BRAIN RESEARCH 1527 (2013) 149–160

Available online at www.sciencedirect.com
SciVerse ScienceDirect
www.elsevier.com/locate/brainres


Brain Research

Research Report

An ERP study of structural anomalies in native and semantic free artificial grammar: Evidence for shared processing mechanisms

Ángel Tabullo[a], *Yamila Sevilla*[b], *Enrique Segura*[a], *Silvano Zanutto*[a], *Alejandro Wainselboim*[a,*]

[a]*Instituto de Medicina y Biología Experimental, Conicet, Argentina*
[b]*Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Argentina*

A R T I C L E I N F O

Article history:
Accepted 11 May 2013
Available online 24 May 2013

Keywords:
Native grammar
Artificial grammar
Sequential learning
ERPs
P600

A B S T R A C T

Artificial grammars have been widely applied to the study of sequential learning in language, but few studies have directly compared the neural correlates of artificial and native grammar processing. In this study, we examined Event Related Potentials (ERPs) elicited by structural anomalies in semantic-free artificial grammar sequences and sentences in the subjects' native language (Spanish). Although ERPs differed during early stages, we observed similar posterior negativities (N400) and P600 effects in a late stage. We interpret these results as evidence of at least partially shared neural mechanisms for processing of language and artificial grammars. We suggest that in both the natural and artificial grammars, the N400 and P600 components we observed can be explained as the result of unfulfilled predictions about incoming stimuli.

1. Introduction

A growing body of evidence supports the claim that the extraction and processing of structural regularities might be relevant for language acquisition and processing (Gómez and Gerken, 2000; Saffran, 2003; Seidenberg et al., 2002). This process has been termed sequential learning, i.e. "…the ability to encode and represent the order of discrete elements occurring in a sequence…" (Conway and Christiansen, 2001), and may provide a mechanism "…for acquiring predictive relationships between sequence elements, independently of whether such regularities are represented in terms of rules, statistical associations, or some combination between the two…" (Christiansen et al., 2012). Several studies have shown that sequential learning may account for a range of linguistic phenomena that goes from phoneme (Maye et al., 2002) and word segmentation (Saffran et al., 1996; Peña et al., 2002); to acquisition of phrase structure and syntactic categories (Gómez and Gerken, 1999; Saffran, 2002; Thompson and Newport, 2007), both in children and adults (see Folia et al., 2010 for a review). In order to control for intrinsically linguistic properties, these studies have used artificial stimuli designed to capture some of the regularities found in natural language. This experimental paradigm is known as *artificial grammar learning* (Reber, 1967). Artificial grammars consist of a series of items (letters, phonemes, words, strings) and a

---

*Correspondence to: Vuelta de Obligado 2490 (C1428ADN), Buenos Aires, Argentina. Fax: +54 11 4786 2564.
E-mail address: awainselboim@ibyme.conicet.gov.ar (A. Wainselboim).

0006-8993/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.brainres.2013.05.022

specific set of rules to combine them. In the absence of semantic information, subjects can rely only on sequential information to learn how the elements can be ordered. After being exposed to a number of artificial grammar sequences, subjects usually show above-chance performance when asked if a new sequence has been built according to the grammar rules (e.g. Reber, 1989, 1993). This marks the ability to detect and extract structural regularities of the stimuli, even in cases where subjects do not evidence an explicit knowledge of them.

Although artificial grammars have been widely applied to language research, few studies have tried to verify the existence of shared neural mechanisms for artificial grammar and natural language processing. So far, neuroimaging studies performed on the matter indicate that language-related brain regions (particularly, Broca's area—BA 44/45) are indeed engaged in artificial grammar processing (Bahlmann et al., 2008; Petersson et al., 2004, 2012; Folia et al., 2011). In addition, the potential overlap between electroencephalographic (EEG) correlates of artificial and native grammar processing has only been addressed recently (Christiansen et al., 2012).

EEG studies of language syntax processing have identified an ERP component that is present in a wide variety of syntactic contexts, a late posteriorly distributed positivity called P600 (also known as *syntactic positive shift*) (Hagoort, et al., 1993; Osterhout and Holcomb, 1992). P600 has been found after morphosyntactic violations (Hagoort et al., 1993; Gunter et al., 1997) and phrase structure violations (Hagoort et al., 1993; Neville et al., 1991), where it is usually preceded by a left anterior negativity (LAN) (Coulson et al., 1998; Friederici et al., 1993; Gunter et al., 2000; Hagoort et al., 2003) or an early left anterior negativity (ELAN) (Friederici et al., 1993; Hahne and Friederici, 1999; Lau et al., 2006) respectively. P600 has also been reported in well-formed sentences, such as dispreferred continuations of syntactically ambiguous sentences (garden path effect) (Osterhout and Holcomb, 1992; Osterhout et al., 1994; Kaan and Swaab, 2003), or long-distance structural dependencies (Felser et al., 2003; Fiebach et al., 2002; Kaan et al., 2000; Phillips et al., 2005). The component was initially considered to reflect syntactic reanalysis (in the case of garden path effects) and repair (in the case of syntax violations) processes (Friederici, 1995; Münte et al., 1997) occurring in a late and more controlled stage, while automatic first parsing processes based on word category information would be indexed by ELAN (Hahne and Friederici, 1999). An alternative and later interpretation claimed that the component was not restricted to reanalysis and repair, but instead reflected the cost of integration of a word with the previous syntactic context (Kaan et al., 2000). According to the authors, and following Gibson (1998), incoming words allow the parser to make predictions about the syntactic category of following input (for instance, a clause – initial "who" – phrase predicts a verb or preposition which can assign a thematic role to it), and syntactic integration is the process of combining the current input with them. When these predictions are unfulfilled by the current word, additional resources are needed to parse the sentence into a cohesive structure, thus resulting in a higher integration cost. The authors proposed that P600 would be an index of this syntactic integration difficulty. Reanalysis and repair situations would be specific cases of increased integration difficulty, since both require additional processing of the unexpected or anomalous structure to combine it with the previous context. Yet according to this account, even in the absence of reanalysis a P600 component should also be elicited by grammatical sentences at points where integration difficulty is increased compared to control cases. In their study, Kaan et al. (2000) found a P600 effect in sentences where a *wh*-phrase (which tends to be immediately integrated at the verb) was not adjacent to the verb, thus supporting their integration difficulty hypothesis. Later accounts have also considered P600 as an index of prediction and integration processes (Hagoort, 2003, 2009).

If artificial grammars constitute a valid paradigm to study language acquisition and processing, then the EEG correlates of structural anomalies in artificial grammar sequences should resemble those of natural language. An artificial language learning study found that after extensive training, highly proficient subjects elicit a native-like pattern of ELAN-P600 after structural anomalies (Friederici et al., 2002). It should be noted that this study intended to replicate the conditions for second language (L2) learning, and thus provided a more ecologically valid setting than most artificial grammar studies (where learning takes place by simple exposure and the sequences have no referent or meaning), including a semantic context for the artificial grammar (the artificial lexicon included nouns, verbs, adjectives and adverbs) and a feedback-based lengthy training. Under these conditions, it is possible that new linguistic-like stimuli might be processed in a native like manner. In addition, Bahlmann et al. (2006) compared ERP responses to violations of local and hierarchical rules in artificial grammars, showing a late posteriorly distributed positivity similar to P600 in both cases. The authors concluded that the late posterior positivity observed after grammar violations could be considered an index of the difficulty to integrate the anomalous stimuli with the previous sequence. Tabullo et al. (2011) also found ERPs compatible with P600 effects after expectancy violations in artificial grammar sequences. Nevertheless, none of the previous works included a direct within-subject comparison of artificial and native grammar ERPs. In a recent work, Christiansen et al. (2012) addressed this specific issue. In this case a similar P600 effect was found for both grammars, while a LAN was observed in natural language sentences only. The authors interpreted this result as evidence of at least partially overlapping neural mechanisms for artificial and natural grammar processing, and proposed that "the P600 component is not language-specific (…) rather, it is a broader index of violations and the cost of integration of expectations based on sequential (statistical) learning processes". This study did not include an explicit feedback-based training like Friederici et al. (2002), although a semantic context for the artificial grammar (which was a simplified version of that of Friederici's study) was present. In this case, the items in the artificial grammar lexicon corresponded to the shapes and colors of visual referents, and the sequences described visual scenes.

Given the aforementioned results, and following Christiansen et al. (2012), we wished to (i) compare the ERP correlates of structural anomalies in natural language and artificial grammar, and (ii) determine the degree of overlap between them.

Even though previous studies found P600 effects both in presence and absence of semantics during artificial grammar training (Mueller et al., 2005,2008), we considered that the most conservative approach was to eliminate semantic content from training, since we wished to focus the analysis on the processing of predictive associations between items. Therefore, a semantic-free artificial grammar was employed. If an analogous P600 effect were to be found in both natural and artificial grammars, it would suggest that violations of expectations based on predictive associations of sequential stimuli (such as artificial grammars) are processed in a similar way to language syntactic expectations and/or involve at least partially overlapping neural mechanisms. As in Christiansen et al. (2012) we did not expect to find an overlapping LAN effect, since both artificial language (Friederici et al., 2002) and second language studies (Hahne, 2001; Rossi et al., 2006) suggest that this ERP is only observed outside native language after long exposure and high proficiency levels have been achieved.

2. Results

2.1. Behavioral data

Participants performance (proportion of correct responses) was significantly better than chance in both the artificial grammar (*Mean*$=65.32\pm8.16\%$; $t(14)=7.27$, $p<0.001$) and native grammar (*Mean*$=97.22\pm3.73$; $t(14)=48.97$, $p<0.001$) tasks. Their proportion of correct responses was significantly higher in the native grammar task ($t(14)=-15.35$, $p<0.001$), which was expected since their exposure to the artificial grammar was brief (ca. 20 min) compared to their experience with their native language.

2.2. ERP analysis

Visual inspection of Grand Average waveforms showed an early positivity between 200 ms to 400 ms after the onset of the critical word in the native grammar. The native grammar positivity was higher for syntax violations and was more prominent at central and posterior sites with a latency and topography similar to that of a P300 (Duncan-Johnson and Donchin, 1982; see Picton, 1992; Polich, 1998, 2007; Verleger, 1997, for reviews). The artificial grammar also showed a positivity that appeared to be larger for grammatical sequences. In both natural and artificial grammars a posterior negativity could also be observed beginning approximately 440/450 ms after the onset of an ungrammatical sentence. However, the effect was more widely distributed in the native grammar, and seemed lateralized to the right and more brief in the artificial grammar (450–550 ms). The topography and latency of these effects is similar to that of the N400 (see Kutas et al., 2006; Kutas and Federmeier, 2011; Lau et al., 2008, for reviews). Finally, a late positivity was found between 600 ms and 1000 ms for syntax violations in both natural and artificial grammars. The effect seemed widely distributed in an early window (600–800 ms) and more posteriorly distributed in a later window (800–1000). The latency and topography of this component were similar to that of P600 (Gouvea et al., 2010; Kuperberg, 2007; Kutas et al., 2006 for a review). Therefore, the three temporal windows that were considered for analysis based on visual inspection of the Grand Average waveforms were congruent with those found in previous literature.

Fig. 1 shows Grand Average ERP waveforms obtained with the native (upper panel) and artificial (lower panel) grammar for grammatical and ungrammatical trials. Fig. 2 displays difference wave topographies for both grammars in the aforementioned time-windows of interest. Statistical analysis of the effects is reported below.

2.2.1. *Early positivity (200–400 ms)*

In the native grammar data, the ANOVA showed a main effect of Grammaticality ($F(1,14)=14.991$, $p=0.002$, $\eta_p^2=0.517$) and a Grammaticality × Region interaction ($F(2,28)=5.787$, $p=0.013$, $\eta_p^2=0.292$). Ungrammatical sentences elicited higher positivities than correct ones. Pairwise comparisons showed that the effect was more significant at posterior ($p<0.001$) than anterior sites ($p<0.013$). In the artificial grammar data, no significant main effects or interactions were found. Comparison between grammars showed a significant Task × Grammaticality interaction ($F(1,14)=17.842$, $p=0.001$, $\eta_p^2=0.560$). The positivity was higher for ungrammatical sentences in the native grammar ($p=0.002$), while no significant differences were observed in the artificial grammar data ($p=0.611$).

2.2.2. *N400 effects*

In the native grammar data, the ANOVA indicated a marginally significant main effect of grammaticality ($F(1,14)=3.823$, $p=0.071$, $\eta_p^2=0.214$) and a significant Grammaticality × Region interaction ($F(2,28)=4.027$, $p=0.029$, $\eta_p^2=0.223$). Pairwise comparisons showed that ungrammatical sentences elicited significantly larger negativities at posterior sites ($p=0.032$), and marginally significant at central sites ($p=0.056$). The ANOVA of the artificial grammar data (450–550 ms) showed significant Grammaticality × Region ($F(2,28)=8.254$, $p=0.008$, $\eta_p^2=0.371$) and Grammaticality × Lateral position ($F(4,56)=4.629$, $p=0.018$, $\eta_p^2=0.248$) interactions. The negativity was larger for ungrammatical sequences in central and posterior sites (p's<0.039) located in the right hemisphere (p's<0.012). The effects in both natural and artificial grammars were compared within the 450–550 ms window of interest by means of a Task (native grammar, artificial grammar) × Region × Lateral position × Grammaticality repeated measures ANOVA. A significant effect of Grammaticality ($F(1,14)=8.153$, $p=0.013$, $\eta_p^2=0.368$) and a marginally significant effect of Task ($F(1,14)=3.571$, $p=0.08$, $\eta_p^2=0.203$) were found, but no Task × Grammaticality interaction ($p=0.440$) was observed, indicating that the N400-like component did not differ significantly between grammars.

2.2.3. *P600 effects*

Within the first window of interest (600–800 ms), the ANOVA of both the native grammar ($F(1,14)=20.249$, $p<0.001$, $\eta_p^2=0.591$) and the artificial grammar ($F(1,14)=6.161$, $p=0.026$, $\eta_p^2=0.306$) yielded a significant main effect of Grammaticality,

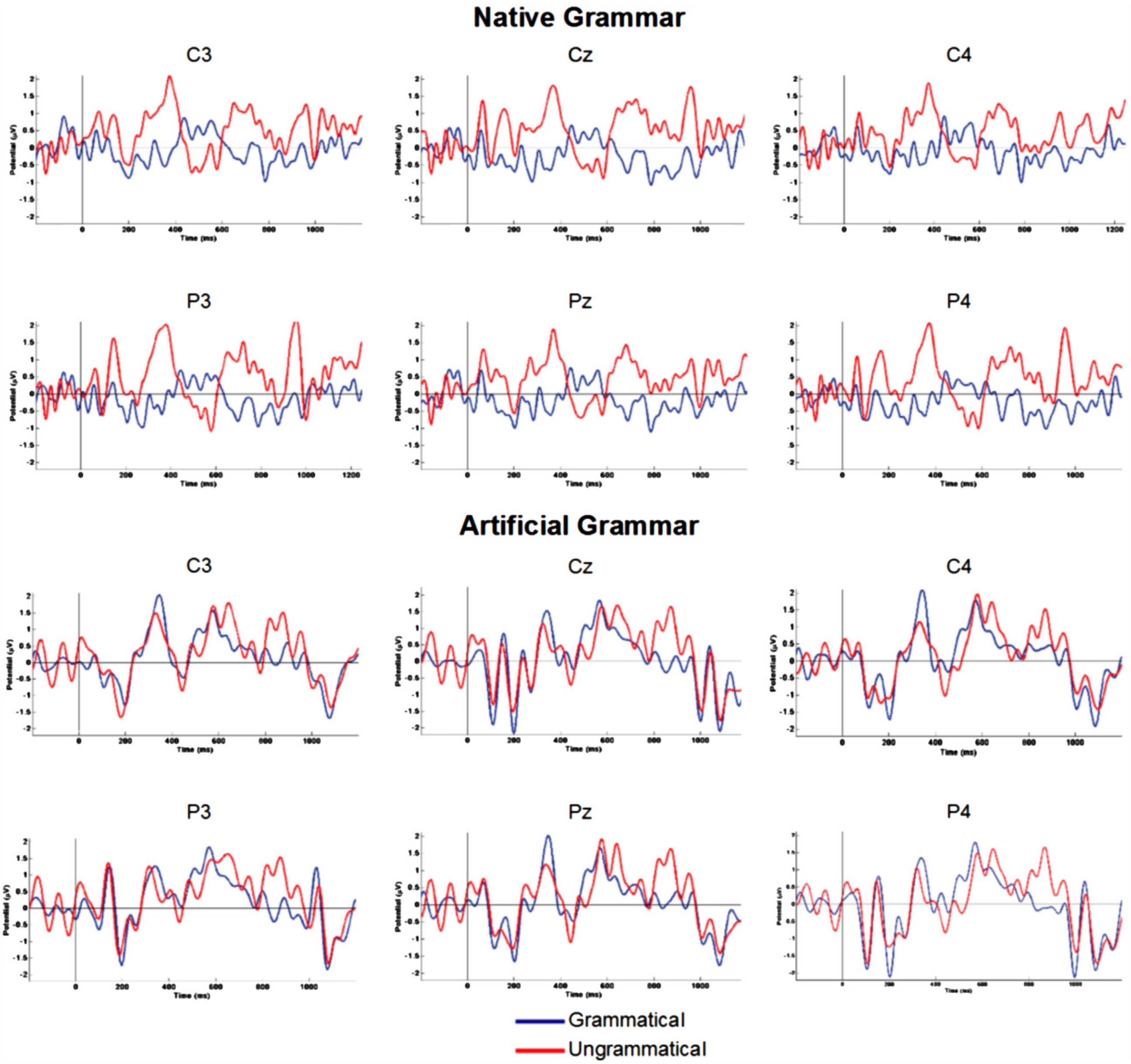


Fig. 1 – Grand average waveforms for the grammatical (solid black) and ungrammatical (dashed gray) trials in the native grammar (top) and artificial grammar (bottom) task. Central and parietal electrodes are shown from left, midline and right scalp sites. Grand average waveforms for the grammatical (blue) and ungrammatical (red) trials in the native grammar (top) and artificial grammar (bottom) data. Central and parietal electrodes are shown from left, midline and right scalp sites.

though its significance and effect size were larger in the native grammar task. The effects were compared by means of a Task (native grammar, artificial grammar) × Region × Lateral position × Grammaticality repeated measures ANOVA. While main effects of Grammaticality ($F(1,14)=27.024$, $p<0.001$, $\eta_p^2=0.659$) and Task ($F(1,14)=10.753$, $p=0.005$, $\eta_p^2=0.434$) were observed, no Task × Grammaticality interaction ($p=0.225$) was found, indicating that the P600 effect did not differ significantly among grammars. The grammaticality main effect indicated that ERPs were more positive for both grammatical and ungrammatical sentences in the native grammar data.

Within the second window of interest (800–1000 ms), significant main effects of Grammaticality (*Native grammar:* $F(1,14)=9.670$, $p=0.008$, $\eta_p^2=0.409$; *Artificial grammar:* $F(1,14)=8.673$, $p=0.011$, $\eta_p^2=0.383$) and Region × Grammaticality interactions (*Native grammar:* $F(2,28)=15.548$, $p=0.001$, $\eta_p^2=0.526$; *Artificial grammar:* $F(2,28)=20.748$, $p<0.001$, $\eta_p^2=0.597$) were observed for both grammars. Pairwise comparisons showed that the effect was significant at central (*Native grammar:* $p=0.006$; *Artificial grammar:* $p=0.012$) and posterior sites (*both grammars:* $p=0.001$). Comparison of P600 effects replicated the Region × Grammaticality interaction ($F(2,28)=25.371$, $p<0.001$, $\eta_p^2=0.644$) and showed a significant Task × Region interaction ($F(2,28)=11.219$, $p=0.003$, $\eta_p^2=0.445$), but no Task × Grammaticality interaction ($p=0.596$). ERPs were more positive in the artificial grammar for both conditions at frontal sites ($p=0.007$), but the P600 did not differ among grammars.

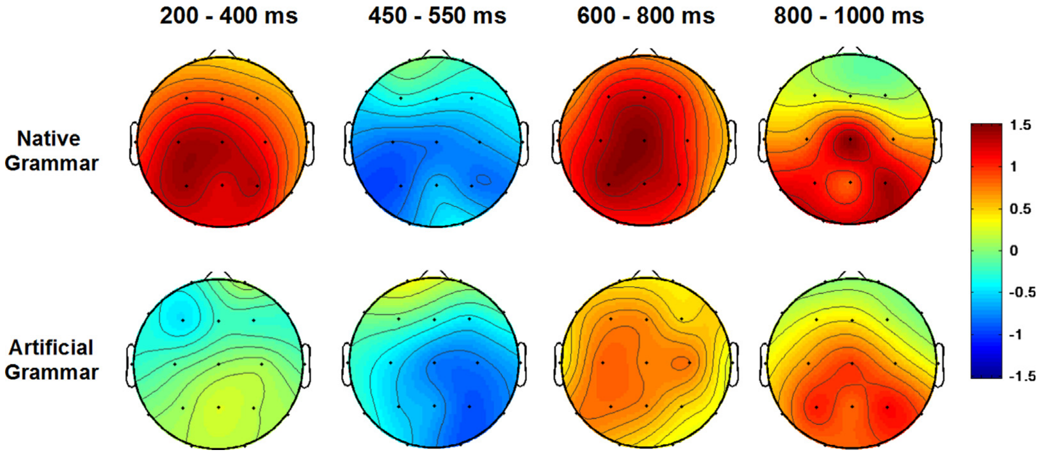


Fig. 2 – Difference wave scalp topographies for the native and artificial grammar tasks, averaged within each time-window of interest. Scale is given in microvolts.

3. Discussion

Our results showed that participants acquired significant knowledge about the semantic-free artificial grammar, and were able to build structural expectations of the stimulus sequences, thus performing above chance in the grammaticality judgment task. Furthermore, when processing of structural anomalies in the artificial grammar sequences was compared with syntax violations of their native grammar, a similar ERP pattern emerged at late processing stages, although the early components differed. Within an early window (200–400 ms), a P300 component was observed for ungrammatical sentences in the native grammar, while no significant effects were found in the artificial grammar instead. These ERPs were followed in both natural and artificial grammars by a similar N400-like posteriorly distributed negativity, and a widely distributed P600 effect. Interpretation and implications of these findings are discussed below.

i. Different ERP patterns at early processing stages

Differences observed between artificial and native grammar ERPs would suggest different neural mechanisms at early processing stages. In the native grammar, we expected to find an early left anterior negativity (ELAN) for syntax violations. The ELAN has been observed after phrase structure violations of native language, and has been proposed to reflect early automatic parsing processes (Friederici et al., 1993; Hahne and Friederici, 1999; Lau et al., 2006; Neville et al., 1991). However, we observed a P300 effect after the onset of the critical word instead. We interpreted this result as a consequence of the type of violation we chose for the ungrammatical sentences. While placing a preposition after the verb renders the sentence ungrammatical (e.g.: *La guía les recomendó a los turistas que observaran **del** paisaje. /The guide recommended the tourists that they observed **of** the landscape*"), a valid (though awkward) structure could still be possible if the sentence had continued including a subject for the subordinate clause (e.g: *La guía les recomendó a los turistas que observaran **del** paisaje los árboles./The guide reccomended the tourists that they observed **of** the landscape the trees*). At the onset of the preposition, the participants could not be certain whether the sentence would be a plain violation or a very infrequent (though valid), form. This syntactic ambiguity may have induced a different path of processing rather than the plain detection of a violation. In fact, the pattern of an early positivity followed by a later P600 effect has been reported in a previous study of native Spanish speakers (Leone-Fernández et al., 2012). The authors compared processing of two forms of the Spanish auxiliary verb "to be" (*ser* and *estar*), following either a concrete object or an event as subject nouns. They observed a typical P600 effect when the verb *ser* followed a concrete object (*"La silla **es** en la cocina"/"The chair [+Obj] is [+Ser] in the kitchen"*), which is considered syntactically ill-formed. However, the verb *estar* following an event ("La fiesta está en la cocina"/"The party [+Ev] is [+Estar]in the kitchen"), which was considered an anomalous, but more acceptable structure according to a previous grammaticality judgment task, elicited an early and centrally distributed positivity followed by a frontal P600 effect. This ERP pattern was interpreted as a P300-related component (Donchin and Coles, 1988; Polich, 2007), signaling the detection of an unexpected element, followed by a P600 reflecting reanalysis and repair operations. While our stimuli did not include the same verbs and sentence structures, they do have in common the fact that, at the onset of the critical word, ungrammatical sentences could still have a meaningful interpretation despite their unusual structure. This might explain the similar ERP pattern observed in Leone-Fernández et al. (2012) and the present work, as well as the absence of a LAN in both cases. In order to verify if our ungrammatical sentences could be considered as syntactically awkward, but potentially meaningful (like those in Leone Fernandez study), we conducted a syntactic acceptability test with 35 new participants. Experimental sentences were: syntactically correct (*El municipio les prohibió a los contrastistas asfaltar la calle*/City hall forbade the contractors to pave the streets), syntactically incorrect (*La guía les recomendó a los turistas que observaran **del** paisaje.*/"The guide recommended the tourists that they observed **of** the landscape) or syntactically anomalous (*La guía les recomendó a los turistas que observaran* **del** *paisaje los árboles./The guide recommended the tourists that they observed* **of** *the landscape the trees*). Sentences were ranked on a 5-point scale, from 1 ("Horrible") to 5 ("Perfect"). As expected, acceptability differed significantly among sentences ($F(2,88)=119.082$, $p<0.001$, $\eta_p^2=0.730$). Syntactically correct sentences were ranked as more acceptable than the rest (p's<0.001) but, more crucially, anomalous sentences were considered more acceptable than syntactically incorrect ones ($p=0.036$). Therefore, at the onset of the preposition, there was still the possibility of encountering a more plausible structure than a plain violation, which could have led to an ERP pattern similar to Leone-Fernandez study. Within this context, it is not surprising that the ERPs had differed between the native and the artificial grammar, since the level of ambiguity that could be found at the onset of the critical word in the real language sentences was absent in the artificial grammar sequences (where there was no possible continuation that could have "repaired" the structure after a violation). Other types of category violations, or even agreement violations may have elicited a more typical ELAN/LAN component in the native grammar. In addition, even though no left anterior negativities could be identified in our artificial grammar, it has been shown that LAN-like ERPs can be obtained after an extensive training in an artificial language (Friederici et al., 2002). In the same way, while the LAN is typically absent in late L2 speakers (Hahne, 2001), it can be found in highly proficient subjects (Rossi et al., 2006).

ii. Similar ERP pattern at later processing stages

The presence of an N400 after a syntax violation in the native grammar might seem surprising, considering that this component has been mostly reported to be elicited by words that are unexpected or incongruent with the previous semantic context. This component results even more counterintuitive in the case of the artificial grammar, where no semantic information was available. Nevertheless, this pattern of results has been previously reported in a study (Mueller et al., 2008) that compared the processing of case violations in a group of native Japanese speakers with participants that learned a semantic-free artificial grammar based on the syntax rules of Japanese. Case violations elicited a biphasic N400–P600 effect in both groups, even though the artificial grammar learners observed sentences with no referents, and could not associate them to any kind of meaning. While the N400 observed in the Japanese native speakers could be interpreted as indicative of a conflict in the domain of thematic processing (Mueller et al., 2005), the absence of semantic content in the artificial grammar learner group made that explanation unlikely. Instead, the authors concluded that the component was reflecting the violation of a lexical expectancy regarding a specific word form (a suffix). The learner group might have been able to acquire these expectancies due to their exposure to the relational structure of the case marked arguments in the training stimuli. The sensitivity to these patterns of co-occurrences would have been enough to generate a native-like ERP pattern when these subjects were faced with case violations. The authors' claim was based on two previous sources of evidence: (1) the N400 component has been reported to be sensitive to earlier and lower levels of lexical processing such as lexical accessing (Deacon et al., 2004; Doyle et al., 1996); and (2) N400 modulations have been reported in natural (McLaughlin et al., 2004) and artificial grammar (Sanders et al., 2002) studies in absence of semantic content. While our artificial grammar did not include case marking, participants in our study may have been able to acquire predictive associations between the individual ordered items of the sequences, allowing them to expect particular items or perceptual properties at different positions. For instance, participants may have learned that both *ducu* and *nili* can only be followed by *bane, cane, fadi, pefa* or *leda* (category A items), and never by *siru, soti, nomi, revu* or *roli* (category B items), even when they did not treat A and B tokens as belonging to an actual syntactic category. These lexical expectancies would have been unfulfilled in sequence violations, thus eliciting an N400 effect. Furthermore, Tabullo et al. (2011) in a previous experiment with the same artificial grammar items and combinatorial possibilities, found a similar N400-like negativity starting 400 ms after the onset of the critical item in ungrammatical sequences. This interpretation is in accordance with previous semantic-free grammar studies that have explained the appearance of N400 effects in terms of word form (phonological and/or orthographical) information processing that triggers a lexical search of the presented pseudo-words (McLaughlin et al. 2004; Mueller et al. 2008; Sanders et al. 2002).

We cannot fully discard that the observed N400 effect might have been the result of processing at the semantic level, since subjects might have developed their own "semantics" during the training phase of the experiment, arbitrarily assigning a *sui generis* meaning to the trained pseudo-words. Nevertheless, we find this possibility unlikely, since none of the participants referred having consciously assigned a meaning to the stimuli when they were debriefed. In addition, as mentioned above, previous studies that employed grammars without semantic content have proposed that the appearance of N400 effects are related to lexical search instead of semantic processes (McLaughlin et al., 2004; Mueller et al., 2008; Sanders et al. 2002).

In the native grammar, the onset of an unexpected lexical item (a preposition instead of an article or a pronoun) may have triggered similar processes. Therefore, our results are congruent with those of Mueller et al. (2008), and suggest that the N400 component observed in both the native and artificial ungrammatical stimuli may be the functional correlate of expectancy violations about formal properties of upcoming stimuli. It should be noted that, even though no significant differences were found between the N400 effects, the negativity observed in the artificial grammar was shorter and more right lateralized than the one in the native grammar. These differences might be at least partly explained by participants' proficiency differences between the artificial and native grammar tasks. A previous experiment has shown that both proficiency and age of acquisition of a language can modulate the latency and distribution of N400 effects (Newman et al., 2012).

Following the N400 effect, a late positivity was observed for structural anomalies in both native and artificial grammars, with a time-course and topography analogous to that of a P600. The presence of this component was quite expected, since P600 has been observed in a wide variety of linguistic and non-linguistic contexts. In native grammar studies, P600 has been found after morphosyntactic violations (Hagoort et al., 1993; Gunter et al., 1997) and phrase structure violations (Friederici et al., 1999; Hahne and Friederici, 1999), disprefered continuations of syntactically ambiguous sentences (Osterhout and Holcomb, 1992; Osterhout et al., 1994; Kaan and Swaab, 2003), or long-distance structural dependencies (Felser et al., 2003; Fiebach et al., 2002; Kaan et al., 2000; Phillips et al., 2005). Outside the language domain, similar late positivities have been observed in nonlinguistic contexts, like harmonic and diatonic anomalies in music (Patel et al., 1998; Patel, 2003) and violations of arithmetic series (Núñez-Peña and Honrubia-Serrano, 2004), mathematical operations (Martín-Loeches et al., 2006), geometric shape sequences (Besson and Macar, 1987), rule-governed letter sequences (Lelekov et al., 2000), and artificial grammars (Bahlmann et al., 2006).

The similarity between the P600 effects in the native and artificial grammar might be interpreted as evidence of (at least partially) overlapping neural mechanisms for language syntax and structured sequence processing. Similar findings regarding P600 had been previously reported in artificial language studies (Friederici et al., 2002; Mueller et al., 2005) that employed more realistic and “language-like” training, and in another study where a semantic-free artificial grammar was trained (Mueller et al., 2008). However, none of these studies provided a direct within-subject comparison of ERPs, but compared groups of native speakers and non-native learners instead. Christiansen et al. (2012) did compare within-subject ERP responses to anomalies in language and artificial grammar sequences. Their protocol included visual forms as referents for the items in the artificial grammar, and training sentences that described visual scenes, thus providing a semantic context for artificial language learning. In the present study, in order to center the analysis on the processing of predictive associations between items, a semantic-free artificial grammar was employed.

We suggest that the presence of the P600 in both the natural and artificial grammar can be interpreted in terms of Kaan's account: an index of the processing cost to integrate upcoming words to the preceding context (Kaan, 2007, 2009; Kaan et al., 2000; see also Hagoort, 2009). In the case of natural language, online sentence comprehension is proposed to be based to a substantial degree on predictive processing (Gibson, 1998; Hagoort, 2009; see also DeLong et al., 2005, van Berkum et al., 2005) where lexical, semantic and syntactic information is combined to predict the characteristics of the incoming word. This would allow for faster processing, since semantic and syntactic information becomes pre-activated before the next stimuli is actually perceived. However, when the input does not match the predictions based on the preceding context, integration becomes more difficult and additional resources are recruited for further processing, thus eliciting a P600. This interpretation accounts for most of the instances of P600 reported so far: syntax violations, dispreferred grammatical continuations in garden path sentences and long distance structural dependencies (see Kaan, 2007; Kaan et al., 2000; also Christiansen et al., 2012 for an extensive review). In particular, it applies to our natural grammar violations, where the preceding syntactic structure (the verb followed by an indirect object) allowed to predict an infinitive or a *that* complement, rendering the appearance of a prepositional phrase unlikely. Therefore, regardless of whether the subjects failed to integrate the preposition to the sentence, or analyzed it as part of an awkward but plausible phrase (as seen on Leone-Fernández et al., 2012), the integration difficulty account predicts a P600 effect due to the higher processing cost of the unexpected linguistic stimuli. Regarding our artificial grammar stimuli, we propose that predictive associations can also be learned between tokens from different categories. In this way, participants might be able to expect *ducu* after a category B token, *nili* after a C token, and a category A token after *ducu* or *nili*. These predictions would have been unfulfilled in the case of sequence violations, and the additional neural resources recruited to process the unexpected stimuli would have been reflected in the P600. Previous research shows that this kind of predictive associations is indeed possible in linear finite-state artificial grammars (Bahlmann et al., 2006). Furthermore, this study also concluded that the late posterior positivity observed after grammar violations could be considered an index of the difficulty to integrate the anomalous stimuli with the previous sequence.

It should be noted that violations in both our native and artificial grammars were, in principle, similar in nature: both involved the intrusion of an item (word or pseudo-word) in an invalid position within the sequence. In addition, it is possible that the prepositions in the native grammar may have been considered as a highly unusual and/or ambiguous element rather than a plain syntax violation (given the possibility of a meaningful, though unlikely, continuation of the sentence). This would explain why we observed an early P300 followed by a P600 instead of the typical ELAN-P600 pattern elicited by language phrase structure violations. The P300 could be signaling the allocation of attentional resources (Johnson, 1988) engaged by the unexpected syntactic item, and the P600 could be the result of a higher structural integration cost (Kaan et al., 2000). As we mentioned in the previous section, this potential ambiguity was absent in artificial grammar violations, where we still observed a similar P600 effect. Therefore, this similarity suggests that P600 is indeed associated with the cost of unfulfilled structural predictions. Similarly, Christiansen et al. (2012) noted that, due to the brief exposure to the artificial grammar, subjects may have failed to acquire actual category information, processing category violations as simple unfulfilled sequential expectations instead. If this was indeed the case, predictions based on the linear ordering of the artificial grammar and the linguistic context in the natural grammar both elicited a P600 response when incoming stimuli failed to meet them.

In addition, evidence from a previous study by our lab (Tabullo et al., 2011), using an artificial grammar with the same items and combinatorial rules, although with differences in the training protocol, suggests that the P600 effect may be modulated by the subject's expectancy of the incoming stimuli. In this experiment, we presented one of the syntax structures more frequently than the other during training, and compared the ERPs evoked by grammatical frequent, grammatical infrequent and ungrammatical sequences. We found that, within 560–660 ms, not only ungrammatical but also correct but *infrequent* sequences elicited a posterior positivity when compared to frequent ones. The effects of probabilistic manipulations and subjective expectancy on P600 has been interpreted as evidence that the component is a particular instance of P300, a family of ERPs related to the detection of unexpected and infrequent stimuli (Gunter et al., 1997; Coulson et al., 1998), although dissociations between the two components have been reported in basal ganglia lesioned patients (Frisch et al., 2003) and aphasics (Wassenaar et al., 2004) (see Osterhout and Hagoort, 1999 for more counterarguments). While the P600/P300 debate exceeds the scope of the present work, we did observe separate P300 and P600 effects in the native grammar, therefore our results do not suggest that these components can be considered identical.

Regarding P600, our results constitute a replication of Christiansen et al. findings, obtained on a different linguistic population (Spanish speakers), using a semantic-free training protocol and presenting a different kind of natural syntax violations (our study included phrase structure violations, while Christiansen et al. (2012) showed subject noun/verb number agreement violations). The fact that we still observed analogue components in artificial and natural grammar violations supports the hypothesis of at least a partial overlap between both kinds of processing. In their study, Christiansen et al. (2012) propose that P600 is not restricted to language, but is"*…a broader index of violations and the cost integration of expectations based on sequential learning processes*". Other researchers also interpret the presence of similar positivities in non-linguistic contexts such as music and mathematics, as evidence that the P600 reflects domain-general processes that are recruited by language (Kaan, 2007, 2009; Hagoort, 2009; Patel et al., 1998). While we acknowledge that our current evidence is insufficient to draw conclusions about the domain-specificity of the component or its relation to statistical learning, we believe this possibility is worth exploring and should be addressed in further research.

4. Conclusions

Our results suggest that processing of structural anomalies in artificial grammar and natural language recruits at least partially overlapping neural mechanisms, which is in agreement with Christiansen et al. (2012) study. We suggest that both the N400 and P600 components we observed can be explained as the result of unfulfilled predictions about incoming stimuli in both the natural and artificial grammars that affect lexical search (N400) and structural integration processes (P600). Therefore, we propose to further investigate the nature of these processes and their domain specificity. Finally, we reckon that our conclusions are based on the interpretation of the lack of significant differences between experimental conditions, which always demands a certain caution. This methodology has led to similar results in previous studies that compared natural language with music (Patel et al., 1998) and artificial grammar (Christiansen et al., 2012).

5. Experimental procedures

5.1. *Participants*

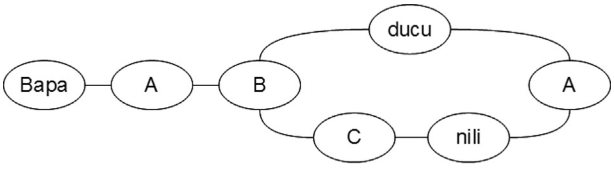
Fifteen subjects (9 female) participated in the study (age range: 19–34, mean: 26.4±4.3 years). All subjects were right-handed, spoke Spanish as first language and had completed college level education. Subjects had no history of neurologic or psychiatric disorders, neither were they under any medication at the moment of the experiment.

5.2. *Stimulus material*

A finite-state artificial grammar with a lexicon of 17 pseudo-words was designed for the study. Pseudo-words were phonotactically correct but nonexistent in Spanish, and they all shared the structure: "vowel-consonant-vowel-consonant". The lexicon was composed of 3 categories of pseudo-words (A: "bane", "cane", "fadi", "pefa", "leda"; B: "siru", "soti", "nomi", "revu", "roli"; C: "pere", "mene", "tese", "sele") and 3 conjunctions: (1: "bapa", 2: "ducu", 3: "nili"). Pseudo-word categories were arbitrarily defined based on formal characteristics in the following way: (i) Category A: pseudo-words without "o" or "u"; (ii) Category B: Pseudo-words without "a", ending in "u" or "i"; Category C: Pseudo-words only with "e".

The artificial grammar admitted two possible sequence types: type (1) conjunction 1 – category A – category B – conjunction 2 – category A (e.g.: *"bapa bane roli ducu fadi"*); type (2) conjunction 1 – category A – category B – category C – conjunction 3 – category A (e.g.: *bapa pefa soti tese nili cane*). Sequence types 1 and 2 were thus equivalent in their first 3 items, and differed after the third item (see Fig. 3). All pseudo-words within a given category were presented the same number of times during training, and none of them were repeated within the same sequence.

A total of 90 sequences from types 1 and 2 were created for the training stage. The test stage consisted of 160 new sequences, 80 were built according to the grammar rules (grammatical sequences) and 80 (ungrammatical sequences) contained one of two category violations: presenting "nili" when "ducu" was expected (and vice versa) or ending the



Category A: bane, cane, fadi, pefa, leda
Category B: siru, soti, nomi, revu, roli
Category C: pere, mene, tese, sele

Fig. 3 – Diagram of the finite-state artificial grammar. Each node represents an item or stimulus category, and the lines indicate valid transitions between them. Grammatical sequences are generated by connecting the nodes from left to right.

sequence with an item from category B instead of A (Table 1). Half of the test sequences were type 1 and the other half were type 2 sequences. Therefore, the test stage consisted of 40 grammatical and 40 ungrammatical sequences from each type.

Native grammar experimental stimuli consisted of 80 sentences, 40 grammatical and 40 with category violations. In half the grammatical sentences, the verb complement was a *that* clause (*La mediadora les permitió a las partes que expresaran sus reclamos/*“The mediator allowed the parts that they expressed their claims”) and in the other half it was an infinitive clause (*El canal les prohibió a sus periodistas mencionar el tema/*“The channel forbade its journalists to mention the topic”). Category violations in the ungrammatical sentences consisted in adding a preposition to the noun phrase within the complement *(La decoradora les recomendó a los jóvenes combinar* ***de*** *los colores/*“The decorator recommended the young men to combine **of** the colors”; *El modisto les permitió a las supermodelos que acortaran* ***de*** *los vestidos/The couturier allowed the supermodels that they shortened* ***of*** *their shorts*) (see Table 2). An additional list of 80 sentences, 40 of them grammatical and 40 containing verb inflection violations (*La radio les prohibió a sus locutoras que* ***entregar*** *las facturas/* “The radio forbade its announcers that **to deliver** the bills”) was used as filler materials. These stimuli were part of an additional experiment and were included as fillers in order to increase the difficulty of the task and prevent participants from focusing only on a specific part of the sentence and/or error type.

5.3. *Procedure*

The experiment was carried out in an acoustically and electrically isolated room. In the artificial grammar training stage, subjects were told they would read sentences from a fictional language, and were instructed to try to learn which combinatorial possibilities were allowed between pseudo-words.

Sequences where displayed at the center of a computer screen, one item at a time. Each item had a duration of 600 ms and interstimulus interval was set at 500 ms. Periodically, the software asked the subject if a specific item had appeared in the previous sequence, in order to ensure that his/her attention was directed to the stimuli. Subjects were exposed to 90 sequences in two blocks of 45 trials, with a short interval between them. After approximately 20 min of training, subjects moved on to the test stage. Subjects were explained that they would see new sequences, some of them congruent with the allowed combinatorial possibilities of the grammar, and some of them containing errors. They were asked to decide if each sequence was correct or incorrect, and to respond by pressing a key, as fast as they could, without making mistakes. No feedback was provided during the test phase, therefore participants did not know whether they had responded correctly or not, ensuring in this way that no further learning of the grammar rules occurred during this phase. 160 sequences were presented in two blocks of 80 trials, with a short break between them. EEG activity was recorded in this stage. The number of presented trials per condition during the test phase was designed to increase the reliability of the obtained ERP recordings.

After a short break following the artificial grammar test stage, participants moved on to the native grammar stage. They were informed that they would see now sentences from their native language and once again, they would have to decide whether they were correct or not. The 160 experimental and filler sentences were displayed in a similar way than artificial grammar sequences, and were divided in two 80-trial blocks with a break between them, while EEG activity was recorded.

All experimental tasks were programmed using the Python software platform (www.python.org).

5.4. *EEG recording*

Electroencephalographic activity was recorded from 19 cap-mounted tin electrodes (international 10/20 system, biauricular reference, Electro-Cap International Inc.). Electrode impedances were kept under 10 kΩ. EEG was sampled at 256 Hz, and bandpass filtered at 0.5–30 Hz. ERP's were time-locked to the onset of the violation in the ungrammatical

Table 1 – Artificial grammar stimuli.

Sequence type	Grammatical sequence	Ungrammatical sequence
(i)	*Bapa pefa soti ducu bare*	*Bapa pefa soti* ***nili bare*** *Bapa cane siru ducu* ***soti***
(ii)	*Bapa fadi revu tese nili lane*	*Bapa fadi revu tese* ***ducu*** *lane* *Bapa pefa soti pere nili* ***siru***

Table 2 – Native grammar stimuli.

Sentence type	Grammatical sentence	Ungrammatical sentence
That clause	*La mediadora les permitió a las partes que expresaran sus reclamos*/The mediator allowed the parts that they expressed their claims	*El modisto les permitió a las supermodelos que acortaran* ***de*** *los vestidos*/The couturier allowed the supermodels that they shortened **of** their dresses
Infinitive clause	*El canal les prohibió a sus periodistas mencionar el tema*/The channel forbade its journalists to mention the topic	*La decoradora les recomendó a los jóvenes combinar* ***de*** *los colores*/The decorator recommended the young men to combine **of** the colors

sequences, and its corresponding item in the grammatical sequences. Epoch length was 2000 ms, with a 200 ms prestimulus interval as baseline. EEG signal processing and ERP analysis were carried out with EEGLAB software (Delorme and Makeig, 2004). Ocular artifacts were removed from data by means of ICA-based artifact correction, applying the ADJUST algorithm (Mognon et al., 2011). Epochs containing other kinds of artifacts were detected by visual inspection and excluded from the analysis (less than 10% of the trials, evenly distributed across conditions).

5.5. *Data analysis*

Mean voltage was calculated for each time window of interest, and analyzed by a $3 \times 5 \times 2 \times 2$ repeated measures ANOVA with the following within-subject factors: Electrode Region (ROIs: Anterior, Central, Posterior), Electrode lateral Position (ROIs: 1 to 5, from left to right: e.g., position 1 includes F7, T3 and T5, while position 5 includes F8, T4 and T6), Grammaticality (ungrammatical, grammatical) and Task (native grammar, artificial grammar). Effect-size was estimated with partial eta-squared coefficient η_p^2 (Cohen, 1973; Haase, 1983). Greenhouse-Geisser correction was applied to sphericity violations, and post hoc comparison p-values were Bonferroni adjusted.

REFERENCES

Bahlmann, J., Gunter, T.C., Friederici, A.D., 2006. Hierarchical and linear sequence processing: an electrophysiological exploration of two different grammar types. J. Cogn. Neurosci. 18, 1829–1842.

Bahlmann, J., Schubotz, R.I., Friederici, A.D., 2008. Hierarchical artificial grammar processing engages Broca's area. NeuroImage 42, 525–534.

Besson, M., Macar, F., 1987. An event-related potential analysis of incongruity in music and other non-linguistic contexts. Psychophysiology 24, 14–25.

Christiansen, M.H., Conway, C.M., Onnis, L., 2012. Similar neural correlates for language and sequential learning: evidence from event-related brain potentials. Lang. Cogn. Processes 27, 231–256.

Cohen, J., 1973. Eta-squared and partial eta-squared in fixed factor ANOVA designs. Edu. Psychological Meas. 33, 107–112.

Conway, C.M., Christiansen, M.H., 2001. Sequential learning in non-human primates. Trends in Cogn. Sci. 5, 539–546.

Coulson, S., King, J., Kutas, M., 1998. Expect the unexpected: event-related brain responses of morphosyntactic violations. Lang. Cogn. Processes 13, 21–58.

Deacon, D., Dynowska, A., Ritter, W., Grose-Fifer, J., 2004. Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. Psychophysiology 41 (1), 60–74.

DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8, 1117–1121.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Methods 134, 9–21.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of cognitive updating?. Behav. Brain Sci. 11, 357–427.

Doyle, M.C., Rugg, M.D., Wells, T., 1996. A comparison of the electrophysiological effects of formal and repetition priming. Psychophysiology 33 (2), 132–147.

Duncan-Johnson, C.C., Donchin, E., 1982. The P300 component of the event-related brain potential as an index of information processing. Biol. Psychol. 14, 1–52.

Felser, C., Clahsen, H., Münte, T.F., 2003. Storage and integration in the processing of fillergap dependencies: an ERP study of topicalization and wh-movement in German. Brain Lang. 87, 345–354.

Fiebach, C.J., Schlesewsky, M., Friederici, A.D., 2002. Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German wh-questions. J. Memory Lang. 47, 250–272.

Frisch, S., Kotz, S.A., von Cramon, D.Y., Friederici, A.D., 2003. Why the P600 is not just a P300: the role of the basal ganglia. Clin. Neuropsychol. 114, 336–340.

Friederici, A.D., 1995. The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. Brain Lang. 50, 259–281.

Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. Cogn. Brain Res. 1, 183–192.

Friederici, A.D., Steinhauer, K., Frisch, S., 1999. Lexical integration: sequential effects of syntactic and semantic information. Memory Cogn. 27, 438–453.

Friederici, A.D., Steinhauer, K., Pfeifer, E., 2002. Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. PNAS 99, 529–534.

Folia, V., Uddén, J., de Vries, M., Forkstam, C., Petersson, K.M., 2010. Artificial language learning in adults and children. Lang. Learn. 60 (2), 188–220.

Folia, V., Forkstam, C., Ingvar, M., Hagoort, P., Petersson, K.M., 2011. Implicit artificial syntax processing: genes, preference, and bounded recursion. Biolinguistics 5, 105–132.

Gibson, E., 1998. Linguistic complexity: locality of syntactic dependencies. Cognition 68, 1–76.

Gómez, R.L., Gerken, L., 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70, 109–135.

Gómez, R.L., Gerken, L.A., 2000. Infant artificial language learning and language acquisition. Trends Cogn. Sci. 4, 178–186.

Gouvea, A., Phillips, C., Kazanina, N., Poeppel, D., 2010. The linguistic processes underlying the P600. Lang. Cogn. Processes 25 (2), 149–188.

Gunter, T.C., Stowe, L.A., Mulder, G., 1997. When syntax meets semantics. Psychophysiology 34, 660–676.

Gunter, T.C., Friederici, A.D., Schriefers, H.J., 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. J. Cogn. Neurosci. 12 (4), 556–568.

Haase, R.F., 1983. Classical and partial eta square in multifactor ANOVA designs. Edu. Psychol. Meas. 43, 35–39.

Hagoort, P., 2003. How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. NeuroImage 20, S18–S29.

Hagoort, P., 2009. Reflections on the neurobiology of syntax. In: Bickerton, D., Szathmáry, E. (Eds.), Biological Foundations and Origin of Syntax. MIT Press, Cambridge, MA, pp. 279–296.

Hagoort, P., Brown, C.M., Groothusen, J., 1993. The syntactic positive shift (SPS) as an ERP-measure of syntactic processing. Lang. Cogn. Processes 8, 439–483.

Hagoort, P., Wassenaar, M., Brown, C., 2003. Syntax-related ERP-effects in Dutch. Cogn. Brain Res. 16, 38–50.

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

Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis. early automatic and late controlled processes. J. Cogn. Neurosci. 11 (2), 194–205.

Johnson Jr., R., 1988. The amplitude of the P300 component of the event-related potential: review and synthesis. In: Ackles, P., Jennings, J.R., Coles, M.G.H. (Eds.), Advances in Psychophysiology: A Research Annual, 3. JAI Press, Inc, Greenwich, CT, pp. 69–137.

Kaan, E., 2007. Event-related potentials and language processing. A brief introduction. Lang. Ling. Compass 1 (6), 572–591.

Kaan, E., 2009. Syntax and its putative relation to the brain. In: Bickerton, D., Szathmáry, E. (Eds.), Biological Foundations and Origin of Syntax, 3. MIT Press, Cambridge, MA Strüngmann Forum Reports.

Kaan, E., Swaab, T.Y., 2003. Electrophysiological evidence for serial sentence processing: A comparison between non-preferred and ungrammatical continuations. Cogn. Brain Res. 17, 621–635.

Kaan, E., Harris, A., Gibson, E., Holcomb, P.J., 2000. The P600 as an index of syntactic integration difficulty 15, 159–201Lang. Cogn. Processes 15, 159–201.

Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: challenges to syntax. Brain Res. 1146, 23–49.

Kutas, M., van Petten, C., Kluender, R., 2006. Psycholinguistics electrified II: 1994–2005. In: Traxler, M., Gernsbacher, M.A. (Eds.), Handbook of Psycholinguistics second ed. Elsevier, New York, pp. 659–724.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.

Lau, E., Stroud, C., Plesch, S., Phillips, C., 2006. The role of structural prediction in rapid syntactic analysis. Brain Lang. 98 (1), 74–88.

Lau, E., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de)constructing the N400. Nat. Rev. Neurosci. 9, 920–933.

Lelekov, T., Dominey, P.F., Garcia-Larrea, L., 2000. Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. Neuroreport 11 (5), 1129–1132.

Leone-Fernández, B., Molinaro, N., Carreiras, M., Barber, H.A., 2012. Objects, events and “to be” verbs in Spanish—an ERP study of the syntax–semantics interface. Brain Lang. 120 (2), 127–134.

Martín-Loeches, M., Casado, P., Gonzalo, R., de Heras, L., Fernández-Frías, C., 2006. Brain potentials to mathematical syntax problems. Psychophysiology 43, 579–591.

Maye, J., Werker, J., Gerken, L.A., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82, 101–111.

McLaughlin, J., Osterhout, L., Kim, A., 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. Nat. Neurosci. 7, 703–704.

Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48, 229–240.

Mueller, J.L., Hahne, A., Fujii, Y., Friederici, A.D., 2005. Native and non-native speakers’ processing of a miniature version of Japanese as revealed by ERPs. J. Cogn. Neurosci. 17 (8), 1229–1244.

Mueller, J.L., Girgsdies, S, Friederici, A.D., 2008. The impact of semantic-free second-language training on ERPs during case processing. Neurosci. Lett. 443 (2), 77–81.

Münte, T.F., Szentkuti, A., Wieringa, B.M., Matzke, M., Johannes, S., 1997. Human brain potentials to reading syntactic errors in sentences of different complexity. Neurosci. Lett. 17 235 (3), 105–108.

Neville, H., Nicol, J.L., Barss, A., Forster, K.I., Garrett, M.F., 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. J. Cogn. Neurosci. 3 (2), 151–165.

Newman, A.J., Tremblay, A., Nichols, E.S., Neville, H.J., Ullman, M.T., 2012. The influence of language proficiency on lexical semantic processing in native and late learners of English.
J. Cogn. Neurosci. 24 (5), 1205–1223.

Núñez-Peña, M.I., Honrubia-Serrano, M.L., 2004. P600 related to rule violation in an arithmetic task. Cogn. Brain Res. 18, 130–141.

Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. J. Memory Lang. 31 (6), 785–806.

Osterhout, L., Hagoort, P.A., 1999. A superficial resemblance does not necessarily mean you are part of the family: counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. Lang. Cogn. Processes 14, 1–14.

Osterhout, L., Holcomb, P.J., Swinney, D.A., 1994. Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. J. Exp. Psychol.: Learn., Memory, Cogn. 20 (4), 786–803.

Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J., 1998. Processing syntactic relations in language and music: an event-related potential study. J. Cogn. Neurosci. 10 (6), 717–733.

Patel, A.D., 2003. Language, music, syntax, and the brain. Nat. Neurosci. 6, 674–681.

Peña, M., Bonatti, L., Nespor, M., Mehler, J., 2002. Signal-driven computations in speech processing. Science 298, 604–607.

Petersson, K.M., Forkstam, C., Ingvar, M., 2004. Artificial syntactic violations activate Broca's region. Cogn. Sci. 28, 383–407.

Petersson, K.M., Folia, V., Hagoort, P., 2012. What artificial grammar learning reveals about the neurobiology of syntax. Brain Lang. 120 (2), 83–95.

Phillips, C., Kazanina, N., Abada, S., 2005. ERP effects of the processing of syntactic long distance dependencies. Cogn. Brain Res. 22, 407–428.

Picton, T.W., 1992. The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9, 456–479.

Polich, J., 1998. P300 clinical utility and control of variability. J. Clin. Neurophysiol. 15, 14–33.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148.

Reber, A.S., 1967. Implicit learning of artificial grammars. J. Verbal Learn. Behav. 6, 855–863.

Reber, A.S., 1989. Implicit learning and tacit knowledge. J. Exp. Psychol.: Gen. 118 (3), 219–235.

Reber, A.S., 1993. Implicit Learning and Tacit Knowledge. Oxford University Press, Oxford.

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

Saffran, J.R., 2002. Constraints on statistical language learning. J. Mem. Lang. 47, 172–196.

Saffran, J.R., 2003. Statistical learning: mechanisms and constraints. Curr. Directions Psychol. Sci. 12, 110–114.

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

Sanders, L.D., Newport, E.L., Neville, H.J., 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat. Neurosci. 5, 700–703.

Seidenberg, M.S., MacDonald, M.C., Saffran, J.R., 2002. Does grammar start where statistics stop?. Science 298, 553–554.

Tabullo, A., Sevilla, Y., Pasqualetti, G., Vernis, S., Segura, E., Yorio, A., Zanutto, B.S., Wainselboim, A.J., 2011. Expectancy modulates a late positive ERP in an artificial grammar task. Brain Res. 1373, 131–143.

Thompson, S., Newport, E., 2007. Statistical learning of syntax: the role of transitional probability. Lang. Learn. Dev. 3, 1–42.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. J. Exp. Psychol.: Learn., Mem. Cogn. 31, 443–467.

Verleger, R., 1997. On the utility of P3 latency as an index of mental chronometry. Psychophysiology 34, 131–156.

Wassenaar, M., Brown, C.M., Hagoort, P., 2004. ERP effects of subject-verb agreement violations in patients with Broca's aphasia. J. Cogn. Neurosci. 16, 553–576.