When A Nonnative Accent Lets You Spot All the Errors: Examining the Syntactic Interlanguage Benefit

Leah Gosselin[1], Clara D. Martin[2,3], Ana González Martín[4], and Sendy Caffarra[2,5,6]

Abstract

■ In our continuously globalizing world, cross-cultural and cross-linguistic communications are far from exceptional. A wealth of research has indicated that the processing of nonnative-accented speech can be challenging for native listeners, both at the level of phonology. However, few online studies have examined the underpinnings of accented speech recognition from the perspective of the nonnative listener, even though behavioral studies indicate that accented input may be easier to process for such individuals (i.e., the interlanguage speech intelligibility benefit. The current EEG study first examined the phonological and syntactic analysis of nonnative-accented speech among nonnative listeners. As such, 30 English learners of Spanish listened to syntactically correct and incorrect Spanish sentences produced in native and nonnative-accented Spanish. The violation in the incorrect sentences was caused by errors that are typical (i.e., gender errors; **la color*) or atypical (i.e., number errors; **los color*) for English learners of Spanish. Results indicated that nonnative listeners elicit a phonological mismatch negativity (PMN) when attending to speech produced by a native Spanish speaker. Furthermore, the nonnative listeners showed a P600 for all grammatical violations, indicating that they repair all errors regardless of their typicality or the accent in which they are produced. Follow-up analyses compared our novel data to the data of native listeners from the methodologically identical precursor study. These analyses showed that native and nonnative listeners exhibit directionally opposite PMN effects; whereas natives exhibited a larger PMN for English-accented Spanish, nonnatives displayed a larger PMN in response to native Spanish utterances (a classic interlanguage speech intelligibility benefit). An additional difference was observed at the syntactic level: Whereas natives repaired only atypical number errors when they were English-accented, nonnative participants exhibited a P600 in response to all English-accented syntactic errors, regardless of their typicality (a syntactic interlanguage speech intelligibility benefit). Altogether, these results suggest that accented speech is not inherently difficult to process; in fact, nonnatives may benefit from the presence of a nonnative accent. Thus, our data provide some of the first electrophysiological evidence supporting the existence of the classic interlanguage speech intelligibility benefit and its novel syntactic counterpart. ■

INTRODUCTION

One of the most difficult aspects of second language (L2) learning is that of phonology. Indeed, when it comes to language acquisition, phonology is theorized to possess an especially restrictive critical period, with some placing its cessation as early as 12 months (Ruben, 1999). This means that even highly proficient L2 learners often possess a noticeable nonnative accent when they speak. In today's globalized society, individuals are constantly confronted with such speech input; this societal change is reflected by an upsurge in the scientific literature examining accented speech processing.

Classic behavioral studies have indicated that native listeners (first language [L1] listeners) are better at understanding other native speakers than nonnative speakers (e.g., Smith et al., 2003; van Wijngaarden, Steeneken, & Houtgast, 2002; van Wijngaarden, 2001; Munro & Derwing, 1999; Munro, 1998). Although little research has focused on the phonological processing of L2 listeners, some results suggest that the presence of a nonnative accent is far less problematic for this group: Listeners with a nonnative accent have little difficulty in decoding and understanding accented speech (Hayes-Harb, Smith, Bent, & Bradlow, 2008; Bent & Bradlow, 2003; van Wijngaarden et al., 2002).

Though this seminal finding has mostly been supported cross-linguistically, it is not known whether these same effects extend to higher levels of linguistic analysis, such as syntax. Are L2 listeners able to detect syntactic violations when these errors are produced in a nonnative accent? Are they better than L1 listeners at doing so? Is this true for all types of syntactic errors? To investigate these questions, the current EEG study extends previous work conducted on native listeners (Caffarra & Martin, 2019) to nonnative listeners. That is, we explore L2 listeners' processing of

[1]University of Ottawa, [2]Basque Center on Cognition, Brain and Language, Donostia, Spain, [3]Ikerbasque–The Basque Foundation for Science, Bilbao, Spain, [4]Universidad del Atlántico Medio, Las Palmas, Spain, [5]Stanford University School of Medicine, [6]University of Modena and Reggio Emilia

Journal of Cognitive Neuroscience 34:9, pp. 1650–1669
https://doi.org/10.1162/jocn_a_01886

nonnative-accented speech, specifically the temporal dynamics of their phonological and syntactic processing for such input.

The Interlanguage Speech Intelligibility Benefit

In nonnative language acquisition, L2 speakers generally preserve some articulatory features of their L1; as a result, there may be common phonological divergences from the native "norm" in L2 speech. This idiolect is classically known as an *interlanguage* (Selinker, 1972). The properties of the interlanguage are believed to depend on learners' specific experiences with their L2 and, potentially, the timeline of the idiolect's fossilization (Selinker, 1972). This means that interlanguages are not entirely uniform across all learners; nonetheless, learners who share a common nonnative status (and especially a common native language) are assumed to retain overlapping interlanguage properties (Dickerson, 1975).

Crucially, interlanguages have proven to be a key intelligibility factor for both native (e.g., Braun, Dainora, & Ernestus, 2011; Floccia, Butler, Goslin, & Ellis, 2009; Clarke & Garrett, 2004; Smith et al., 2003; van Wijngaarden et al., 2002; van Wijngaarden, 2001; Munro & Derwing, 1999; Munro, 1998) and nonnative (e.g., Xie & Fowler, 2013; Weber, Broersma, & Aoyagi, 2011; Bent & Bradlow, 2003; van Wijngaarden et al., 2002) listeners. For instance, L1 listeners confronted with an interlanguage may experience word identification and sentence comprehension difficulties (e.g., Munro & Derwing, 1995), though exposure to nonnative accents may abate these problems, potentially because of habituation or adaptation mechanisms (e.g., Porretta, Tucker, & Järvikivi, 2016; Witteman, Weber, & McQueen, 2013; Clarke & Garrett, 2004; Gass & Varonis, 1984). However, in a seminal paper, Bent and Bradlow (2003) showed that this obstacle does not necessarily hold for other L2 individuals. This line of research indicates that L2 listeners are not hindered by the presence of accented speech. In fact, they may even outperform native listeners when it comes to detecting linguistic cues that are produced in a phonologically deviated interlanguage (e.g., Bent & Bradlow, 2003). This phenomenon was coined the "interlanguage speech intelligibility benefit." It is theorized that the benefit arises as speaker–listener pairs who share a nonnative status experience facilitation for deviations that are interlanguage errors; it is also notable that L2 listeners generally procure more exposure to L2 speech (particularly, their own variety) than do native listeners (e.g., Harding, 2012). The benefit appears to hold when the nonnatives possess the same L1 ("matched" interlanguage benefit: L1 Korean individuals show a benefit when listening to Korean-accented English sentences; Bent & Bradlow, 2003) and, often, when they do not ("mismatched" interlanguage benefit: for instance, L1 Korean individuals show a benefit when listening to Chinese-accented English; Bent & Bradlow, 2003; though see Wang & van Heuven, 2015). Some researchers suggest that the mismatched interlanguage benefit only holds when the "accent is perceptually confusable with the standard pronunciation for the L2 listener" (Weber et al., 2011, p. 479).

Since Bent and Bradlow's (2003) influential paper, the interlanguage speech intelligibility benefit has commonly been examined by utilizing within-group and between-group designs (see also similar themes in code-switching research: Phillips & Pylkkänen, 2021). Within-group designs showed that L2 listeners may find nonnative-accented speech as intelligible or even more intelligible than native-accented speech (Korpal & Sobkowiak, 2020; Podlipský, Šimáčková, & Petráž, 2016; Xie & Fowler, 2013; Harding, 2012; Weber et al., 2011; Leikin, Ibrahim, Eviatar, & Sapir, 2009; Bent & Bradlow, 2003; Smith & Rafiqzad, 1979; though see Lee, Kang, & Nam, 2020; Smith, Hayes-Harb, Bruss, & Harker, 2009; Hayes-Harb et al., 2008; Stibbard & Lee, 2006). Between-group designs showed that nonnative speech is more intelligible for nonnative listeners than for native listeners (Chu, 2013; Xie & Fowler, 2013; Lee & Xue, 2011; Hayes-Harb et al., 2008; Bent & Bradlow, 2003; though see Podlipský et al., 2016; Algethami, Ingram, & Nguyen, 2011; Smith et al., 2009; Stibbard & Lee, 2006). This between-group interlanguage benefit is also known as benefit for listeners. For ease of comprehension, the mise-en-scène in Item 1 distinguishes the within-group and between-group interlanguage speech intelligibility benefit (ISIB).

1. **Mise-en-scène**: John, Marco, and Giulia are discussing in English. While John is a native speaker of English, Marco and Giulia are native speakers of Italian and possess a noticeable Italian accent when they speak English.
   a. **Within-group ISIB ("for talkers")**: Marco comprehends English better when it is spoken by Giulia than by John.
   b. **Between-group ISIB ("for listeners")**: Marco comprehends Giulia's English better than John does.

Note, however, that the bulk of the research examining the interlanguage speech benefit utilizes behavioral tasks (e.g., transcription, phoneme recognition, accent/comprehensibility ratings, word identification). Though valuable, these techniques fail to give us a complete understanding of accented speech processing. Accent-related behavioral differences, though informative, are unidimensional (see Kaan, 2007). For instance, finding a lower transcription accuracy for nonnative-accented speech compared with native-accented speech may be taken as an indication that nonnative input is more difficult to comprehend; however, the cause, functional interpretation, or temporal dynamics of this effect cannot be gleaned from the data. Furthermore, behavioral measures may be impacted by external factors, such as participants' strategic processing or their attitudes toward the stimuli (Kaan, 2007). This is especially important for the current study, as nonnative accents are known to be subjected to covert (but not necessarily overt) negative biases (Roessel, Schoel, & Stahlberg, 2020). For these reasons, the utilization of implicit processing measures, such as EEG, are essential in understanding the interlanguage speech intelligibility benefit and in complementing the existing behavioral findings in the literature.

ERP Studies

ERPs exemplify a suitable technique to examine implicit and real-time language processing. ERPs are obtained by segmenting continuous EEG recordings and time locking each segment to the presentation of a specific event. This technique is particularly useful for linguistic analysis, as ERPs are highly temporally precise and are not reliant upon secondary behavioral tasks. Indeed, electrophysiology is the only technique that currently allows researchers to track the temporal course of processing for single words in continuous speech. Furthermore, ERPs are multidimensional (Kotz, 2009; Kaan, 2007). This means that possible accent-related ERP differences can be interpreted quantitatively and qualitatively; the polarity, amplitude, and timing of the electrophysiological peaks allow us to disentangle the participants' various levels of linguistic processing. In effect, distinct ERP components have been identified and associated to phonological processing or syntactic processing. These components will be reviewed below.

Phonological Processing

The P2 and the PMN are ERP components classically related to perceptual processing. During speech processing, these components may thus be modulated according to the accent in which an utterance is produced.

First, the P2 (or P200) is a positive-going ERP component; the peak in neural activity occurs approximately 200 msec after the target stimulus onset, usually on centrofrontal sensors. It is believed that the P2 indexes a listener's ease of extraction of the acoustic features within a given speech input; the P2 is reduced when feature extraction is more challenging or laborious (Reinke, He, Wang, & Alain, 2003). Though no known ERP studies have directly examined the interlanguage speech benefit, accented speech studies conducted on L1 listeners show that nonnative-accented speech perception leads to a decrease in the P2 as compared with native-accented speech (Foucart & Hartsuiker, 2021; Romero-Rivas, Martin, & Costa, 2015; Thomas, Martin, & Caffarra, under review). This suggests that, for L1 listeners, it is more difficult to extract the acoustic information from a nonnative interlanguage than from native speech (see Romero-Rivas et al., 2015).

Second, the PMN is a negative-deflecting ERP component, which typically peaks centrofrontally at approximately 300 msec poststimulus onset (see Lewendon, Mortimore, & Egan, 2021, for a review). The PMN is a prelexical effect; an increased negativity occurs when there is an inconsistency between the sensory input of a given target and its expected phonological representation (Newman & Connolly, 2009; Connolly, D'Arcy, Kujala, & Alho, 2001; Connolly & Phillips, 1994). Past research shows that the presence of a nonnative accent leads to an increased PMN response among L1 listeners (Porretta, Tremblay, & Bolger, 2017; Thomas et al., under review; see also Schiller et al., 2020; Goslin, Duffy, & Floccia, 2012), especially for those who report little or no experience with the accent under study (Porretta et al., 2017). This appears to suggest that, for L1 listeners, lexical recognition (based on acoustic features) is arduous for words originating from an L2 interlanguage.

Up to this point, it has not been examined whether these same accent-related ERP effects are observed when it is L2 listeners who are being tested. If L2 listeners veritably possess an interlanguage speech intelligibility benefit, nonnative-accented speech may not be indexed by the ERP effects associated to phonological processing difficulties. Specifically, L2 listeners may not generate a reduced P2 or an increased PMN for interlanguage input compared with standard native input. The current study thus contributes to the original—largely behavioral—interlanguage speech benefit literature by providing electrophysiological data.

Syntactic Processing

Most of the accented speech research examines prosody, phonetics, and phonology (e.g., Korpal & Sobkowiak, 2020; Porretta et al., 2017; Wang & van Heuven, 2015; Xie & Fowler, 2013; Braun et al., 2011; Lee & Xue, 2011; Weber et al., 2011; Floccia et al., 2009; Leikin et al., 2009; Smith et al., 2003, 2009; Hayes-Harb et al., 2008; Stibbard & Lee, 2006; Clarke & Garrett, 2004; Bent & Bradlow, 2003; van Wijngaarden et al., 2002; van Wijngaarden, 2001; Munro, 1998; Munro & Derwing, 1995, 1999; Thomas et al., under review). However, it is still unclear whether the interlanguage benefit may extend to high-level properties of L2 speech analysis, such as syntax and morphosyntax.

Akin to the level of phonology, learning the morphosyntax of an L2 can be challenging (e.g., Flege, Yeni-Komshian, & Liu, 1999) and L2 learners may thus preserve morphosyntactic parameters of their L1. This can create common syntactic interlanguage errors. It is possible that speaker–listener pairs who share a nonnative status may be very effective at managing these syntactic deviations (i.e., a syntactic interlanguage benefit).

Once again, ERPs represent a suitable technique to examine whether the interlanguage speech intelligibility benefit extends to the level of syntax: A classic ERP component, the P600, is mainly associated to syntactic processing. The P600 component is a positive inflection in the electrical activity of the brain. It occurs approximately

600 msec after a target stimulus and is generally distributed over posterior sensors. The P600 is most often associated to repair mechanisms provoked by grammatical errors (Kaan & Swaab, 2003; Osterhout & Holcomb, 1992). For instance, in native-accented speech, sentences containing a syntactic violation elicit an increased positivity compared with grammatical sentences (the P600 effect); this is true of both written and spoken stimuli (e.g., Hagoort & Brown, 2000).

A limited number of studies have examined L1 listeners' processing of syntactic errors produced in nonnative-accented speech (see also Gosselin, Martin, Navarra-Barindelli, & Caffarra, 2021; Sabo, 2021; Holt, Kung, & Demuth, 2018; Grey & van Hell, 2017; Hanulíková, Van Alphen, Van Goch, & Weber, 2012; etc., for the processing of nonnative-accented semantic errors). The results yielded by this line of research are relatively consistent: When grammatical errors are produced by nonnative speakers of a target language, L1 listeners exhibit a reduced (or absent) P600 effect as compared with errors produced by native speakers (Xu, Abdel Rahman, & Sommer, 2020; Caffarra & Martin, 2019; Hanulíková et al., 2012; see also Grey & van Hell, 2017, for similar results involving the Nref, an ERP component associated to pronoun resolution). Thus, it appears as though native listeners do not repair violations that are nonnative-accented. This may be a result of L1 listeners' reduced intelligibility of nonnative input; perhaps this group is simply not detecting (and therefore not repairing) the critical syntactic violations.

However, recent results obtained by Caffarra and Martin (2019) complexify the picture. These authors examined whether the "typicality" of certain syntactic violations in nonnative-accented speech may modulate the way in which they are processed. Native speakers of Spanish listened to sentences produced in a native Spanish accent or in English-accented Spanish. Some of these sentences contained errors that were typical for English learners of Spanish (i.e., gender errors); others contained errors that were atypical for English learners of Spanish (i.e., number errors; see the Materials section for further details). Crucially, Caffarra and Martin (2019) observed that the L1 listeners exhibited a reduced P600 for nonnative English-accented errors, but only when the error was typical in that accent (i.e., only for gender errors). In other words, in the English accent, L1 listeners only repaired atypical grammatical errors. These findings point toward the idea that native listeners' expectations about accented speech modulate their underlying processing of this input (see also Grey, Cosgrove, & van Hell, 2020; Fairchild & Papafragou, 2018; Bosker, Quené, Sanders, & De Jong, 2014); L1 listeners may adapt their syntactic processing according to their expectations and overlook typical errors. Though it is not clear whether this capability is functionally advantageous or detrimental (see the Discussion section), such results demonstrate that error typicality is a dimension that must be considered in research focusing on the syntactic interlanguage speech benefit.

To our knowledge, only one study examined the processing of nonnative-accented syntactic errors among L2 listeners: Grey, Schubel, McQueen, and Van Hell (2019) tested native listeners of Dutch (in the Netherlands) who were also highly proficient in English. These participants listened to grammatically correct and incorrect English sentences that were produced in an American English accent or in Chinese-accented English. Overall, the participants exhibited an Nref for the American English accent, but not the Chinese accent. The authors concluded that the participants only repaired the grammatical errors that were produced in a familiar accent (i.e., American English). It becomes difficult to interpret these results within the scope of the syntactic interlanguage benefit, as both accents in the reviewed study (including the native American English accent) were "foreign," and no comparison with native-accented speech was available.

Nonetheless, once we pair the original interlanguage speech benefit literature with the syntax-related findings from L1 listeners, it becomes possible to conceptualize the potential syntactic interlanguage benefit. A syntactic interlanguage benefit would entail that (1) L2 listeners are as precise (or even more precise) at detecting and repairing syntactic violations when they are produced in a nonnative accent compared with when they are produced in a native accent and/or (2) L2 listeners are better than L1 listeners when it comes to detecting and repairing nonnative-accented syntactic errors. To our knowledge, we are the first to examine whether either of these exemplifications of the syntactic interlanguage benefit take place.

This Study

The current EEG study utilizes the methodology of Caffarra and Martin (2019) and extends it to nonnative listeners. First, we examined L2 listeners' phonological and syntactic analysis of nonnative-accented speech. Next, the novel data from the present experiment and the data of native listeners from the precursor study (Caffarra & Martin, 2019) were statistically compared to determine whether L2 listeners display a classic phonological interlanguage benefit and/or a novel syntactic interlanguage benefit.

As such, L2 Spanish participants (whose L1 was English) listened to Spanish sentences produced in a native Spanish accent or in English-accented Spanish while their EEG was recorded. Analogously to the Caffarra and Martin (2019) study, some of these sentences contained errors in terms of grammatical gender agreement (**la color*, "the$_{f}$ color$_{m}$") and number agreement (**los color*, "the$_{pl}$ color$_{s}$"). Gender violations and number violations differ in frequency of attestation in English-accented Spanish; gender errors are quite typical for this group, whereas number errors are atypical (Sabourin, Stowe, & De Haan, 2006; Franceschina, 2001; Sabourin, 2001).

First, we ask the following general research question: How do L2 listeners process nonnative-accented sentences in real time? Regarding phonological processing, we expect to find evidence of the interlanguage speech intelligibility benefit. That is, we anticipate that L2 listeners might exhibit a facilitation effect during the perceptual analysis of nonnative-accented speech compared with native-accented speech. This hypothesis would be upheld by the following observations: (1) an increased P2 response for nonnative-accented sentences as compared with native-accented sentences and/or (2) a decreased PMN response for nonnative-accented sentences as compared with native-accented sentences. Both the P2 (Reinke et al., 2003) and PMN (Lewendon et al., 2021) typically display centrofrontal distributions; this is the pattern we also expect to observe.

Second, we ask the following more specific research question: Do L2 listeners repair all syntactic violations (regardless of error typicality) when these errors are produced in a nonnative accent? Crucially, we expect that the L2 listeners will display a syntactic interlanguage benefit. Unlike the L1 listeners from previous studies (Caffarra & Martin, 2019) who appeared to overlook typical errors when they were accented, we predict that our participants will be effective at detecting and repairing all accented errors (indexed by a typical posterior P600 effect), regardless of their typicality in English-accented Spanish. This prediction also stands for the syntactic violations produced in native-accented Spanish. If the L2 listeners' processing is purely impacted by their expectations (i.e., the distributional properties of the input), we should instead observe error typicality effects among nonnatives (i.e., only typical gender errors being repaired), as has been observed for native listeners (Caffarra & Martin, 2019).

METHODS

Participants

Thirty-nine L1 speakers of English, who were proficient L2 speakers of Spanish living in the Basque Country (M_{age} = 28.26, *SD* = 9.79; 21 women, 18 men), gave written informed consent and were paid 10 € per hour for their participation. The project was approved by the ethics committee from the Basque Center on Cognition, Brain and Language (Study #051817b). The participants completed three off-line Spanish proficiency tasks: a fill-in-the-blank task (90 Spanish questions), a gender assignment task (300 Spanish nouns), and a Spanish-to-English translation task (80 words). Nine participants were eventually excluded: One participant scored below 50% on the gender assignment task, and five individuals did not actively attend to the experiment (as shown in their accuracy on the online comprehension questions; see the Comprehension Questions and Offline Questionnaire section). Three additional participants were excluded due to noisy EEG data (i.e., more than 40% of trials rejected during preprocessing; see the EEG Recording and Analysis section).

The final participant sample included 30 right-handed L2 Spanish speakers between 19 and 60 years of age (M_{age} = 28.70, *SD* = 10.31; 16 women, 14 men). These individuals had no history of head trauma, psychiatric disorders, and language or hearing problems. Assuming a modest effect size of 0.25, a post hoc power analysis (G*Power 3.1; Heinrich Heine University Düsseldorf) established a statistical power of 96.7% for the sample size of n = 30. The participants' age of acquisition (AoA) of Spanish and their off-line proficiency task results are indicated in Table 1.

Materials

The stimuli consisted of Spanish sentences taken from Caffarra and Martin (2019). This set includes 60 experimental sentence frames with three versions each (180 experimental sentences total). The three versions varied according to a critical determiner phrase (see Table 2). In the first condition, determiner phrases were syntactically correct (e.g., *el color*, "$\text{the}_{\text{M/S}}$ $\text{color}_{\text{M/S}}$"). In the second condition, the sentence contained a gender violation between the critical determiner and target noun (e.g., **la color*, "the_{F} color_{M}"). Finally, in the last condition, the critical determiner phrase carried a number violation, wherein the target noun lacked the plural inflection required by its determiner (e.g., **los color*, "the_{PL} color_{S}"). Note that, similar to number errors, gender errors could not be processed as violations until the end of the critical determiner phrase, because an opposite-gendered synonymic noun was also possible (e.g., ***la color** *del cuadro* "***the**$_{\text{F}}$ **color**$_{\text{M}}$ of the painting" vs. ***la coloración*** *del cuadro* "**the**$_{\text{F}}$ **coloring**$_{\text{F}}$ of the painting"). In other words,

Table 1. Language Background Details for Participant Sample (n = 30)

	Spanish AoA (years)	*Spanish Fill-in-the Blank Accuracy (%)*	*Spanish Gender Assignment Accuracy (%)*	*Spanish-to-English Translation Accuracy (%)*
Mean	15.54	83.33%	84.12%	97.08%
SD	7.46	15.08%	10.28%	5.25%

Table 2. Example of the Experimental Conditions

Condition	*Sentence*
Correct	*Me gusta mucho* ***el***$_{M/S}$ ***color***$_{M/S}$ *del cuadro.*
Gender violation (typical error)	*Me gusta mucho* ****la***$_{F/S}$ ***color***$_{M/S}$ *del cuadro.*
Number violation (atypical error)	*Me gusta mucho* ****los***$_{M/PL}$ ***color***$_{M/S}$ *del cuadro.*

The critical determiner phrases are **bolded**. The English translation of the correct sentence is "I love **the color** of the painting." M = masculine; F = feminine; S = singular; PL = plural.

the identification point for each sentence's grammatical correctness was equivalent across all conditions.

All of the experimental sentences were semantically low constraining; that is, the critical determiner phrase could not be predicted according to the preceding content of the sentence. Furthermore, half of the critical determiner phrases contained a feminine noun, and half contained a masculine noun. None possessed transparent gender suffixes or semantic gender (e.g., *amigo/amiga,* "friend$_M$/friend$_F$"), so superficial processing strategies were not possible on the participants' behalf. The lexical properties of the 60 target nouns can be found in Caffarra and Martin (2019). An additional 160 syntactically correct sentences (e.g., *La historia tuvo un final feliz*. "The story had a happy ending.") were included as fillers. The full stimulus set was thus composed of 340 sentences (65% correct, 17.5% gender violations, and 17.5% number violations).

As previously discussed, gender violations and number violations differ in frequency of attestation in English-accented Spanish. In particular, gender errors are much more common than number errors in English-accented Spanish[1] (Sabourin et al., 2006; Franceschina, 2001; Sabourin, 2001; White, Valenzuela, Kozlowska-Macgregor, Leung, & Ayed, 2001); omission number errors (i.e., errors in which the nominal plural morpheme is missing, like those in the current study) are particularly uncommon in nonnative Spanish speech production (Franceschina, 2001). For these reasons, in this study, gender errors are considered the typical or expected errors, and number errors are designated as atypical errors.

The 340 sentences were recorded by three male native Spanish speakers and by three male native speakers of British English, fluent in L2 Spanish, in an order counterbalanced for syntactic correctness. In each accent, individual speakers produced one third of the sentences (approximately 115 sentences each). The English-accented speakers recorded the stimuli first. Thus, for every individual sentence, the native Spanish speakers were instructed to listen to the English-accented version of the given recording and to match its prosody and speech rate in their own production. Despite this directive, stimuli produced by native speakers were significantly shorter than those produced by English-accented speakers (see Table 3), as has been reported in similar past accented speech research (Grey & van Hell, 2017; Romero-Rivas et al., 2015; Goslin et al., 2012; Hanulíková et al., 2012). However, there were no durational differences between syntactically correct sentences and sentences containing gender or number violations in the English accent or the native accent ($ps > .17$).

As detailed in Caffarra and Martin (2019), the recordings were normed by 60 native speakers of Spanish (M_{age} = 24.0 years, SD = 4.5 years, 35 women, 25 men; see Table 3). These informants were asked to assess the strength of the accent in the recordings (from 1 = *very weak* to 5 = *very strong*). They also provided intelligibility transcriptions of the penultimate word in each sentence and indicated whether they detected the gender and number violations in the incorrect sentences. In brief, the norming procedure confirmed that the English speakers

Table 3. Norming Information of the Recordings from Caffarra and Martin (2019)

	Native Accent	*English Accent*	*p*
Full sentence duration (msec)	2198.68 (533.12)	2476.81 (600.58)	<.001
Target noun duration (msec)	390.03 (109.58)	422.77 (117.13)	<.001
Accent strength rating (1–5)	1.27 (.11)	3.89 (.55)	<.001
Transcription accuracy (%)	98.36 (2.89)	98.34 (3.25)	.95
Error detection accuracy (%)	96.79 (4.88)	93.30 (7.40)	<.001
Prediction accuracy (%)	47.04 (13.93)	36.33 (6.95)	<.001

p Values indicate differences between accents. Standard deviations are in parentheses.

had a strong, detectable nonnative accent, but that they were as intelligible as the Spanish speakers. Though the informants were highly accurate in detecting syntactic errors overall, native-accented errors were identified at a higher rate than English-accented errors (see Gosselin et al., 2021; Grey & van Hell, 2017; Hanulíková et al., 2012, for similar results).

Recall that the speakers recorded both correct and incorrect versions of the sentences in real time. In line with similar past ERP research (Gosselin et al., 2021; Grey et al., 2019, 2020; Xu et al., 2020; Caffarra & Martin, 2019; Grey & van Hell, 2017; Hanulíková et al., 2012), incorrect critical determiner phrases were not cross-spliced onto grammatically correct sentences to prioritize naturalistic, co-articulated speech. To ensure that the speakers had not unconsciously introduced prosodic "markers" of an upcoming error in the stimulus (e.g., differences in speech rate, pitch, pauses), each sentence was trimmed immediately before the target noun and presented to 30 additional native speakers of Spanish (M_{age} = 23.8 years, SD = 4.7, 20 women, 10 men). These informants listened to the stimulus fragments and were required to predict whether they believed (yes or no) that a speech error was going to occur in the remaining sentence. The informants tended to predict that an error was forthcoming in English-accented fragments (M = 46.37%, SD = 17.59%) much more often than in native Spanish-accented fragments (M = 7.88%, SD = 8.95%), $F(1, 179) = 818.45$, $p < .001$, regardless of the syntactic correctness of the complete sentence (English accent: $ps > .090$; native accent: $ps > .53$). This suggests that listeners implicitly expect an influx of errors when speakers are nonnative (see also Gosselin et al., 2021). However, prediction accuracy (i.e., accurately identifying a correct sentence as errorless or an incorrect sentence as containing an error before encountering the critical noun phrase) was below chance for both accents (see Table 3), indicating that the recordings did not give reliable prosodic cues for the eventual syntactic correctness of the sentence.

The 180 experimental sentences were divided into six lists, such that each participant listened to only one version of each sentence frame (i.e., target nouns only appeared once per list), with half produced in the native Spanish accent and half produced in the nonnative English accent. Each participant also listened to all 160 correct filler sentences (also split evenly according to accent). Note that none of the stimulus-related external factors (target duration, accent strength, intelligibility, error detection, or error prediction) interacted with the experimental variables of correctness and/or accent (all Fs < 2.50, all $ps > .050$). As such, analyses yielding significant ERP effects are fully interpretable in terms of the target experimental conditions (i.e., accent, correctness) and are not confounded by potential differences in the stimulus properties.

Procedure

After giving informed consent, the participants were fitted with the EEG cap (see EEG Recording and Analysis section) and then brought into a sound-attenuated chamber equipped with a desktop computer and speakers. The experiment was run through Version 14.4 of Presentation. Participants started by listening to brief introductions from the native and nonnative speakers who produced the recordings; each speaker stated their name and their city and country of origin. This was accomplished so that, from the onset of the experiment, the participants would be aware of the nonnative status of some of the speakers.

The experiment was divided into four blocks of 85 trials (following 12 practice sentences). Each individual trial (as depicted in Figure 1) began with a screen containing

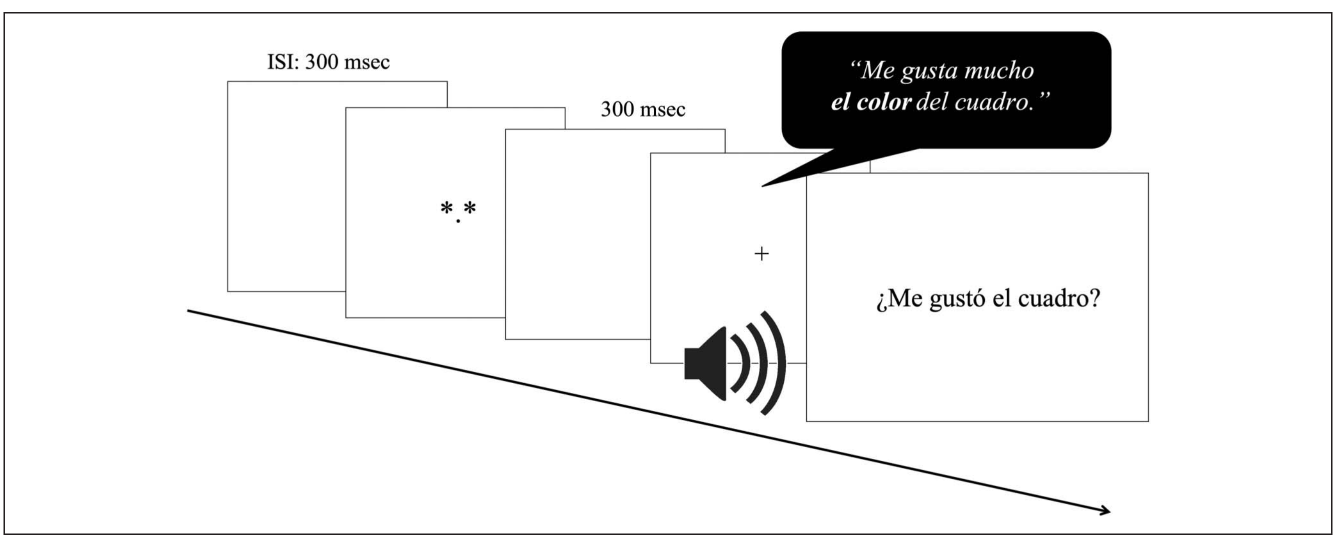

Figure 1. Graphic depiction of the task; a single trial including a comprehension question is shown (Translation: "I love the color of the painting."; "Did I like the painting?").

symbols that resembled a blinking face (*.*). Participants were told that they were free to move their eyes when they saw this screen but that they should limit their ocular movements as much as possible during the other parts of the experiment. The *.* display was participant-controlled. A blank screen was subsequently presented on the computer monitor for 300 msec, followed by a central fixation cross. The auditory sentence played through external speakers ($M_{duration}$ = 2337.74 msec, SD = 584.40) while the fixation cross was displayed. A yes/no comprehension question followed 20% of sentences (equally distributed across experimental conditions; half required a "yes" answer). At the end of the auditory stimuli (or if the trial had a comprehension question, once a keyboard response was selected), the next trial began after an ISI of 300 msec. All participants completed the experiment in its entirety; it lasted approximately 30 min.

Following the EEG portion of the study, participants completed the Spanish knowledge tests (a Spanish proficiency task that consisted of fill-in-the-blank exercises, a gender assignment task for Spanish nouns, and a Spanish-to-English translation task for the critical nouns included in the experiment). They also filled out an off-line questionnaire in which they were asked about their familiarity and exposure to accented speech, their impressions about the occurrence of gender and number errors in English-accented Spanish, as well as their feelings of trust, reliability and perceptibility of the speakers from the recordings.[2]

EEG Recording and Analysis

The participants' EEG was recorded via a 27-channel EasyCap (Brain Products); the electrodes were distributed according to the standard 10–20 configuration (see Figure 2). External electrodes were also applied above and below the participants' right eye, as well as to both their temples and mastoids. During the system setup, experimenters ensured that the facial electrodes maintained impedances below 10 kΩ; electrodes on the scalp were kept below 5 kΩ. The data were recorded at a sampling rate of 500 Hz with a left-mastoid online reference and amplified with a BrainAmp DC amplifier (Brain Products, GmbH). During off-line data processing, the data were re-referenced to the average of both mastoids; a low-pass filter of 20 Hz and a high-pass filter of 0.01 Hz were applied to the data.

Filtered EEG peaks exceeding ±70 μV were automatically rejected. Components accounting for the highest variance of vertical and horizontal eye movements were isolated and corrected via the independent components analysis in the Brain Vision Analyzer software (see Gosselin et al., 2021; Caffarra & Martin, 2019, for similar methods). Three of the participants originally tested were excluded due to excessive data loss (over 40% of trials rejected). Within the final participant sample, 13.80% of trials were discarded, with no differences across conditions (ps > .35). ERPs were time-locked to the acoustic onset of the target noun in the critical determiner phrase (e.g., *color*). Epochs were established according to a 200-msec prestimulus corrected baseline; they extended to 1500 msec postonset of the target noun.

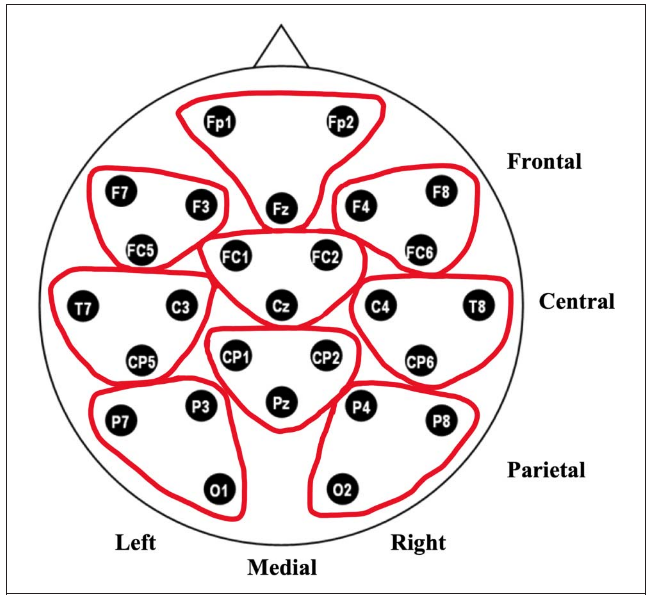

Figure 2. Electrode distribution. The ROIs according to the topographical factors of laterality and longitude are indicated.

Using Version 0.13.1 of the JASP software (University of Amsterdam), a 2 × 3 × 3 × 3 repeated-measures ANOVA was conducted on condition-averaged ERPs for each participant. The ANOVA included the experimental factors of Accent (native Spanish, nonnative English) and Correctness (correct, gender error, number error). The electrodes were separated into nine ROIs with three electrodes each (see Figure 2). The topographic factors of Laterality (left, medial, right) and Longitude (frontal, central, parietal) were also included in the ANOVA.

A conservative (cp = 16) point-by-point split-plot ANOVA (Guthrie & Buchwald, 1991) was computed across the entire epoch (see Figure 3) to determine the time windows of interest for the accent and correctness effects (for similar identification methods, see Molinaro, Giannelli, Caffarra, & Martin, 2017). Two time windows were thus identified: 500–750 msec and 900–1200 msec. Separate ANOVAs were performed on these time windows. In the event of the sphericity violation assumption, the Greenhouse–Geisser correction is reported. For planned post hoc tests, p values are Holm-adjusted to avoid familywise errors (see Eichstaedt, Kovatch, & Maroof, 2013).

Bayesian repeated-measures ANOVAs were also performed in JASP to support the validity of potential null effects or interactions. The structure of these Bayesian ANOVAs follows the design of the frequentist analyses, as detailed above. We include information on the Bayesian analysis of effects across matched models and specifically

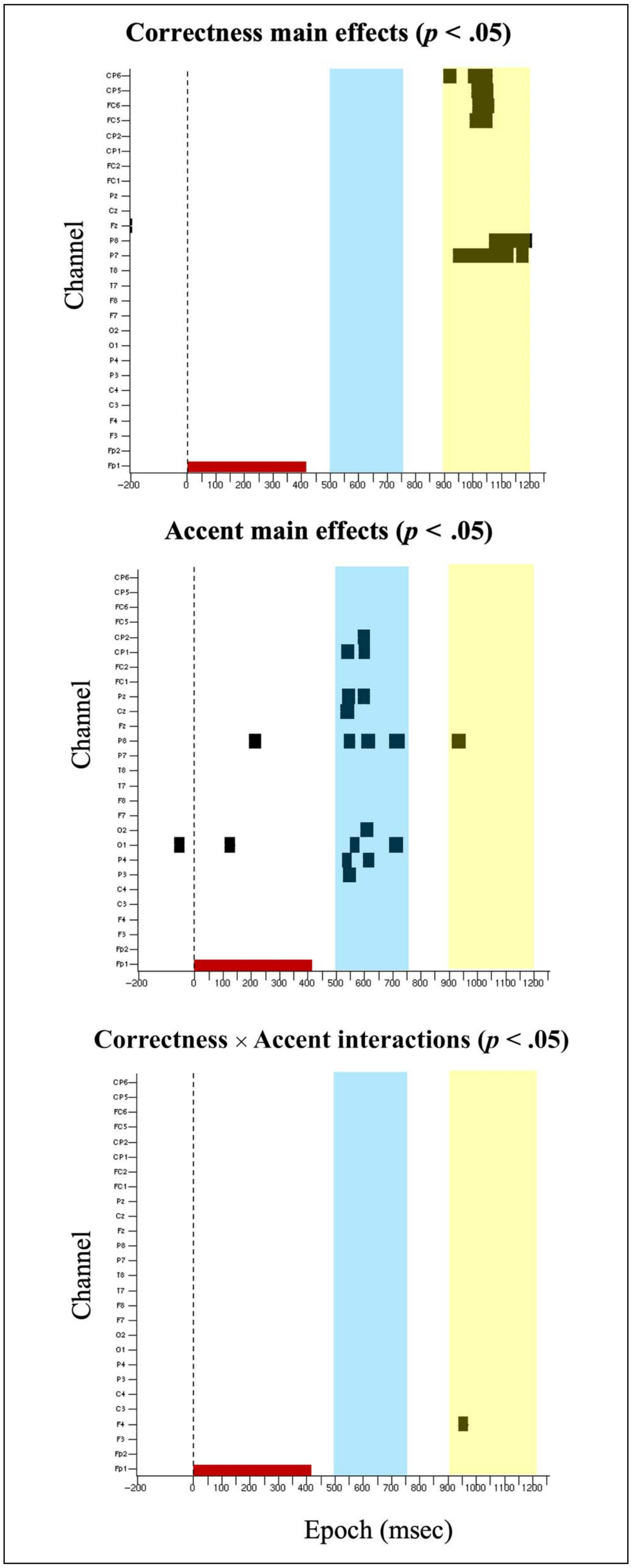

Figure 3. Channel-specific point-by-point split plot ANOVA for the experimental factors; p values are corrected with the Guthrie and Buchwald (1991) adjustment (cp = 16). The dashed line indicates the onset of the target noun; the red interval along the x-axis reflects the approximate length of the target noun (M = 407 msec). The identified time windows (500–750 and 900–1200 msec) are highlighted in blue and yellow, respectively.

report the Bayes inclusion factor (BF_{incl}), which is interpreted as "the evidence in the data for including a predictor" (van den Bergh et al., 2020, p. 11).

Finally, exploratory Pearson correlation matrices were conducted between the elicited ERPs and the participant responses to the off-line questionnaire and their scores on the Spanish knowledge tasks. The dependent and independent variables entered into exploratory correlation matrices were motivated by the findings of main ANOVAs; we therefore discuss their design in the Results section.

RESULTS

Comprehension Questions and Offline Questionnaire

As previously discussed, five participants were found not to have actively attended to the auditory stimuli during the actual experiment. Because their accuracy to the yes/no comprehension questions was below 60%, these individuals were excluded from subsequent analyses. The final participant sample of 30 listeners achieved an average accuracy of 84% (SD = 8.2%). There were no differences in accuracy for the native-accented and English-accented trials, $F(1, 29) = 0.13$, $p = .72$, suggesting that the recordings were similarly understandable for both accents (and thus confirming the off-line pretest of intelligibility; see the Materials section).

Indeed, when asked to assess the perceptibility of the speakers at the end of the study (i.e., "From 1 to 10, how easy was it to identify word-by-word what the English/Spanish native speakers said?"), the participants gave similar ratings to both the native Spanish speakers (M = 4.96, SD = 2.95) and the English-accented speakers (M = 5.18, SD = 2.13), $t(27) = 0.49$, $p = .63$. This finding points toward an interlanguage speech benefit for our L2 listeners; it contrasts similar past research conducted on native participants, wherein accented speakers are typically rated as less comprehensible than native speakers on postexperimental questionnaires (Jiang, Gossack-Keenan, & Pell, 2020; Caffarra & Martin, 2019; Grey et al., 2019; Grey & van Hell, 2017). Interestingly, however, the L2 listeners in this study rated the English-accented speakers as significantly less reliable, $t(27) = 6.65$, $p < .001$, and trustworthy, $t(27) = 3.29$, $p = .003$, than the native Spanish speakers.[3]

To determine whether the participants had the predicted typicality expectations for gender and number violations, they were asked to rate the frequency at which they heard both of these errors in English-accented Spanish in general. Gender errors were indeed encountered far more often than number errors in this accent, $t(27) = 6.36$, $p < .001$, though participants reported that they were more likely to "let pass" speech errors (M = 5.52, SD = 2.77 out of 10, where 0 = *never* and 10 = *always*) than to correct them (M = 2.64, SD = 2.03) in their daily life.

Finally, as exposure to accents or familiarity with accented speech may impact its processing (Caffarra & Martin, 2019; Holt et al., 2018; Grey & van Hell, 2017; Porretta et al., 2017), the postexperimental questionnaire also assessed the participants' experience with this type of input. All but one individual reported having regular contact with English-accented Spanish (M = 13.42 hr/week, SD = 21.02). They also indicated being friends with a large proportion of non-Spanish individuals (M = 59.19%, SD = 29.53%). With this information in mind, it is evident that the current participant sample possessed a high degree of familiarity with nonnative-accented speech.

Main ERP Results (See Figure 4)

First Time Window (500–750 msec)

In the first time window of interest, syntactically correct and incorrect sentences elicited similar brain responses (main effect of Correctness: $F(2, 58) = 0.14$; $p = .87$; $BF_{\text{incl}} = .02$; with topographic factors Fs < 1.10, ps > .39). Though Accent did not yield a main effect, $F(1, 29) = 2.34$; $p = .14$, $BF_{\text{incl}} = 92.26$, an interaction between Accent and Longitude was observed, $F(2, 58) = 5.53$, $p = .015$. Simple main effects in each level of the longitudinal factor indicated that trials in the native accent elicited more negative responses than trials in the nonnative accent across parietal sensors, $F(1, 29) = 5.88$, $p = .022$ (see Figure 5), but not frontal or central sensors (ps > .13). There were no interactions between Accent and Correctness, $F(2, 58) = 0.19$, $p = .79$, $BF_{\text{incl}} = .05$ (with topographic factors, Fs < 1, ps > .46).

Second Time Window (900–1200 msec)

In the second time window of interest, a main effect of Correctness was observed, $F(2, 58) = 3.76$, $p = .026$, $BF_{\text{incl}} = 1.75^{e+7}$ (for raw condition averages, see the Appendix; for critical effects, see Figures 4 and 6). Post hoc comparisons indicated that gender errors, $t(65) = 2.46$, $p_{\text{holm}} = .046$, and number errors, $t(65) = 2.28$, $p_{\text{holm}} = .048$, elicited significantly more positive responses than correct sentences, but that there were no differences between the two types of violations, $t(65) = 0.18$, $p_{\text{holm}} = .86$. This effect was significant across central, $F(2, 58) = 3.56$, $p = .031$, and parietal, $F(2, 58) = 6.39$, $p = .002$, sensors (Correctness × Longitude: $F(4, 260) = 4.68$, $p = .007$). The ERP responses were not modulated according to the accent of the speaker (main effect of Accent: $F(1, 65) = 0.37$, $p = .55$, $BF_{\text{incl}} = .10$, with topographic factors Fs < 1.80, ps > .13) and there were no interactions between Accent and Correctness, $F(2, 58) = 0.63$, $p = .53$, $BF_{\text{incl}} = .24$ (with topographic factors Fs < 1, ps > .40). Note that the Bayesian statistics for the main effect of Accent and the Accent × Correctness interaction are below .25, indicating that the data are substantially more likely under the null hypothesis than the respective alternative hypotheses for these effects of interest.

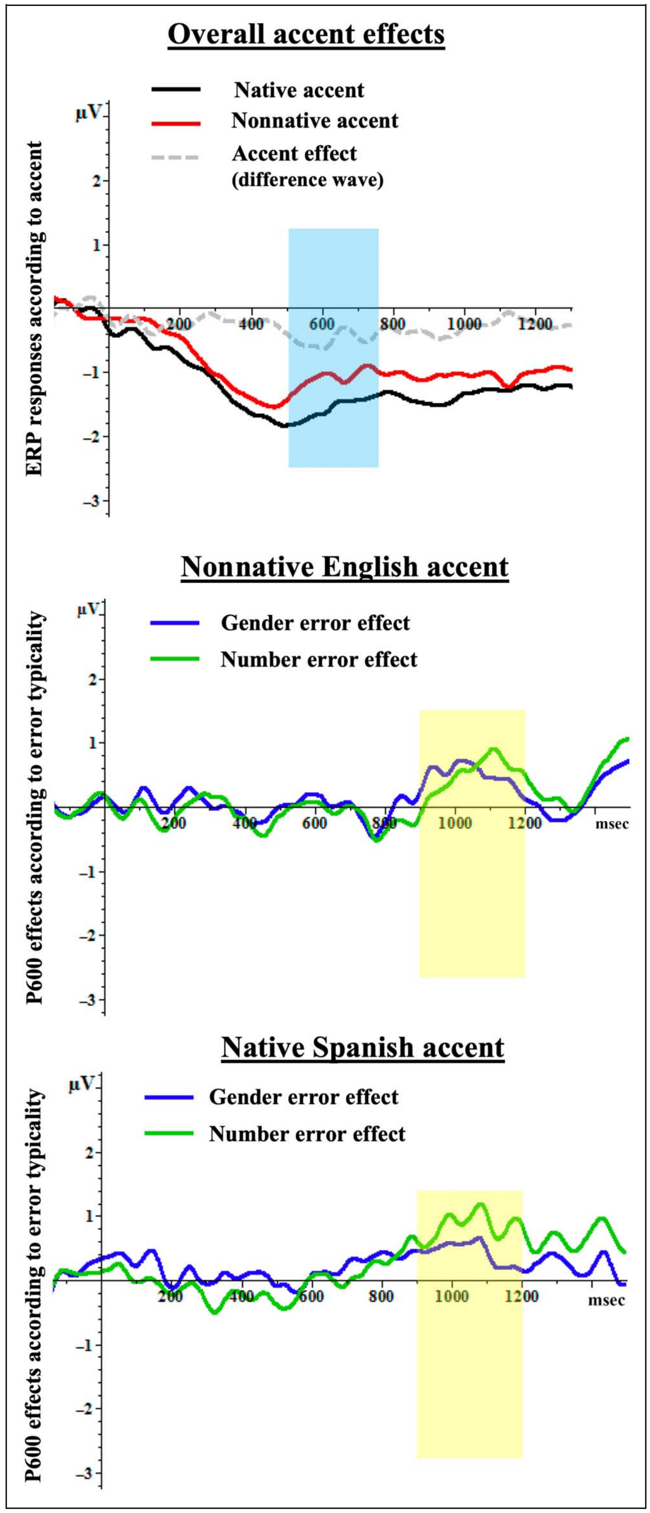

Figure 4. Top: Accent conditions and accent effect (ERPs from the three parietal ROIs are averaged across levels of correctness). The critical time window for this effect (500–750 msec) is highlighted in blue. Middle, bottom: Gender and number error effects within both accents (ERP responses are averaged across the three parietal ROIs). The critical time window for this effect (900–1200 msec) is highlighted in yellow. The ERP waves for all three correctness conditions in each accent are included in the Appendix.

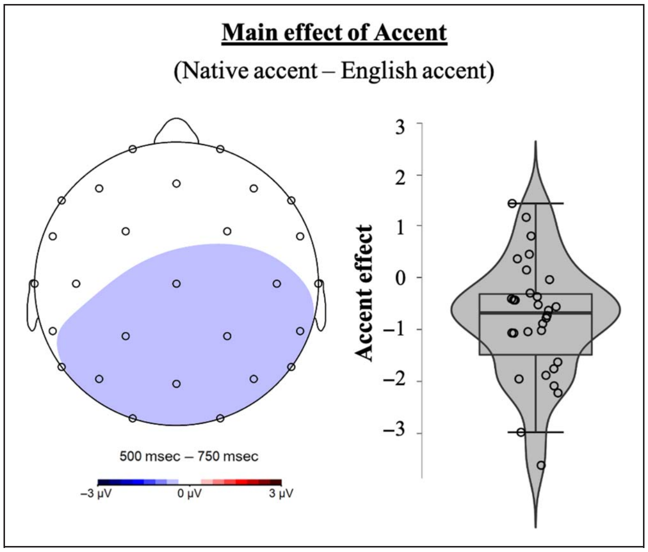

Figure 5. 500–750 msec. Left: The topographic distribution of the negativity for the main effect of accent (obtained by subtracting the average of all English-accented trials from the average of all native-accented trials). Right: Variation in the accent effect across parietal sensors.

Summary of Main Results

From 500 to 750 msec, native-accented trials were indexed by overall more negative responses over parietal ROIs. An increased positivity was subsequently observed (900–1200 msec) for syntactic violations, in both native-accented and nonnative-accented sentences. Given the polarity, timing, distribution, and type of violation examined, this increased positivity is interpreted as a P600. The relative typicality of the violations did not modulate the late P600: Gender and number errors generated similar effects in both accents. Furthermore, there were no overall differences in the magnitude of the P600 effect between the native- and nonnative-accented trials.

Main Correlation Results

Exploratory two-tailed Pearson *r* correlations were conducted on the significant effects found in the main ANOVA results.

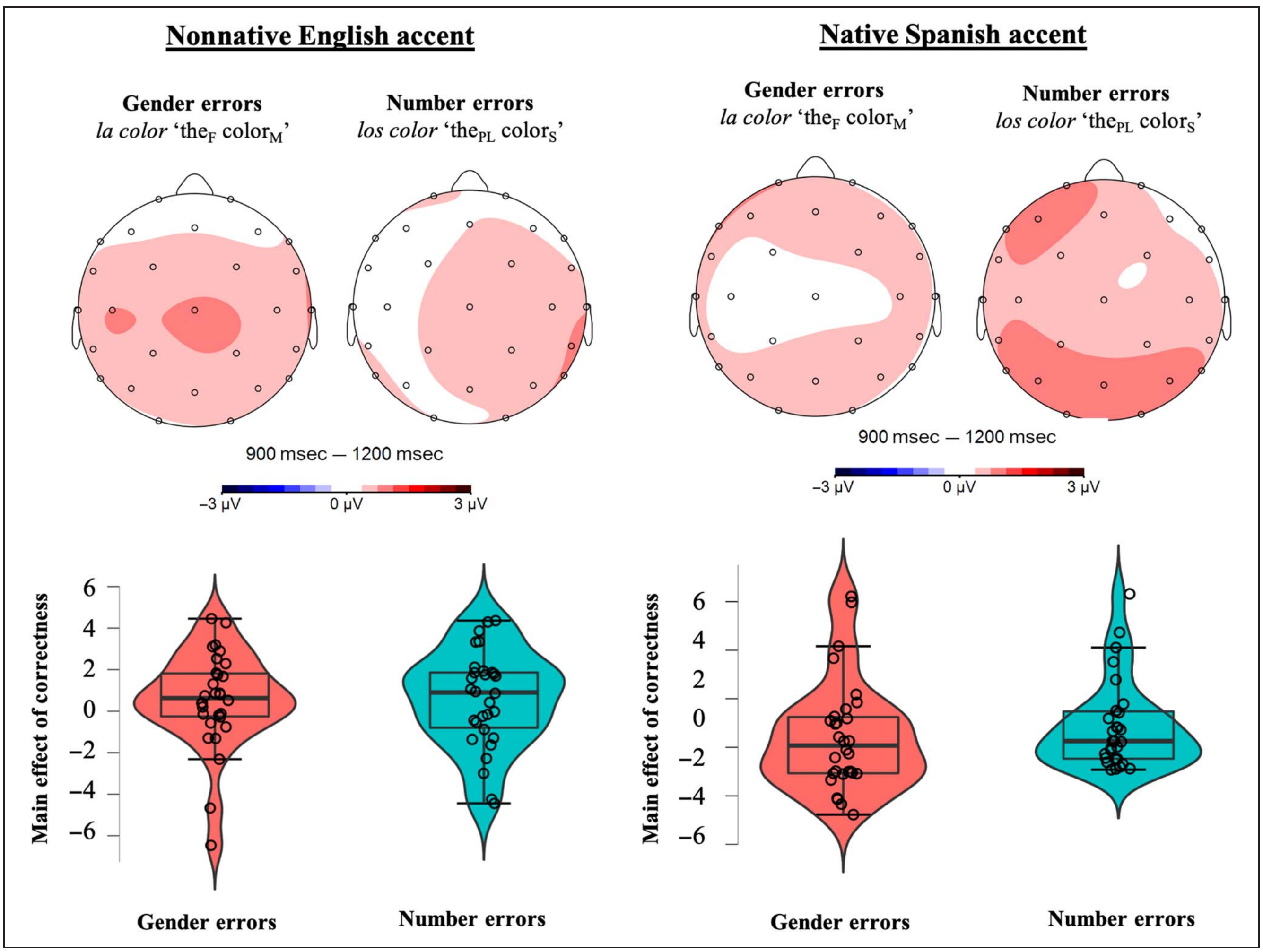

Figure 6. Top: The topography of the increased positivity for gender errors and number errors in both accents (obtained by subtracting the correct condition from both incorrect conditions individually in the 900–1200 msec time window). Bottom: Individual variation in the gender and number error effects across parietal sensors.

500–750 msec

Recall that the main ERP results yielded a main effect of accent from 500–750 msec, wherein native-accented trials were indexed by more negative responses (over parietal sensors). We thus subtracted the average parietal response to nonnative-accented trials (collapsed across correct sentences, gender errors, and number errors) from the average parietal response to native-accented trials (also collapsed across levels of correctness). We refer to this difference as the overall accent effect (see Figure 4, top); correlations with the participants' behavioral and demographic measures were computed.

The accent effect was not significantly correlated to the participants' score on the Spanish knowledge tests; their current age or Spanish AoA; their ratings of trust, reliability, or perceptibility to the recorded speakers from the experiment; their exposure to accented speech; nor their propensity to correct or let pass speech errors (ps ≥ .09).

900–1200 msec

As the main ERP results yielded an overall increased positivity in the final time window, with no differences across accents and error type (gender vs. number error), the effect was collapsed across conditions and correlated to the participants' behavioral and demographic measures. We refer to this collapsed P600 as the overall error effect.

The magnitude of the error effect was not correlated to the participants' score on the Spanish knowledge tests, nor to their current age or Spanish AoA ($p \geq .10$). Similarly, there were no relations between the error effect and the participants' ratings of trust, reliability, or perceptibility to the recorded speakers from the experiment (ps > .18). There were no effects of accent familiarity (rs < .10, ps > .62), perhaps because nearly all of the L2 listeners were quite experienced with English-accented Spanish. However, the overall error effect was found to be highly correlated ($r = .51, p = .006$) to the participants' propensity to correct speech errors (i.e., How often do you correct Spanish mistakes made by foreigners? 1 = *never*, 10 = *always*). Individuals who reported correcting errors more frequently exhibited larger (more positive) error effects (Figure 7). Note that further follow-up analyses indicated that this was true for both gender error effects ($r = .44, p = .018$) and number error effects ($r = .49, p = .008$) independently.

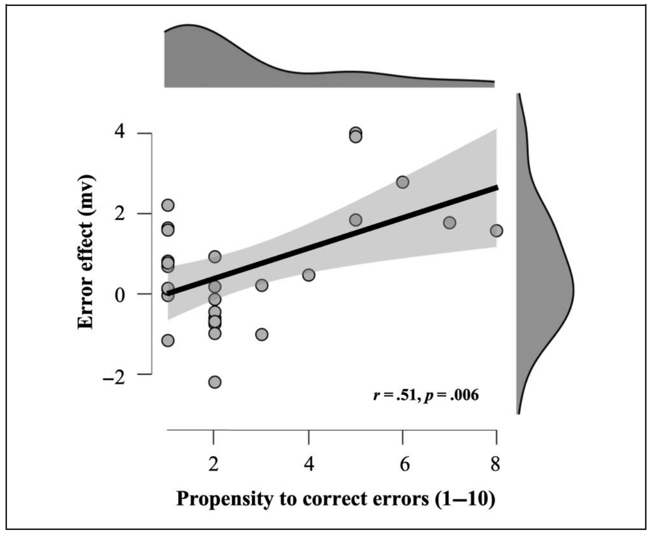

Figure 7. Significant correlation between the overall P600 error effect (obtained by averaging the correctness effect for both error types and accents) and the participants' self-rated propensity to correct speech errors. Shaded areas surrounding the regression line represent the 95% confidence interval. The distribution density of each axis is indicated above and on the right of the scatter plot.

Post Hoc EEG Group Comparisons

The main analyses indicated that there were no overall differences in the magnitude of the P600 effect according to accent. Given that the L2 listeners exhibited similar responses to syntactic violation produced in native Spanish accent and nonnative English accent, these results do not support a syntactic interlanguage benefit for talkers. Nonetheless, the main results also show that L2 listeners are not sensitive to nonnative error typicality: They repaired both gender and number errors. This contrasts previous results conducted on L1 listeners, wherein this group overlooked (i.e., did not repair) typical grammatical errors when they were accented (Caffarra & Martin, 2019). As such, there is reason to believe that the L2 listeners from the current study exhibited a syntactic interlanguage benefit for listeners.

As the materials and design from the current study were identical to those used by Caffarra and Martin (2019), we opted to compare the groups (L1 and L2 Spanish listeners) to verify whether there was statistical (rather than purely descriptive) evidence of a syntactic interlanguage benefit for listeners. That is, we examined post hoc whether the processing differences between the L2 participants in the current sample ($n = 30$) and the L1 participants ($n = 36$) from the precursor study (Caffarra & Martin, 2019) were statistically significant, particularly in terms of their sensitivity to error typicality in the nonnative accent.

As such, the ERP data from Caffarra and Martin (2019) were exported based on the significant spatial and temporal clusters reported in this article. One-tailed t tests and Pearson correlations were computed, as detailed in the subsequent sections. These statistical tests were one-tailed, as the expected direction of the group differences were known in advance (based on the general findings from both the main results of this study and the precursor study).

500–750 msec (Figure 8)

The overall accent effect in the first time window of interest (i.e., the difference between native-accented and

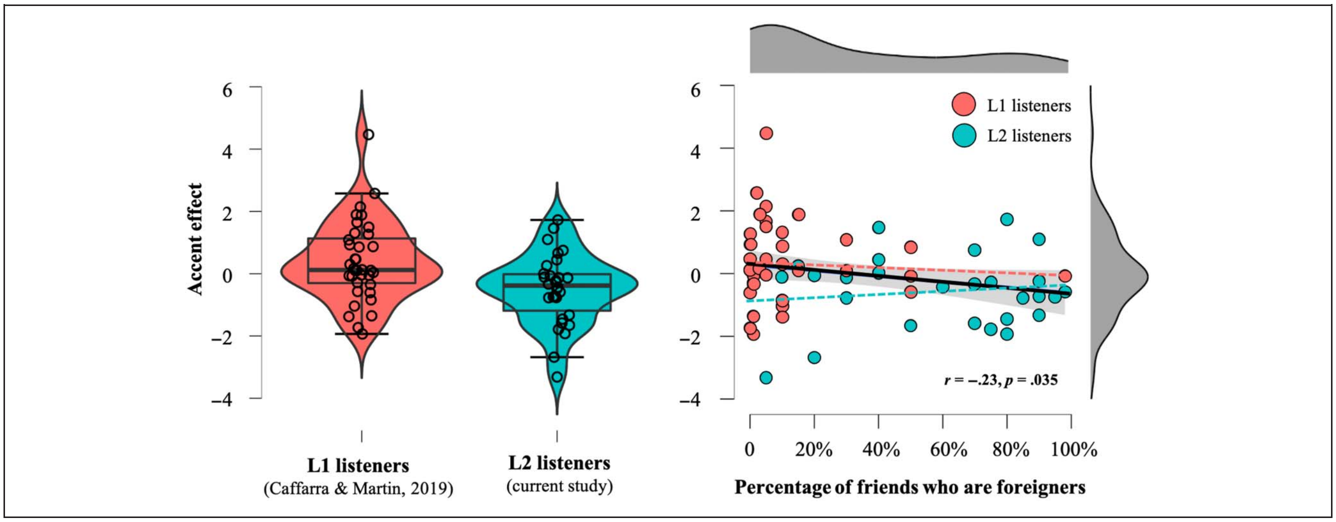

Figure 8. Left: Individual variation in the overall accent effect (native accent – nonnative accent) for both groups of listeners; L2 listeners' accent effect differs significantly from zero. Right: Significant correlation between the accent effect and the proportion of foreign friends (black line). The shaded area represents the 95% confidence interval surrounding the overall regression line. The regression of each individual group is also depicted (dotted blue and orange lines). The distribution density of each axis is indicated above and on the right of the scatter plot.

English-accented trials) was compared across the two groups of listeners. Note that an accent effect near zero indicates that both accents are processed similarly; a more negative accent effect reveals that the native accent produced more negative brain responses than the nonnative accent.

Independent-samples t tests yielded important group-based differences in the accent effect, $t(64) = 2.95$, $p = .004$. In fact, these groups showed directionally opposite effects: Whereas the L2 listeners exhibited a significant negative accent effect (i.e., a more negative response to the native Spanish accent than the nonnative English accent; $M = -0.50$, $SD = 1.14$), L1 listeners displayed a marginal positive accent effect (i.e., a more positive response to the native Spanish accent than the nonnative English accent; $M = .39$, $SD = 1.29$). Note that when gender and number error effects in each accent were compared individually across the groups, no significant differences were observed (ps > .060). These findings suggest that, in the 500–750 msec time window, L1 and L2 listeners show a different sensitivity for indexical (i. e., phonological) cues such as accent, but that they are similar in the way they treat grammatical information.

The overall accent effect for L1 and L2 listeners was correlated to the questionnaire responses available for both groups. No relations were observed between the accent effect and any of the participants' responses to the accent experience questionnaire (ps > .18). However, the participants' AoA of Spanish was negatively correlated to the accent effect ($r = -.36$, $p = .002$); this analysis essentially replicated the independent-samples t test, as the L2 listeners have more latent Spanish AoAs than the L1 listeners. Furthermore, a significant negative correlation involved accent exposure ($r = -.23$, $p = .035$): Participants who reported being friends with a higher proportion of foreigners (i.e., fewer native speakers of Spanish) exhibited a more negative accent effect. This effect was not solely carried by either of the groups independently (L1 listeners: $r = -.07$, $p = .69$; L2 listeners: $r = .11$, $p = .58$).

900–1200 msec (Figure 9)

An error typicality effect was obtained by subtracting the P600 for gender errors from the P600 for number errors. An error typicality effect close to zero indicates that gender errors and number errors are processed similarly (i.e., in this context, that all errors are repaired). A larger (more positive) error typicality effect indicates that there are processing differences according to the typicality of the syntactic error (i.e., in this context, that atypical number errors are repaired and typical gender errors are not).

The L2 listeners' error typicality effect did not differ significantly from zero (native accent: $t(35) = 1.20$, $p = .12$; nonnative: $t(35) = 0.07$, $p = .47$). In contrast, the L1 listeners showed a significant error typicality effect for the nonnative accent, $t(35) = 1.86$, $p = .036$, but not for the native accent, $t(35) = 0.54$, $p = .70$. Note, however, that the direct between-group comparisons did not yield significant differences between the L1 listeners and L2 listeners in terms of error typicality effects, $t(64) = 1.29$, $p = .10$. This may suggest that there is a smooth transition in error sensitivity from one group to the other, rather than an abrupt shift in tolerance to typical errors.

As such, the gender error and number error effects for English-accented trials were correlated to the demographic and behavioral measures available for both groups of listeners. A significant relation was found between the P600 for gender errors and the participants' proportion of

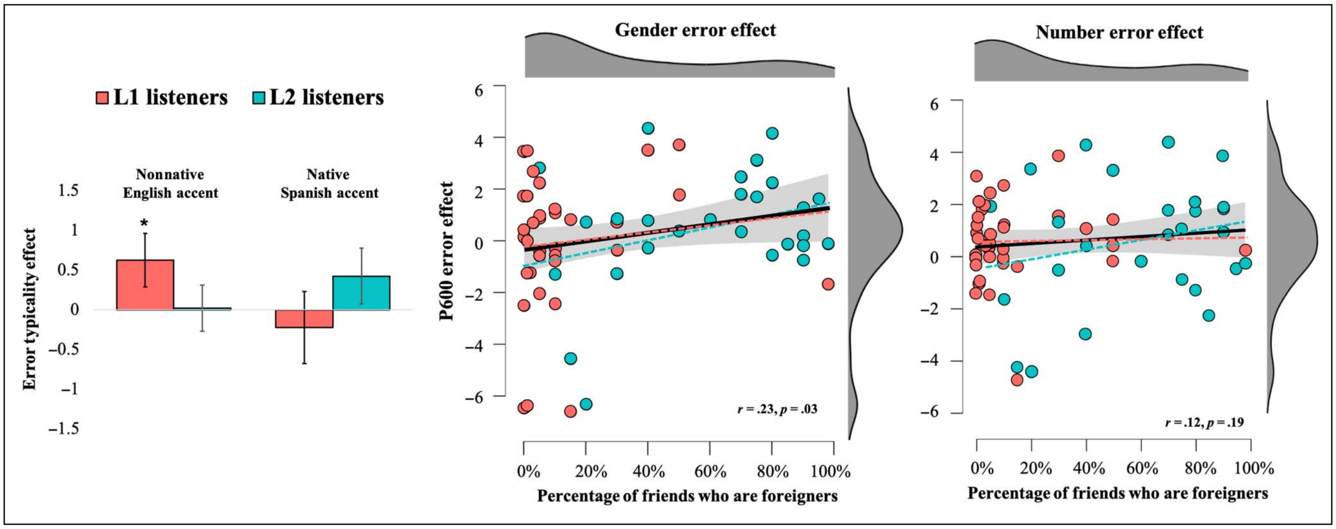

Figure 9. Left: The error typicality effect (gender error effect – number error effect) for both groups of listeners. Error bars represent 1 *SE*. Center, right: Correlation between the P600 for gender errors and the P600 for number errors, and the participants' proportion of foreign friends (black line). The shaded area represents the 95% confidence interval surrounding the overall regression line. The regression of each individual group is also depicted (dotted blue and orange lines). The distribution density of each axis is indicated above and on the right of the scatter plot.

non-Spanish friends ($r = .23, p = .034$); this same factor was not significantly related to the P600 for the number errors ($r = .12, p = .19$). That is, individuals who reported being friends with more foreigners exhibited the largest P600 effects for the typical gender errors. A visual inspection of the scatter plots (see Figure 9) and a Shapiro–Wilk normality test ($W(66) = 0.84, p < .001$, skewness = 0.68, kurtosis = −1.08) suggested that this correlation may have been disproportionately carried by the group of L2 listeners. Follow-up analyses indeed revealed that the correlation between the gender error effect and the proportion of foreign friends was nearly significant for L2 listeners ($r = .32, p = .052$), but not for L1 listeners ($r = .11, p = .26$). This suggests that the observed between-group correlation does not veritably represent a gradual shift in error typicality across listeners, but rather, that it is the result of a confound between listener group and the amount of foreign friends: L2 listeners have a larger proportion of non-Spanish friends than do L1 listeners.

No other demographic or behavioral measures yielded significant correlations (ps ≥ .12).

In summary, the post hoc group comparisons revealed significant phonological and syntactic processing differences between the L2 listeners from the current study and the L1 listeners from Caffarra and Martin (2019). Divergences at the level of phonology support the classic interlanguage speech benefit. Divergences at the level of syntax (i.e., error typicality sensitivity) support the novel syntactic interlanguage speech benefit for listeners.

DISCUSSION

The current study investigated L2 listeners' online processing of native and nonnative-accented speech. First, early phonological processing was examined to assess the original conjectures of the interlanguage speech benefit. Second, syntactic processing was examined to verify whether the potential interlanguage benefit extends to high-level linguistic properties, such as syntax. L2 Spanish participants (whose L1 was English) listened to Spanish sentences produced in a native Spanish accent or in English-accented Spanish while their EEG was recorded. Some of these sentences contained errors that are typical (gender errors: **la color*, "the$_F$ color$_M$") or atypical (number errors: **los color*, "the$_{PL}$ color$_S$") in English-accented Spanish.

The main findings indicated that listeners exhibited an increased PMN for the accent that was not part of their interlanguage (i.e., for L2 listeners, the native accent). The results also demonstrated that L2 listeners (unlike L1 listeners) displayed error repair mechanisms (P600 effects) for all grammatical violations, regardless of their typicality or the accent in which they were produced. Each of these main findings will be discussed in turn below.

Accent-Dependent Phonological Processing

In this study, the L2 listeners displayed an increased parietal negativity for native Spanish-accented trials compared with nonnative English-accented trials. This effect did not appear to reflect syntactic or lexical processing, as it was consistent for grammatically correct and grammatically incorrect sentences. As such, this effect likely suggests that the prelexical phonological processing of the target words changed as a function of speaker accent.

Crucially, the results of the L2 Spanish participants from the current study diverge from the literature conducted on native listeners: Although L1 listeners tend to show increased PMN responses for nonnative-accented speech (confirmed by our between-group comparisons; see also

Thomas et al., under review; Goslin et al., 2012, for regional accents; Porretta et al., 2017, for accent-unfamiliar L1 listeners), our L2 listeners instead displayed directionally opposite effects. Keeping in mind that the PMN is generated when there is a dissemblance between sensory input and lexical representation (Connolly & Phillips, 1994), it makes sense that the L2 listeners exhibited an increased negativity for native Spanish-accented trials compared with English-accented trials. In their case, it is the native accent that is most acoustically distinctive from their interlanguage lexical representations; for L2 listeners, phonological–lexical mapping is therefore more difficult in a native accent.

Altogether, it appears that listeners recruit greater cognitive resources during phonological extraction (as indexed by the PMN) when the accent they are perceiving deviates from their own interlanguage. For L2 listeners, this means that native-accented speech recognition is more arduous than nonnative-accented speech recognition. This appears to be especially true for individuals who have extensive exposure to nonnative accents (i.e., a high proportion of foreign friends). Thus, our data provide some of the first electrophysiological evidence supporting the existence of an interlanguage speech intelligibility benefit. Let us caution, however, that there are still wide individual differences to be appreciated within each of L1 and L2 listener groups. Thus, even if our central tendency analyses showed group-based differences, future studies should continue to investigate this phenomenon.

Syntactic Interface of the Interlanguage Speech Benefit

When confronted with Spanish gender and number errors, the L2 listeners in the current study exhibited the expected ERP signature for grammatical violations. Given the polarity, timing, and type of stimuli examined, we interpret the observed increased positivity as a slightly delayed P600.[4] Crucially, the P600 response exhibited by the L2 listeners (but not the L1 listeners) remained homogeneous across accent types (native Spanish or English-accented Spanish) and error types (typical gender or atypical number errors).

Though L2 listeners themselves appear to adopt accent-independent syntactic processing, between-group processing differences were observed for nonnative-accented utterances. In particular, it appears that L1 listeners, but not L2 listeners, are sensitive to nonnative-accented error typicality. Although L1 listeners show a reduced sensitivity to errors that are more typical in nonnative-accented speech (i.e., they may overlook or abstain from repairing such errors; Caffarra & Martin, 2019; see also Xu et al., 2020; Grey & van Hell, 2017; Hanulíková et al., 2012), L2 listeners (especially those who self-report receiving more exposure to nonnative-accented speech) analogously repair both typical and atypical accented errors. Indeed, the post hoc group comparisons included in the current paper show that L1 and L2 listeners apply two qualitatively different strategies when it comes to nonnative-accented speech processing, even when experimental stimuli and design are matched.

In brief, L1 listeners do not invariably process nonnative-accented grammatical errors as syntactic violations. Contrastively, L2 listeners detect and process all grammatical errors, regardless of their typicality. This divergence falls into line with the interlanguage speech benefit for listeners: L2 listeners appear to outperform natives when it comes to repairing syntactic errors in nonnative speech. The former group displays the expected neural signature to syntactic violations (for which native speakers sometimes do not). Indeed, unlike L1 listeners, the L2 listeners from the current study were able to decode, analyze, and understand nonnative-accented speech just as well as native-accented speech (for comparable results, see Hayes-Harb et al., 2008; Bent & Bradlow, 2003).

Error Typicality Effects: Asset or Liability?

It may be debated whether the L2 listeners ability to detect and repair typical gender errors is veritably a functional advantage, as the label of "interlanguage speech benefit" implies. That is, is it truly beneficial for L2 listeners to detect and repair all grammatical errors, or is it more economic to overlook violations that are frequently encountered?

The between-group comparisons from the current study, along with previous literature (Xu et al., 2020; Grey & van Hell, 2017; Hanulíková et al., 2012), indicate quite consistently that native listeners detect typical accented errors (they display early ERP effects), but that they abstain from repairing them (they do not display P600 effects). Given that the P600 is linked to processes of cognitive effort (e.g., Spotorno, Cheylus, Van Der Henst, & Noveck, 2013; Brouwer, Fitz, & Hoeks, 2012; Kaan & Swaab, 2003), perhaps L1 listeners operate on the basis of processing efficiency; they are able to filter out typical (expected) errors that do not obstruct their linguistic comprehension and thus ultimately economize their limited cognitive resources. Under this perspective, it may be that the native listeners' reduced sensitivity to typical errors constitutes a functional benefit.

However, it appears that L2 listeners adopt an altogether different strategy: They detect and repair every single syntactic violation, regardless of its typicality. Assuming that the ability to overlook errors may be more cognitively economic, the question stands as to whether L2 listeners also possess the capacity to be tolerant to inconsequential violations (i.e., to filter out typical errors). If error repair mechanisms are inflexibly applied by the nonnative listener, such processing may be a liability rather than an asset. The data from the current study suggest that L2 listeners do in fact retain the capability to overlook errors. Indeed, the results demonstrate that, for

English-accented trials, the amplitude of the error effect (i.e., the P600 indexed by gender and number errors) was highly correlated to the participants' propensity to correct speech errors; individuals who reported correcting errors more frequently exhibited the largest P600 effects. Thus, it does not appear to be the case that all L2 listeners stringently and automatically repair grammatical violations. Rather, the positivity is linked to more controlled mechanisms; individual listeners were even self-aware of their propensity to overlook errors. This assumption falls into line with the common supposition that the P600 component is indicative of controlled—rather than automatic—processes (Hahne & Friederici, 1999; see also Gosselin et al., 2021).

Why might some L2 listeners be more willing to correct speech errors than others? We speculate that this difference may be related to some participants' increased metalinguistic awareness: Certain listeners may be particularly critical of ungrammatical nonnative speech input, as they are self-monitoring their own output for similar errors. Countless studies indicate that metalinguistic awareness positively impacts L2 acquisition (e.g., Golonka, 2006; Thomas, 1988; Tunmer & Myhill, 1984). Interestingly, in this study, the L2 listeners who self-reported correcting speech errors more often were also those who scored highest on the Spanish proficiency assessment (Pearson $r = .44, p = .019$). Furthermore, L2 listeners who reported being friends with more foreigners exhibited the largest P600 effects for the typical gender errors. Participants with more foreign friends are likely exposed to greater speech variability in their environment and are conceivably more familiar with nonnative-accented speech, in general. Thus, it appears that accent-familiar nonnative listeners are more likely to detect and repair all syntactic violations (i.e., to display a syntactic interlanguage benefit) as they possess the most metalinguistic knowledge of English-accented Spanish. The potential role of metalinguistic awareness is further bolstered by the finding that the ERP responses of the L1Spanish listeners from Caffarra and Martin (2019) exhibited no relation between self-reported error correction and proficiency (Pearson $r = .12, p = .509$). As native listeners, these participants possess intuitive and implicit judgments about Spanish (Davies, 2003) and would thus not need to rely on external metalinguistic knowledge. Future research should continue to examine whether language learners are particularly critical of errors they are at risk of producing (i.e., an English learner of Spanish who is weary of gender errors) or whether they are simply critical of errors in general.

To summarize, it is clear that L1 and L2 listeners exhibit differences when it comes to the processing of nonnative-accented syntactic errors. However, these differences appear to be functionally advantageous for both groups, in their own respect. That is, L1 listeners operate on the basis of processing efficiency by overlooking errors in nonnative-accented speech when they are expected (typical) and thus unproblematic. Contrastively, L2 listeners (especially those with high metalinguistic awareness) may develop a diagnostic-like attitude toward nonnative productions in an effort to avoid making similar errors in their own speech. The strategy of repairing all syntactic violations (whether typical or atypical) may thus be advantageous in increasing L2 listeners' underlying proficiency.

Remarkably, both groups have adapted to their idiosyncratic linguistic circumstances and appear to reap the benefits of their processing differences. With this in mind, it is advisable to approach the topic of the interlanguage speech benefit in a more nuanced way, rather than viewing it as dichotomous or mutually exclusive (i.e., if one group is advantaged, the other is necessarily disadvantaged). Instead of referring to the phenomenon as a benefit, it may be constructive to refer instead to the phenomenon as an interlanguage speech *contribution* or interlanguage speech *supplement*.

Conclusion

This study investigated (1) whether online data support the original interlanguage speech intelligibility benefit and (2) whether nonnative Spanish speakers may also display a "syntactic" interlanguage benefit when they listen to typical and atypical grammatical violations in English-accented Spanish.

First, the main findings indicated that listeners recruit greater cognitive resources (as indexed by a PMN) when the accent they are perceiving deviates from their own accent; for L2 listeners, this means that the native accent requires heightened processing efforts. Consequently, our data provide some of the first electrophysiological evidence supporting the existence of an interlanguage speech intelligibility benefit.

Second, the results from the current study indicated that L2 listeners detect and repair all syntactic errors, regardless of their typicality or the accent in which they were produced. This processing is distinctive from that of the L1 listeners in a similar precursor study (Caffarra & Martin, 2019), who did not repair typical gender errors when they were produced in English-accented Spanish. Although these results fall into line with a syntactic interlanguage benefit for listeners, we posit that such group-based differences are actually beneficial for both parties: L1 listeners operate on the basis of cognitively efficient processing, and L2 listeners repair all violations, probably as a result of the advantageous implementation of metalinguistic awareness. Thus, the results from this study illustrate that language users adapt to their idiosyncratic linguistic circumstances.

APPENDIX

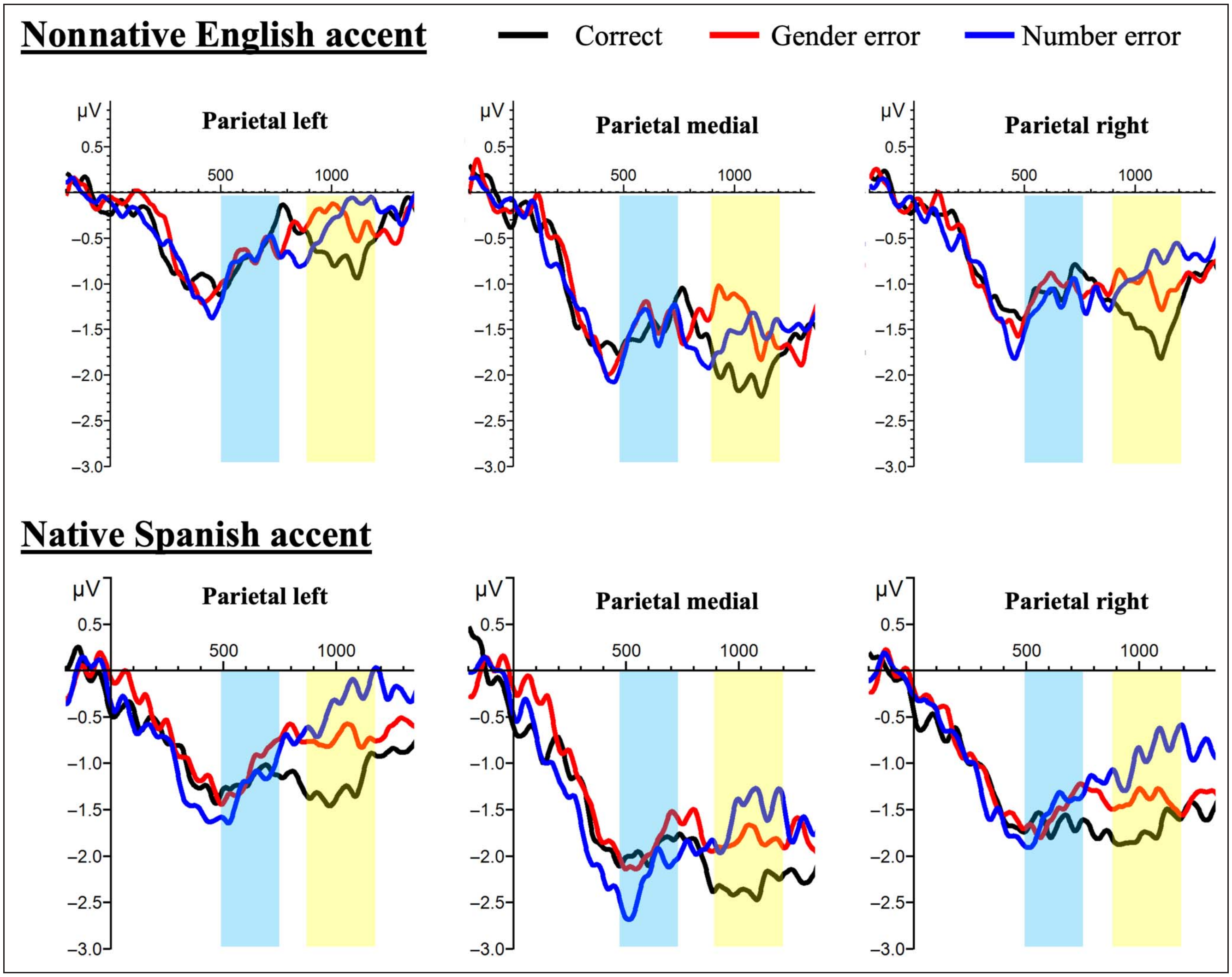

Figure A1. L2 listener averages for the levels of correctness in each accent. The three parietal ROIs are displayed: parietal left (P7, P3, O1), parietal medial (CP1, CP2, Pz), and parietal right (P8, P4, O2). The critical time windows are highlighted (in blue, PMN: 500–750 msec; in yellow, P600: 900–1200 msec).

Acknowledgments

The authors are grateful for the help of all research assistants at the BCBL who contributed to data collection.

Reprint requests should be sent to Leah Gosselin, Department of Linguistics, University of Ottawa, 70 Laurier Ave. E., Ottawa, Ontario, Canada, K1N 6 N5, or via e-mail: lgoss035@uottawa.ca.

Author Contributions

Leah Gosselin: Formal analysis; Methodology; Validation; Visualization; Writing—Original draft; Writing—Review & editing. Clara D. Martin: Conceptualization; Funding acquisition; Investigation; Project administration; Resources; Supervision; Writing—Review & editing. Ana González Martín: Data curation; Investigation; Methodology; Writing—Review & editing. Sendy Caffarra: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Supervision; Writing—Review & editing.

Funding Information

This international collaboration was made possible by a Globalink Research Abroad Award (https://dx.doi.org/10.13039/501100004489) conferred to the first author. The research was supported by the Spanish Ministry of Economy and Competitiveness (CEX2020-001010-S to the BCBL; PID2020-113926GB-I00 to C. D. M.), the Basque government (PIBA18_29 to C. D. M.). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (https://dx.doi.org/10.13039/100010663), grant number: 819093 to C. D. M., the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie (https://dx.doi.org/10.13039/100010663), grant number: 837228 (H2020-MSCA-IF-2018-837228) and the Rita Levi Montalcini fellowship to S. C.

Data Availability Statement

A full stimuli list, as well as the off-line questionnaires can be found online at: https://github.com/leahgosselin/SNAP2.git. All data are available upon request.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: M/M = .286; W/M = .143; M/W = .179; W/W = .393.

Notes

1. This may be due to the learners' specific language background characteristics (e.g., grammatical gender is not part of the native English grammar; setting the parameter within a gendered language is thus arduous). However, it is also possible that grammatical gender is inherently more difficult to learn because of its arbitrary nature (e.g., grammatical gender is an inherent lexical feature; unlike number, it cannot be gleaned from a concept's external characteristics) or its lack of acoustic salience. We remain agnostic as to the explanation underlying the distributional difference.
2. Two participants did not complete this questionnaire.
3. This result has been observed in similar past behavioral research (Hanzlíková & Skarnitzl, 2017; Podlipský et al., 2016). Thus, even if L2 listeners experience a benefit in terms of the intelligibility of nonnative speech, there is no "interlanguage speech *credibility* benefit" (Podlipský et al., 2016); similar to native listeners (e.g., Lev-Ari & Keysar, 2010), L2 listeners tend to rate nonnative speech as less credible.
4. Note that the latency of this P600 (e.g., 900–1200 msec) conforms to the literature; indeed, ERP components are typically slightly delayed when nonnative-accented speech is examined (see Gosselin et al., 2021; Caffarra & Martin, 2019; Grey & van Hell, 2017).

REFERENCES

Algethami, G., Ingram, J., & Nguyen, T. (2011). The interlanguage speech intelligibility benefit: The case of Arabic-accented English. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd pronunciation in second language learning and teaching conference*, September 2010 (pp. 30–42). Ames, IA: Iowa State University.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, *114*, 1600–1610. https://doi.org/10.1121/1.1603234, PubMed: 14514213

Bosker, H. R., Quené, H., Sanders, T., & De Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but nonnative 'um's do not. *Journal of Memory and Language*, *75*, 104–116. https://doi.org/10.1016/j.jml.2014.05.004

Braun, B., Dainora, A., & Ernestus, M. (2011). An unfamiliar intonation contour slows down online speech comprehension. *Language and Cognitive Processes*, *26*, 350–375. https://doi.org/10.1080/01690965.2010.492641

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. https://doi.org/10.1016/j.brainres.2012.01.055, PubMed: 22361114

Caffarra, S., & Martin, C. D. (2019). Not all errors are the same: ERP sensitivity to error typicality in foreign accented speech perception. *Cortex*, *116*, 308–320. https://doi.org/10.1016/j.cortex.2018.03.007, PubMed: 29657069

Chu, P. C. K. (2013). *Interlanguage speech intelligibility benefit: Implications for a model of second language word production and recognition [doctoral dissertation]*. New South Wales University.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*, 3647–3658. https://doi.org/10.1121/1.1815131, PubMed: 15658715

Connolly, J. F., D'Arcy, R. C., Kujala, A., & Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *NeuroReport*, *12*, 237–243. https://doi.org/10.1097/00001756-200102120-00012, PubMed: 11209927

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, *6*, 256–266. https://doi.org/10.1162/jocn.1994.6.3.256, PubMed: 23964975

Davies, A. (2003). *The native speaker: Myth and reality*. Clevedon, United Kingdom: Multilingual Matters. https://doi.org/10.21832/9781853596247

Dickerson, L. J. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, *9*, 401–407. https://doi.org/10.2307/3585624

Eichstaedt, K. E., Kovatch, K., & Maroof, D. A. (2013). A less conservative method to adjust for familywise error rate in neuropsychological research: The Holm's sequential Bonferroni procedure. *NeuroRehabilitation*, *32*, 693–696. https://doi.org/10.3233/NRE-130893, PubMed: 23648625

Fairchild, S., & Papafragou, A. (2018). Sins of omission are more likely to be forgiven in nonnative speakers. *Cognition*, *181*, 80–92. https://doi.org/10.1016/j.cognition.2018.08.010, PubMed: 30149264

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*, 78–104. https://doi.org/10.1006/jmla.1999.2638

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research*, *38*, 379–412. https://doi.org/10.1007/s10936-008-9097-8, PubMed: 19117134

Foucart, A., & Hartsuiker, R. J. (2021). Are foreign-accented speakers that 'incredible'? The impact of the speaker's indexical properties on sentence processing.

Neuropsychologia, *158*, 107902. https://doi.org/10.1016/j.neuropsychologia.2021.107902, PubMed: 34052231
Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, *17*, 213–247. https://doi.org/10.1177/026765830101700301
Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*, 65–87. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x
Golonka, E. M. (2006). Predictors revised: Linguistic knowledge and metalinguistic awareness in second language gain in Russian. *The Modern Language Journal*, *90*, 496–505. https://doi.org/10.1111/j.1540-4781.2006.00428.x
Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*, 92–102. https://doi.org/10.1016/j.bandl.2012.04.017, PubMed: 22694999
Gosselin, L., Martin, C. D., Navarra-Barindelli, E., & Caffarra, S. (2021). The presence of a foreign accent introduces lexical integration difficulties during late semantic processing. *Language, Cognition and Neuroscience*, *36*, 1086–1106. https://doi.org/10.1080/23273798.2021.1909084
Grey, S., Cosgrove, A. L., & van Hell, J. G. (2020). Faces with foreign accents: An event-related potential study of accented sentence comprehension. *Neuropsychologia*, *147*, 107575. https://doi.org/10.1016/j.neuropsychologia.2020.107575, PubMed: 32738253
Grey, S., Schubel, L. C., McQueen, J. M., & Van Hell, J. G. (2019). Processing foreign-accented speech in a second language: Evidence from ERPs during sentence comprehension in bilinguals. *Bilingualism: Language and Cognition*, *22*, 912–929. https://doi.org/10.1017/S1366728918000937
Grey, S., & van Hell, J. G. (2017). Foreign-accented speaker identity affects neural correlates of language comprehension. *Journal of Neurolinguistics*, *42*, 93–108. https://doi.org/10.1016/j.jneuroling.2016.12.001
Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*, 240–244. https://doi.org/10.1111/j.14698986.1991.tb00417.x, PubMed: 1946890
Hagoort, P., & Brown, C. M. (2000). 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. https://doi.org/10.1016/S00283932(00)00053-1, PubMed: 10906378
Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*, 194–205. https://doi.org/10.1162/089892999563328, PubMed: 10198134
Hanulíková, A., Van Alphen, P. M., Van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, *24*, 878–887. https://doi.org/10.1162/jocn_a_00103, PubMed: 21812565
Hanzlíková, D., & Skarnitzl, R. (2017). Credibility of native and nonnative speakers of English revisited: Do nonnative listeners feel the same? *Research in Language*, *15*, 285–298. https://doi.org/10.1515/rela-2017-0016
Harding, L. (2012). Accent, listening assessment and the potential for a shared-L1 advantage: A DIF perspective. *Language Testing*, *29*, 163–180. https://doi.org/10.1177/0265532211421161
Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, *36*, 664–679. https://doi.org/10.1016/j.wocn.2008.04.002, PubMed: 19606271
Holt, R., Kung, C., & Demuth, K. (2018). Listener characteristics modulate the semantic processing of native vs. foreign-accented speech. *PLoS One*, *13*, e0207452. https://doi.org/10.1371/journal.pone.0207452, PubMed: 30517122
Jiang, X., Gossack-Keenan, K., & Pell, M. D. (2020). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Quarterly Journal of Experimental Psychology*, *73*, 55–79. https://doi.org/10.1177/1747021819865833, PubMed: 31293191
Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and Linguistics Compass*, *1*, 571–591. https://doi.org/10.1111/j.1749-818X.2007.00037.x
Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*, 98–110. https://doi.org/10.1162/089892903321107855, PubMed: 12590846
Korpal, P., & Sobkowiak, M. (2020). The perception of native vs. nonnative Danish speech: Bent and Bradlow's matched interlanguage speech intelligibility benefit revisited. *Scandinavian Philology*, *18*, 284–296. https://doi.org/10.21638/11701/spbu21.2020.204
Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*, 68–74. https://doi.org/10.1016/j.bandl.2008.06.002, PubMed: 18657314
Lee, S., Kang, J., & Nam, H. (2020). Identification of English vowels by nonnative listeners: Effects of listeners' experience of the target dialect and talkers' language background. *Second Language Research*, *38*, 449–475. https://doi.org/10.1177/0267658320965648
Lee, J. K., & Xue, X. (2011). The Interlanguage Speech Intelligibility Benefit for Listeners (ISIB-L): The case of English liquids. *Phonetics and Speech Sciences*, *3*, 51–65. https://doi.org/10.13064/KSSS.2019.11.1.051
Leikin, M., Ibrahim, R., Eviatar, Z., & Sapir, S. (2009). Listening with an accent: Speech perception in a second language by late bilinguals. *Journal of Psycholinguistic Research*, *38*, 447–457. https://doi.org/10.1007/s10936-009-9099-1, PubMed: 19259819
Lev-Ari, S., & Keysar, B. (2010). Why don't we believe nonnative speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*, 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025
Lewendon, J., Mortimore, L., & Egan, C. (2021). The phonological mapping (mismatch) negativity: History, inconsistency, and future direction: Corrigendum. *Frontiers in Psychology*, *11*, 619241. https://doi.org/10.3389/fpsyg.2020.619241, PubMed: 33584454
Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*, 71–102. https://doi.org/10.1016/0010-0277(87)90005-9, PubMed: 3581730
Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition*, *164*, 61–73. https://doi.org/10.1016/j.cognition.2017.03.012, PubMed: 28384491
Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, *20*, 139–154. https://doi.org/10.1017/S0272263198002022
Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*, 289–306. https://doi.org/10.1177/002383099503800305, PubMed: 8816082
Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310. https://doi.org/10.1111/0023-8333.49.s1.8
Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom–up and top–down effects on spoken word

processing. *Biological Psychology*, *80*, 114–121. https://doi.org/10.1016/j.biopsycho.2008.04.008, PubMed: 18524453
Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z
Phillips, S. F., & Pylkkänen, L. (2021). Composition within and between languages in the bilingual mind: MEG evidence from Korean/English bilinguals. *eNeuro*, *8*, 1–12. https://doi.org/10.1523/ENEURO.0084-21.2021, PubMed: 34732542
Podlipský, V. J., Šimáčková, Š., & Petráž, D. (2016). Is there an interlanguage speech credibility benefit? *Topics in Linguistics*, *17*, 30–44. https://doi.org/10.1515/topling-2016-0003
Porretta, V., Tremblay, A., & Bolger, P. (2017). Got experience? PMN amplitudes to foreign-accented speech modulated by listener experience. *Journal of Neurolinguistics*, *44*, 54–67. https://doi.org/10.1016/j.jneuroling.2017.03.002
Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics*, *58*, 1–21. https://doi.org/10.1016/j.wocn.2016.05.006
Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*, 781–791. https://doi.org/10.1016/S0926-6410(03)00202-7, PubMed: 14561463
Roessel, J., Schoel, C., & Stahlberg, D. (2020). Modern notions of accent-ism: Findings, conceptualizations, and implications for interventions and research on nonnative accents. *Journal of Language and Social Psychology*, *39*, 87–111. https://doi.org/10.1177/0261927X19884619
Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech. *Frontiers in Human Neuroscience*, *9*, 167. https://doi.org/10.3389/fnhum.2015.00167, PubMed: 25859209
Ruben, R. J. (1999). A time frame of critical/sensitive periods of language development. *Indian Journal of Otolaryngology and Head and Neck Surgery*, *51*, 85–89. https://doi.org/10.1007/BF02996542, PubMed: 23119557
Sabo, E. (2021). *Social factors in the production, perception and processing of contact varieties: Evidence from bilingual corpora, Nativeness evaluations, and real-time processing (EEG) of Spanish-accented English [doctoral dissertation]*. University of Michigan. https://doi.org/10.7302/1544
Sabourin, L. L. (2001). L1 effects on the processing of grammatical gender in L2. *EUROSLA Yearbook*, *1*, 159–169. https://doi.org/10.1075/eurosla.1.13sab
Sabourin, L., Stowe, L. A., & De Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *22*, 1–29. https://doi.org/10.1191/0267658306sr259oa
Schiller, N. O., Boutonnet, B. P. A., De Heer Kloots, M. L., Meelen, M., Ruijgrok, B., & Cheng, L. L. (2020). (not so) great expectations: Listening to foreign-accented speech reduces the brain's anticipatory processes. *Frontiers in Psychology*, *11*, 2143. https://doi.org/10.3389/fpsyg.2020.02143, PubMed: 32982877
Selinker, L. (1972). Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching*, *10*, 209–232. https://doi.org/10.1515/iral.1972.10.1-4.209
Smith, B. L., Bradlow, A. R., Bent, T., Solé, M. J., Recasens, D., & Romero, J. (2003). Production and perception of temporal contrasts in foreign-accented English. In *Proceedings of the XVth international congress of phonetic sciences* (pp. 519–522). Barcelona, Spain.
Smith, B. L., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, *37*, 257–275. https://doi.org/10.1016/j.wocn.2009.03.001
Smith, L. E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly*, *13*, 371–380. https://doi.org/10.2307/3585884
Spotorno, N., Cheylus, A., Van Der Henst, J. B., & Noveck, I. A. (2013). What's behind a P600? Integration operations during irony processing. *PLoS One*, *8*, e66839. https://doi.org/10.1371/journal.pone.0066839, PubMed: 23826155
Stibbard, R. M., & Lee, J. I. (2006). Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. *Journal of the Acoustical Society of America*, *120*, 433–442. https://doi.org/10.1121/1.2203595, PubMed: 16875239
Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual & Multicultural Development*, *9*, 235–246. https://doi.org/10.1080/01434632.1988.9994334
Thomas, T., Martin, C. D., & Caffarra, S. (under review). An ERP investigation of accented isolated single word processing. https://doi.org/10.1101/2022.07.07.499100
Tunmer, W. E., & Myhill M. E. (1984). *Metalinguistic awareness and bilingualism*. In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic awareness in children*. (Springer series in language and communication, Vol. 15). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-69113-3_10
van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K., et al. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année Psychologique*, *120*, 73–96. https://doi.org/10.3917/anpsy1.201.0073
van Wijngaarden, S. J. (2001). Intelligibility of native and nonnative Dutch speech. *Speech Communication*, *35*, 103–113. https://doi.org/10.1016/S0167-6393(00)00098-4
van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for nonnative listeners. *Journal of the Acoustical Society of America*, *111*, 1906–1916. https://doi.org/10.1121/1.1456928, PubMed: 12002873
Wang, H., & van Heuven, V. J. (2015). The interlanguage speech intelligibility benefit as bias toward native language phonology. *Perception*, *6*, 2041669515613661. https://doi.org/10.1177/2041669515613661, PubMed: 27551352
Weber, A., Broersma, M., & Aoyagi, M. (2011). Spoken-word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics*, *39*, 479–491. https://doi.org/10.1016/j.wocn.2010.12.004
White, L., Valenzuela, E., Kozlowska-Macgregor, M., Leung, I., & Ayed, H. B. (2001). The status of abstract features in interlanguage: Gender and number in L2 Spanish. In *BUCLD 25 proceedings* (pp. 792–802). Somerville, MA: Cascadilla Press.
Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. *Attention, Perception, & Psychophysics*, *75*, 537–556. https://doi.org/10.3758/s13414-012-0404-y, PubMed: 23456266
Xie, X., & Fowler, C. A. (2013). Listening with a foreign-accent: The interlanguage speech intelligibility benefit in mandarin speakers of English. *Journal of Phonetics*, *41*, 369–378. https://doi.org/10.1016/j.wocn.2013.06.003, PubMed: 24293741
Xu, J., Abdel Rahman, R., & Sommer, W. (2020). Perceived language competence modulates criteria for speech error processing: Evidence from event-related potentials. *Language, Cognition and Neuroscience*, *35*, 752–765. https://doi.org/10.1080/23273798.2018.1562558