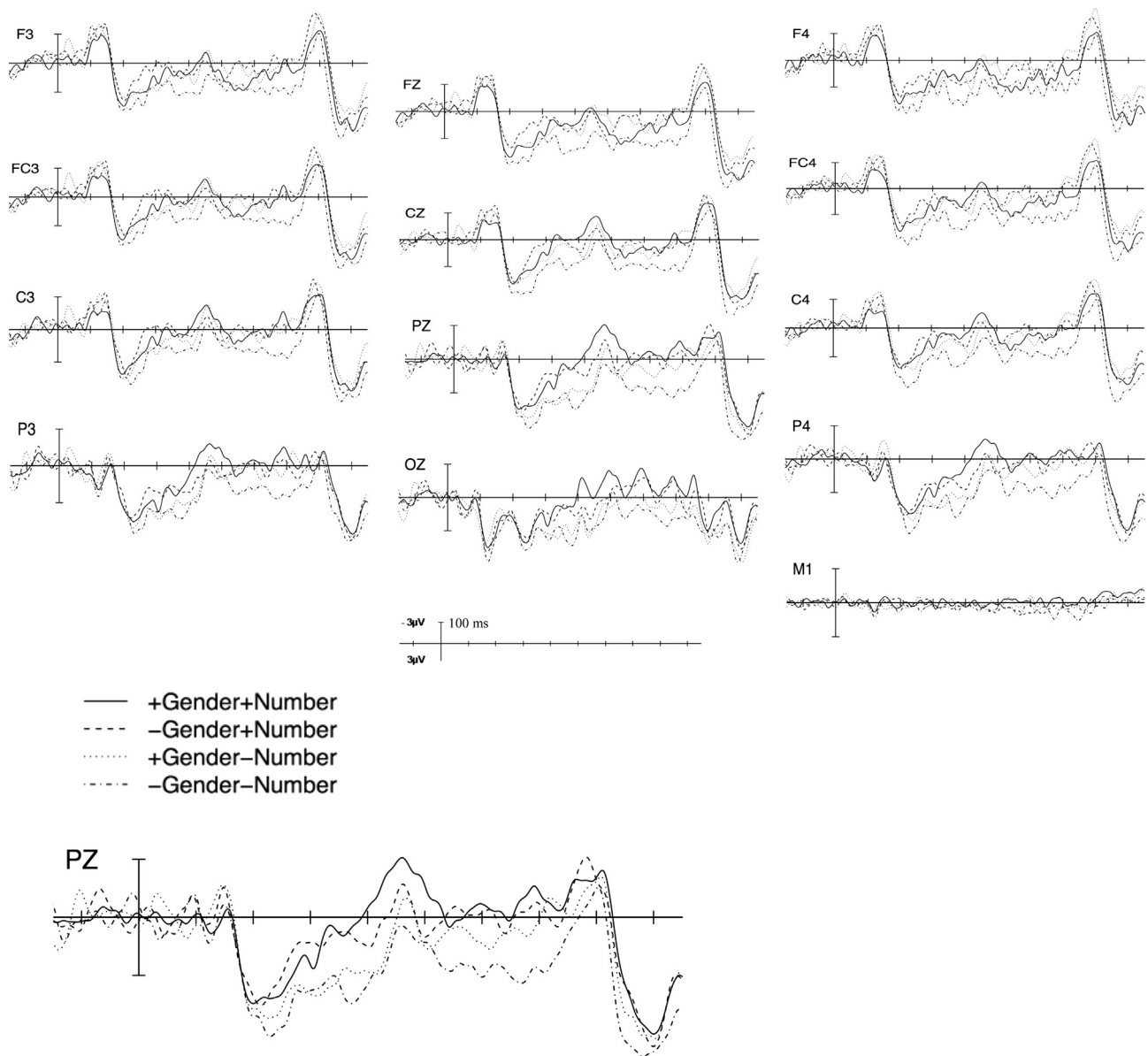


Fig. 7. Grand mean waveforms for all conditions and the most representative electrode for the P600 component (PZ) for the high proficiency group.

processes when gender and/or number features are violated on the clitic. This type of ERP signature not only patterns with the one observed for other closed-class referential structures such as reflexives and subject pronouns (Osterhout & Mobley, 1995), but is aligned with the more general finding that gender violations elicit a P600 component in a variety of grammatical structures and also for grammatical structures that are in a referential connection (Hammer et al., 2005; Lamers et al., 2006, 2008; Schmitt et al., 2002). Overall, previous ERP results on grammatical gender processing marked on closed-class words have been used to support the claim that grammatical gender is syntactically represented and that its online processing is syntactically driven, particularly when presented in a sentence context. The current results validate this view, as grammatical gender violations using clitic pronouns elicit a P600. For number violations marked on clitics, a P600 component was observed suggesting that number violations on clitics also require syntactic reanalysis. The P600 effect observed for pure gender violations was however the strongest effect. The present results could reflect a previous behavioral literature showing that gender concord might be

cognitively more complex to process than number (Antón-Méndez, 1999; Antón-Méndez et al., 2002).

The central question addressed by Experiment 2 was whether proficient L2 learners of Spanish would show sensitivity to violations of grammatical gender and number marked on clitic pronouns. Clitic pronouns were specifically chosen as a grammatical structure not encoded in English, and as such it is not a structure that is represented in the L2 speakers' L1. Clitic pronouns are also an example of a grammatical structure that has been shown to be particularly complex to acquire and master during bilingual language acquisition (Pérez-Leroux et al., 2011) and for late L2 learners both in comprehension and production (Rossi & Dussias, *in preparation*), especially when the L1 and the L2 do not overlap in the presence of this structure. Results in the ERP literature which have looked at gender and number processing in a variety of grammatical structures and languages have yielded mixed results, some revealing that late L2 speakers are sensitive to those violations and show a similar ERP signature than native speakers (Foucart & Frenck-Mestre, 2011, 2012) while others fail to find similar results. Past studies have also shown that cross-language

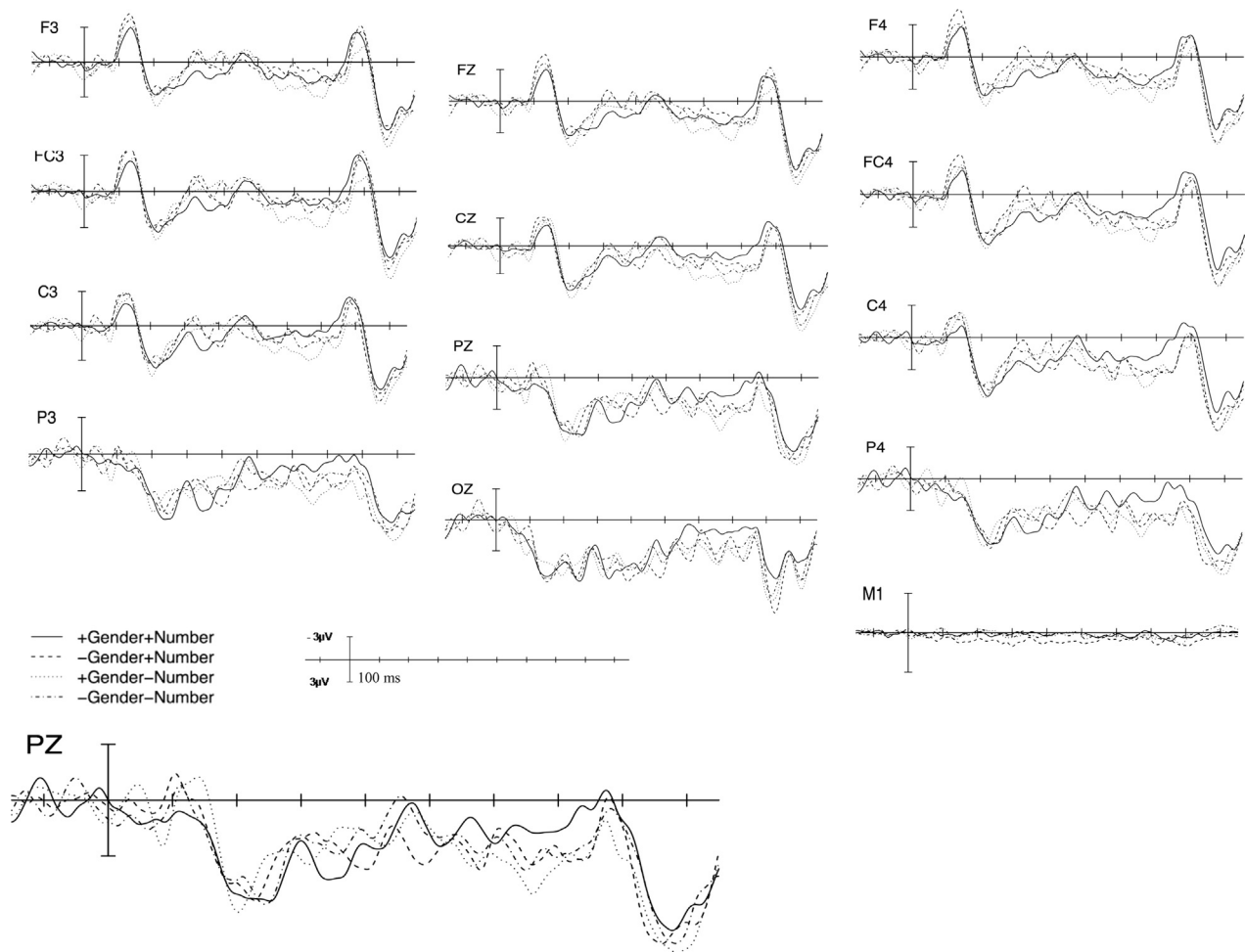


Fig. 8. Grand mean waveforms for all conditions and the most representative electrode for the P600 component (PZ) for the very-high proficiency group.

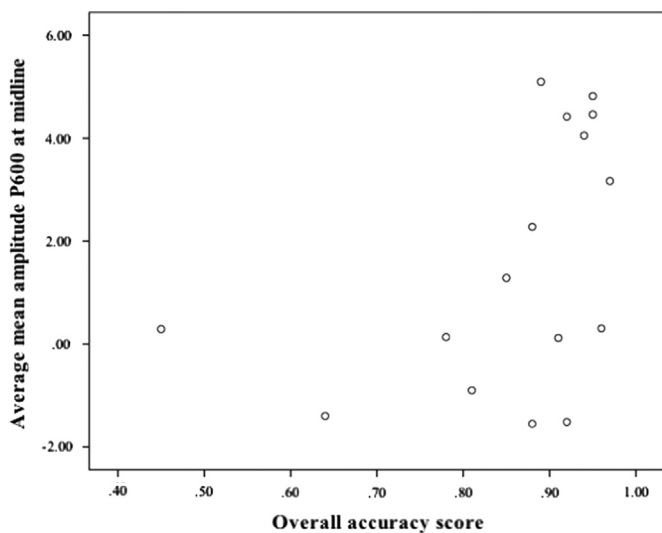


Fig. 9. Plot showing the correlation between the average mean P600 amplitude and the overall accuracy score.

similarity and proficiency appear to be factors that determine whether late L2 learners show a native-like neurophysiological response to violations of gender and number (Sabourin & Stowe, 2008; Foucart & Frenck-Mestre, 2011, 2012). The results of the present study show that native speakers of English, who are

proficient learners of Spanish, are sensitive to number violations marked on the clitic, but not to pure gender violations. These initial results can be interpreted in support of a deficit processing account (Clahsen & Felser, 2006; Hawkins & Chan, 1997), and more generally in line with representational deficit account. However, the fact that L2 speakers showed a sensitivity to number (which is present in the L1, but not the L2) at the clitic pronoun is in itself suggestive that English L2 speakers of Spanish were able to process morphosyntactic information marked on a grammatical structure which is unique to the L2. As such, an extreme version of representational accounts does not seem viable to explain this pattern of performance. A post-hoc analysis carried out on two subgroups of the L2 speakers suggested that sensitivity to grammatical gender could indeed emerge in the ERP signature as a degree of increased proficiency. The emergence of sensitivity to gender in the very-high proficient subset of L2 speakers is in line with representational accessibility accounts supporting the idea that L2 attainment is possible as a function of increased proficiency, even for grammatical structures and features that are not present in the L1 (Schwartz & Sprouse, 1996). The present analysis for the two subgroups of L2 speakers will necessarily need to be validated by further data, with larger data sets, and with the support of standardized measures of language proficiency.

Although ERP methods provide a temporally sensitive measure of the earliest stages of language processing, the requirement to use RSVP presentation in the ERP paradigm with words presented one at a time also makes it difficult to test other aspects of language processing that are more available in methods such as

eye tracking. In particular, the present results suggest that both native Spanish speakers and late English L2 learners of Spanish were sensitive to violation of grammatical gender and number, but it was not possible to determine whether the two groups differ in the strategies they use to repair the processed violations. In a study in progress, Rossi and Dussias (in preparation) are taking this approach by using the same experimental materials in an eye-tracking paradigm. Measures such as first and second pass regressions will provide important information about which strategies native and L2 speakers utilize once they encounter gender and number violations marked on clitic pronouns. Recent data reported by Hopp (2013) showing correlations between accuracies in lexical gender assignment in production and speed of utilizing grammatical gender information predictively in an visual eye-tracking paradigm strongly suggest that further research is needed linking comprehension and production strategies to have a broader understanding of gender and number processing in L2 speakers.

In sum, the reported data suggest that at least for a subset of late bilinguals resemble native speakers in their sensitivity to violations of grammatical gender and number. However, our results show also that processing grammatical gender at the clitic remains a difficult task for less proficient L2 speakers and does indeed suggest that these are difficult structures to acquire but that the constraints associated with complete acquisition are not insurmountable. In future research it will be important to exploit converging sources of evidence across different methods to identify the locus of differences between native and L2 speakers.

Acknowledgments

This research was supported by the Marica De Vincenzi grant to Eleonora Rossi, and by the NIH Grant HD053146 to Judith F. Kroll, NSF Grants BCS-0955090 and OISE-0968369 to Judith F. Kroll and Paola E. Dussias, and NSF Grant BCS-0821924 and NIH grant R21 DO71758-01 to Paola E. Dussias. We thank the anonymous reviewers and Holger Hopp for helpful comments on this manuscript.

References

- Abutalebi, J., Cappa, S. F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, 4, 179–190.
- Ainsworth-Darnell, K., Shulman, R., & Boland, J. (1998). Dissociating brain responses to syntactic and semantic anomalies: evidence from event-related brain potentials. *Journal of Memory and Language*, 38, 112–130.
- Antelmi, D. (1997). *La prima grammatica dell'italiano: indagine longitudinale sull'acquisizione della morfologia italiana*. Bologna: Il Mulino.
- Andrews, G., Birney, D., & Halford, G. S. (2006). Relational processing and working memory capacity in comprehension of relative clause sentences. *Memory & Cognition*, 34, 1325–1340.
- Antón-Méndez, M. I. (1999). *Gender and number agreement processing in Spanish* (Unpublished Doctoral dissertation). Tucson: University of Arizona.
- Antón-Méndez, M. I., Nicol, J. L., & Garrett, M. (2002). The relation between gender and number agreement processing. *Syntax*, 5, 1–25.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number in Spanish: an ERP comparison. *Journal of Cognitive Neuroscience*, 17, 137–153.
- Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement?. In: M. Carreiras, & C. Clifton, C., Jr. (Eds.), *The on-line study of sentence comprehension: eye-tracking and beyond* (pp. 309–327). New York: Psychology Press.
- Birdsong, D. (1999). *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Erlbaum.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44, 235–249.
- Blackwell, A., & Bates, E. (1995). Inducing agrammatical profiles in normals: evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience*, 7, 228–257.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42.
- De Vincenzi, M. (1999). Differences between the morphology of gender and number: evidence from establishing coreferences. *Journal of Psycholinguistic Research*, 28, 537–554.
- De Vincenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L., et al. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language*, 85, 280–296.
- Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review*, 108, 759–788.
- Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese–English second language speakers. *Second Language Research*, 26, 443–472.
- Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, 30, 321–338.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. *Brain and Language*, 50, 259–281.
- Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, 14, 379–399.
- Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66, 226–248.
- Frenck-Mestre, C., Foucart, A., Carrasco, H., & Herschensohn, J. (2009). Processing of grammatical gender in French as a first and second language: evidence from ERPs. *EUROSLA Yearbook*, 9, 76–106.
- Frenck-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in French: evidence from ERPs. *Acta Psychologica*, 128, 528–536.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the “critical period” hypothesis. *Proceedings of the National Academy of Science of the United States of America*, 99, 529–534.
- Friederici, A. D., Garrett, M. F., & Jacobsen, T. (1999). Editorial and introduction to Part I of the special issue processing of grammatical gender. *Journal of Psycholinguistic Research*, 28, 455–456.
- Gillon Dowens, M. G., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish—evidence from event related potentials. *Neuropsychologia*, 49, 1651–1659.
- 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.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–111.
- Grüter, T. (2006). Object (clitic) omission in L2 French: Missetting or missing surface inflection. In *Proceedings of the 8th generative approaches to second language acquisition conference (GASLA 2006)* (vol. 66973, pp. 63–71). Somerville, MA: Cascadia Proceedings Project.
- Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: a web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, 45, 765–771.
- Guasti, M. T. (1994). Verb syntax in Italian child grammar: finite and non-finite verbs. *Language Acquisition*, 3, 1–40.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556–568.
- Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. (2012). On the time course of accessing meaning in a second language: an electrophysiological and behavioral investigation of translation recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1165–1186.
- Hagoort, P., & Brown, C. M. (1999). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, 38, 1531–1549.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439–483.
- Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 30, 251–266.
- Hahne, A., Eckstein, K., & Friederici, A. D. (2004). Brain signatures of syntactic and semantic processes during children's language development. *Journal of Cognitive Neuroscience*, 16, 1302–1318.
- 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, 123–140.
- Hammer, A., Jansma, B., Lamers, M., & Münte, T. F. (2005). Pronominal reference in sentences about persons or things: an electrophysiological approach. *Journal of Cognitive Neuroscience*, 17, 227–239.
- Herschensohn, J. (2000). *The second time around: minimalism and L2 acquisition*. Amsterdam/Philadelphia: John Benjamins.
- Herschensohn, J. (2001). Missing inflection in L2 French: accidental infinitives and other verbal deficits. *Second Language Research*, 17, 273–305.
- Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage*, 15, 647–660.

- Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. *Second Language Research*, 13, 187–226.
- Hopp, H. (2007). *Ultimate attainment at the interfaces in second language acquisition: grammar and processing* (Unpublished Doctoral thesis). Groningen (The Netherlands): University of Groningen.
- Hopp, H. (2010). Ultimate attainment in L2 inflection: performance similarities between non-native and native speakers. *Lingua*, 120, 901–931.
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: relations between lexical and syntactic variability. *Second Language Research*, 29, 33–56.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Johnson, J. S., Shenkman, K., Newport, E., & Medin, D. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language*, 35, 335–352.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 98, 122–149.
- Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*, 274, 114–116.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15, 98–110.
- Kayne, R. (1991). Romance clitics, verb movement and PRO. *Linguistic Inquiry*, 22, 647–686.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: evidence from event-related potentials. *Journal of Memory and Language*, 52, 205–225.
- King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 376–395.
- Kotz, S. (2009). A critical review of ERP and fMRI evidence on L2 processing. *Brain and Language*, 109, 68–74.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic anomaly. *Science*, 207, 203–205.
- Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: event-related potential investigations. In: M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego, CA: Academic Press.
- Lamers, M. J. A., Jansma, B. M., Hammer, A., & Münte, T. F. (2006). Neural correlates of semantic and syntactic processes in the comprehension of case marked pronouns: evidence from German and Dutch. *BMC Neuroscience*, 7, 23.
- Lamers, M. J. A., Jansma, B. M., Hammer, A., & Münte, T. F. (2008). Differences in the processing of anaphoric reference between closely related languages: neurophysiological evidence. *BMC Neuroscience*, 9, 55.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.
- Linck, J. A. (2013). Working memory and second language comprehension and production: a meta-analysis. *Psychonomic Bulletin & Review*, 21, 1–75.
- Lonzi, L., & Luzzatti, C. (1993). Relevance of adverb distribution for the analysis of sentence representation in agrammatic patients. *Brain and Language*, 45, 306–317.
- McDonald, J. L. (2000). Grammaticality judgments in a second language: influences of age of acquisition and native language. *Applied Psycholinguistics*, 21, 395–423.
- McDonald, J. L. (2006). Beyond the critical period: processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381–401.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, 7, 703–704.
- McLaughlin, J., Tanner, D., Pitkanen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., et al. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, 60, 123–150.
- MacWhinney, B. (1997). Second language acquisition and the competition model. In: A. M. B. de Groot, & J. Kroll (Eds.), *Tutorials in bilingualism: psycholinguistic perspectives* (pp. 113–144). Mahwah, NJ: Erlbaum.
- MacWhinney, B. (2005). A unified model of language acquisition. In: J. Kroll, & A. M. B. de Groot (Eds.), *Handbook of bilingualism: psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.
- Miceli, G., Silveri, C., Romani, C., & Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic patients. *Brain and Language*, 36, 447–492.
- Miceli, G., & Mazzucchi, A. (1990). Agrammatism in Italian: Two case studies. In: Menn L., & Obler L. (Eds.), *Agrammatic aphasia: A cross linguistic narrative sourcebook* (pp. 717–757). Amsterdam: John Benjamins.
- Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, 67, 224–237.
- Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, 11, 671–717.
- Molinero, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 47, 908–930.
- Molinero, N., Kim, A., Vespignani, F., & Job, R. (2008). Anaphoric agreement violation: an ERP analysis of its interpretation. *Cognition*, 106, 963–974.
- Montrul, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition*, 33, 163–192.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: an event-related potential study. *Language Learning*, 60, 154–193.
- Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3, 151–165.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1995). On the language specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*, 8, 507–526.
- Osterhout, L., McLaughlin, J., Pitkanen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Language Learning*, 56, 199–230.
- Osterhout, L., & Mobley, L. A. (1995). Event-related potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 14, 283–317.
- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., et al. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics*, 21, 509–521.
- Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, 22, 2728–2744.
- Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L2 and L2 sentence processing: a study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, 25, 501–528.
- Perani, D., Paulesu, E., Sebastián-Gallés, N., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: proficiency and age of acquisition of the second language. *Brain*, 121, 1841–1852.
- Pérez-Leroux, A. T., Cuza, A., & Thomas, D. (2011). Clitic placement in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, 14, 221–232.
- Reznik, M., Dubrovsky, S., & Maldonado, S. (1995). Agrammatism in Spanish: A case study. *Brain and Language*, 51, 355–368.
- Rizzi, L. (1982). *Issues in Italian syntax*. New York: Foris.
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 18, 2030–2048.
- Rossi, E. (2007). (Ph.D. dissertation). *Clitic production in Italian agrammatism*. Groningen: Grodil (Series number 64).
- Rossi, E. (2013). Modulating the sensitivity to syntactic factors in production: Evidence from syntactic priming in agrammatism. *Applied Psycholinguistics*. available on CJO, <http://dx.doi.org/10.1017/S0142716413000374>.
- Rossi, E., Diaz, M., Kroll, J. F. & Dussias, P. E. (under review). When sentence comprehension in the native language is not enough: What second language learners and bilinguals tell us about the processing of morphosyntax.
- Rossi, E., & Dussias, P. E. (in preparation). Using eye tracking to examine sensitivity to gender and number in L2 sentence processing.
- Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2004). Antecedent priming at trace positions in children's sentence processing. Unpublished manuscript, University of Essex.
- Sabourin, L., & Haverkort, M. (2003). Neural substrates of representation and processing of a second language. In: R. van Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The lexicon–syntax interface in second language acquisition* (pp. 175–195). Amsterdam: Benjamins.
- Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research*, 24, 397–430.
- Sabourin, L., Stowe, L., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22, 1–29.
- Sagarra, N. (2007). Online processing of gender agreement in low proficient English–Spanish late bilinguals. In: M. J. Cabrera, J. Camacho, V. Déprez, N. Flores, & L. Sánchez (Eds.), *Current issues in linguistic theory series* (pp. 240–253). Amsterdam: John Benjamins.
- Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 120, 2022–2039.
- Santor, M. (2007). Second language acquisition of Italian accusative and dative clitics. *Second Language Research*, 23, 37–50.
- Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12, 40–72.
- Schmitt, B. M., Lamers, M., & Münte, T. F. (2002). Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. *Cognitive Brain Research*, 14, 333–346.
- Silva-Pereyra, J., Gutierrez-Sigut, E. V. A., & Carreiras, M. (2012). An ERP study of coreference in Spanish: semantic and grammatical gender cues. *Psychophysiology*, 49, 1401–1411.
- Stavrakaki, S., & Kouvava, S. (2003). Functional categories in agrammatism: Evidence from Greek. *Brain and Language*, 86, 129–141.
- Steinhauer, K., & Drury, J. E. (2011). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, 120, 135–162.

- Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Language Research*, 25, 13–41.
- Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: a brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L1. *Applied Linguistics*, 1–26.
- Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2011). Language related ERP components. In: S. J. Luck, & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 397–440). New York: Oxford University Press.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Studies in Second Language Acquisition*, 27, 173–204.
- Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: an electrophysiological investigation. *Psychophysiology*, 38, 41–63.
- Weber-Fox, C. M., & Neville, H. J. (1996). maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–256.
- Weber-Fox, C. M., & Neville, H. J. (1999). Functional neural subsystems are differentially affected by delays in second-language immersion: ERP and behavioral evidence in bilingual speakers. In: D. Birdsong (Ed.), *The critical period hypothesis and second language acquisition* (pp. 23–38). Mahwah, NJ: Erlbaum.