Why pitch sensitivity matters: event-related potential evidence of metric and syntactic violation detection among Spanish late learners of German

Maren Schmidt-Kassow[1]*, M. Paula Roncaglia-Denissen[2] and Sonja A. Kotz[2]

[1] *Institute of Medical Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany*
[2] *Minerva Research Group Neurocognition of Rhythm in Communication, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*

Event-related potential (ERP) data in monolingual German speakers have shown that sentential metric expectancy violations elicit a biphasic ERP pattern consisting of an anterior negativity and a posterior positivity (P600). This pattern is comparable to that elicited by syntactic violations. However, proficient French late learners of German do not detect violations of metric expectancy in German. They also show qualitatively and quantitatively different ERP responses to metric and syntactic violations. We followed up the questions whether (1) latter evidence results from a potential pitch cue insensitivity in speech segmentation in French speakers, or (2) if the result is founded in rhythmic language differences. Therefore, we tested Spanish late learners of German, as Spanish, contrary to French, uses pitch as a segmentation cue even though the basic segmentation unit is the same in French and Spanish (i.e., the syllable). We report ERP responses showing that Spanish L2 learners are sensitive to syntactic as well as metric violations in German sentences independent of attention to task in a P600 response. Overall, the behavioral performance resembles that of German native speakers. The current data suggest that Spanish L2 learners are able to extract metric units (trochee) in their L2 (German) even though their basic segmentation unit in Spanish is the syllable. In addition Spanish in contrast to French L2 learners of German are sensitive to syntactic violations indicating a tight link between syntactic and metric competence. This finding emphasizes the relevant role of metric cues not only in L2 prosodic but also in syntactic processing.

Keywords: auditory language processing, P600, speech segmentation, trochee, L2

INTRODUCTION

Speaking and hearing a second language (L2) may pose a particular challenge for a non-native speaker. Imagine yourself in a situation talking to people in a different language: if you are not a fluent L2 speaker you may feel lost in an overwhelming mass of sounds, and may have no idea where a word begins or ends. However, if you read a book or a paper in a non-native language, word recognition may be rather simple, as blank spaces between letter strings signal the end or beginning of words (see **Figure 1**). Hence, the correct translation of extracted words is a residual challenge when reading in a non-native language and thus may affect lexical access.

In *auditory* language processing this form of extraction is more complicated. The perceiver has to identify those acoustic cues in a continuous speech stream that indicate when a new word begins. This fundamental capacity in speech perception is called *segmentation*. Various segmentation cues are used in parallel to speed up word recognition. These cues span distributional cues (Saffran et al., 1996), allophonic and phonotactic cues, lexical cues (Mattys et al., 2005), and prosodic cues (e.g., Kim et al., 2008). Here we primarily focus on prosodic cues in speech segmentation. If speech segmentation is not successful, lexical access, grammatical encoding,, or semantic integration becomes impossible. Furthermore, if speech segmentation from a continuous speech stream were not hard enough, prosodic cues signaling the beginning or end of a

word vary from language to language (e.g., Cutler, 1994; Tyler and Cutler, 2009). Hence, the interpretation of segmentation cues that are seemingly effortlessly acquired in infancy (Jusczyk, 2002) may not account too much in second language acquisition. This results in persistent problems in late L2 phonology acquisition even when L2 learners have been immersed in a second language (Flege et al., 1997; MacKay et al., 2001; Piske et al., 2002).

In this context, two interesting questions have been intensively discussed: (1) Are L2-specific prosodic segmentation cues used during L2 acquisition, or is L2 segmentation primarily lexically driven (Sanders et al., 2002; White et al., 2009)? (2) Is the successful use of L2 specific prosodic cues dependent on the phonological proximity of L1 and L2 (Sanders et al., 2002; Toro et al., 2009)?

Concerning the second question, speech segmentation in L2 should be easier the closer L1 and L2 phonological proximity is (Flege et al., 1995; Best et al., 2001; Yu and Andruski, 2010). For example, several studies have shown that listeners stick to their native segmentation strategies when exposed to an L2 (Cutler et al., 1989, 1992; Otake et al., 1996; Cutler and Otake, 1999; Tyler and Cutler, 2009). However, different approaches explaining phonological proximity exist. The so-called *rhythm-based segmentation account* (Nazzi et al., 2006) states that segmentation strategies differ as a function of the rhythmic class a language belongs to (Murty et al., 2007). This account differentiates three rhythmic language
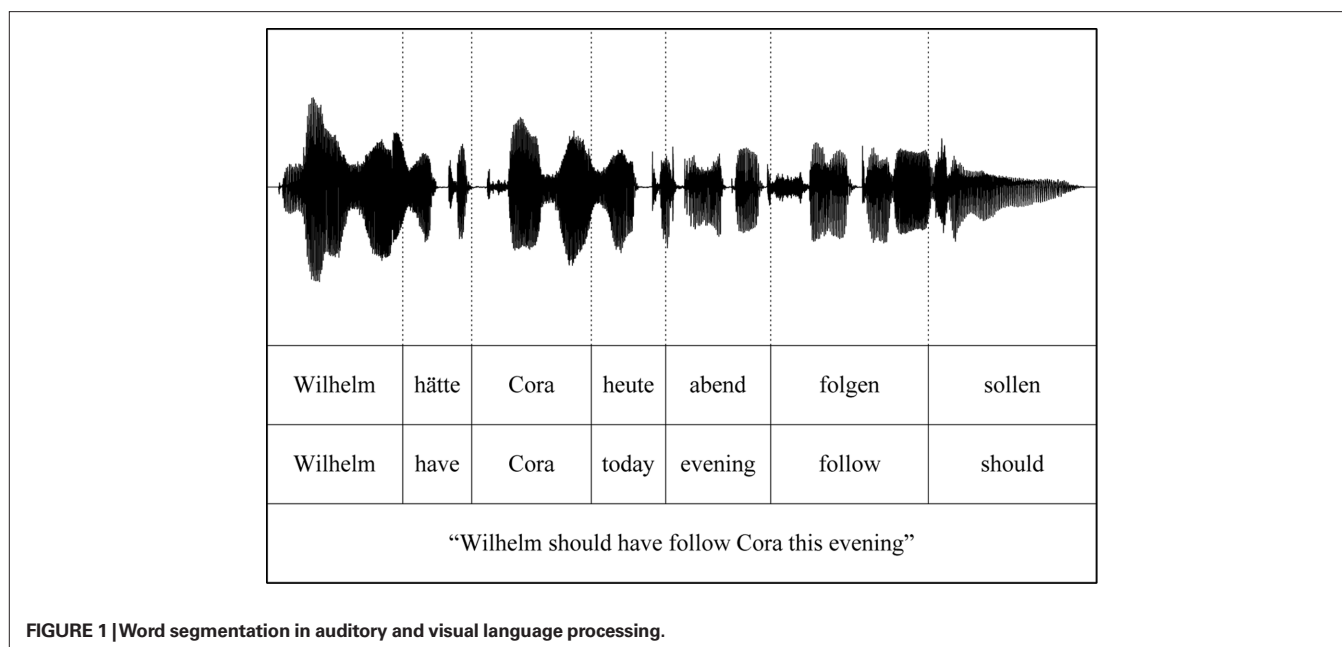
Wilhelm	hätte	Cora	heute	abend	folgen	sollen
Wilhelm	have	Cora	today	evening	follow	should

"Wilhelm should have follow Cora this evening"

FIGURE 1 | Word segmentation in auditory and visual language processing.

classes, i.e., syllable-timed, stress-timed, and mora-timed languages (e.g., Auer, 1993; Nazzi and Ramus, 2003). Syllables of stress-timed languages such as German or English show greater variability in auditory prominence (influenced by intensity, duration, and frequency) than syllables of syllable-timed languages such as French or Spanish (Lee and Todd, 2004). German, for instance, relies on a prominent pattern of stressed and unstressed syllables, i.e., trochaic units (Eisenberg, 1991; Féry, 1997). These trochaic units have been shown to play an important role in speech segmentation. Evidence in support for this concept comes from language acquisition (Sansavini, 1997; Nazzi and Ramus, 2003), or discrimination of languages purely based on prosodic information (Ramus et al., 2000; Ramus, 2002).

However, looking closer at the use of prosodic cues in L1 segmentation (i.e., vowel lengthening or pitch movement), it becomes evident that the use of segmentation cues may vary beyond speech rhythm classes (Tyler and Cutler, 2009). Toro et al. (2009) suggested that pitch is the relevant segmentation cue in Spanish and English, but not in French. This is interesting insofar as Spanish and French belong to the same rhythmic group, i.e., syllable-timed languages, but do not share the same prosodic cues in segmentation. Toro et al. (2009) proposed that the English, German, and Spanish stress systems differ from the French stress system as the former but not the latter use contrastive and lexical stress. This results in a diminished phonological representation of stress use in the L1 acquisition of French and may result in the so-called stress deafness phenomenon (Dupoux et al., 2008). This, in turn, involves insensitivity to pitch cues in speech segmentation (Toro et al., 2009). It is a central feature of stress-timed languages that meter, i.e., the succession of stressed and unstressed syllables, plays an important role in speech processing. Stressed syllables frequently signal the beginning of a new word and hence are a relevant cue in speech segmentation (as discussed in the *metrical segmentation hypothesis*, Cutler and Norris,

1988; Norris et al., 1995). In previous electrophysiological studies, we provided evidence that monolingual German speakers extract and process metric information early on in sentence processing (Schmidt-Kassow and Kotz, 2009a,b). Expectancy violations in a trochaic metric pattern elicit a biphasic event-related potential (ERP) response comparable to the well-known syntactic pattern consisting of an early anterior negativity and a late positivity (P600, Schmidt-Kassow and Kotz, 2009a). Furthermore, metric errors are detected earlier than syntactic errors and metric and syntactic processes interact in the late positive component (P600). As the extraction of metric patterns is prominent in German, but not in French, we investigated to what extent French L2 learners of German make use of these cues, and whether a potential insensitivity toward these cues impacts other linguistic functions such as syntax. We conducted a study with French native but highly proficient L2 speakers of German and reported a lack of detecting metric expectancy violations in German (Schmidt-Kassow et al., 2011). However, when focusing attention to syntactic correctness French speakers display comparable syntactic ERP effects to monolingual German speakers. This pattern changed when attention was focused on the metric structure. Here, French L2 learners of German did not display a P600 in the metric nor the syntactic conditions. In sum, this study provides evidence that French L2 learners of German are insensitive to deviations from a regular metric structure and furthermore have difficulties with implicit syntactic processing (i.e., in cases when attention is not directed to the syntactic structure).

We consider that French native speakers are not able to detect metrically regular stress patterns in German due to acquired pitch insensitivity. However, it remains unclear whether this result is a consequence of French being a syllable-timed language. French native speakers could thus be insensitive to trochaic units in stress-timed languages as the syllable is the fundamental segmentation unit in syllable-timed languages (Sebastián-Gallés et al., 1992;

Goyet et al., 2010). Therefore, these initial L2 results may not just be motivated by acquired pitch insensitivity, but may also result from different segmentation strategies in linguistic rhythmic groups. Native speakers of syllable-timed languages may adhere to the syllable as a segmentation unit in their L2, while native speakers of stress-timed languages apply the trochaic segmentation unit.

We therefore investigated whether Spanish L2 learners of German are sensitive to trochaic units in German sentence processing using the same paradigm as in French L2 learners of German. Spanish is comparable to French as both languages belong to syllable-timed rhythm languages. Hence, the syllable but not the trochaic foot is the fundamental segmentation unit in Spanish and in French. However, in contrast to French, Spanish utilizes lexical and contrastive stress comparable to stress-timed languages. It has also been shown that Spanish native speakers are sensitive to pitch cues (Toro et al., 2009). Based on the assumption that our previous evidence from native French speakers is primarily caused by the insensitivity to pitch cues we therefore formulated the following hypotheses:

1. Spanish L2 learners of German should show an electrophysiological response to syntactic as well as metric expectancy violations during auditory sentence processing (biphasic pattern consisting of an early negativity and a late positivity) comparable to German native speakers.
2. Behavioral performance of Spanish L2 learners of German should resemble the performance of German monolingual speakers under both task conditions, i.e., they should be able to judge metric and syntactic correctness in German sentences without any problems.

MATERIALS AND METHODS
PARTICIPANTS
Thirteen (11 female) right-handed native speakers of Spanish, aged 19–30 years (mean age = 24) participated in the experiment. All were proficient speakers of German. They had learned German at high school or at the university (mean age of acquisition = 17.0, SD = 4.7) and had not previously lived in a German-speaking country. At the time of testing all participants had spent 6 months in Germany and they mainly used German in daily communication (mean "percentage of German per day" = 53.46, SD = 17.0). A self-assessment questionnaire revealed high proficiency in production (median = 7) and perception (median = 8) on a 10 point rating scale ranging from 1 (very low) to 10 (very high). None of the participants reported any neurological impairment or hearing deficit.

MATERIALS
We selected 52 German sentence quadruplets with a consistent trochaic pattern (regular succession of stressed and unstressed syllables) containing a metric violation, a syntactic violation, a metric and syntactic violation, and a correct control (see **Table 1**). These sentence quadruplets have already been used in previous experiments and are known to elicit an early negativity and a P600 in German monolinguals (Schmidt-Kassow and Kotz, 2009a).

Table 1 | Experimental conditions.

Condition	Example
Correct	'Vera 'hätte 'Christoph 'gestern 'morgen **'duzen** 'können *Vera could have addressed Christoph informally yesterday morning*
Metric violation	'Vera 'hätte 'Christoph 'gestern 'morgen **du'ZEN** 'können *Vera could have addressed Christoph informally yesterday morning*
Syntactic violation	'Vera 'hätte 'Christoph 'gestern 'morgen **'duzte** 'können *Vera could have address Christoph informally yesterday morning*
Double violation	'Vera 'hätte 'Christoph 'gestern 'morgen **duz'TE** 'können *Vera could have address Christoph informally yesterday morning*

The critical item is printed in bold. Quotation marks are placed to indicate stressed syllables.
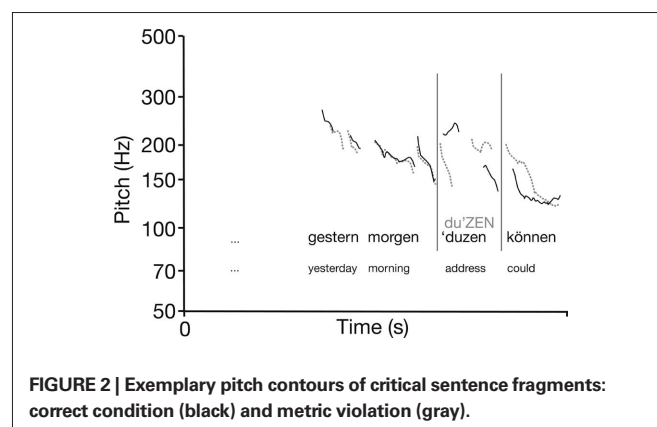

FIGURE 2 | Exemplary pitch contours of critical sentence fragments: correct condition (black) and metric violation (gray).

Acoustic analyses of the material revealed that pitch was the most reliable acoustic cue in the metric violation condition. **Figure 2** illustrates that pitch patterns were almost identical for metrically incorrect and correct conditions up to the critical item, but then diverge diametrically.

PROCEDURE
Participants were tested individually in a sound-attenuating booth. They were seated in a comfortable chair and were informed that they were going to listen to acoustically presented sentences and to move and blink as little as possible. Subjects were asked to participate in two sessions. In the first session they were instructed to evaluate the metrical homogeneity of each sentence, whereas in the second session (2 months later) they judged the grammatical correctness of sentences. The order of the tasks was counterbalanced across participants. Each trial started with a visual cue (asterisk) in the center of a computer screen. Two thousand milliseconds after the offset of the presented trial, participants were asked to perform the respective judgment. The next trial started 2000 ms after the participant's button press.

After a short practice session, 208 experimental sentences (52 per condition) were presented acoustically via two loudspeakers in pseudo-randomized order. The experimental trials were presented in four blocks of approximately 8 min each. After the second block participants were offered a break of 5 min.

ELECTROPHYSIOLOGICAL RECORDINGS

The EEG was recorded from 59 scalp sites by means of Ag/AgCl electrodes mounted in an elastic cap (Electro-Cap Inc., n.d.) according to the 10–20 International System (cf. Pivik et al., 1993). The Sternum served as ground, the left mastoid as on-line reference (recordings were re-referenced to averaged mastoids off-line). Electrode impedances were kept below 3 kΩ. In order to control for eye movements, a horizontal and a vertical EOG was recorded. EEG and EOG signals were digitized on-line with a sample frequency of 250 Hz. An anti-aliasing filter of 67.5 Hz was applied during recording.

DATA ANALYSES

Individual EEG recordings were scanned for artifacts such as electrode drifting, amplifier blocking, muscle artifacts, eye movements, or blinks by means of a rejection algorithm as well as on the basis of visual inspection. Epochs lasted from 100 ms before onset of the critical item (main verb) up to 1800 ms after the critical item. All contaminated trials were rejected and the remaining trials (*syntactic task:* correct condition = 56%, metric condition = 54%, syntactic condition = 58%, double condition = 56%; *metric task:* correct condition = 50%, metric condition = 52%, syntactic condition = 52%, double condition = 54% were averaged per participant, condition, and electrode site. For graphical display only, data were filtered off-line with a 7-Hz low pass filter.

RESULTS

BEHAVIORAL DATA

In the syntactic correctness task, correct response rates for all sentence types were above 98% (see **Figure 3**, correct: 99.4%, SD: 1.2; syntactic: 98.1%, SD: 2.7, metric: 98.8%, SD: 2.3, double: 99.1%, SD: 0.9). There were no significant differences between conditions (all p > 0.1).

A repeated-measures ANOVA in the *metric task* (see **Figure 3**) revealed a significant main effect of *condition* [$F_{(3,36)} = 8.56$, $p < 0.001$]. Planned comparisons of the factor condition (correct/metrically violated/syntactically violated/doubly violated) revealed no significant differences between the syntactic (mean: 94.8%, SD: 6.9) and the correct (98.1%; SD: 2.9) condition, but significant differences between the double violation (78.8%; SD: 19.6) and the correct condition [$F_{(1,12)} = 12.42$, $p < 0.01$], and between

the metric violation (86.7%; SD: 11.4) and the correct condition [$F_{(1,12)} = 15.33$, $p < 0.01$]. Due to restricted degrees of freedom a Bonferroni-adjusted α-level of 0.025 was applied.
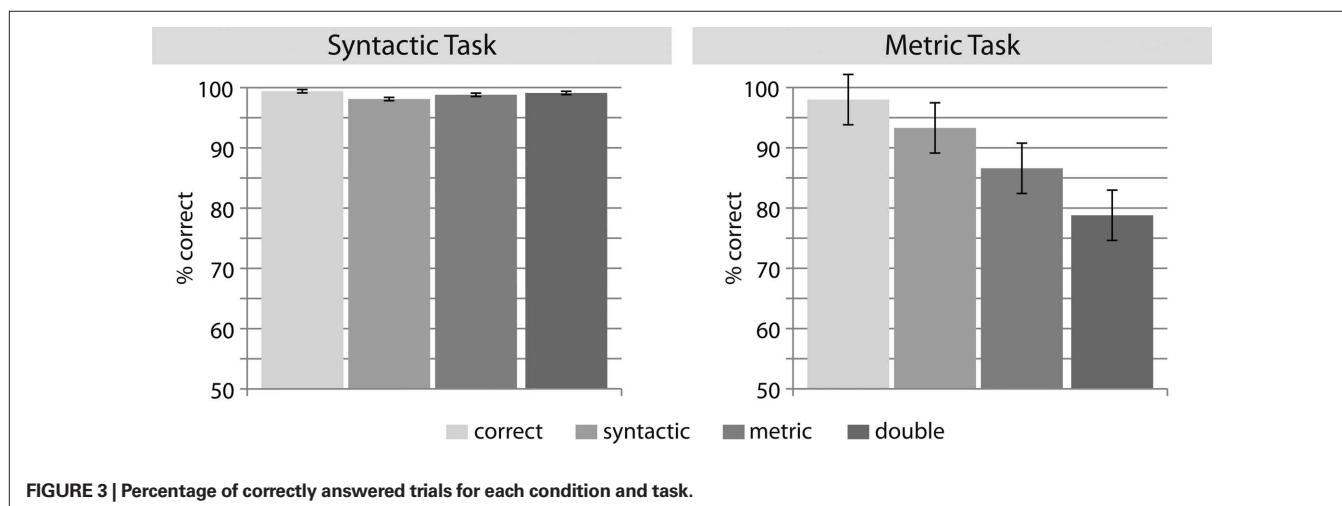
ERP DATA

Syntactic task

In the syntactic correctness task visual inspection of the data revealed a late *positive component* in response to syntactic and double violations and a posteriorly distributed *negativity* elicited by metric violations (see **Figure 4**).
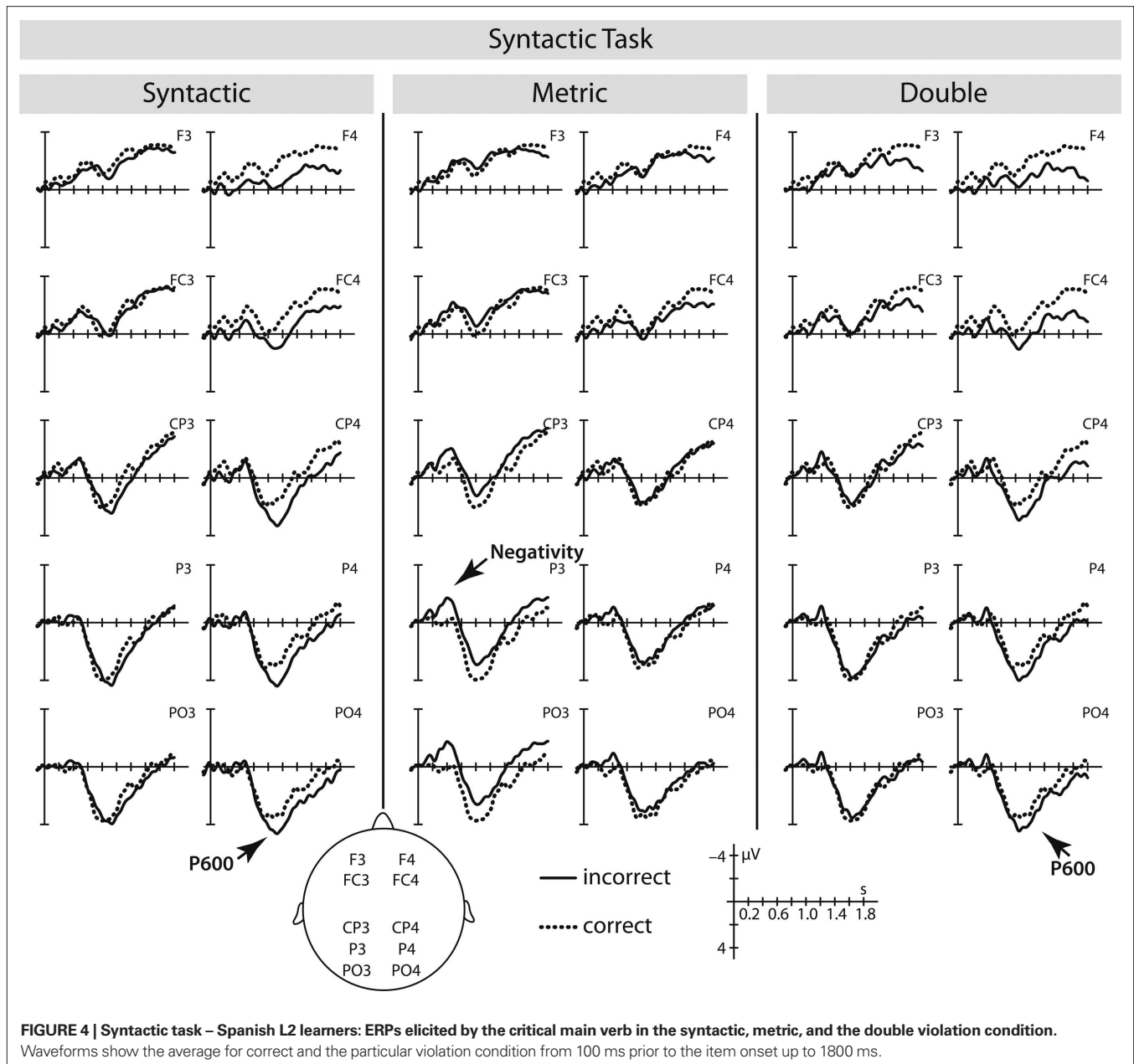
This visual impression was confirmed by a 50-ms-time-line-analysis in each task ranging from 0 to 1800 ms with the following regions of interest: *anterior left* (AF7, AF3, F7, F5, F3, FT7, FC5, FC3), *anterior right* (AF4, AF8, F4, F6, F8, FT8, FC6, FC4), *posterior left* (TP7, CP5, CP3, P7, P5, P3, PO7, PO3, O1), and *posterior right* (TP8, CP4, CP6, P8, P4, P6, PO4, PO8, O2). A repeated-measures ANOVA including the factors *condition* (correct/metrically incorrect/syntactically incorrect/doubly violated), *region* (anterior/posterior), and *hemisphere* (left/right) for each 50 ms segment was computed. The Greenhouse and Geisser (1959) Correction was applied for effects with more than 1 degree of freedom. Based on visual inspection and the described time-line-analysis, we established different time-windows for the critical conditions for statistical evaluation. Hence, we computed separate ANOVAs for the syntactic, metric, and double violation condition.

For the *syntactic violation* condition, we computed a repeated-measures ANOVA including the factors *condition* (correct/syntactically incorrect), *hemisphere* (left/right), and *region* (anterior/posterior) in a 900 to 1500-ms time window. This resulted in a significant interaction of *condition × hemisphere* [$F_{(1,12)} = 25.12$, $p < 0.001$], confirming a condition effect over right electrode-sites [$F_{(1,12)} = 15.44$, $p < 0.01$].

In the *double violation* condition an ANOVA with the same factors, but in a slightly later time window, i.e., 1050–1800 ms, was computed. Here, the ANOVA yielded a significant positive main effect of *condition* [$F_{(1,12)} = 14.48$, $p < 0.01$], but no interaction with the factor *hemisphere*.

Lastly, we applied the same statistical model to the *metric violation* condition in a time window from 300 to 550 ms measured from the onset of the critical word. However, based on the results from the

FIGURE 3 | Percentage of correctly answered trials for each condition and task.

FIGURE 4 | Syntactic task – Spanish L2 learners: ERPs elicited by the critical main verb in the syntactic, metric, and the double violation condition.
Waveforms show the average for correct and the particular violation condition from 100 ms prior to the item onset up to 1800 ms.

time-line-analysis we restricted our analysis to posterior electrode-sites. This analysis yielded a significant negative deflection for the metric violation compared to the correct condition [$F(1,12) = 7.13$, $p < 0.02$], but no interaction with the factor *hemisphere*.
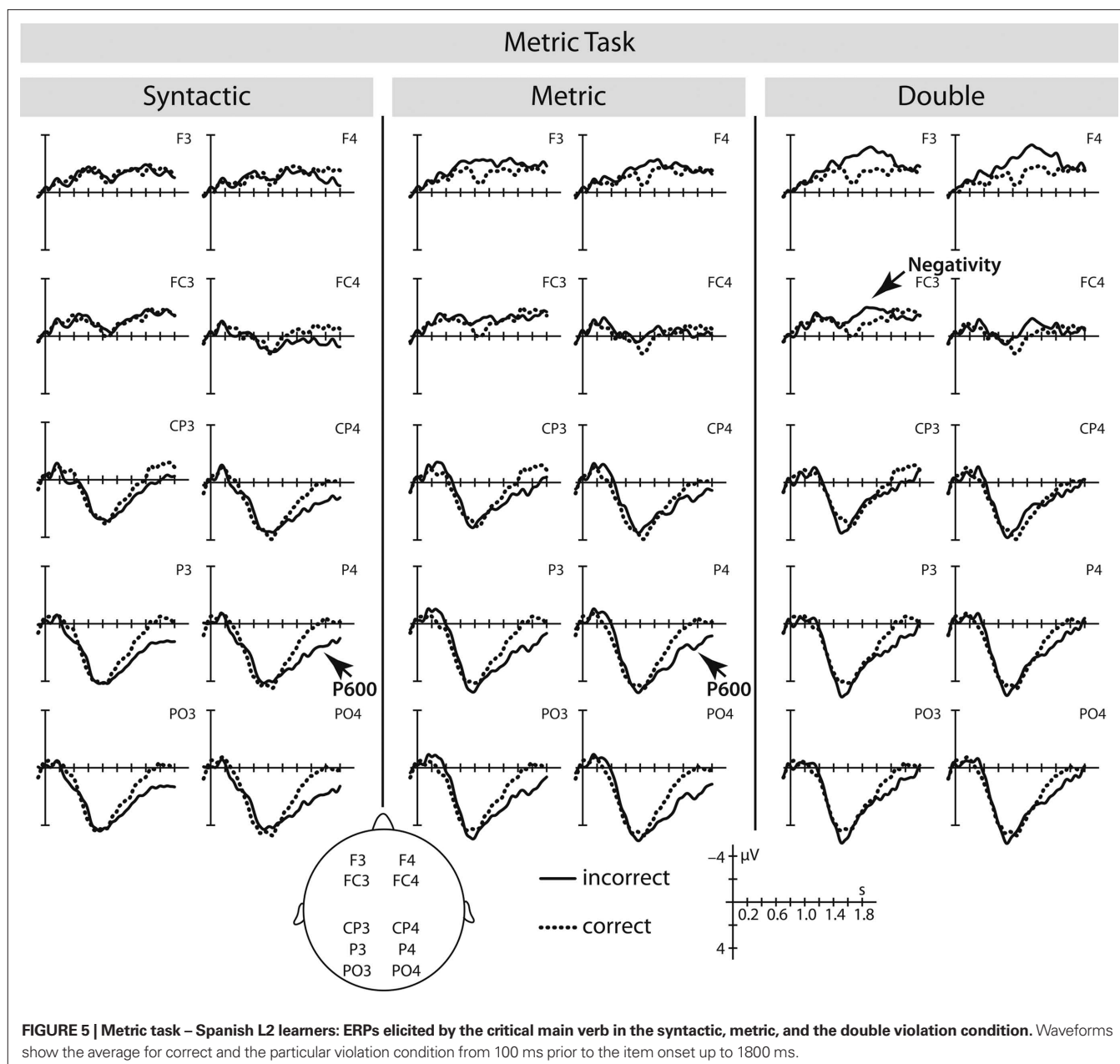
Metric task

In the metric homogeneity judgment three ERP components were elicited, i.e., a late *positive component* in response to syntactic and metric violations and an anteriorly distributed *negativity* elicited by double violations (see **Figure 5**).

The same 50 ms-time-line-analysis as described above was applied to define those time-windows that entered the final repeated-measures ANOVA. The analysis revealed different time-windows as well as different distributions across condition.

Hence, we computed a repeated-measures ANOVA including the factors *condition* (correct/syntactically incorrect) and *hemisphere* (left/right) over posterior electrode-sites in a time window from 1300 to 1800 ms for the *syntactically violated* condition. This resulted in a posterior positivity [$F(1,12) = 8.33$, $p = 0.01$].

In the *metric violation* condition, an ANOVA with the same factors but in a time window from 1500 to 1700 ms also yielded a significant effect of *condition* [positivity; $F(1,12) = 5.25$, $p = 0.04$].

In the *double violation* condition, the prior time-line-analysis revealed an anteriorly distributed negative effect. Hence, we restricted the final ANOVA to anterior electrode-sites, and included the same factors as described above. We found a significant negatively polarized *condition* effect in a time window from 750 to 1400 ms [$F(1,12) = 10.83$, $p < 0.01$].

FIGURE 5 | Metric task – Spanish L2 learners: ERPs elicited by the critical main verb in the syntactic, metric, and the double violation condition. Waveforms show the average for correct and the particular violation condition from 100 ms prior to the item onset up to 1800 ms.

DISCUSSION

In the current experiment, Spanish late learners of German listened to German sentences that included syntactic, metric, or combined violations. They were instructed to judge either the metric homogeneity or the syntactic correctness of a spoken sentence. The study was motivated to find out whether Spanish late L2 learners of German are sensitive to trochaic units in a stressed-timed L2 language even if the syllable is the main segmentation cue in their L1. Furthermore, we aimed to follow-up the question whether sensitivity to metric cues is prerequisite for implicit syntactic processing.

Previous studies with German native speakers (Schmidt-Kassow and Kotz, 2009a) have shown high performance (above 80% correct responses) and a biphasic ERP pattern (early negativity followed by a late positivity) in all critical conditions of the same paradigm

independent of task instructions. However, French late learners of German (Schmidt-Kassow et al., 2011) performed at chance level, when asked to judge the metric homogeneity of spoken sentences. Concerning the ERP results, only a P600 in response to syntactic and double violations in the *syntactic task* was found for French participants, while they failed to show any effect in the metric task.
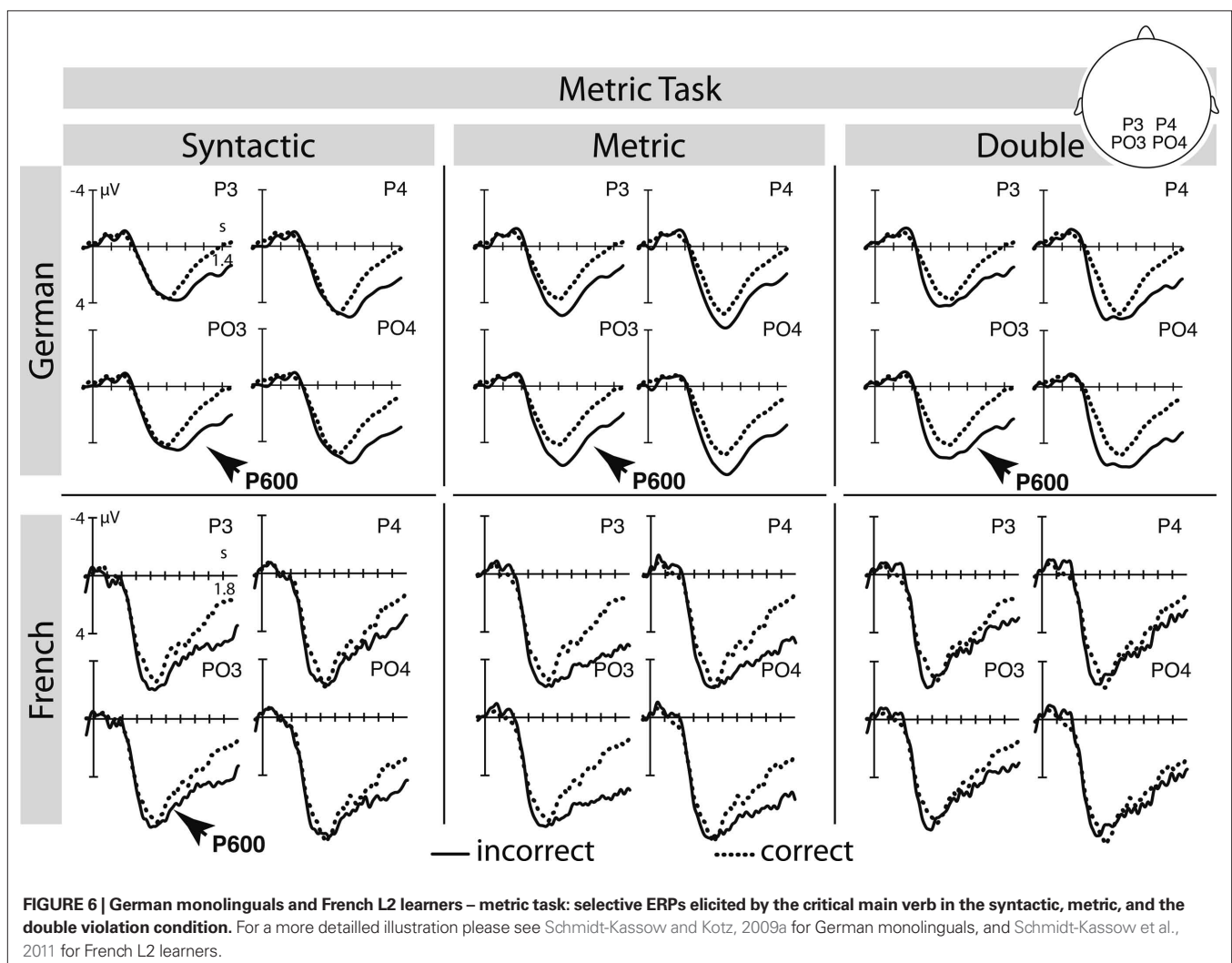
The Spanish late L2 learners of German in the current study had learned German approximately 5 years later than the previously tested French–German L2 learners (12.3 years, Schmidt-Kassow et al., 2011). However, their performance was similar to German native speakers (above 78%) and their ERP pattern indicates that they are sensitive toward metric and syntactic expectancy violations. Spanish L2 learners of German showed a late positivity (P600) in response to syntactic and double violations and an early posteriorly
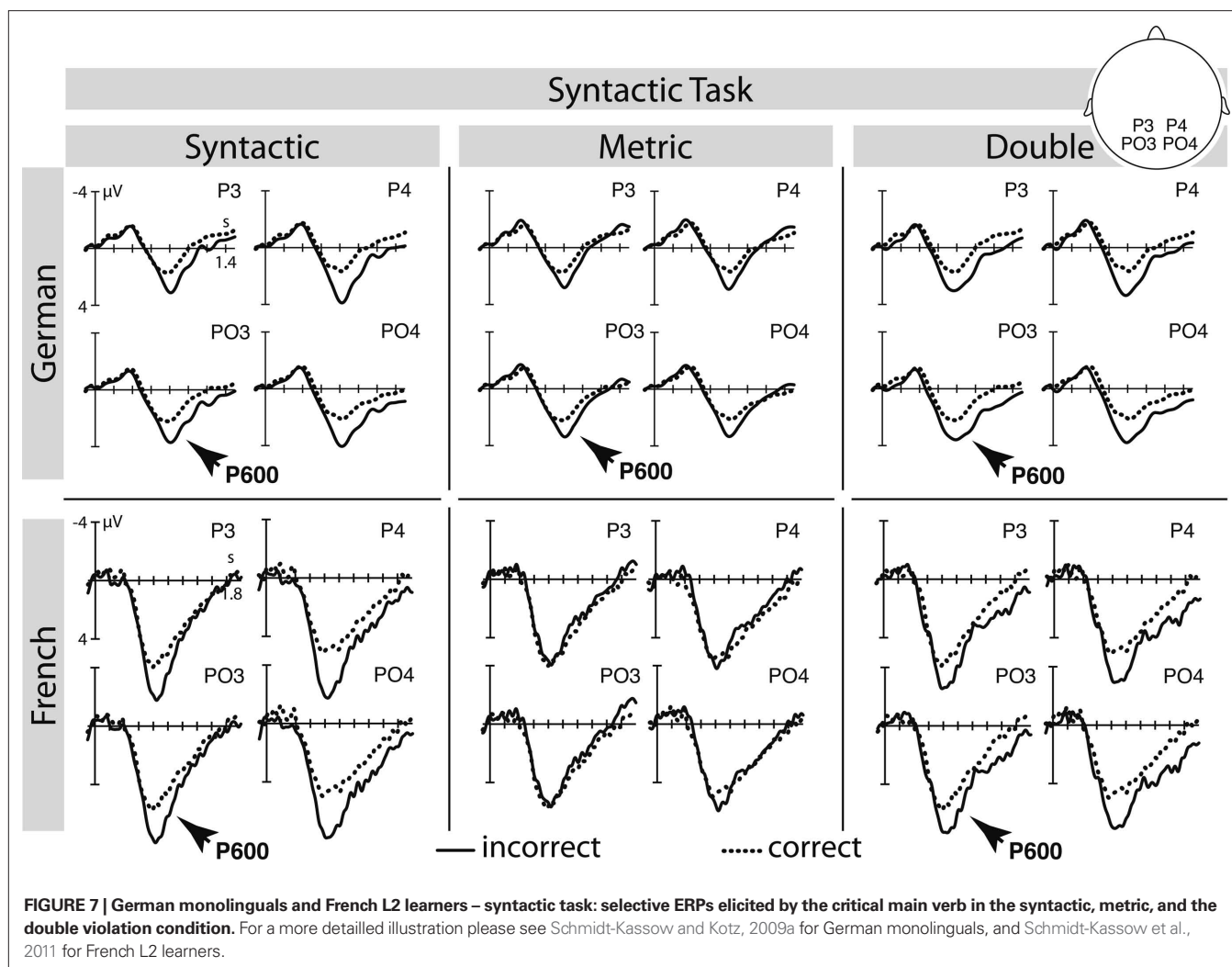
distributed negativity in response to metric violations in the *syntactic task*. This is in contrast to French L2 learners of German, who failed to show a response to metric violations. Hence, Spanish L2 learners were sensitive to trochaic units although they focused on the grammaticality of the sentences.

In the *metric homogeneity task*, a late posterior positivity (late P600) was evoked in response to syntactic and metric violations in the Spanish L2 learners, while double violations elicited an anteriorly distributed long-lasting negative shift. Under this task instruction French L2 learners of German had failed to show any significant ERP effect.

Given the current ERP evidence, Spanish L2 learners of German show comparable results to German monolinguals. However, there are noticeable latency and distributional differences in the implicit metric condition in the syntactic task and the double violation condition in the metric task. In the implicit metric condition we found a posterior negativity although a P600 was expected. We interpret this result as a first recognition of metric expectancy violations although attention was not directed to metric processing. Even in German native speakers, the amplitude of the metric P600 is larger when attention is directed to the metric pattern rather than to syntactic structure (see **Figures 6 and 7**). It is therefore remarkable that Spanish L2 learners of German show an ERP in response to implicit metric processing at all. Interestingly, Kotz et al. (2008) provided ERP evidence on early Spanish learners of English that in parts resembles our pattern. Even though they were early learners of English, the early negativity in response to syntactic violations in Spanish natives had a centro-parietal maximum (as it is the case for metric violations in the current experiment). A second somewhat unexpected ERP component concerns the anterior negativity in response to the double violation under metric task conditions. However, a closer look at the behavioral data reveals that participants performed worst in evaluating this particular condition (78.8% correct). This might be due to interfering cues in this condition. Here, metric and syntactic incongruencies are combined, however, subjects were asked to ignore syntactic violations and focus on metric incongruencies. Obviously, subjects had difficulties in ignoring syntactic violations, given that syntactic violations are particularly salient. Hence, syntactic violations under metric task conditions could have been irritating. We thus interpret the anterior negativity as a electrophysiological correlate of conflicting cues in the signal. In line with Schröger and Wolff (1998) we argue that

FIGURE 6 | German monolinguals and French L2 learners – metric task: selective ERPs elicited by the critical main verb in the syntactic, metric, and the double violation condition. For a more detailed illustration please see Schmidt-Kassow and Kotz, 2009a for German monolinguals, and Schmidt-Kassow et al., 2011 for French L2 learners.

FIGURE 7 | German monolinguals and French L2 learners – syntactic task: selective ERPs elicited by the critical main verb in the syntactic, metric, and the double violation condition. For a more detailled illustration please see Schmidt-Kassow and Kotz, 2009a for German monolinguals, and Schmidt-Kassow et al., 2011 for French L2 learners.

this negativity is a correlate of re-orientation from task-irrelevant to task-relevant aspects of a stimulus. In the current experiment re-orientation seems to spend cognitive resources as participants selectively failed to show a significant P600 in response to the double violation condition although a P600 is present in the metric violation condition.

Furthermore, it is striking that in Spanish L2 learners all P600 components evoked in the current experiment are considerably later compared to the previously reported results from German monolinguals. Snijders et al. (2007) have provided neurophysiological evidence that English adults show similar but delayed and reduced segmentation responses to Dutch stimuli compared to Dutch L1. Hence, L2 segmentation is qualitatively different although Dutch and English are highly similar languages. Hence, we argue that our current ERP results may reflect similar delayed segmentation effects in a foreign language. As elaborated in the introduction successful segmentation is a precondition for lexical access or syntactic processing. If segmentation is slowed down in L2 processing it seems plausible that other linguistic processes are likewise delayed (e.g., syntactic processing). Our syntactic ERP results are comparable to results from low proficient Italian learners

of German as reported by Rossi et al. (2006). The authors reported that low proficient show a delayed P600 component in response to agreement violations in a time window comparable to our results (900–1500 ms). Hence, delayed latencies seem to be a common phenomenon in L2 research (e.g., Hahne, 2001; Moreno and Kutas, 2005; Dowens et al., 2010).

One might also argue, that metric violations, particularly in the ears of late L2 learners, may be less noticeable than syntactic violations and hence latency differences between the metric and the syntactic task are due to saliency differences between conditions. Indeed based on the participants' performance, it seems that both violation conditions are equally salient to German monolinguals (detection of syntactic violations: 99% correct, detection of metric violations: 96% correct; Schmidt-Kassow and Kotz, 2009a), but not to Spanish L2 learners of German. Although Spanish participants were confident in making judgments in both tasks performance was lower in the metric than the syntactic task (detection of syntactic violations: 98% correct, detection of metric violations: 86% correct). Hence, one may suspect that ERP latency differences may be driven by differences in saliency and behavior. However, latency differences as a function of task are also found in German

monolinguals but here in the reversed order (Schmidt-Kassow and Kotz, 2009a): The P600 in response to metric violations deflected about 300 ms earlier compared to the syntactic P600. Furthermore, we found extremely varying latencies across language groups in the syntactic task (syntactic P600: German monolinguals: 750–1050, French L2: 700–1100, Spanish L2: 900–1500) despite comparable behavioral performance in this particular condition. We thus argue that saliency differences alone cannot explain the observed latency differences. This issue has undoubtedly to be followed in future studies.

As elaborated in the introduction we aimed to investigate whether Spanish L2 learners show comparable ERP results to French L2 learners given that both languages belong to the same rhythm class and may therefore lead to insensitivities to trochaic units in speech. The following table (**Table 2**) provides an overview of the results from the current study compared to the results from French L2 learners of German and German monolinguals.

The detailed summary reveals that electrophysiological data of Spanish L2 learners of German dramatically differ from that observed in French L2 learners of German. Although the ERP pattern differs in latency and distribution from German monolinguals, Spanish late learners are responsive to the same events as native German speaker (i.e., to metric as well as syntactic violations). Compared to the German and the Spanish data, French native speakers showed an altogether different pattern (selective reduction for certain ERP components). In those instances where the French speakers showed an electrophysiological P600 response the latency was similar to native speakers of German. Hence, Spanish L2 learners of German are sensitive to L2 sentence violations across the board, while French L2 learners are not. This is even true given the fact that Spanish L2 learners had acquired German at a later stage in life. Thus, we have reasons to believe that Spanish early L2 learners of German should show a German native-like ERP patter. Given that French does not carry contrastive stress, French L2 learners may

potentially develop selective "stress deafness" (see Dupoux et al., 2008) during first language acquisition. If stress deafness is indeed developed during L1 acquisition, it seems to persist and affect L2 acquisition, resulting in an insensitivity toward pitch cues in the L2 perception of stress-timed languages. However, Spanish learners are able to use pitch cues and to detect metric irregularities in German. Furthermore, they also show electrophysiological responses under implicit task demands, i.e., they process metric and syntactic errors even if attention is not directed to these processes via task demands.

One has to keep in mind, that the current participants learned German relatively late in high school. This is in contrast to our previous experiments with French L2 learners of German who acquired German at around 12 years of age. Hence, it is not surprising, that ERPs elicited in French natives deflected earlier compared to the ERP responses reported in the current experiment (for a recent review see Kotz, 2009). In the context of the current and our previous studies, the latency of ERP responses seems to reflect L2 onset: French learners showed only a very selective ERP response but with native-like latency, while Spanish learners showed responses to all violation conditions, but with a delayed latency. Other factors that have been intensively studied in the literature do not vary systematically between French and Spanish late learners of German, i.e., comparable cross-linguistic similarity (e.g., Tokowicz and MacWhinney, 2005) and both groups have similarly high proficiency (e.g., Ojima et al., 2005; Rossi et al., 2006). We particularly controlled proficiency levels by means of an extensive questionnaire in which French as well as Spanish participants were asked to indicate how confident they are in reading, listening, writing, or speaking German (on a 10 point Likert scale). Here, medians of Spanish natives do not differ significantly from French natives neither in reading (Spanish = 7, French = 8; $p = 0.077$), writing (Spanish = 7, French = 7; $p = 0.437$), listening (Spanish = 8, French = 9; $p = 0.270$), or speaking (Spanish = 7, French = 7; $p = 0.347$). However, our Spanish group provided better behavioral

Table 2 | Comparison of ERP data from previous studies with recent results.

Task	Condition	German (Schmidt-Kassow and Kotz, 2009a)	Spanish (current study)	French (Schmidt-Kassow et al., 2011)
Syntactic	Syn	Early anterior negativity (500..650 ms)	Delayed positivity (900..1500 ms)	Late posterior positivity (700..1100 ms)
		Late posterior positivity (750..1050 ms)		
	Met	Early anterior negativity (250..650 ms)	Posterior negativity (300..550 ms)	/
		Late posterior positivity (750..1050 ms)		
	Doub	Early anterior negativity (250..650 ms)	Delayed positivity (900..1500 ms)	Late posterior positivity
		Late posterior positivity (750..1050 ms)		
Metric	Syn	Early anterior negativity (600..850 ms)	Delayed posterior positivity (1300..1800 ms)	/
		Late posterior positivity (850..1150 ms)		
	Met	Early anterior negativity (200..850 ms)	Delayed posterior positivity (1500..1700 ms)	/
		Late posterior positivity (550..1150 ms)		
	Doub	Early anterior negativity (200..850 ms)	Late anterior negativity (750..1400 ms)	/
		Late posterior positivity (550..1150 ms)		

Syn, syntactic violation; Met, metric violation; Doub, double violation; ms, milliseconds.
/ indicates missing significant differences between conditions.

performance in the metric homogeneity judgment task than French late learners who again showed shorter ERP latencies. Hence, there is a conflict between behavioral performance and the electrophysiological response (longer latencies though better performance in Spanish natives). The ERP latency seems to be influenced by other factors than proficiency or cross-linguistic similarity, such as for instance the age of L2 onset (Dowens et al., 2010).

Furthermore, we provide evidence for implicit syntactic knowledge in Spanish L2 learners of German in contrast to French L2 learners. This is particularly interesting as (i) our results from German monolinguals show a strong interaction between meter and syntax processing, and (ii) French native speakers failed to show an ERP response to both, meter and implicit syntax. It would be premature to conclude that metric competence in a second language is a prerequisite for implicit syntactic processing capacity, though. In future studies we have to follow-up on this question by successively testing L2 learners of German at different learning stages and different ages of acquisition. Hence, we would be able to distinguish between effects that are based on later ages of acquisition and effects that are primarily based on the time interval a second language has been spoken.

The main goal of the current study was to investigate whether Spanish native speakers are able to detect trochaic units in a continuous speech stream that are primarily realized via pitch. Previous studies have already shown that Spanish native speakers use pitch as a cue for speech segmentation (Toro et al., 2009). However, as Spanish is a syllable-timed language, the syllable but not the trochaic unit is the primary segmentation unit (Sebastián-Gallés et al., 1992). Our

results indicate that Spanish L2 learners of German detect metric regularity in sentences, as evidenced by the current ERP and behavioral data. Hence, they differ from French L2 learners, who were not able to extract trochaic units (Schmidt-Kassow et al., 2011). We thus suggest that the insensitivity to trochaic units in French is not specific to speakers from syllable-timed languages, but rather results from stress deafness in French natives compared to German or Spanish natives. In this context, it may be particularly interesting to test French natives which are not completely insensitive to stress such as simultaneous French–German bilinguals (Dupoux et al., 2010). If these bilinguals are in fact sensitive to pitch cues they should provide a similar ERP pattern as Spanish L2 learners tested here.

We conclude that pitch sensitivity is a critical feature to learn a stress-timed language such as German. Based on the current and previous results the successful use of trochaic units seems to be a precondition for native-like (implicit) syntactic processing. Hence, the reported data provide further evidence for the close connection of meter and syntax.

REFERENCES

Auer, E. T. (1993). *Dynamic Processing in Spoken Word Recognition: The Influence of Paradigmatic and Syntactic States.* Ph.D. thesis, States University of New York at Buffalo, Buffalo.

Best, C. T., McRoberts, G. W., and Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *J. Acoust. Soc. Am.* 109, 775–794.

Cutler, A. (1994). The perception of rhythm in language. *Cognition* 50, 79–81.

Cutler, A., Mehler, J., Norris, D., and Segui, J. (1989). Limits on bilingualism. *Nature* 340, 229–230.

Cutler, A., Mehler, J., Norris, D., and Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cogn. Psychol.* 24, 381–410.

Cutler, A., and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 113–121.

Cutler, A., and Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *J. Acoust. Soc. Am.* 105, 1877–1888.

Dowens, M. G., Vergara, M., Barber, H. A., and Carreiras, M. (2010).

Morphosyntactic processing in late second-language learners. *J. Cogn. Neurosci.* 22, 1870–1887.

Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: stress "deafness" in simultaneous French-Spanish bilinguals. *Cognition* 114, 266–275.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., and Peperkamp, S. (2008). Persistent stress "deafness": the case of French learners of Spanish. *Cognition* 106, 682–706.

Eisenberg, P. (1991). Syllabische Strukturen und Wortakzent: Prinzipien der Prosodik deutscher Woerter. *Z. Sprachwissenschaft* 10, 37–64.

Féry, C. (1997). Uni und Studis: Die besten Wörter des Deutschen. *Linguistische Ber.* 172, 461–490.

Flege, J. E., Munro, M. J., and MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *J. Acoust. Soc. Am.* 97(Pt 1), 3125–3134.

Flege, T. J. E., Frieda, E. M., and Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *J. Phon.* 25, 169–186.

Goyet, L., de Schonen, S., and Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-

learning infants: an ERP study. *Brain Res.* 1332, 75–89.

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

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *J. Psycholinguist. Res.* 30, 251–266.

Jusczyk, P. W. (2002). Some critical developments in acquiring native language sound organization during the first year. *Ann. Otol. Rhinol. Laryngol. Suppl.* 189, 11–15.

Kim, J., Davis, C., and Cutler, A. (2008). Perceptual tests of rhythmic similarity: II. Syllable rhythm. *Lang. Speech* 51(Pt 4), 343–359.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain Lang.* 109, 68–74.

Kotz, S. A., Holcomb, P. J., and Osterhout, L. (2008). ERPs reveal comparable syntactic sentence processing in native and non-native readers of English. *Acta Psychol. (Amst.)* 128, 514–527.

Lee, C. S., and Todd, N. (2004). Towards an auditory account of speech rhythm: application of a model of the auditory "primal sketch" to two multi-language corpora. *Cognition* 93, 225–254.

MacKay, I. R., Meador, D., and Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica* 58, 103–125.

Mattys, S. L., White, L., and Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *J. Exp. Psychol. Gen.* 134, 477–500.

Moreno, E. M., and Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish-English bilinguals. *Brain Res. Cogn. Brain Res.* 22, 205–220.

Murty, L., Otake, T., and Cutler, A. (2007). Perceptual tests of rhythmic similarity: I. Mora rhythm. *Lang. Speech* 50(Pt 1), 77–99.

Nazzi, T., Iakimova, G., Bertoncini, J., Fredonie, S., and Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: emerging evidence for crosslinguistic differences. *J. Mem. Lang.* 54, 283–299.

Nazzi, T., and Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Commun.* 4, 233–243.

Norris, D., McQueen, J. M., and Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *J.*

Exp. Psychol. Learn. Mem. Cogn. 21, 1209–1228.

Ojima, S., Nakata, H., and Kakigi, R. (2005).An ERP study of second language learning after childhood: effects of proficiency. *J. Cogn. Neurosci.* 17, 1212–1228.

Otake, T., Yoneyama, K., Cutler, A., and van der Lugt, A. (1996). The representation of Japanese moraic nasals. *J. Acoust. Soc. Am.* 100, 3831–3842.

Piske, T., Flege, J. E., MacKay, I. R. A., and Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica* 59, 49–71.

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

Ramus, F. (2002). Language discrimination by newborns: teasing apart phonotactic, rhythmic, and into national cues. *Annu. Rev. Lang. Acquis.* 2, 85–115.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., and Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288, 349–351.

Rossi, S., Gugler, M. F., Friederici, A. D., and Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *J. Cogn. Neurosci.* 18, 2030–2048.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.

Sanders, L. D., Neville, H. J., and Woldorff, M. D. (2002). Speech segmentation by native and non-native speakers: the use of lexical, syntactic, and stress-pattern cues. *J. Speech Lang. Hear. Res.* 45, 519–530.

Sansavini, A. (1997). Neonatal perception of the rhythmical structure of speech. *Early Dev. Parent.* 6, 3–13.

Schmidt-Kassow, M., and Kotz, S. A. (2009a). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *J. Cogn. Neurosci.* 21, 1693–1708.

Schmidt-Kassow, M., and Kotz, S. A. (2009b). Attention and perceptual regularity in speech. *Neuroreport* 20, 1643–1647.

Schmidt-Kassow, M., Rothermich, K., Schwartze, M., and Kotz, S. A. (2011). Did you get the beat? Late proficient French-German learners extract strong-weak patterns in tonal but not in linguistic sequences. *Neuroimage* 54, 568–576.

Schröger, E., and Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related potentials. *Neuroreport* 9, 3355–3358.

Sebastián-Gallés, N., Dupoux, E., Segui, J., and Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *J. Mem. Lang.* 31, 18–32.

Snijders, T. M., Kooijman, V., Cutler, A., and Hagoort, P. (2007). Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Res.* 1178, 106–113.

Tokowicz, N., and MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Stud. Second Lang. Learn.* 27, 173–204.

Toro, J. M., Sebastian-Galles, N., and Mattys, S. L. (2009). The role of perceptual salience during the segmentation of connected speech. *Eur. J. Cogn. Psychol.* 21, 786–800.

Tyler, M. D., and Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *J. Acoust. Soc. Am.* 126, 367–376.

White, L., Melhorn, J., and Mattys, S. L. (2009). Segmentation by lexical subtraction in Hungarian speakers of second-language English. *Q. J. Exp. Psychol. (Colchester)* 63, 544–554.

Yu, V. Y., and Andruski, J. E. (2010). A cross-language study of perception of lexical stress in English. *J. Psycholinguist. Res.* 39, 323–344.