The Nature of Morphosyntactic Processing during Language Perception. Evidence from an Additional-Task Study in Spanish and German.

Annette Hohlfeld[a], Manuel Martín-Loeches[a], and Werner Sommer[b]

[a]Center for Human Evolution and Behavior, UCM-ISCIII, Madrid

[b]Humboldt-University at Berlin

Address correspondence to:

Annette Hohlfeld

Neurologic Rehabilitation Center Wolletzsee

Language Therapy Department

Zur Welse 2, 16278 Wolletz

Tel.: ++49-(0)33337-49-453

E-Mail: annette.hohlfeld@yahoo.de

Abstract (250 words)

The present study investigates in how far morphosyntactic processing is affected by an additional non-verbal task and whether this effect differs between German and Spanish, two languages with differences in processing grammatical gender (lexical vs. cue-based processing). By manipulating task load and language we aimed at getting an insight into subprocesses of morphosyntax and their dependence on resources of general and verbal working memory, respectively. In more general terms, this study contributes to the debate on the modularity of morphosyntax. Written German or Spanish sentences with or without gender violations were presented word by word to native speakers. The critical words temporally overlapped in different degrees with a non-linguistic stimulus (a high or low tone). In a single task (Experiment 1) participants judged sentence acceptability and ignored the tones. Experiment 2 required a response to the tones. Left-anterior negativity (LAN) and P600 components were analyzed in the ERPs to critical words. Whereas the LAN was not affected by any of the experimental manipulations, the P600 was modulated as a function of language during the single task conditions (Experiment 1). In Experiment 2 the additional task did not add up with this effect; instead, the differences between language groups vanished. This may indicate that the processes reflected in the P600 draw on resources of general working memory. The LAN data seem to be in line with modularity of first pass morphosyntactic processing, although this interpretation might contradict findings from other studies. The P600 results may highlight the flexibility of sentence-based syntactic processing.

Keywords: additional task, gender congruency, modularity, interfaces, language perception, LAN, P600

Declarations of interest: none

The modularity of language

There is a continuing debate on the nature of morphosyntactic knowledge and processing with respect to the question in how far it is modular or more strongly encapsulated than other parts of linguistic knowledge. The notion of modularity of mental functions goes back to Fodor (1983) and was defined by a number of criteria. As summarized by Jackendoff (2002) these involve specialized content, automaticity, susceptibility to focal brain damage and evidence for innateness. Indeed, our language system seems to meet most, if not all, of these criteria (Pinker, 1994). Phonological and syntactic structures have to be processed to get from external signals to concepts and vice versa. Jackendoff (2002) considers our ability to process these structures as qualitatively different from anything else in the mind. Furthermore, there is a large amount of empirical evidence about the automaticity of language perception, e.g. the persistence of the Stroop effect (MacLeod, 1991; Zahedi et al., 2019), or the accuracy and efficiency of language production (cf. Garrett, 2000). Susceptibility to focal brain damage, causing severe language production and perception deficits, was first reported by Broca (1865) and Wernicke (1874), respectively; for a recent review of an involvement of the left inferior frontal gyrus in morphosyntactic processing cf. Kemmerer (2015, Chapt. 13). With respect to innateness it was Chomsky (1981) who suggested that there is a kind of innate toolkit, termed *Universal Grammar*, a language acquisition device, which enables every normal child to learn any human language.

The processing of syntax or morphosyntax has always been a particular testcase. With respect to the role of verbal working memory during sentence perception, Caplan and Waters (1999) suggested a separate verbal working memory subsystem that stores syntactic structures during language perception. This idea was based on studies in which participants had to process sentences of varying syntactic complexity while keeping a verbal load (digit recall task) within short-term memory. Interference effects between the two tasks only occurred when storing the digits interrupted the sentence task, hence requiring an attention shift. When digits were presented prior to the sentence, no interference effects occurred. The authors concluded that syntactic processing and the recall task did not compete for the same resources and suggested an independent subsystem for processing syntactic structures. In contrast, Fedorenko, Gibson and Rohde (2006) provided evidence against a specific pool of resources for sentence processing. In their experiment participants had to read sentences of varying complexity and additionally memorized one, two, or three nouns that either occurred within the sentence they were just reading or not. An interaction between syntactic complexity and difficulty of the additional task (memorize dissimilar nouns) was observed in the hard condition when three nouns had to be memorized. From these findings the authors concluded that sentence comprehension and the additional task draw on the same resource pool in working memory.

More evidence for the special status of syntax comes from second-language learning. Whereas syntactic analysis is very sensitive to age of acquisition (Weber-Fox and Neville, 1996, 2001), semantic processing appears to be independent of acquisition age. Once critical age periods are passed, it is impossible for learners to achieve native speaker competence with respect to syntax. Scherag, Demuth, Rösler, Neville, and Röder (2004) reported that native speakers who do not use their native language due to immigration into a different country still retain their morphosyntactic abilities.

The aim of the present study

The present study aims at exploring which phases of morphosyntactic processing are affected by an additional non-linguistic cognitive operation. If morphosyntactic processing is modular and encapsulated in the language system, we would not expect any modulations by the additional task. However, if there are subprocesses that interfere with general cognition, the additional task should have a modulating effect. This was investigated in speakers of two different languages - German and Spanish. These languages differ in their grammatical gender systems, i.e. in the overtness of gender cues. Whereas grammatical gender in Spanish is morphologically marked on each word in the noun phrase, involving a close succession of gender markers, grammatical gender in German is lexicalized and not morphologically marked (for further explanation see below). In a single task experiment (Experiment 1) participants were asked to read sentences and judge the acceptability of the sentence. Experimental sentences contained violations in gender congruency. An additional non-verbal stimulus (a synthetically generated high or low tone) was acoustically presented, but no responses were required. In the additional non-linguistic operation (dual task, Experiment 2) a response to the tone was required. The temporal overlap of the sentence acceptability task and the additional task was manipulated by varying the stimulus onset asynchrony (SOA), allowing for a precise control of the time-course of interference, similar to a Psychological Refractory Period (PRP) paradigm (e.g. Pashler, 1994). Morphosyntactic processing was assessed by recording syntax-relevant event-related brain potential (ERP) components – the left anterior negativity (LAN) and the P600. In the following we will discuss the interpretations of these components and review relevant literature. Finally, the present study and its manipulations are introduced in more detail and predictions are made with respect to possible modulations of the LAN and P600 by the experimental manipulations.

Syntax-related ERP components and their interpretations

Fine-grained measures of syntactic processing have been derived from ERPs, which were also employed in the present study. There is wide agreement that syntactic processing during sentence perception is reflected by at least two distinct components in ERPs - LAN and the P600. The LAN was first reported by Kluender and Kutas (1993) as an amplitude difference between ERPs to syntactically correct and incorrect sentence elements. Usually, the LAN is observed between 200 and 500 ms after word onset and shows a left anterior scalp topography (e.g. Coulson, King, and Kutas, 1998; Osterhout and Holcomb, 1992; Steinhauer and Drury, 2012). The P600 component is a long-lasting positivity, sometimes beginning as early as 200 ms, and commonly reaching its maximum between 600 and 800 ms after stimulus onset. Like the LAN it is elicited by syntactically incorrect sentence elements, but also by other configurations.

There is great diversity in the literature with respect to the interpretation of both the LAN and P600 components. Friederici, Hahne, and Mecklinger (1996) suggested the LAN to reflect an initial processing stage, during which syntactic structure is built up. Other researchers relate the LAN to verbal working memory load. For instance, Urbach (1993) showed that the distance between a main verb (predicate) and a noun phrase (subject) affects the occurrence of a LAN. A LAN was elicited by the main verb WAS following a long ungrammatical relative clause (e.g. The cook (helped in the kitchen) **was** busy), whereas no anterior negativity was measured in a similar sentence with a short ungrammatical relative clause like "The instructor (taught) **was** Spanish.". Similar results come from studies that investigated filler-gap constructions ("What did you put the book on __.": Kluender, Kutas, 1993; King and Kutas, 1995) or establishing a relation between a pronoun and a referent (Coulson, King, and Kutas, 1998). As proposed by Kluender and Kutas (1993) the LAN may reflect the parsers' backward search through verbal working memory to find a recently activated and appropriate entity. The further the parser has to go back, the larger the LAN. Coulson et al. (1998) pointed out that the LAN may index operations specific to verbal working memory. Kluender and Kutas (1993) mention that manipulations of working memory load in non-linguistic tasks (mental roatation, mental arithmetics, visio-spacial processing) elicit slow potential shifts with topographies different from the LAN´s (also cf. Martin-Loeches, Casado, Gonzalo, De Heras, and Fernández-Frías, 2006).

A completely different opinion about the origin of the LAN has been suggested by Osterhout (1997), Osterhout, McLaughlin, Kim, Greenwald, and Inoue (2004), Tanner and van Hell (2014), Tanner (2015, 2019); these authors propose that the LAN is an artifact of the spatio-temporal overlap between N400, generally held to indicate semantic processing (Kutas and Hillyard, 1984) and P600 due to variability in language processing between individuals. Whereas certain violations cause an N400 in some individual, these violations cause a P600 in others. In the grand means this looks like a biphasic response, which however does not correspond to any participant's individual response. However, as pointed out by Molinaro, Barber, and Carreiras (2011) anterior negativities are independent components that can be observed without the occurrence of the P600 and according to the authors are associated with an increased cost of morphosyntactic processing. Furthermore, it is hard to understand how two posterior components can artificially produce an anterior one.

Likewise, different interpretations have been suggested for the P600. Friederici, Hahne, and Mecklinger (1996) suggested the P600 to reflect a second stage of reanalysis (see also Kutas, Van Petten, and Kluender, 2006) following the LAN. Kolk, Chwilla, Van Herten, and Oor (2003) and Kolk, Chwilla (2007) basically attribute the P600 to monitoring. Once an unexpected linguistic event is encountered, it is re-attended to check upon its veridicality. According to Kolk and colleagues this monitoring process depends on executive control. They showed that syntactic complexity affects the P600. However, it was pointed out that the monitoring process is not restricted to unexpected syntactic events but it likewise applies to unexpected semantic or even orthographic (i.e. spelling errors) unexpected events.

The impact of discourse on the P600 has been investigated by van Berkum, Brown, and Hagoort (1999) and van Berkum, Brown, Hagoort, and Zwitserlood (2005). Larger P600 were observed for sentences that were embedded in a discourse that allowed two referents. This was interpreted as a more extensive use of short- or long-term memory resources. In case of referential ambiguity two candidate referents must be maintained for an unresolved single referential slot. This imposed an additional load on working memory, which was reflected in the P600.

Coulson, King, and Kutas (1998) argued that the P600 belongs to the P3-family assuming that the P600 is an index of a domain-general mechanism. In their study the amplitude of the P600 varied as a function of the proportion of grammatical sentences relative to ungrammatical ones. Kuperberg (2007) suggests a multi-stream processing account when verb-argument relationships have to be understood. A semantic memory- based stream operates in parallel with a stream that combines the analysis of morphosyntactic structure and semantic-thematic constraints such as animacy. It is noted that these streams are affected by task and working memory. Moreover, these comprehension streams are believed not to be unique to language, but to also guide, for instance, visual comprehension. As recognized by the authors, this account is compatible with the notion of interfaces between different language compartments as well as between language and general cognition as assumed by Jackendoff (2002, 2007). Moreover Kuperberg (2007) agrees with Jackendoff on the position that sentence comprehension is not purely syntactically driven. Another multi-stream account was suggested by Hagoort (2009), who discussed that the P600 amplitude is modulated by the competition between alternative options, which are generated by the semantic and syntactic processing streams (cf. Bornkessel, Bornkessel-Schlesewski, 2008 for another two-stream model).

In contrast, Brouwer, Fitz, and Hoeks (2012) argued that there is no need to assume multiple processing streams. According to these authors the P600 reflects a single process of updating new information with a mental representation of what is to be communicated. This updating process is not language specific but is invoked also in other situations.

The question of modularity of syntax was investigated by using combined semantic and syntactic violations in the same target word (Hagoort, 2003), where an amplitude increase of the N400 component in the condition with double violations as compared to single semantic violations was observed. In contrast, there were no amplitude differences between single syntactic and double violations in the P600. Thus, whereas semantic processing was affected by syntactic processing, the converse was not the case. It was suggested that semantic processing is open (altruistic) whereas syntactic processing is encapsulated (egoistic). However, Martín-Loeches, Nigbur, Casado, Hohlfeld, and Sommer (2006) could not replicate these findings. In their study, where violations occurred in non-final sentence positions, the N400 amplitude was not boosted by syntactic violations. The P600 amplitude was reduced in case of double violations compared to single syntactic violations. The authors suggested that semantics plays a prevailing role during sentence interpretation. Once the parser has detected the semantic anomaly of a word no further resources are invested in assessing its syntactic role.

Another paradigm to investigate whether the underlying cognitive operations proceed independently and in parallel or whether the tasks interact due to, for example, sharing of resources or requiring simultaneous access to the same cognitive process, is the dual task paradigm (e.g. Navon and Miller, 2002, Pashler, 1994). This paradigm was applied, for example, in a study by Martín-Loeches, Abdel-Rahman, Casado, Hohlfeld, Schacht, and Sommer (2009), where participants had to make both acceptability judgements for Spanish sentences and to memorize a spoken word, which was presented in temporal overlap with the sentence. In the syntactic condition sentence acceptability was manipulated by violating gender agreement between the noun and the succeeding adjective. The acoustically presented word was an adjective that was meaningless with respect to the written sentence and could be gender congruent or incongruent with the noun in the sentence. This linguistic sentence-extraneous information did not modulate the LAN but affected the P600, supporting the idea that the processes underlying LAN and P600 are different. As mentioned above the present study also applied the dual task paradigm to investigate the nature of morphosyntactic processing. Next the paradigm of the present study will be introduced in more detail.

<u>The current study: manipulations and predictions</u>

The present study comprised two experiments – Experiment 1 was a single task experiment (only judging sentence acceptability), which served as a baseline with respect to the language sensitive ERP components and Experiment 2, the actual dual task experiment. The dual task paradigm employed in Experiment 2 varied the temporal overlap (stimulus onset asynchrony, SOA) of the sentence acceptability task and the additional task, to control the time-course of interference. In order to hold constant the intervals between all words within a sentence, all inter-stimulus intervals (ISIs) were adjusted to the SOA between the tone (S1) and the word, in which the gender violation could occur (S2, see Figure 1). Thus, the ISIs between the words of a given sentence varied according to the SOA, being either 200, 500, or 800 ms (in the single and dual task). These factors were manipulated within participants. Language, i.e. differences in the gender marking systems, to be explained below, was a group factor.

German nouns fall into three gender classes – masculine, feminine and neuter gender (*masc*, *fem*, and *neut*). Although attempts have been made to determine formal gender cues in terms of suffixes (Köpcke, 1982), German nouns are generally believed to be inherently marked for gender (Eisenberg, 1994). Thus, gender information is part of the lexical entry and retrieved from the mental lexicon when needed, for instance, when selecting the appropriate definite determiner *der, die*, or *das* on which gender is marked in the singular (Hohlfeld, 2006; Levelt, Roelofs, Meyer, 1999; Shantz and Tanner, 2017; van Turenout, 1998). In contrast, there are only two genders in Spanish – masculine and feminine – and the majority of nouns can be assigned to the feminine or masculine gender class, respectively, on the basis of their terminations in *–a* or *–o* (rule-based assignment rather than lexical retrieval). However, in Spanish there are also some exceptions to these rules. As shown by Gollan and Frost (2001) for Hebrew and by Desroches and Paivio (1990) as well as Taft and Meunier (1998) for French, formal gender cues can facilitate gender assignment or gender identification if those cues are reliable indicators of a particular gender in a language. For

Spanish one might suggest that due to the large number of reliable cues, gender can be computed on the basis of these cues and is lexically stored only in case of exceptions.

Whereas gender congruency in German has to be established between determiner and noun in singular only, in Spanish not only determiners but also adjectives are morphologically inflected for gender in both singular and plural (*Die*$_{\text{fem}}$ *rote Rose verblüht./ The*$_{\text{fem}}$ *red rose withers.* vs. *La*$_{\text{fem}}$ *rosa*$_{\text{fem}}$ *roja*$_{\text{fem}}$ *se marchita./ The*$_{\text{fem}}$ *rose*$_{\text{fem}}$ *red*$_{\text{fem}}$ *withers..// Der*$_{masc}$ *rote Hut passt/ The*$_{masc}$ *red hat fits.* vs. *El*$_{masc}$ *sombrero*$_{masc}$ *rojo*$_{masc}$ *encaja./ The*$_{masc}$ *hat*$_{masc}$ *red*$_{masc}$ *fits.*). Both languages also differ in word order as illustrated in the previous examples. Whereas in German adjectives precede the noun, the canonical word order in Spanish is adjective following the noun. In the experimental sentences used in the present study gender congruency violation in German sentences occurred in the noun (*Der*$_{masc}$ *rote Rose*$_{\text{fem}}$ *verblüht.);* in Spanish gender violations occurred in the adjective (*La*$_{\text{fem}}$ *rosa*$_{\text{fem}}$ *roj-o*$_{masc}$ *se marchita.).* Thus, words in which gender violations occurred (target words) differed in word category – nouns versus adjectives. Words preceding the target also differed in word category and gender transparency – whereas German adjectives do not carry formal gender cues, Spanish nouns do. Furthermore, the distance between formally gender-marked words varies between languages, which poses a different load on verbal working memory. In German the adjective is formally unmarked for gender. Thus gender information, which is encoded by the definite article (the first word in the noun phrase), has been maintained in verbal working memory until the occurrence of the noun (the third word in the noun phrase). In Spanish the word immediately preceding the critical word also carries a formal gender cue, which poses less effort on verbal working memory.

Which predictions can be made with respect to the LAN? The predictions strongly depend on the interpretation of the component to which one adheres. If one sticks to the interpretation of the LAN as to reflect verbal working memory load during parsing one might assume SOA effects on the LAN already in the single task (Experiment 1) with small or absent LAN at SOA 100 and larger LANs at SOA 700 (please keep in mind that the ISI was adopted according to the SOA). Moreover, in the single task there should be an effect of language on the LAN because German and Spanish vary with respect to the distance between the gender-marked words. To check for congruency in Spanish the parser has to search backwards only one word (*La*$_{\text{fem}}$ *rosa*$_{\text{fem}}$ *roj-o*$_{masc}$ *se marchita.)*, whereas in German it has to go back two words until it reaches the determiner (*Der*$_{masc}$ *rote Rose*$_{\text{fem}}$ *verblüht.)*. According to the idea of Kluender and Kutas (1993) this should affect the LAN amplitude – yielding larger amplitudes in the German compared to the Spanish speakers. One might even imagine an interaction of Language and SOA, with smallest LAN-amplitudes for Spanish at SOA 100 in the single task. According to the literature the LAN might depend on verbal working memory, but it is generally not believed to depend on resources of non-linguistic working memory. Consequently, effects observed in the single-task Experiment 1 should not change when an additional non-linguistic task has to be performed as in Experiment 2. In contrast, if one adheres to the interpretation that the LAN reflects an automatic phase of building up syntactic structure (Friederici et al., 1996), no noticeable effects of Language, SOA in the single task and no effects of the additional task are to be expected.

The predictions on the P600 are more straightforward. Whereas the literature cited above disagrees on the question, whether there is more than one processing stream underlying the P600 it agrees that the component depends on working memory capacity and is not restricted to language, i.e. to verbal working memory. Therefore, the following predictions can be made for the present experiments: Effects of Language and SOA should be measurable already in Experiment 1 (the single task) as these factors manipulate verbal working memory load. Processing the additional non-linguistic task (effects of SOA in experiment 2) poses extra costs on general working memory. If verbal working memory effects from Experiment 1 call on a separate resource, effects of the additional task should be additive on the effects of the single tasks. If there is not a separate resource pool for verbal working memory the additional task should modulate the pattern of results from the single task.

In sum, we expect different modulations of LAN and P600 by the additional task. There might be no differences between both components in the single task, if one assumes that both components are sensitive to verbal working memory load. In the additional task both components should dissociate, because this task increases load in general working memory. To our knowledge nothing is known with respect to effects of an additional non-linguistic task on processing grammatical gender in different gender systems. If the detection of gender incongruencies in German and Spanish differs in processing costs (due to the different gender systems, as discussed above) and if the sentence acceptability and the additional task draw on a common resource pool, the processing of an additional task might differentially modulate this effect in German and Spanish. If tasks draw on separate resource pools we expect additive effects on the P600.

Experiment 1: Single Task

It was one aim of this single task experiment to explore whether the present material (sentences with gender incongruencies) would induce a LAN and a P600. Moreover it was of interest whether these components would differ between German and Spanish due to the differences in the gender systems. Because the same material and stimulation conditions were used as for the dual-task Experiment 2 (where responses had to be given to the additional tone) Experiment 1 also provided baseline data for Experiment 2 with respect to the question, whether LAN and P600 are affected by a variable SOA between the tone and the critical word, which yielded a variable ISI between words. As shown by Hohlfeld, Martín-Loeches, and Sommer (2015) the late parts of the P600 (induced by semantic violations) were modulated by SOA already in a similar single task paradigm.

Method

Participants

In this experiment 16 native German speakers (12 women; mean age: *M* = 23.8, range 20 to 42 years) and 18 native Spanish speakers (17 women; mean age: *M* = 21.83, range 18 to 34 years) took part. The experiments were conducted in Madrid with Spanish students or incoming German exchange students or with Spanish post graduates. Participation was remunerated with money for the German and with credit points for the Spanish speakers. According to the Oldfield handedness inventory (Oldfield, 1971) all participants but one were right-handed.

Materials

The non-linguistic stimuli (S1) were synthetically generated tones of either 600 or 300 Hz and of 100 ms duration. The critical material for the language task consisted of 240 correct sentences. In case of the German sentences the structure was always determiner-adjective-noun-verb - [Det]+[Adj]+[N]+[V]. Spanish sentences always followed the word order [Det]+[N]+[Adj]+[V]. From these correct sentences 240 incorrect versions were generated by violating congruency of grammatical gender.

In German sentences, gender violations always occurred between singular determiner and noun (e.g. Der virtuose Pianist probt./$\textit{The}_{\textit{masc}}$ *talented piano* $\textit{player}_{\textit{masc}}$ *is practicing*. – **Die** virtuose **Pianist** probt./$\textit{The}_{\textit{fem}}$ *talented piano* $\textit{player}_{\textit{masc}}$ *is practicing;* see Appendix for further examples). The gender violation became evident at the moment the noun was processed, which therefore is the critical word (CW) to which ERPs were synchronized. Consequently, CWs did not differ between correct and incorrect versions; therefore, we refrain from reporting any frequency measures or other lexical parameters. The proportion of masculine, feminine and neuter gender in our 240 correct experimental sentences corresponds to the distribution of the three genders in the German language, namely 50 %, 30%, and 20%, respectively, with 120, 72, and 48 corresponding sentences (Bauch, 1971). Incorrect versions of these sentences were derived by assigning the two other possible genders in equal proportion. From the 120 correct sentences containing masculine nouns 60 incorrect sentences with neuter and 60 incorrect sentences with feminine nouns were created. This logic was applied to the correct feminine and neuter sentences as well.

In the Spanish sentences congruency in grammatical gender was violated between noun and adjective (e.g. La ardilla lista trepa./*The*$_{fem}$ *squirrel*$_{fem}$ *clever*$_{fem}$ *climbs*. – **La** ardill**a** list**o** trepa./*The*$_{fem}$ *squirrel*$_{fem}$ *clever*$_{masc}$ *climbs.*), in either singular or plural. Consequently, ERPs were synchronized to the adjective. Nouns and adjectives marked for masculine and feminine gender by *–o* and *–a*, respectively, were preferred to unmarked nouns (terminating in *–e*). Due to the fact that the root adjective was the same in acceptable and unacceptable versions of the experimental sentences, the average number of letters in the adjectives was constant (*M* = 7.61, *SE* = 0.12). Word frequency for acceptable and unacceptable words was also held constant with *M* = 17.19 occurrences per million in case of acceptable words and *M* = 15.28 occurrences per million in case of unacceptable words (*SE*s = 1.94 and 1.78, respectively). Unacceptable adjectives were derived from acceptable ones, for example, by changing its masculine into its feminine form.

Although Spanish and German differed with respect to word category and overtness of gender in the second and third word of our experimental sentences, these sentences were structurally the same. When we compared brain potentials, we compared responses to the third word in the sentence, in which readers realized a violation in both languages.

In addition, 140 filler sentences with four or five words were constructed for both the German and the Spanish materials. Seventy filler sentences were correct, of the 70 incorrect filler sentences, 35 contained syntactic violations, whereas the other 35 contained semantic violations. The semantic violations were included to induce deep processing during reading and to prevent participants from scanning the sentences only for discrepant determiner-adjective-noun pairs. Syntactic violations in the filler sentences occurred either in number congruency between sentence subject and predicate or in word category. Sentences with semantic violations followed the structure of an experimental sentence and contained adjectives that were semantically incongruent with the noun. German and Spanish materials were kept as similar as possible. In fact, most German sentences were derived from the Spanish ones by translation.

Procedure

In the paradigm used here a tone (S1) and a visually presented CW (S2), that is, either the noun (in German material) or the adjective (in the Spanish material) of a sentence were presented at three specified SOAs (100, 400 or 700 ms). The CW was always the penultimate word in the experimental sentences. In Experiment 1 no response was required to the tone (no R1). In both experiments sentences required acceptability judgments after sentence termination (R2). The basic structure of an experimental trial is illustrated in Figure 1. A trial started with a fixation point in the centre of the computer screen. After an interval of 2 s the experimental sentence was visually presented word by word, with each word being shown for 300 ms. In order to hold constant the intervals between all words within a sentence, all inter-stimulus intervals (ISIs) were adjusted to the SOA between S1 and S2 for this sentence (see Figure 1). Therefore, the ISIs between the words of a given sentence varied according to the SOA, being either 200, 500, or 800 ms.

=============== Please insert Figure 1 about here =============

From the pool of 240 acceptable sentences and their 240 unacceptable counterparts two sets of 240 experimental stimuli each were created, containing 120 acceptable and 120 unacceptable sentences. Within each set, none of the sentences was repeated and appeared only in one of its versions (acceptable vs. unacceptable). SOA was randomly assigned to the sentences within a set, taking care that there was an equal amount of acceptable and unacceptable sentences within each SOA. Furthermore, the assignments of high and low tones to each sentence were counterbalanced. All responses had to be given within 2.5 s after the sentence final word, which was always terminated with a period. A new experimental trial began either immediately after the response, or if 2.7 s had passed since sentence offset.

Participants were seated in a dimly lit, sound attenuated chamber. Participants were instructed to fixate the centre of the screen. Half of the participants responded to acceptable sentences with their right hand and with their left to unacceptable ones; for the others this assignment was reversed. Preceding the experiment proper, 24 practice trials were repeatedly presented until the participant responded correctly in all trials. After every 38 trials (comprising 24 experimental and 14 filler trials) there was a short break.

EEG Recording and Data Analysis

The electroencephalography was recorded by means of 27 tin electrodes mounted within an electrode cap (ElectroCap International), using the Brain Vision Recorder. Scalp locations for electrodes according to the revised version of the 10/20 International System (American Electroencephalographic Society, 1991) were Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FC4, T7, C3, Cz, C4, T8, TP7, CP3, CP4, TP8, P7, P3, Pz, P4, P8, PO7, PO8, O1, O2, and the left mastoid (M1). All electrodes were initially referenced to the right mastoid (M2). Bipolar horizontal and vertical electrooculograms (EOG) were recorded for artifact monitoring. Electrode impedances were kept below 5 kOhms, using ECI™ electrode gel. The signals were recorded continuously with a band pass from 0.01 to 30 Hz and a sampling rate of 250 Hz.

By using the Brain Vision Analyzer the continuous recordings were first segmented into 2.2-s epochs starting 200 ms before the onset of the adjectives in the experimental materials. Artifacts were automatically rejected by eliminating epochs where any of the signals exceeded a range of 200 μV during the epoch. Offline, ocular corrections for blinks, vertical, and horizontal eye movements were made using the method described by Gratton, Coles, and Donchin (1983). A visual inspection of the epochs was also carried out, eliminating those epochs that still presented artifacts after the previous procedures, which were based on a rather wide range of criteria. We also eliminated epochs with errors, that is, acceptable sentences judged as unacceptable or unacceptable sentences judged as acceptable. After recalculating the data to an average mastoids reference, ERP averages were generated for each participant, electrode and experimental condition. All ERP waveforms were referred to a baseline, starting 100 ms prior to the target word, and depicted as well as analyzed in a time window of 1.5 s from target onset. For further analyses of the P600 a 7-Hz low pass filter was applied to reduce residual noise. Analyses with respect to the LAN were performed on the basis of the filter setting at recording (i.e. at a low pass filter of 30 Hz).

ERP amplitude measures were submitted to ANOVAs including Language (German vs. Spanish) as between-subjects factor, and repeated measures on SOA, Acceptability and Electrode site (taking into account the 27 above-mentioned electrodes). By applying the subtraction procedure of Osman and Moore (1993), between ERPs to unacceptable and acceptable sentences at each SOA, the pure ERP response to syntactic violations was extracted, devoid of any overlapping activity, for example, due to the tone stimulus. Error rates were also submitted to ANOVAs comprising the factors SOA and Acceptability as well as the group-factor Language. When appropriate, degrees of freedom were corrected according to Huynh and Feldt (1976). Due to the fact that responses were to be given after the sentences had ended, reaction times (RTs) cannot be interpreted as representing processing times for the critical words (CWs) and were therefore not considered here.

Results

Performance

Judgments of experimental sentences were significantly affected by the experimental factors SOA, $F(2, 64) = 3.55$, $p < .05$, and Acceptability, $F(1, 32) = 44.49$, $p < .001$, with mean error rates for short to long SOA of 5.81, 4.80, and 6.24% (*SE*s 0.62, 0.56, 0.53). Acceptable sentences were harder to judge than unacceptable ones (*M*s 7.72 vs. 3.50%, *SE*s 0.66 vs. 0.46, respectively). Furthermore SOA and Acceptability interacted significantly, $F(2, 64) = 3.48$, $p < .05$. Post hoc tests, comparing the acceptable and unacceptable conditions within each SOA revealed significant differences between error rates only for SOA 400 and 700, *F*s (1, 33) = 22.08 and 37.53, *p*s < .001, respectively, but not for SOA 100 (*F* = 2.85). The group-factor Language did not yield any significant results. The reader is referred to Table 1 for more descriptive data.

===============Table 1=====================

ERP Data

As depicted in Figures 2 and 3, in both language groups neuronal activity differed between the acceptable and unacceptable conditions in both a first time window of 400 to 600 ms as well as in later time windows between 650 to 1250 ms. According to the left-anterior and posterior scalp distributions of these effects (see Fig. 4) they can be identified as LAN and P600 components, respectively.

=============Insert Figure 2, 3 and 4 about here please=================

Statistical analyses (ANOVAs including all electrodes) were performed within the above-mentioned time ranges in steps of 100 ms. As shown in Table 2, Acceptability was significant between 400 and 600 ms as well as from 650 to 1250 ms, both as main effects and in interactions with Electrode. From Figure 3 one might assume effects of SOA on the LAN, but statistical analyses did not corroborate this impression. Moreover, these analyses confirmed that the LAN was not modulated by any of the experimental factors. In contrast, later parts of the P600 (starting at 950 ms) were modulated by SOA, visible in Figure 3 as increasing duration with increasing SOA, independent of language group. Significant main effects of the group-factor Language were observed in two consecutive time segments between 750 and 950 ms. This effect seems to be due to slightly smaller amplitudes in the German group. Significant interactions of the group-factor Language with either SOA or Acceptability, or both, occurred from 850 ms onwards. By taking a look at Figure 3 this interaction can be accounted for by smaller amplitudes in the German than in the Spanish group, particularly at short SOA.

================Table 2======================

Discussion

This experiment compared effects of the different gender marking systems in German and Spanish. Moreover, we obtained baseline data for the subsequent additional-task study (Exp. 2). It was of interest whether the chosen material would induce a LAN and a P600 in the ERP and whether the variable SOA between stimuli would affect these components.

At performance level, differences between gender marking systems did not lead to a difference in error rates. The variable SOA, however, affected performance as an interaction of Acceptability and SOA. As the post hoc tests – comparing the acceptable and unacceptable conditions within each SOA – revealed, it was easier to detect unacceptable (gender incongruent) sentences than acceptable sentences at the two longer SOAs. At the short SOA, when the preceding and succeeding words as well as the tone appeared shortly before the CW, the acceptable and unacceptable conditions were similarly difficult; this held true in both languages.

The gender-incongruent words successfully induced a LAN and a P600, which were clearly dissociated not only in terms of timing and topography but also functionally. The LAN was not modulated by Language or SOA. As mentioned in the Introduction a verbal working memory interpretation of the LAN would predict effects of SOA and Language already in this single task. The absence of effects might support the view of an encapsulation of the underlying processes. This will be further discussed in the context of the additional task (Experiment 2).

Various modulations emerged for the P600 component. The P600 was larger for Spanish than for German as a main effect and in interaction with SOA. The direction of effects needs some discussion. If one assumes that there is smaller load on working memory to keep track on morphosyntactic congruency in the Spanish group due to overt gender markers, one might expect smaller P600 amplitudes in the Spanish group. To explain why we found the reverse pattern we refer to Kolk et al. (2003), who suggested that the P600 is the response to an unexpected event. As these authors point out a strong syntactic bias due to a certain initial syntactic or morphosyntactic structure establishes a preference for the continuation of a particular sentence or phrase structure. According to this reasoning the syntactic bias in Spanish might be stronger than in German due to overt gender markers. Thus, the gender violation in the target word might come more unexpected in Spanish than in German, which could be interpreted as a gender interference effect (cf. Schriefers, Meyer, and Levelt, 1990). Additionally the violation may be just more salient in Spanish. In German there is not even a morphological incongruency between the target word and the immediately preceding word. As put by Hagoort (2009) P600 amplitude is modulated by competition between alternative options. The competition between two words that carry different morphological gender markers as in Spanish seems to be harder (reflected as an increase of P600 amplitude in the Spanish group) than between two words that carry inherent gender as in German. One might also argue that Germans do not rely on the formal, i.e. morphosyntactic route of processing, because gender is lexically stored. According to Kaan, Harris, Gibson, and Holcomb (2000) an increase in P600 amplitude reflects an increase in the difficulty of syntactic integration. It is suggested that the larger efforts of the Spanish group during this phase of syntactic integration was rewarded at the performance level with similar error rates as in the German group. Thus, Spanish participants seem to successfully overcome the strong syntactic bias and carry on with sentence processing like German participants.

Modulations of the P600 by SOA during later time windows as observed here, that is, longer component durations at longer SOAs, had already been reported by Hohlfeld et al. (2015), and were interpreted either as an index of limited processing time at short SOAs or as an interference effect of the additional stimulus. Gouvea, Phillips, Kazanina, and Poeppel (2010) compared the P600 to ungrammatical sentences (e.g. The patient met the doctor while the nurse with the white dress **show** the chart during the meeting.), garden-path sentences (e.g. The patient met the doctor and the nurse with the white dress showed the chart during the meeting.) and to long distance wh-dependencies (e.g. The patient met the doctor to whom the nurse with the white dress showed the chart during the meeting.). They found variability with respect to onset, latency and duration of the component and conclude that the underlying processes begin as soon as sufficient information has been accumulated and last until a complete structure is built by establishing syntactic relations. Thus, in the present experiment the parser has to wait longer for incoming information when words are presented at a slower rate. In contrast, structure building is speeded up when presentation rate of words is fast.

Experiment 2: Dual Task

The aim of our second experiment was to extend findings from Experiment 1, which revealed susceptibility of the P600 to manipulations in verbal working memory. Experiment 2 tackled the question of modularity of morphosyntax and investigated whether increased demands in general working memory by processing an additional, non-linguistic task would affect morphosyntactic processing. We were interested whether the effect of the additional task would be additive to the pattern of results from Experiment 1 or whether the observed pattern would change. One could imagine even stronger differences between German and Spanish due to a stronger increase of the P600 amplitude in the Spanish group. Possibly, additional task load might also induce differences between language groups even at the behavioral level. With respect to the LAN it was of interest whether there would be any modulations at all.

The same procedure was used as in the single task experiment above, except that now, the previously irrelevant tones also required a (high-priority) foot response in addition to the sentence acceptability task. The different intervals between tones and target words, together with the instructed priority of the tone discrimination task, turned the paradigm into a regular overlapping task paradigm with variable SOAs (see Fig. 1). The SOA variation manipulated the temporal overlap between the processes required by tone discrimination and the processing of morphosyntactically acceptable and unacceptable sentences.

Method

Participants

Sixteen native German speakers (9 women; mean age: *M* = 22.56, range 20 to 27 years) and 18 native Spanish speakers (17 women; mean age: *M* = 20.00, range 18 to 28 years) participated. Participants were university students or had graduated from university and differed from those in Experiment 1. Again, participation was remunerated with money for the German and credit points for the Spanish group. According to the Oldfield handedness inventory (Oldfield, 1971) all participants were right-handed.

Materials

Exactly the same sentences were used as in the single task version of this study. Foot responses were recorded with two keys, embedded into a footrest. The keys were pressed with the big toes, shoes being taken off.

Procedure

Apart from the additional task, the procedure was the same as in Experiment 1. As before, manual acceptability decisions had to be made to the visually presented sentences but now, also responses to the tones were required. High-pitched and low-pitched tones had to be responded to with the left or right foot. The assignment of foot to tone was counterbalanced. Foot responses to the tones were to be executed with priority over the manual sentence acceptability decisions. Task load was manipulated by means of the temporal overlap (SOAs) between tone discrimination and processing the adjectives in the sentences.

Again, after every 38 trials there was a brief resting period. Sixteen versions of the experiment were generated balancing acceptable and unacceptable versions of the CWs, mapping of hand and acceptability, and stimulus-response assignment of the tone task. High and low-pitched tones occurred equi-probably in each condition combination and independent of the acceptability condition. To familiarize participants with the dual task requirements, there were three blocks of practice before the experiment proper. Foot as well as hand responses were first practiced in single task blocks and then in combination in a dual task block. For practice trials other language stimuli were used than in the main experiment.

Data Analysis

Data analysis of the behavioral and ERP responses to the visually presented target words was performed as in the single task Experiment 1 with the same repeated measures factors. Furthermore, also performance (reaction times and error rates) in the additional tone pitch discrimination task was analyzed with repeated measures on SOA and Acceptability, and Language as a group factor. With respect to the tone task, response times are meaningful because responses had to be given immediately.

Results

Additional Task Performance

As to be expected in a PRP-paradigm, there were few effects in the additional task data. Mean error rate over all condition was 5.76 % (*SE* = 1.90). Neither SOA nor Acceptability yielded significant main effects (*Fs* < 1) nor did these factors interact (*F* = 1.40). The factor Language however was significant, F (1, 32) = 4.56, $p < .05$, with fewer errors in the German than in the Spanish group (*M*s = 2.89 vs. 8.63%, *SE*s = 1.95 vs. 1.84, respectively). Reaction times (overall mean = 788.74 ms, *SE* = 85.11) were unaffected by SOA, Acceptability, and their interaction. The group-factor Language failed to reach significance (*M*s = 677.8 vs. 899.6 ms, *SE*s = 87.6 vs. 82.6, respectively), F (1, 32) = 3.39, $p = .075$). In sum, the German group performed better than the Spanish with respect to error rates. In reaction times the factor Language failed significance but there was a trend for the German group to respond faster than the Spanish group.

Sentence Acceptability Task Performance

Table 3 presents the error rates for the sentence acceptability task in both German and Spanish participants. The accuracy of acceptability judgments was significantly affected by SOA (*F* (2, 64) = 15.02, *p* < .001), with the largest number of errors occurring at short SOA (*M*s = 14.58, 10.46, and 10.08%, *SE*s = 1.43, 1.14, 0.91, from SOA 100 to 700). Acceptability did not yield a main effect (*F* < 1) but interacted significantly with Language (*F* (1, 32) = 3.99, *p* = .054). Whereas in the German group there were slightly fewer erroneous responses to acceptable than to unacceptable sentences (*M*s = 7.81 vs. 9.11%; SEs = 1.60 vs. 1.80), the Spanish group committed more errors in the acceptable than in the unacceptable condition (*M*s = 16.34 vs. 13.56%; *SE*s = 1.51 vs. 1.70). Acceptability strongly interacted with SOA (*F* (2, 64) = 16.82, *p* < .001). At SOA 100 fewer errors occurred in the acceptable as compared to the unacceptable condition (*M*s = 12.25 vs. 16.91%; *SE*s = 1.38 vs. 1.95). This pattern was reversed at the two longer SOAs (SOA 400: *M*s = 11.60 vs. 9.32%; *SE*s = 1.39 vs. 1.15; SOA 700: *M*s = 12.37 vs. 7.78%; *SE*s = 1.06 vs. 1.08). Language showed a significant main effect, *F* (1, 32) = 9.46, *p* < .01, with fewer errors in the German than in the Spanish group (*M*s = 8.46 vs. 14.95%, *SE*s = 1.53 vs. 1.44, for German and Spanish, respectively). Moreover, the three-way interaction Language x Acceptability x SOA was significant, *F* (2, 64) = 4.29, *p* < .05 (for details, see Table 3).

===============Table 3=====================

To obtain a better understanding of these interactions post-hoc pairwise comparisons were performed for the effect of SOA within the acceptable and unacceptable conditions for each language. The latter analyses revealed that high temporal overlap with the additional task affected the unacceptable conditions (*p*s <.01 for both German and Spanish) but not the acceptable conditions (*p*s = .12 and .61 for German and Spanish, respectively). Moreover, this effect was much stronger in Spanish than in German (*F*s = 22.0 vs 7.7). In a further test we compared the unacceptable conditions between groups within each SOA. At long SOAs groups did not differ (*p*s > .05), but at short SOA there was a strong trend (*p* = .06; one-tailed

= .03) for German speakers to commit fewer errors in detecting gender violations than for Spanish speakers.

ERP Data

Figure 5 depicts grand average ERPs elicited by acceptable and unacceptable target words for the three SOA conditions at the F3 and Pz electrodes in German and Spanish participants. Differences in amplitude between acceptable and unacceptable target words were evident both in the LAN as well as in the P600 time windows (cf. difference waves in Fig. 6). Figure 7 shows the scalp topographies of the LAN and P600 components in the time windows 400 to 500 and 700 to 900 ms, averaged across all SOAs.

============== Please insert Figures 5, 6 and 7 about here ============

In a first step, ANOVAs were run as in the single task Experiment 1, including all 27 electrodes, for 100-ms time windows (Table 4). Acceptability yielded significant effects in the earlier time windows (400 – 600 ms) in one-tailed analyses, whereas it was significant in two-tailed analyses throughout a long period from 650 to 1150 ms, as a main effect or in interaction with electrode, corroborating the presence of a P600. Within the LAN time windows we did not find any significant effects of Language or SOA. Importantly, in this dual-task experiment, the interaction of Acceptability x SOA was significant from 650 to 850 ms, implying a significant modulation of this part of the P600 by the additional task. Like in Experiment 1, we found a modulation of the P600 by SOA in late time segments (from 1050 ms onwards), which we do not interpret as a proper dual task effect, but as an effect of the ISI. Within the P600 time range the group-factor Language did not yield any main effects, but interacted with SOA between 750 and 950 ms. Interactions of Language x SOA x Acceptability x Electrode occurred between 950 and 1150 ms.

To corroborate the existence of a LAN, in a second step of analysis average amplitude measures were analyzed in a region-of-interest analysis, which included electrodes F7, F3, Fz, and FC3 from 400 to 500 and 500 to 600 ms. From 400 to 500 ms there were significant main effects of Acceptability, F (1, 32) = 5.83, $p < .05$ (indicating the presence of a LAN), of SOA, F (2, 64) = 41.90, $p < .001$; and SOA x Electrode, F (6, 192) = 23.88, $p < .001$. From 500 to 600 ms the following significant effects occurred: SOA (F (2, 64) = 57.07. $p < .001$); and SOA x Electrode (F (6, 192) = 27.93. $p < .001$). There was however no effect of acceptability in this time segment.

================Insert Table 4 about here, please==================

We refrained from a ROI-analysis in the P600-time window, because the analysis including all electrodes yielded stable results.

Discussion

To contribute to the question of modularity of morphosyntax the effect of an additional non-linguistic task on morphosyntactic processing was of major interest in this second experiment. Morphosyntactic processing was tapped by comparing the processing of grammatical gender in German and Spanish.

At the performance level we observed the typical findings in PRP paradigms that reaction times and error rates in the additional tone task, which had priority over the language task per instruction, were not affected by SOA. In contrast, error rates in the language task were significantly affected by task overlap, with more errors at short SOA, that is, at high temporal overlap between tasks. This pattern was present in both language groups and can be interpreted as an index of processing costs due to high task overlap. Moreover, Spanish speakers committed more errors in detecting gender violations when task overlap was high and the interval between gender incongruent items was short. Interestingly, at longer SOAs this pattern reversed and the Spanish speakers committed more errors when they judged acceptable sentences. The German speakers also committed more errors while judging unacceptable sentences at high task overlap, but at longer SOAs there was no difference in error rates between acceptable and unacceptable sentences. When error rates for unacceptable sentences were compared between language groups, the groups only differed at the short SOA (with more errors made by the Spanish group). The implications of these findings with respect to processing differences for gender information are discussed in the General Discussion.

On the level of ERPs we found a small LAN in a one-tailed analysis, justified by the LAN induced in the single task Experiment 1 by the same linguistic material. Consequently, we had a directed expectation about the LAN effect. An additional region-of-interest-analysis also confirmed the existence of a LAN. However, we did not observe any modulations of this LAN neither by task overlap nor by language.

In contrast to the LAN, several effects were observed in the P600 time window. The P600 between 650 and 850 ms was affected by the additional task. Interestingly, the effect of Language observed in Experiment 1, that is larger amplitudes in the Spanish group between 750 and 950 ms, particularly at short SOA, was absent in Experiment 2. Here, within this time window, the factor Language was not significant, neither as main effect nor in interaction with Acceptability or SOA.

In the ERPs of the present experiment we also observed late effects of SOA from 1050 ms onwards. These late effects were also observed in Experiment 1, where they started from 950 ms onwards. The effects of SOA seem to be due to limited processing time for sentences at short ISIs. Hohlfeld et al. (2015) showed a similar delay of these effects by 100 ms when going from a single task to a dual task experiment. A wrap-up mechanism is suggested, which flexibly assists structure building – depending on factors such as processing speed. In Experiment 2, processing of the additional task seems to delay the onset of this wrap-up. Differences between the Spanish and German group occurred in the late P600 (950 – 1250 ms) in both experiments. Particularly at short SOA 100, the processes underlying the P600 seem to have been speeded up in the German group, whereas it took the Spanish group a bit longer. When there is less time constraint at long SOA 700, this difference between German and Spanish disappears.

The late P600 was affected by SOA in both experiments. During this phase lexical-semantic integration might take place with the aim of continuing sentence processing (violations were always followed by a subsequent word). Responses in the language task had to be given at the end of the sentence. When the next word followed rather quickly, as in SOA 100, the processes underlying the P600 seem to have been speeded up. This effect of sentence wrap-up has already been reported by Hohlfeld et al. (2015) for the P600 induced by semantic violations. As argued, the processing at sentence level does not only speed up processing but also seems to have a protective effect in the additional task situation. Although the additional task situation was highly demanding, strong postponements of the P600 peak amplitude or complete deterioration were not observed. In contrast, in a comparable additional-task study applying word pairs, the N400 component was delayed by more than 200 ms (Hohlfeld, Sangals, and Sommer, 2004). Thus, the sentence context might stabilize the linguistic processing pathway. It is suggested that during sentence processing our cognitive system flexibly copes with interfering influences of the environment and protects on-going processes as long as possible. It might be of interest for further research, whether this protective aspect of sentence processing is part of the wrap-up mechanism suggested above that is sensitive to speed of processing.

General Discussion

In order to investigate the modularity of morphosyntactic processing the present study explored the effects of an additional non-linguistic task. Specifically, processing differences due to the formal and inherent gender systems of Spanish and German respectively were investigated under different degrees of concurrent task load, imposed by temporal overlap.

The ERP components LAN and P600 served as markers of early and late morphosyntactic processing, respectively. At first, a single task experiment was performed, which employed the same materials and presentation conditions as the subsequent dual-task experiment. We investigated whether the varying SOA affects LAN and P600 already in a singe task condition. In Experiment 2 responses to the additional tone stimulus were required and thus the paradigm resembled a PRP design, which allowed measuring the effects of additional task processing on LAN and P600, in both German and Spanish.

In consonance with suggestions by others (cf. Friederici, Mecklinger, Spencer, Steinhauer, and Donchin, 2001), our results support the notion of a multi-layered process or different syntactic processing stages; these stages seem to be differentially susceptible to interference with a non-linguistic additional task. The LAN in the dual task experiment was small and statistically confirmed in a ROI and a one-tailed analysis. Also in the single task (Experiment 1) neither Language nor SOA affected the LAN. Hence, what we observed in the LAN time-window in both experiments seems to be in concord with the idea that the LAN reflects an early morphosyntactic process that might be encapsulated. This coalesces with the idea of an automatic first stage of parsing, during which syntactic structure is built (Jiménez-Ortega, García-Milla, Fondevila, Casado, Hernández-Gutiérrez, and Martín-Loeches, 2014). However, this interpretation is in contrast to findings from other studies relative to the encapsulation of the processes indexed by the LAN. For instance, Martín-Loeches, Fernandez, Schacht, Sommer, Casado, and Jimenez-Ortega (2012) demonstrated modulations of the LAN by emotional valence, as its amplitude increased in case of negative and decreased in case of positive relative to neutral words. For more evidence against the notion of encapsulation during the LAN time window cf. also Shao and Neville, 1998; Anderson and Holcomb, 2005; Barkley, Kluender and Kutas, 2015. It is necessary to clarify exactly which experimental conditions or circumstances vary the LAN pattern.

According to Jackendoff (2002, 2007) there is a component in our language system that has no direct interfaces with other linguistic and non-linguistic processing streams, which is called *syntactic formation rules*. If one adheres to the notion that the processes underlying the LAN might be encapsulated, at least under certain conditions, they might reflect this type of linguistic knowledge, which has to be accessed during sentence perception – in the present case, for instance, language-specific knowledge about grammatical gender congruency.

For a contrasting view, following the reasoning of Tanner and colleagues (Tanner and van Hell, 2014) one might speculate that the LAN in the present study and its underlying processes is much more susceptible to individual variability when the additional task has to be processed as compared to the P600. Processing two tasks in parallel might increase individual variability particularly in the LAN time window with the consequence that it is statistically corroborated only in a region-of-interest analysis including four electrodes or in a one-tailed analysis including all electrodes.

The early stage of processing is followed by a stage that seems to be more open to other processing streams as indexed by the effects of the additional task on the P600. As suggested by Kolk et al. (2003) the processes underlying the P600 reflect monitoring with the aim to detect an unexpected event. Our finding that the effect of the additional task was not additive on the effect of Language observed in the single task, supports the notion that the process underlying the P600 draws on domain-general resources such as working memory and might reflect an interface between language comprehension and general cognition.

Our findings in the LAN and P600 time windows give room to two different models of sentence comprehension. If the absence of any effects on the LAN in both experiments is taken as an index of a encapsulated automatic first stage of morphosyntactic processing (which has to be handled with caution as pointed out above), the subsequent processes underlying the P600 may then be interpreted as processes of domain-general monitoring (Kolk et al., 2003, 2007). On the other hand if one adheres to the notion of interindividual variability in the LAN, which puts doubts on the existence of this component (cf. Tanner et al, 2014), one needs to shift morphosyntactic processing into the P600 time-window as has been suggested, for instance, by Kuperberg (2007). Both accounts fit the data at hand and further research is needed to distinguish between them.

What can we infer from the findings with respect to gender processing in the two different languages? If one looks at the performance data in the single task there is no difference between groups. As the main effect of acceptability reveals (with more errors made in case of acceptable sentences) both groups might apply a strategy of monitoring for errors, which might help in detecting unacceptable sentences. On performance level, the overtness of gender cues does not seem to facilitate this process in the Spanish group. Data from the electrophysiological level reveal that there even might be some extra processing resources needed in the Spanish group to overcome the strong morphosyntactic bias posed by the first two words of the sentence that carry formal congruent gender markers. Or, as has been suggested above, particularly at short SOA the incongruent gender cue might interfere with the preceding gender cues. In Experiment 2, when an additional task had to be processed, group differences showed up at both electrophysiological and performance levels. The P600 in the Spanish group lasted longer than in the German group, particularly at short SOA. Error rates at short SOA significantly differed between the groups in the unacceptable condition, i.e. when sentences with gender incongruencies had to be judged, Spanish participants committed more errors. At longer SOAs this group difference disappeared. This pattern might again be a reflection of the syntactic bias that the Spanish group has to cope with, i.e. with the preference for the built-up of a specific syntactic structure due to the input. In the single task this did not reach the behavioral level, but in the dual task it did. When looking at error rates at SOA 400 and 700 in the dual task further group differences become evident. Whereas error rates in judging acceptable and unacceptable sentences did not differ in the German group, there were fewer errors for unacceptable sentences in the Spanish group. One might speculate that this pattern reflects the different ways of gender processing – the lexical route in German and the formal, cue-based route in Spanish. At longer SOAs the cue-based route seems to help in detecting gender incongruencies, whereas at short SOA it poses extra costs, particularly when an additional task has to be processed.

Finally, we turn to the mechanisms of interference. One interpretation of interference effects between the additional task and the P600 is based on the assumption of a central processing bottleneck, that is, a single processing channel for certain “central” mental operations (e.g. response selection, memory retrieval). This central stage can only handle one task at a time (e.g. Pashler, 1994). While the central mental operation is performed for a given task, such operations in the service of other tasks have to wait. Data from the present study is only partly in accordance with the predictions made by the single channel model. Error rates and response time in the tone task (the primary task) were not affected by SOA. Thus, according to the model, this operation is handled by the central channel, while the secondary task (sentence processing) has to wait. For the processes underlying sentence processing as reflected by the LAN and P600 the single channel account would predict postponements (latency shifts) rather than amplitude manipulations. In the present study however, we did not see strong latency shifts. An alternative model explains interference effects by assuming limited central resources that have to be shared between tasks (Navon and Miller, 2002). Applied to the present case, to perform the non-linguistic Task 1, participants have to draw on central resources, which are also necessary to perform the sentence acceptability Task 2. The vanishing of the language effect observed in Experiment 2 appears to be more in accordance with such a redistribution of resources than with a central bottleneck. The assumption that the processes underlying the P600 draw on central resources is supported by Kolk and colleagues (2003), who suggest that the P600 depends on executive control. Furthermore, this is in line with the strong task-dependence of P600 effects (Schacht, Sommer, Shmuilovich, Casado, and Martín-Loeches, 2014; Xu, Abdel Rahman and Sommer, 2019).

In conclusion, we did not observe any effects on the LAN – neither by SOA nor by the additional task. It has been suggested that this finding is in line with the notion of a modular first-parsing process (e.g. the application of syntactic formation rules), although this contradicts evidence from other studies. For further research it might be of interest to disentangle which conditions modulate the LAN and which conditions leave the component unaffected. Only that would allow a universally valid interpretation of the component. The effect of language and the corresponding gender marking systems, indexed by an increased P600 amplitude in Experiment 1 in the Spanish group, disappeared when an additional task had to be processed (Experiment 2). This is interpreted as a sign of the flexibility of the cognitive system to withdraw resources from one – linguistic – processing stream to cope with another task under time pressure. The effects of the additional task on the processes underlying the P600 further support the idea that these processes belong to central processes and are controlled rather than automatic. Wrap-up effects (protective effects) of the sentence were replicated. Hence, processing at sentence level (in contrast to processing at word level) may speed up processing and protect parsing against disturbing influences from the environment.

In sum, with respect to processing during the P600 time window our data support the notion of flexibility of the language system. If necessary - as for instance in a dual task situation - resources are shared between language processing and other cognitive operations. In addition, a wrap-up mechanism might prevent a complete break down of sentence processing.

References

Anderson, J. E., & Holcomb, P. J. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and language*, *94*(2), 200-216.

Barkley, C., Kluender, R., and Kutas, M. (2015). Referential processing in the human brain: An Event-Related Potential (ERP) study. *Brain research*, *1629*, 143-159.

Bauch, H. J. (1971). Zum Informationsgehalt der Kategorie Genus im Deutschen, Englischen und Polnischen. *Wissenschaftliche Zeitschrift der Universität Rostock; Gesellschafts- und Sprachwissenschaftliche Reihe, Heft 6*. Rostock, Germany: Universität Rostock.

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2008). An alternative perspective on “semantic P600” effects in language comprehension. *Brain research reviews*, *59*(1), 55-73.

Broca, P. (1865). Sur la faculté du langage articulé. *Bulletins de la Société Anthropologique de Paris*, 493-494.

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain research*, *1446*, 127-143.

Caplan, D. and Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77-126.

Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht.

Coulson, S., King, J. and Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*, 21-58.

Desroches, A., and Paivio, A. (1990). Le phonem initial des noms inanimes et son effet sur l’identification du genre grammatical. *Canadian Journal of Psychology, 44*, 44-57.

Eisenberg, P. (1994). *Grundriss der Deutschen Grammatik*. Stuttgart, Weimar.

Fedorenko, E., Gibson, E., and Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of memory and language*, *54*(4), 541-553.

Fodor (1983). *The modularity of mind*. Cambridge, Mass.: MIT Press.

Friederici, A.D., Hahne, A., Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22 (5),* 1219-1248.

Friederici, A.D., Mecklinger, A., Spencer, K.M., Steinhauer, K., and E. Donchin (2001). Syntactic parsing preferences and their on-line revisions: aspatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*, 303-323.

Gollan, T., and Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research, 30*, 627-651.

Gouvea, A. C., Phillips, C., Kazanina, N., and Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and cognitive processes*, *25*(2), 149-188.

Gratton, G., Coles, M.G.H., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468-484.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15*, 883-899.

Hagoort, P. (2009). Reflections on the Neurobiology of Syntax. In: Bickerton, D., and Szathmáry, E. (Eds.). (2009). *Biological foundations and origin of syntax* (Vol. 3). MIT Press. 279-296.

Hohlfeld, A. (2006). Accessing grammatical gender in German: The impact of gender-marking regularities. *Applied Psycholinguistics, 27*, 127-142.

Hohlfeld, A., Martín-Loeches, M., and Sommer, W. (2015). Is semantic processing during sentence reading autonomous or controlled? Evidence from the N400 in a dual task paradigm. *Advances in Cognitive Psychology, 11* (2), 42-55.

Hohlfeld, A., Sangals, J., and Sommer, W. (2004). Effects of Additional Tasks on Language Perception: An ERP Investigation. *Journal of Experimental Psycholog: Learning, Memory, and Cognition, 30 (5),* 1012-1025.

Huynh, H., and Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1,* 69-82.

Jackendoff, R. (2002). *Foundations of language. Brain, Meaning, Grammar, Evolution.* Oxford: Oxford University Press.

Jackendoff, R. (2007). A Parallel Architecture perspective on language processing. *Brain Research, 1146,* 2-22.

Jiménez-Ortega, L., García-Milla, M., Fondevila, S., Casado, P., Hernández-Gutiérrez, D., and Martín-Loeches, M. (2014). Automaticity of higher cognitive functions: Neurophysiological evidence for unconscious syntactic processing of masked words. *Biological psychology*, *103*, 83-91.

Kaan, E., Harris, A., Gibson, E., and Holcomb, P. J. (2000). P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15 (2)*, 159-201.

King, J. W., and Kutas, M. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of cognitive neuroscience*, *7*(3), 376-395.

Kluender, R. and Kutas, M. (1993). Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*, 196-214.

Kemmerer, D. (2015). *Cognitive Neuroscience of Language*. New York: Taylor and Francis.

Kolk, H. H., Chwilla, D. J., Van Herten, M., and Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and language, 85*(1), 1-36.

Kolk, H. H., and Chwilla, D. J. (2007). Late positivities in unusual situations. *Brain and Language, 100*, 257-261.

Köpcke, K. M. (1982). *Untersuchungen zum Genussystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23-49.

Kutas, M., and Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature 307*, 161-163.

Kutas, M., Van Petten, C., and Kluender, R. (2006). Psycholinguistics Electrified II. In M. A. Gernsbacher and M. Traxler (Eds.), *Handbook of Psycholinguistics* (2nd ed.). New York: Elsevier Press.

Levelt, W.J.M., Roelofs, A., Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22 (1)*, 1-75.

MacLeod, C.M., 1991. Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin 109*, 163-203.

Martín-Loeches, M., Abdel-Rahman, R., Casado, P., Hohlfeld, A., Schacht, A., and Sommer, W. (2009). Rules and heuristics during sentence comprehension: Evidences from a dual-task brain potential study. *Journal of Cognitive Neuroscience, 21*, 1365-1379.

Martín-Loeches, M., Casado, P., Gonzalo, R., De Heras, L., and Fernández-Frías, C. (2006). Brain potentials to mathematical syntax problems. *Psychophysiology, 43*(6), 579-591.

Martín-Loeches, M., Fernandez, A., Schacht, A., Sommer, W., Casado, P., Jimenez-Ortega, L., Fondevila, S. (2012). The influence of emotional words on sentence processing: electrophysiological and behavioral evidence. *Neuropsychologia, 50*, 3262-3272.

Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A. and Sommer,W. (2006). Semantics prevalence over syntax during sentence processing: A brain potential study of noun-adjective agreement in Spanish. *Brain Research, 1093* (1), 178-189.

Molinaro, N., Barber, H. A., and Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *cortex*, *47*(8), 908-930.

Navon, D.A., and Miller, J. (2002). Queuing or Sharing? A Critical Evaluation of the Single-Bottleneck Notion. *Cognitive Psychology*, *44*, 193-251.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97-113.

Osman, A. and Moore, C. M. (1993). The locus of dual-task interference: Psychological refractory effects on movement-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1292-1312.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and language*, *59*(3), 494-522.

Osterhout, L. and Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., and Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond*, 271-308.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220-244.

Pinker, S. (1994). *The Language Instinct*. New York: Morrow.

Schacht, A., Sommer, W., Shmuilovich, O., Casado, P., and Martín-Loeches, M. (2014). Differential Task Effects on N400 and P600 Elicited by Semantic and Syntactic Violations. *PLoS ONE, 9(3),* e91226. doi: 10.1371/journal.pone.0091226 *Joint first-authorship.

Scherag, A., Demuth, L., Rösler, F., Neville, H. J., and Röder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, *93*(3), B97-B108.

Schriefers, H., Meyer, A., Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Learning*, *29*, 86-102.

Shantz, K., and Tanner, D. (2017). Talking out of order: task order and retrieval of grammatical gender and phonology in lexical access. *Language, Cognition and Neuroscience*, *32*(1), 82-101.

Shao, J., and Neville, H. (1998). Analyzing semantic processing using event-related brain potentials. *The Newsletter of the Center for Research in Language.* University of California, San Diego, La Jolla CA 92039, 11 (5). Available: http://www.crl.ucsd.edu/newsletter.html.

Smith, N., and Tsimpli, I. M. (1991). Linguistic modularity? A case study of a ‘savant’linguist. *Lingua*, *84*(4), 315-351.

Steinhauer, K. and Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and language*, 120, 135-162.

Tanner, D., and Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289-301.

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: a commentary on" grammatical agreement processing in reading: ERP findings and future directions" by Molinaro et al., 2014. *Cortex; a journal devoted to the study of the nervous system and behavior*, *66*, 149.

Tanner, D. (2019). Robust neurocognitive individual differences in grammatical agreement processing: A latent variable approach. *Cortex*, *111*, 210-237.

Taft, M., and Meunier, F. (1998). Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research, 27,* 23-45.

Urbach, T. P. (1993). Access to non-syntactic information during parsing: Evidence from event-related potentials (ERPs) in garden-path sentences. In *Poster session presented at the 6th annual CUNY conference on human sentence processing, Amherst, MA*.

Van Berkum, J. J., Brown, C. M., and Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of memory and language*, *41*(2), 147-182.

Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 443.

van Turennout, M., Hagoort, P., Brown, C. (1998). Brain activity during speaking. From syntax to phonology in 40 milliseconds. *Science, 280,* 572-574.

Weber-Fox, C. M., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8* (3), 231-256.

Weber-Fox, C. M., and Neville, H. J. (2001). Sensitive periods differentiate processing of open- and closed-class words: An ERP study of Bilinguals. *Journal of Speech, Language, and Hearing Research, 44* (6), 1338-1354.

Wernicke, C. (1874). *Der aphasische Symptomenkomplex*. Breslau, Germany: Cohn and Weigert.

Xu, J., Abdel Rahman, R., and Sommer, W. (2019). Perceived language competence modulates criteria for speech error processing: evidence from event-related potentials. *Language, Cognition and Neuroscience*, 1-14. doi:10.1080/23273798.2018.1562558

Zahedi, A., Abdel Rahman, R., Stürmer, B., Sommer, W. (in press). Common and Specific Loci of Stroop Effects in Vocal and Manual Tasks, Revealed by Event-Related Brain Potentials and Post-Hypnotic Suggestions. *Journal of Experimental Psychology: General.*

Author Note

This study was made possible by a grant from the German Research Council to A.H. and the cooperation contract between Complutense University of Madrid and Humboldt University at Berlin. The German Research Concil had no involvement in study-design, data-collection, analysis, interpretation and writing of the report. We wish to thank Pilar Casado who supported the elaboration of the material as well as data collection. In addition we thank Eva Köppen, Sokrates-student in Madrid, who supported the recruiting of German participants in Madrid and registration.

Tables

Table 1.

Error Rates for the Sentence Acceptability Task of Experiment 1 for German and Spanish participants, respectively (Mean Values in Percent with Standard Errors in Parentheses).

Target	German	Spanish
SOA 100		
Acceptable	5.78 (1.33)	7.92 (1.25)
Unacceptable	4.68 (1.26)	4.86 (1.19)
SOA 400		
Acceptable	6.56 (1.37)	7.50 (1.29)
Unacceptable	2.5 (0.69)	2.64 (0.65)
SOA 700		
Acceptable	7.50 (1.38)	11.11 (1.30)
Unacceptable	2.19 (0.64)	4.16 (0.60)

Table 2. F-values and Significance Levels from the ANOVAs of ERP Amplitudes in Experiment 1.

	Df	400[a]	500	650	750	850	950	1050	1150
SOA	2, 64	6.69**	4.52*	18.82***	42.89***	28.63***	21.05***	16.43***	28.41***
SOA*Electr	52, 1664	3.92**	2.95**	8.36***	5.98***	6.88***	6.73***	7.98***	7.78***
Acc	1, 32	4.39*	9.78**	84.34***	83.81***	54.28***	21.40***	--	--
Acc*Electr	26, 832	4.50***	6.00***	30.01***	29.88***	31.55***	26.56***	17.49***	9.99***
SOA*Acc	2, 64	--	--	--	--	--	4.26*	15.64***	12.82***
SOA*Acc*Electr	52, 1664	--	--	--	--	--	1.92*	2.34**	2.10**
Language	1, 32	--	--	--	5.80*	4.99*	--	--	--
SOA*Language*Electr	52, 1664	--	--	--	--	--	--	2.16*	--
Acc*Language*Electr	26, 832	--	--	--	--	--	--	2.78**	3.01**
SOA*Acc*Language	2, 64	--	--	--	--	--	3.60*	--	--
SOA*Acc*Lang*Electr	52, 1664	--	--	--	--	2.05*	2.03*	1.90*	--

*p < .05. **p < .01. ***p < .001.

[a]Segments were of 100 ms duration

Table 3.

Error Rates for the Sentence Acceptability Tasks of Experiments 2, for German and Spanish participants, respectively (Mean Values in Percent with Standard Errors in Parentheses).

Target	German	Spanish
SOA 100		
Acceptable	9.21 (1.01)	15.27 (2.44)
Unacceptable	13.12 (2.32)	20.69 (3.03)
SOA 400		
Acceptable	6.41 (1.31)	16.80 (2.35)
Unacceptable	7.81 (1.09)	10.83 (1.94)
SOA 700		
Acceptable	7.81 (1.44)	16.94 (1.54)
Unacceptable	6.41 (1.21)	9.16 (1.72)

Table 4. F-values and Significance Levels from the ANOVAs of ERP Amplitudes in Experiment 2.

	Df	400[a]	500	650	750	850	950	1050	1150
SOA	2, 64	20.75**	18.17***	14.95***	25.91***	21.01***	16.53***	10.46***	27.88***
SOA*Electr	52, 1664	27.12***	27.55***	17.91***	19.97***	27.71***	16.88***	17.72***	15.48***
Acc	1, 32	(.057[b])	(.05[b])	88.94***	127.07***	111.48***	42.29***	8.91**	--
Acc*Electr	26, 832	--	--	9.56***	15.34***	18.15***	13.25***	8.09***	6.30***
SOA*Acc	2, 64	--	--	4.95*	5.00*	--	--	7.68**	12.92***
SOA*Acc*Electr	52, 1664	--	--	--	--	--	--	2.05*	2.16*
Language	1, 32	--	--	--	--	--	--		--
SOA*Language	2, 64	--	--	--	3.1*	(.047[b])	--	--	--
SOA*Language*Electr	52, 1664	--	--	--	--	--	--	--	--
Acc*Language*Electr	26, 832	--	--	--	--	--	--	--	--
SOA*Acc*Language	2, 64		--	--	--	--	--	--	--
SOA*Acc*Lang.*Electr	52, 1664	--	--	--	--	--	1.76*	1.82*	--

*p < .05. **p < .01. ***p < .001.

[a]Segments were of 100 ms duration

[b]one-tailed

Figure Captions

Figure 1. Chronometric depiction of an experimental trial (arbitrary scaling) showing a syntactically unacceptable Spanish sentence. The single task experiment required manual responses only to the sentences` acceptability. In the dual task experiment both, foot responses to the tones and manual responses to the sentences had to be given.

Figure 2. Single task Experiment 1: Panel A and B depict ERP wave shapes at the F3 and Pz electrode in response to acceptable and unacceptable German and Spanish target words, respectively, at each SOA (with 15 Hz low pass filtering for the purpose of presentation only).

Figure 3. Single task Experiment 1: Difference waves between ERPs to acceptable and unacceptable target words at F3 (15 Hz) and Pz (7 Hz) electrode are depicted for SOA 100, 400, and 700 (with panel A and B for German and Spanish, respectively).

Figure 4. Single task Experiment 1: Topographies of difference wave amplitudes between 400 to 500 ms and between 700 to 900 ms after target onset. Data were collapsed across the SOA conditions (with panel A and B for German and Spanish, respectively).

Figure 5. Dual task Experiment 2: Panel A and B depict ERP wave shapes at the F3 and Pz electrode in response to acceptable and unacceptable German and Spanish target words, respectively, at each SOA (with 15 Hz low pass filtering for the purpose of presentation only).

Figure 6. Dual task Experiment 2: Difference waves between ERPs to acceptable and unacceptable target words at F3 (15 Hz) and Pz (7 Hz) electrode are depicted for SOA 100, 400, and 700 (with panel A and B for German and Spanish, respectively).

Figure 7. Dual task Experiment 2: Topographies of difference wave amplitudes between 400 to 500 ms and between 700 to 900 ms after target onset. Data were collapsed across the SOA conditions (with panel A and B for German and Spanish, respectively).

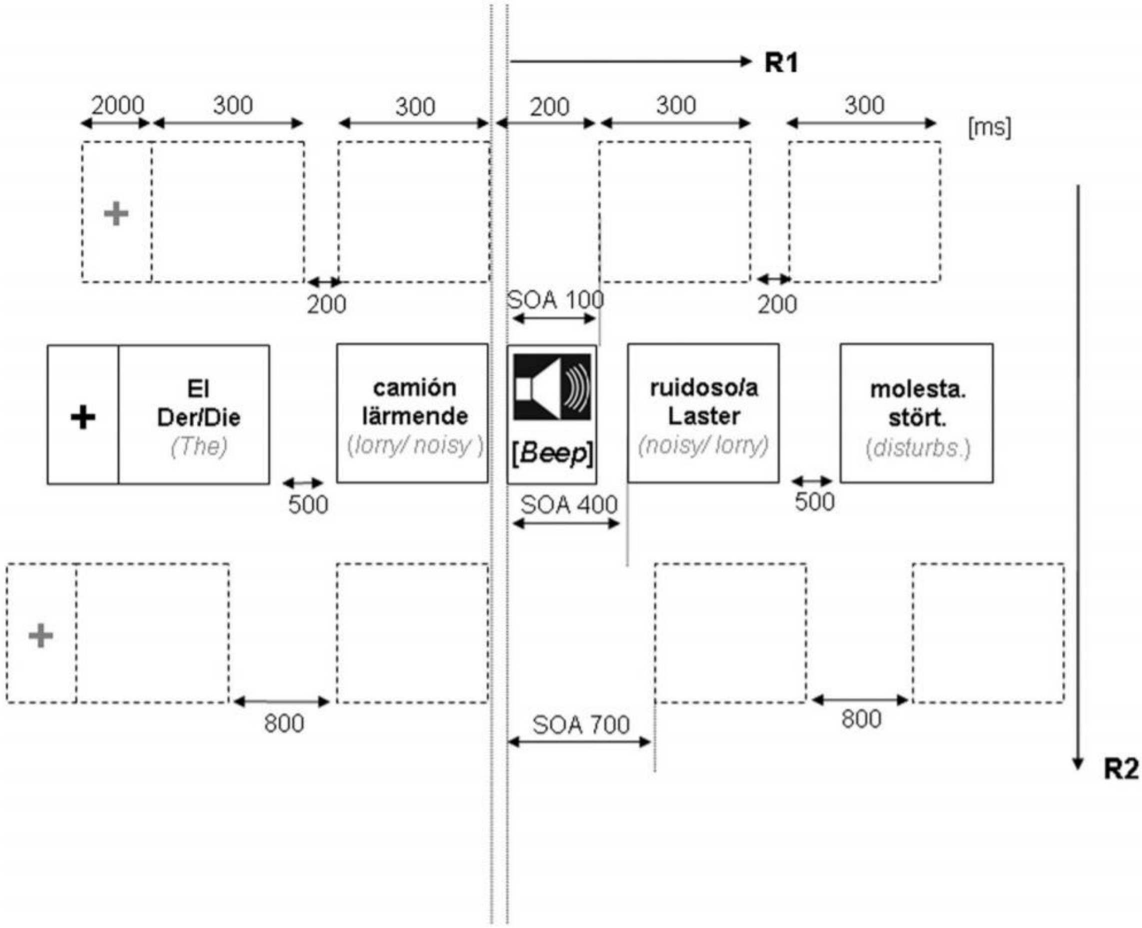
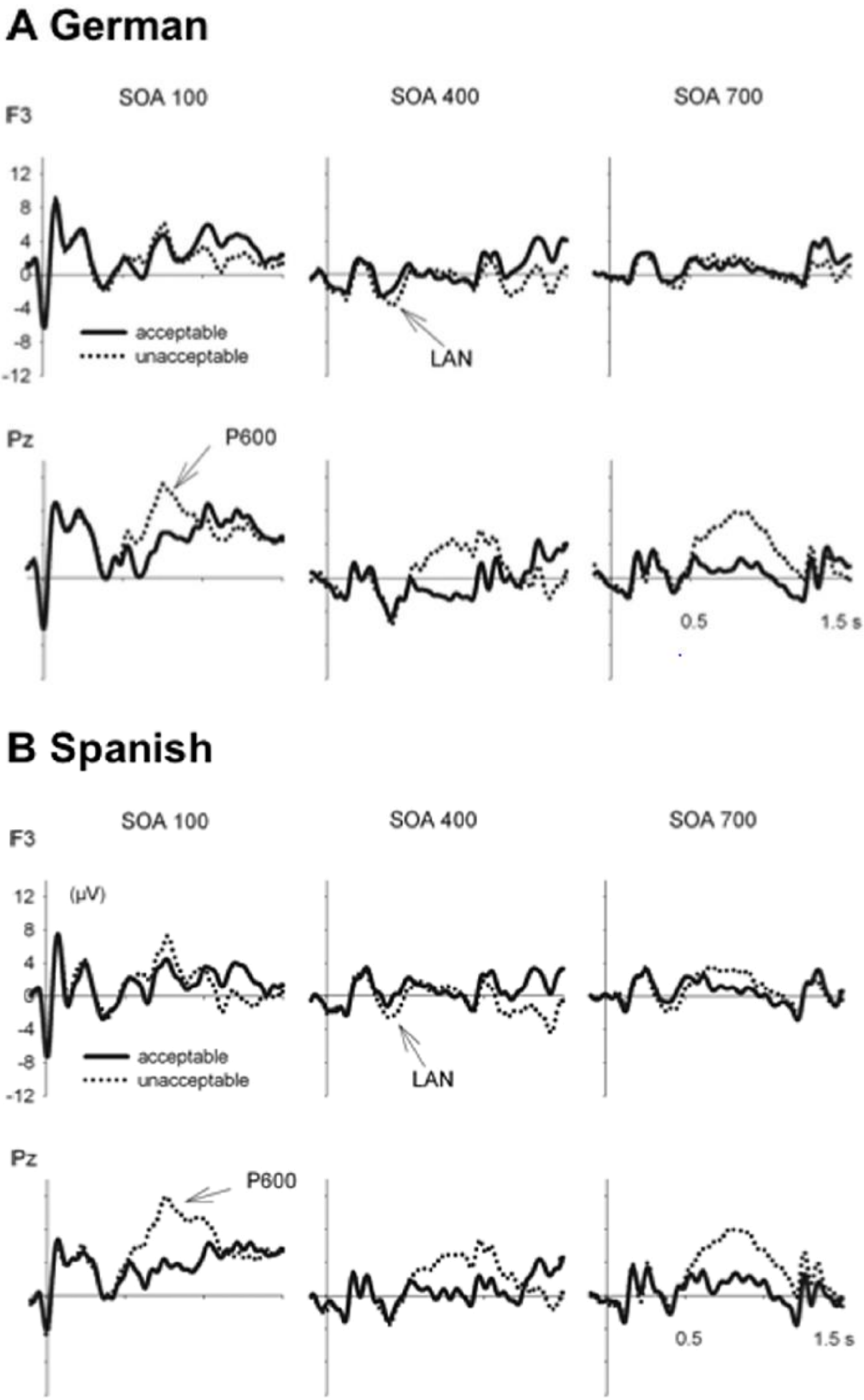

R1
2000
300
300
200
300
300
[ms]
200
SOA 100
200
El
Der/Die
(The)
camión
lärmende
(lorry/ noisy)
[Beep]
ruidoso/a
Laster
(noisy/ lorry)
molesta.
stört.
(disturbs.)
500
SOA 400
500
800
SOA 700
800
R2

Figure 1.



A German
SOA 100
SOA 400
SOA 700
F3
12
8
4
0
-4
-8
-12
acceptable
unacceptable
LAN
Pz
P600
0.5
1.5 s
B Spanish
SOA 100
SOA 400
SOA 700
F3
12
(μV)
8
4
0
-4
-8
-12
acceptable
unacceptable
LAN
Pz
P600
0.5
1.5 s

Figure 2

.

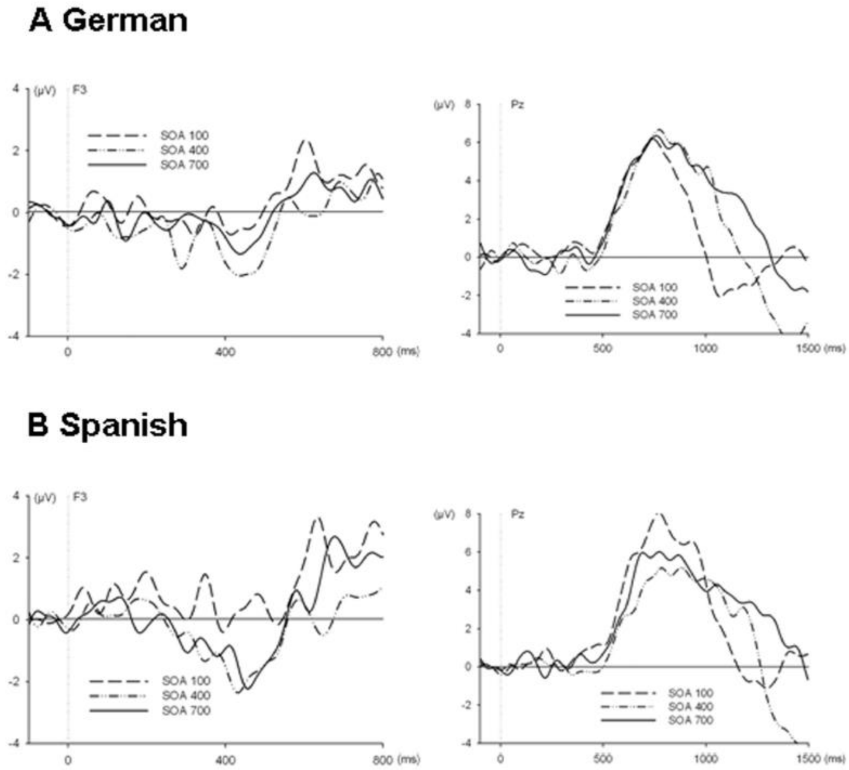


Figure 3

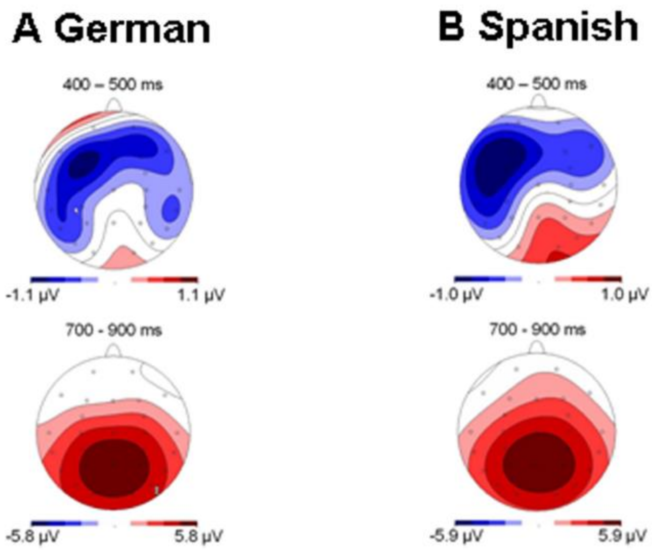


Figure 4

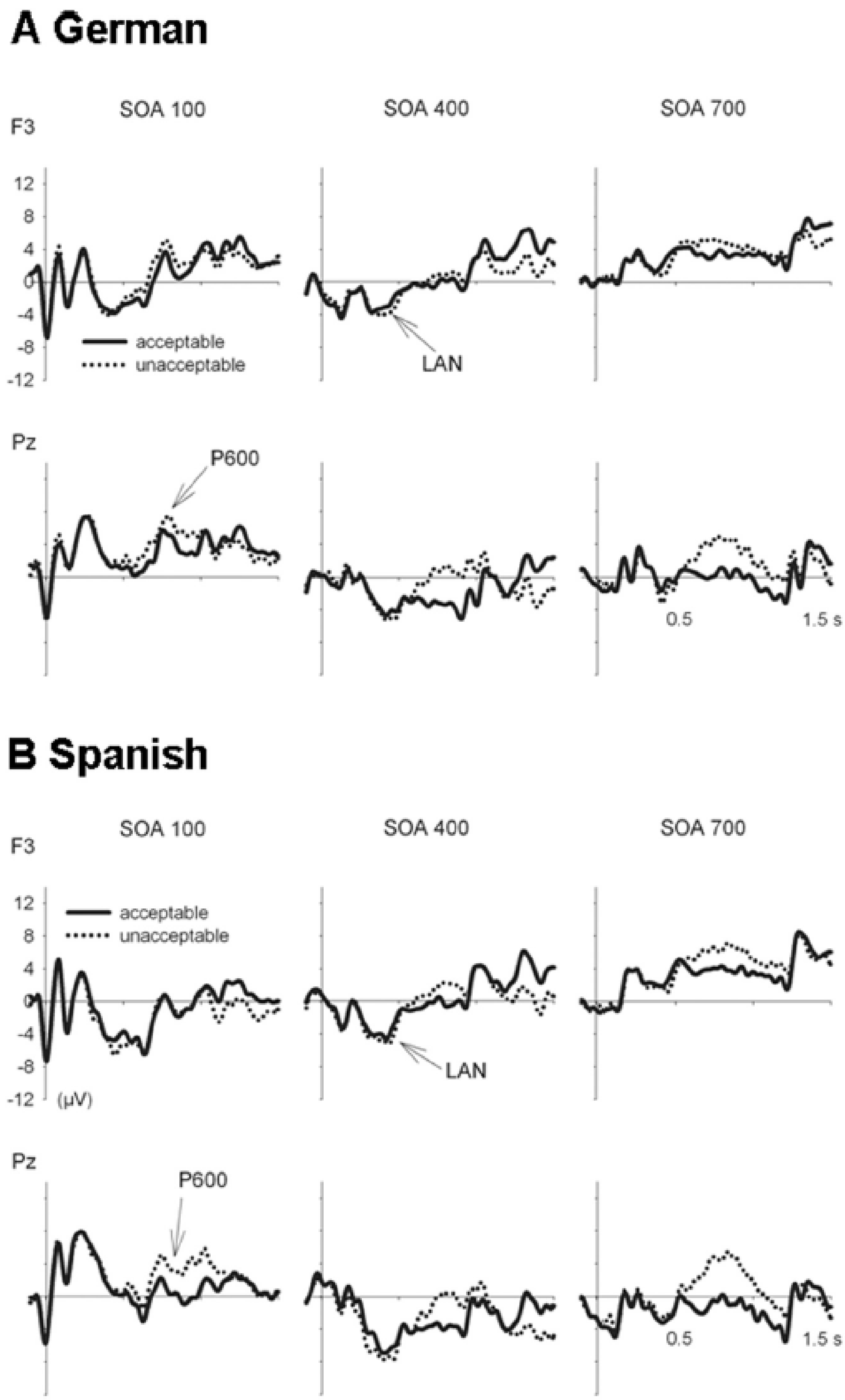

A German
SOA 100
SOA 400
SOA 700
F3
12
8
4
0
-4
-8
-12
acceptable
unacceptable
LAN
Pz
P600
0.5
1.5 s
B Spanish
SOA 100
SOA 400
SOA 700
F3
12
8
4
0
-4
-8
-12
(μV)
acceptable
unacceptable
LAN
Pz
P600
0.5
1.5 s

Figure 5

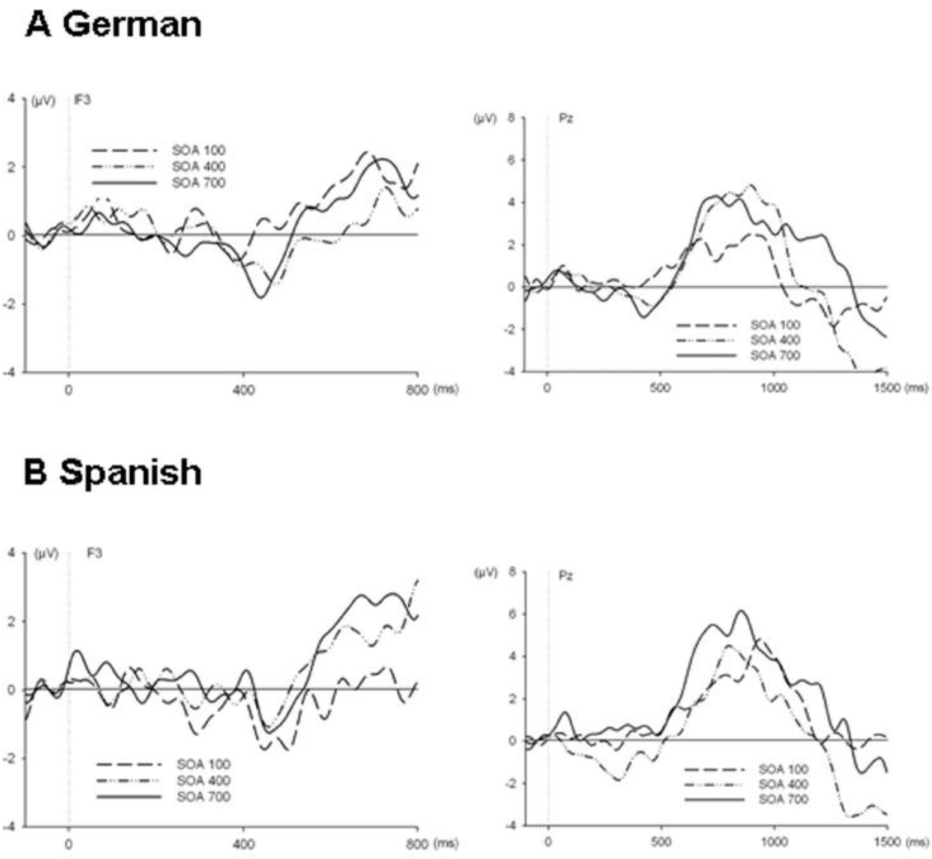


Figure 6

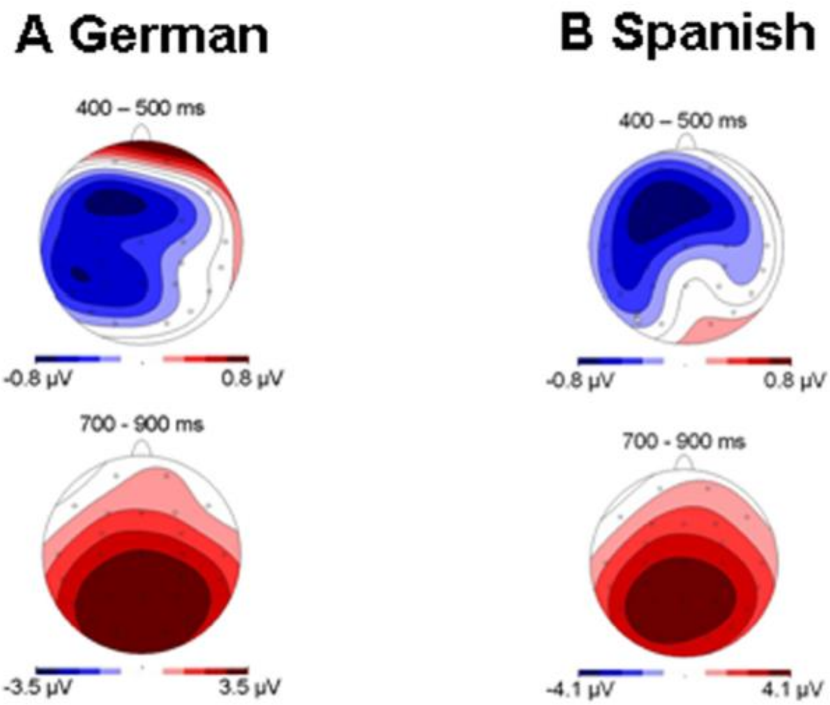


Figure 7

Appendix

List of 12 example sentences (out of a set of 240 for German and Spanish each) with gender congruencies and incongruencies in determiners or adjectives, for German and Spanish respectively. Corresponding English non-literal translations are given in parentheses.

German	Spanish
Das/Der alkoholische Getränk ermüdet.	La bebida alcohólica/o cansa.
(The alcoholic drink makes tired.)	
Das/Der angestochene Rad zischt.	Las ruedas pinchadas/os chillan.
(The pinched wheel squeals.)	*(The pinched wheels squeal.)*
Das/Die hübsche Mädchen flirtet.	La chica guapa/o liga.
(The pretty girl flirts.)	
Das/Die lästige Insekt summt.	Los insectos venenosos/as aletean.
(The annoying insect hums.)	*(The poisonous insects flutter.)*
Der/Das aggressive Feind attackiert.	Los enemigos agresivos/as luchan.
(The aggressive enemy attacks.)	*(The aggressive enemies fight.)*
Der/Das aktive Vulkan qualmt.	El volcán rabioso/a erupciona.
(The active volcano smokes.)	*(The raging volcano erupts.)*
Der/Die lärmende Laster stört.	El camión ruidoso/a molesta.
(The noisy lorry disturbs.)	
Der/Die leckere Wein betäubt.	El vino tinto/ a emborracha.
(The tasty wine is narcotic.)	*(The red wine makes drunk.)*
Die/Der abenteuerliche Reise prägt.	El viaje aventurero/a concluye.
(The adventurous journey shapes.)	*(The adventurous journey terminates.)*

Die/Der gute Zeit stärkt.	El tiempo bueno/a anima.
(The good time makes strong.)	*(The good time makes strong/gives hope.)*
Die/Das humanitäre Hilfe wächst.	La ayuda solidaria/o aumenta.
(The humanitarian help increases.)	*(The solidarian help increases.)*
Die/Das idyllische Stunde vergeht.	Las horas idílicas/os pasan.
(The idyllic hour passes by.)	*(The idyllic hours pass by.)*