ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Acta Psychologica xxx (2007) xxx–xxx

acta psychologica

www.elsevier.com/locate/actpsy

ERPs reveal comparable syntactic sentence processing in native and non-native readers of English

Sonja A. Kotz [a,*], Phillip J. Holcomb [b], Lee Osterhout [c]

[a] Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany
[b] Department of Psychology, Tufts University, Medford, USA
[c] Department of Psychology, University of Washington, Seattle, USA

Received 1 October 2007; accepted 16 October 2007

Abstract

L2 syntactic processing has been primarily investigated in the context of syntactic anomaly detection, but only sparsely with syntactic ambiguity. In the field of event-related potentials (ERPs) syntactic anomaly detection and syntactic ambiguity resolution is linked to the P600. The current ERP experiment examined L2 syntactic processing in highly proficient L1 Spanish-L2 English readers who had acquired English informally around the age of 5 years. Temporary syntactic ambiguity (induced by verb subcategorization information) was tested as a language-specific phenomenon of L2, while syntactic anomaly resulted from phrase structure constraints that are similar in L1 and L2. Participants judged whether a sentence was syntactically acceptable or not. Native readers of English showed a P600 in the temporary syntactically ambiguous and syntactically anomalous sentences. A comparable picture emerged in the non-native readers of English. Both critical syntactic conditions elicited a P600, however, the distribution and latency of the P600 varied in the syntactic anomaly condition. The results clearly show that early acquisition of L2 syntactic knowledge leads to comparable online sensitivity towards temporal syntactic ambiguity and syntactic anomaly in early and highly proficient non-native readers of English and native readers of English.

PsycINFO Classification: 2720; 2530

Keywords: ERPs; P600; L2 syntactic processing; Syntactic ambiguity; Syntactic anomaly

1. Introduction

To fully account for a language user's competence of a language it is necessary to explain how different types of linguistic knowledge are used during sentence comprehension. A speaker of two languages presents a particular challenge as it is necessary to describe comprehension in the non-native and the native language. Previous psycholinguistic research investigating L2 acquisition has been primarily concerned with the representation of languages in bilinguals, and has focussed on two issues: the organization of the lexical and conceptual system, and the access of lexical-semantic information (see Kotz & Elston-Guettler, 2007 for a recent review).

The question of how non-native speakers comprehend syntactic structure in their L2 has received less attention, and previous research had only sparsely provided detailed and conclusive comparisons of native and non-native online syntactic comprehension. More recently though, L2 syntactic comprehension of, for example, phrase structure constraints and verb agreement has been investigated with online measures such as ERPs (Hahne, 2001; Rossi, Gugler, Hahne, & Friederici, 2006; Tokowicz & MacWhinney, 2005; Weber-Fox & Neville, 1996; Weber-Fox & Neville, 2001). The use of such a temporally fine-grained measure allows investigating how L2 speakers of a language process syntactic information while sentence comprehension unfolds in time. Most importantly, such an

* Corresponding author. Tel.: +49 341 9940231; fax: +49 341 9940113.
E-mail address: kotz@cbs.mpg.de (S.A. Kotz).

0001-6918/$ - see front matter
doi:10.1016/j.actpsy.2007.10.003

online measure can reveal implicit processing of syntactic knowledge that may be distinct from explicit syntactic knowledge as evidenced in grammaticality judgment (see for a similar argument, Tokowicz & MacWhinney, 2005). However, there is no ERP evidence on a second syntactic phenomenon namely syntactic ambiguity in the L2. Thus, the focus of the current L2 study was to investigate syntactic ambiguity and syntactic anomaly in parallel in non-native and native readers of English.

Monolingual research on syntactic processing has reported early syntactic components such as the early left anterior negativity (ELAN) and left anterior negativity (LAN) (for an overview see Friederici, 2002), and a late centro-parietal positivity called the P600 (e.g., Osterhout & Holcomb, 1992). The P600 is evoked by a number of different syntactic manipulations such as violations of agreement (Hagoort, Brown, & Groothusen, 1993; Osterhout & Mobley, 1995; Vos, Gunter, Kolk, & Mulder, 2001), of verb inflection (Friederici, Pfeifer, & Hahne, 1993; Gunter, Stowe, & Mulder, 1997), of case inflection (Münte, Heinze, Matzke, Wieringa, & Johannes, 1998), of pronoun inflection (Coulson, King, & Kutas, 1998), of phrase structure (Friederici et al., 1993; Hahne & Friederici, 1999; Neville, Nicol, Barss, Forster, & Garrett, 1991), of case marking (Coulson et al., 1998; Friederici & Frisch, 2000; Frisch & Schlesewsky, 2001, 2005), and of verb-argument structure (Friederici & Frisch, 2000; Frisch, Hahne, & Friederici, 2004; Osterhout, Holcomb, & Swinney, 1994).

In addition, a P600 is also elicited by syntactic ambiguity (Friederici, Hahne, & Mecklinger, 1996; Frisch, Schlesewsky, Saddy, & Alpermann, 2002; Kaan & Swaab, 2003; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Osterhout & Holcomb, 1992; Osterhout & Holcomb, 1993), and by increased syntactic complexity (Friederici, Hahne, & Saddy, 2002; Kaan, Harris, Gibson, & Holcomb, 2000; Kaan & Swaab, 2003). Some authors have argued that the P600 elicited by outright syntactic anomaly runs a different time course and displays a different scalp distribution than the P600 elicited by syntactic ambiguity (see Kaan & Swaab, 2003). While Hagoort and colleagues (Hagoort, Brown, & Osterhout, 1999; Van Berkum, Brown, & Hagoort, 1999) have linked the frontally distributed P600 to the revision of non-preferred ambiguous structures, Friederici et al. (2002) extended the frontal P600 to sentence complexity. On the other hand, Kaan and Swaab (2003) disputed the fact that the processing of sentence complexity and non-preferred syntactic structure is restricted to a frontal P600. Their results showed that the integration of syntactically complex and non-preferred structures also affects the posterior P600 that has been linked to repair processes of anomalous syntactic structure. Furthermore, the posterior P600 amplitude rise and latency was larger for anomalous than non-preferred structures. These results suggest that posterior positivities may reflect both revision and repair and can be distinguished by amplitude rise and latency.

ERP reports on L2 syntactic processing have not been very consistent. Weber-Fox and Neville (2001) investigated open/closed class word processing and reported an age of acquisition (AoA) effect (latency delays) for closed class words (N280), but not for open class words (N350). A P600 was only elicited by phrase structure violations if syntactic rules were acquired before the age of 10 (1996), while an early left anterior negativity (N125) was not found in any of the tested bilingual groups. A second and later left-lateralized negativity (300–500 ms) showed a bilateral distribution after an AoA of 10 years. Hahne and Friederici (2001) studied phrase structure violations in late (after the age of 10) L2 learners and did not find any syntax related ERP effects. In another study Hahne (2001) reported P600 effects (but no ELAN) for late L2 Russian learners of German. A similar result was reported by Müller and colleagues (Mueller, Hahne, Fujii, & Friederici, 2005). Native adult German speakers acquired a miniature grammar of Japanese. Phrase structure violations resulted in a P600. In a recent study by Rossi et al. (2006) the authors propose that the variable results in L2 syntactic processing may not just depend on AoA, but also on the level of achieved L2 proficiency. Comparing the ERPs of late high and low proficient German and Italian learners of Italian and German, respectively, they reported a comparable syntactic processing pattern for both phrase structure and agreement violations in late high proficient learners. Late low proficient learners showed a bi-phasic pattern (ELAN/P600) to phrase structure violations with a reduced and delayed P600, but only a delayed P600 to agreement violations. The authors claim that these differences across groups and syntactic structures suggest deviant neural processes in online syntactic and thematic processing in L2 learners despite highly advanced behavioural skills. However, the fact that native-like syntactic processing profiles can be seen in late proficient L2 learners calls into question whether AoA is the sole and driving force in L2 attainment of syntactic knowledge. Lastly, Tokowicz and MacWhinney (2005) addressed the critical issue of syntactic transfer effects in late L2 learners. Comparing tense-marking that was comparable between L1 and L2, determiner number agreement that differed between L1 and L2, and determiner gender agreement that was specific to L2, they reported P600 effects for similar sentence structures in L1/L2 and L2-specific syntax. In addition, they reported dissociations of online P600 effects and offline behavioural responses. While L2 learners displayed online sensitivity to grammatical violations, they responded at chance level during grammatical judgment. The authors concluded that behavioural task demands could cause such a dissociation (see for a similar argument McLaughlin, Osterhout, & Kim, 2004).

Next to considering the effects of cross-linguistic syntactic similarities, the effect of syntactic ambiguity resolution is a completely understudied issue in L2 online sentence processing, and has not been investigated with ERPs so far. However, eye-movement recordings show that highly proficient L2 readers can be "garden-pathed" (i.e., go by the preferred reading of a sentence structure (Frazier & Rayner,

1982) in the same manner as L1 readers (i.e., Frenck-Mestre & Pynte, 1997; Juffs & Harrington, 1996)). However, sensitivity to syntactic ambiguity appears to be critically influenced by AoA, proficiency, and cross-linguistic syntactic similarity (for a review see Frenck-Mestre, 2005). As reported in Dussias (2003), Fernàndez (1999) presented results that late learners of English (L1 Spanish) used L1 parsing strategies while processing syntactic ambiguity (relative clause attachment (RC)) in English), while early learners varied in their attachment preference with some showing a monolingual profile and others showing late learner profiles. Similarly, Dussias (2003) compared proficient L2 speakers of Spanish or English to monolingual speakers of the respective L2 while they were reading temporarily ambiguous sentences (RCs). This approach allowed testing whether L1 parsing strategies influence L2 parsing. Based on respective Spanish and English monolingual results Dussias predicted that if L2 speakers use the same structural and contextual constraints as a native speaker of the respective L2, then L2 speakers of Spanish should use a high attachment strategy and interpret a RC noun phrase (NP) as referring to the first noun of a complex NP, while L2 speakers of English should apply low attachment, that is, should attach the RC to the second noun of a complex NP. Results from an offline questionnaire did not confirm the main predictions. Data from an online behavioural study supported low attachment preference in L1 Spanish-L2 English speakers, but not high attachment for L1 English-L2 Spanish speakers. The seemingly contradictory offline and online behavioural results were explained by processing delays that may result from a dual syntactic system that prefers late closure, or by language exposure that may favour the syntactic attachment used in the current language context.

Taken together, the L2 ERP results on syntactic anomaly detection and the offline and online behavioural results on L2 syntactic ambiguity point to the necessity to study L2 syntactic comprehension online, to directly compare L1 and L2 syntactic processing similarities and differences, and to directly compare L2 syntactic anomaly and syntactic ambiguity processing. Such an approach can deliver critical results and implications for the theoretical formation of the human parsing system.

Here, we tested L2 syntactic knowledge in early and highly proficient L2 readers of English by directly comparing the processing of temporary syntactic ambiguity and syntactic anomaly. The choice of early high proficient L2 readers of English rather than late L2 readers was to ensure that if the L2 syntactic processing system is calibrated to processing syntactic knowledge in a native-like manner, detection of temporary syntactic ambiguity and of syntactic anomaly should result in similar ERP patters as in native readers of English. Thus, in addition to early and highly proficient L2 readers, a monolingual English group was tested to confirm this hypothesis.

Furthermore, we wanted to find out if (1) phrase structure violation that adheres to the same syntactic principles in English and Spanish results in a comparable posterior P600 effect in non-native and native readers of English. For example, unacceptable English sentences such as "*The broker hoped to sell the stock was sent to jail.*" and the Spanish translation equivalent "*El broker deseaba vender el stock fue enviado a la carcel.*" would result in a syntactic violation. In English the auxiliary "*was*" violates phrase structure rules if the initial reading of the sentence is an active clause reading. In Spanish the sentence requires a relative pronoun "*que*" that introduces a relative clause. However, the violation position in the respective sentences varies between English and Spanish.

With respect to (2) temporary syntactic ambiguity (here induced by verb subcategorization information in English) we predicted that if temporary syntactic ambiguity is distinct in L2, then highly proficient L2 readers of English should show a similar P600 pattern as monolingual readers of English (see Hoover and Dwivedi (1998), Juffs (1998), Tokowicz and MacWhinney (2005) for similar arguments on L2 syntactic specificity). To illustrate, temporary syntactic ambiguity in an English sentence such as "*The broker persuaded to sell the stock* ..." is not possible in a Spanish translation equivalent "*El broker convenció para vender el stock* ...". Here, the transitive verb (*convencer*) requires an accusative object (*a alguien*) and the following complement needs to be preceded by a preposition (*para*) to render the sentence acceptable. Thus, the Spanish translation equivalent "*El broker convencido para vender el stock fue enviado a la carcel.*" of the acceptable English sentence "*The broker persuaded to sell the stock was sent to jail.*" does not resolve temporary syntactic ambiguity as both the short and the long Spanish sentence examples require (i) a prepositional phrase to introduce the complement, and (ii) the inflection of the main verb "*convencer*" is marked differently in the two sentences (*convenció vs. convencido*).

Following the results of Osterhout and Holcomb (1992) and Kaan and Swaab (2003), the current study should also allow to critically evaluate (3) whether a P600 elicited by temporary syntactic ambiguity shows a more frontal or a similar posterior P600 effect as outright syntactic violations (i.e., phrase structure violations). If the P600 elicited by temporary syntactic ambiguity shows a more posterior distribution, but a different latency and amplitude rise than the P600 elicited by syntactic phrase structure violation then both revision and repair of syntactic structure coincide in a posterior P600 effect. In addition, we expected N400 effects on final words based on semantic constraints imposed by the sentence context (Osterhout & Holcomb, 1992). If indeed, L2 readers of English process semantic constraints in a similar fashion as L1 readers of English there should be comparable N400 effects in the two groups.

Lastly, given the findings of Weber-Fox and Neville (1996), Osterhout, McLaughlin, Inoue, and Loveless (2000), and Tokowicz and MacWhinney (2005), we predicted (4) a possible dissociation of ERP results and behavioral acceptability judgment.

2. ERP experiment

As previously reported, the P600 is not just sensitive to outright syntactic anomaly, but also to processing differences between sentences that are syntactically legal. Osterhout and Holcomb (1992) reported P600 effects in temporary syntactically ambiguous sentences as well as for phrase structure violations. Consider the short sentences (1) "*The broker hoped to sell the stock*." and (2) "*The broker persuaded to sell the stock*". The infinitival marker "*to*" in sentences such as (2) should elicit a P600 as the transitive verb "*to persuade*" in an active sentence requires a noun phrase (direct object). Thus, a clausal complement can only be attached to the sentence in form of a passivized reduced relative clause (*The broker* [*who was*] *persuaded to sell the stock*...). As this is not the case in (2) the sentence is rendered unacceptable. However, sentence type (2) can become acceptable providing a reduced relative clause extension such as (4) "*The broker persuaded to sell the stock was sent to jail.*" On the other hand, sentence extensions to short sentences including an intransitive verb such as "*hoped*" in (1) renders a long sentence such as (3) "*The broker hoped to sell the stock was sent to jail.*" unacceptable in comparison to (4) as the auxiliary verb "*was*" violates phrase structure rules in English if an initial simple active reading of the sentence is pursued. Summarizing the predictions presented in Section 1 we hypothesize that if early proficient non-native readers of English have native-like syntactic knowledge, both groups should show P600s elicited by the two types of unacceptable syntactic structures as compared to acceptable syntactic structures.

2.1. Methods

2.1.1. Participants

Two groups of Tufts undergraduate students (12 per group, 6 female per group) participated for class credit (L1 group) or were paid (15 $, L2 group). The age range was 17 to 24 (mean age: 19). All participants were right handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). All participants had normal or corrected-to normal vision. L2 participants were selected on the basis of a language questionnaire (see Table 1), age of L2 acquisition (mean: 5.3 years), the same L1 language background (Puerto Rican Spanish), similar social and educational background, and comparable use of L1 (speaking in social contexts) and L2 (reading and writing, speaking with English friends), as well as two behavioural comprehension tests, a Grammaticality Judgment Task (Linebarger, Schwartz, & Saffran, 1983) and the Test for Reception of Grammar (T.R.O.G; Bishop, 1982).

2.1.2. Material

We created four lists that each contained 120 sentences (see example types (1)–(4) above). A total of 15 intransitive verbs and 15 transitive verbs were used. Sentences such as (1) exemplify that the verb can be used without a noun phrase (intransitive verbs), while verbs in sentences such as (2) require a noun phrase (transitive verbs). Transitive and intransitive verbs did not differ in mean frequency (Kučera & Francis, 1967) in the present tense (intransitive = 74, transitive = 64, $t(14) < 1$) or in the past tense (intransitive = 80, transitive = 37, $t(14) = 2.00$, $p > .05$), or in mean length (intransitive: 7.6; transitive: 7.7). Verbs were repeated approximately the same number of times (6–9 times) in different sentence contexts, and acceptable and unacceptable sentence types across the four lists. Sentences were counterbalanced across the four lists such that each list contained only one version of each sentence and 30 exemplars of each sentence type. The entire set of experimental sentences is provided in Osterhout and Holcomb (1992). The four lists were counterbalanced across participants.

Table 1
An Overview of self-rating scores in the early high proficient L2 group

L2 learner participants	High proficiency/early AOA ($N = 12$)
Age (yr)	19.0
AOA (yr)	5.3
Years abroad	8.3
Grammar Comprehension (E/S)	9.2/8.5
Reading (E/S)	9.2/8.5
Speaking (E/S)	9.3/8.8
Writing (E/S)	9.2/8.5

Abbreviations: yr: years; E/S: respective scores for English and Spanish.

Filler sentences were not included for the following reasons: (i) number and length of the sentences, (ii) upkeep of data quality (reduced physiological and cognitive artefacts etc), (iii) active and reduced relative clause sentences appear equally often, and each verb appears equally often in acceptable and unacceptable sentences, and (iv) replication (see Osterhout & Holcomb, 1992).

2.1.3. Procedure

Participants saw the following trial sequence: A fixation cross (500 ms) in the centre of a computer screen was replaced by single words (300 ms) in a word-by-word manner. The inter-stimulus interval was 350 ms. Sentence-ending words appeared with a period, and were followed by a 1450-ms blank screen. Then a prompt signalled the participants to judge a sentence as acceptable or unacceptable by pressing one of two buttons. The response hand (left or right) was counterbalanced across participants. Participants were seated in a sound-attenuating chamber and instructed to carefully read each sentence before making the required judgment. Participants were provided with a few examples of syntactically acceptable, syntactically unacceptable sentences. If they did not understand the criteria for acceptability, additional examples were provided. The total ERP session lasted approximately 2 h including preparation time.

2.1.4. EEG recording and analysis

EEG activity was recorded from 13 scalp locations, using tin electrodes attached to an elastic cap (Electrocap

International). Electrode placement included International 10–20 system locations (Jasper, 1958) of homologous positions over the left and right occipital (O1/2), frontal (F7/8) regions, and frontal (Fz), central (Cz), and parietal (Pz) midline sites. In addition, several non-standard sites such as anterior temporal (FT7/8), temporo-parietal (TP7/8), and centro-parietal (CP5/6) were applied. Vertical eye movements and blinks were monitored by means of an electrode placed beneath the left eye, and horizontal eye movements were monitored by an electrode positioned to the right of the right eye. The 15 channels were referenced to the left mastoid electrode, and were amplified with a bandpass of 0.01 to 100 Hz (3db cut off) by a Grass Model 12 amplifier system. The right mastoid electrode was recorded actively and provided a channel to re-reference data after data collection.

Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was performed online by a Data Translation 2801-A board and an AT-compatible computer, at a sampling frequency of 200 Hz. Epochs comprised 1180 ms following the critical stimulus presentation aligning the epochs to a 100 ms pre-stimulus baseline. Trials characterized by excessive eye movement (vertical or horizontal) or amplifier blocking were rejected. For analyses involving words embedded within sentences, less than 10% of the trials were removed per condition. ERPs were quantified by computer as the mean voltage within a latency range following presentation of words of interest, relative to baseline of activity.

2.2. Results

2.2.1. Behavioural data analyses

2.2.1.1. Offline grammaticality judgment. Separate ANOVAs with repeated measures on the Grammaticality Judgment Task (Linebarger et al., 1983) and the Test for Reception of Grammar (T.R.O.G; Bishop, 1982) revealed no significant differences of reader type (all $p > 1$). Native readers of English scored 95% correct on a variety of grammatical structures, while non-native readers of English scored 94% correct. The T.R.O.G led to similar results; native English readers: 99% correct, non-native English readers: 98% correct.

2.2.1.2. Delayed online acceptability judgment. A $2 \times 2 \times 2$ analysis of variance with the between-subjects factor Group (native/non-native), and two within-subjects factors, Type (acceptable/unacceptable), and Length (short/long) was applied. Overall, native readers of English judged the acceptability of sentences more securely than non-native readers of English ($F(1, 22) = 14.01$, $p < .001$), and short sentences were judged more securely than long sentences ($F(1, 22) = 6.90$, $p < .01$). A three-way interaction of Group, Type, and Length ($F(1, 22) = 15.38$, $p < .001$) showed that native readers of English judged 94% of the short sentences (1), and 80% of the long sentences (4) as acceptable, while they judged 86% of the short sentences (2) and 91% of the long sentences (3) as unacceptable. Non-native readers of English also showed differences in the acceptability judgment of short and long sentence types. Short sentences such as (1) were judged highly correct (96%), while long acceptable sentences such as (4) were judged at chance level (56%). Short sentences of type (2) and long sentences of type (3) were judged as unacceptable at (66%) and 91%, respectively (see also Fig. 1).

2.2.2. ERP data analyses

ANOVAs of repeated measures were applied separately to midline and lateral sites. Midline analysis included the between-subjects factor Group (native/non-native), and the within-subjects factors Type (acceptable/unacceptable), and electrode-site (FZ, CZ, and PZ). Lateral analyses included the between-subjects factor Group (native/non-native), and the within-subjects factors Type (acceptable/unacceptable), electrode-site (frontal, anterior temporal, temporo-parietal, centro-parietal, and occipital), and hemi-

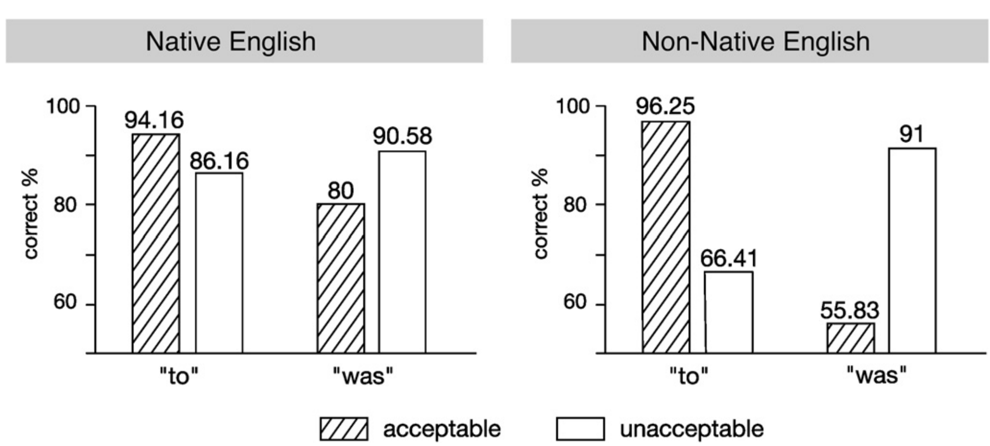

Fig. 1. Mean percent correct (correct%) in the delayed online acceptability judgment comparing native (left) and non-native (right) readers of English in the critical syntactic conditions ("*to*" = temporary syntactic ambiguity; "*was*" = phrase structure constraints). Scores for acceptable judgment are presented in marked bars, those for unacceptable judgment are presented in white bars.

sphere (left/right). For data reduction sentence types (1) and (3) (acceptable) as well as sentence types (2) and (4) (unacceptable) were combined for the analysis of temporary syntactic ambiguity (at the position of the infinitival marker "*to*") as the sentences were comparable up to sentence-ending noun "*stock*" in the respective sentence types. Analysis of phrase structure violation (at the position of the auxiliary verb "*was*") was calculated for sentence types (3/unacceptable) and (4/acceptable) only.

To protect against excessive Type I error due to violations of the assumption of equal variances of differences between conditions of within-subject factors (Huynh & Feldt, 1976; Keppel, 1982), the Geisser-Greenhouse correction (Geisser & Greenhouse, 1959) was applied when evaluating effects with more than one degree of freedom. We followed significant interactions with planned pair-wise comparisons. Based on visual analysis of the averaged waveforms in both groups and a conventional latency range of the P600 (e.g., Osterhout & Holcomb, 1992: 500–800 ms) two time-windows were calculated: 500–650 ms (*early*) and 650–800 ms (*late*). These two time-windows were defined on the basis of onset and offset differences of the critical words "*to*" in sentence types (1), (2), (3), and (4) and "*was*" in sentence types (3) and (4) in native and non-native readers of English. In addition, Friederici and colleagues (Friederici et al., 1996) and Kaan and Swaab (2003) reported latency differences between a P600 elicited to outright violations and a P600 elicited to

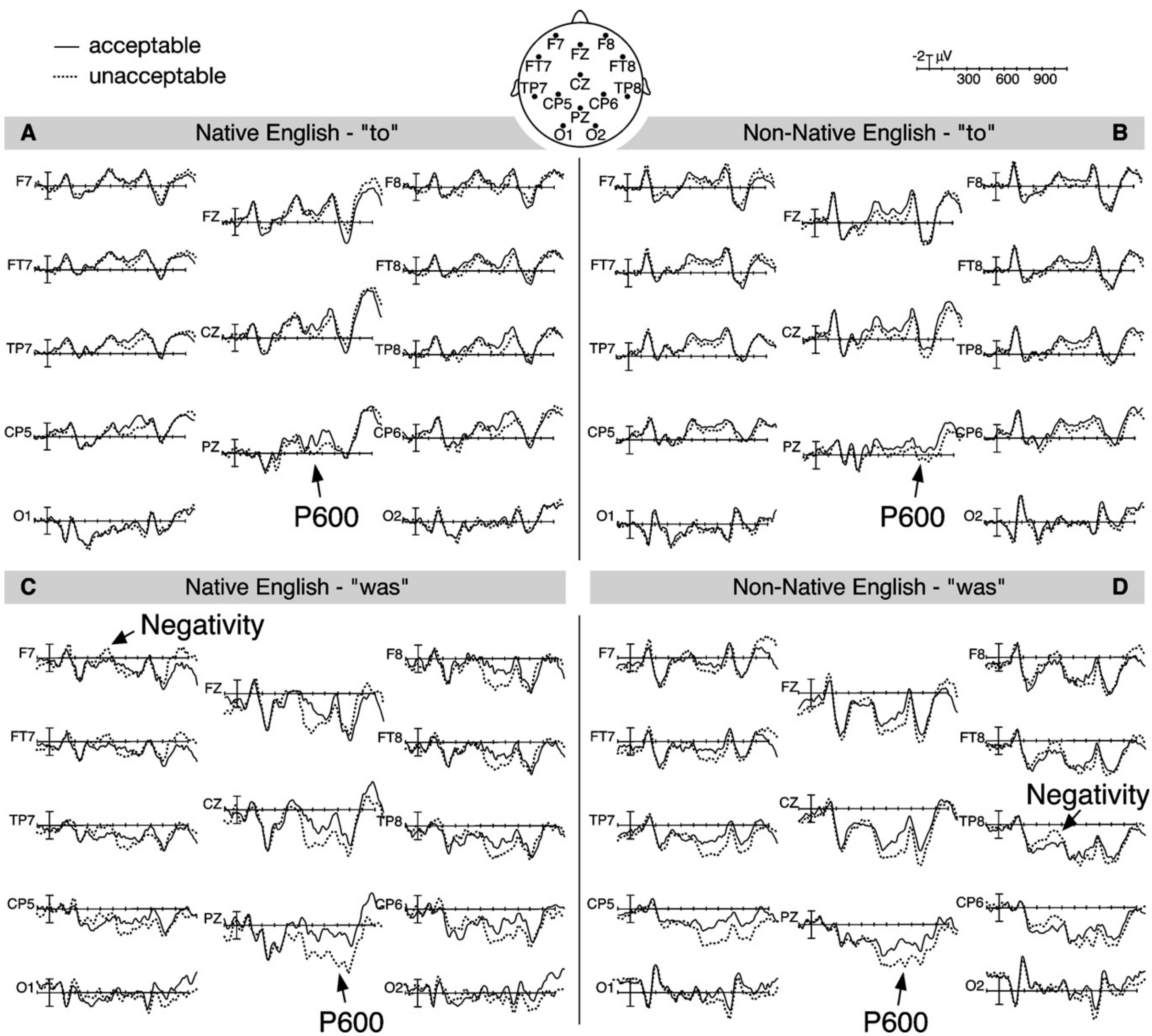

Fig. 2. (A–D): ERPs elicited by the critical target words in the temporary syntactic ambiguity condition ("*to*") for native readers (left-A) and non-native readers (right-B), and in the phrase structure violation condition ("*was*") for native readers (left-C) and non-native readers (right-D). In black lines we show the brain responses to acceptable, in dotted lines the brain responses to unacceptable critical target words are shown. Waveforms show averages for critical target words from 100 ms prior to stimulus onset up to 1100 ms post-stimulus onset. Positivity is plotted down, negativity is plotted up.

Please cite this article in press as: Kotz, S. A. et al., ERPs reveal comparable syntactic sentence processing in native ..., *Acta Psychologica* (2007), doi:10.1016/j.actpsy.2007.10.003

ambiguous/non-preferred sentence structures. Thus, to capture latency differences between the potentially different P600s and differences between native and non-native readers of English, a separation into two time-windows was thought to be warranted. In addition, a negativity preceding the P600 in phrase structure violations was quantified based on visual inspection and previously reported negativities (Osterhout & Holcomb, 1992; Weber-Fox & Neville, 1996). Lastly, to quantify a previously reported end of sentence negativity (see Osterhout & Holcomb, 1992), acceptable and unacceptable nouns in the four sentences types ("*stock*" in short active sentences and "*jail*" in long passive sentences) were evaluated in a classical N400 time-window between 300–500 ms.

2.2.2.1. The P600 and temporary syntactic ambiguity. Brain responses to the infinitival marker "*to*" in acceptable and unacceptable sentences are displayed in Fig. 2A and B for native and non-native speakers of English, respectively. In this and subsequent analyses, the general shape of the obtained waveforms was consistent with that previously reported by others (e.g., Osterhout & Holcomb, 1992).

Analysis of midline sites did not reveal differences in Group in either time-window (p >.6, respectively), but a main effect of Type ($F(1, 22) = 5.54$, $p < .03$) in the *late* time-window showing a more positive-going brain wave elicited to "*to*" in unacceptable (2 and 4) than in acceptable (1 and 3) sentences.

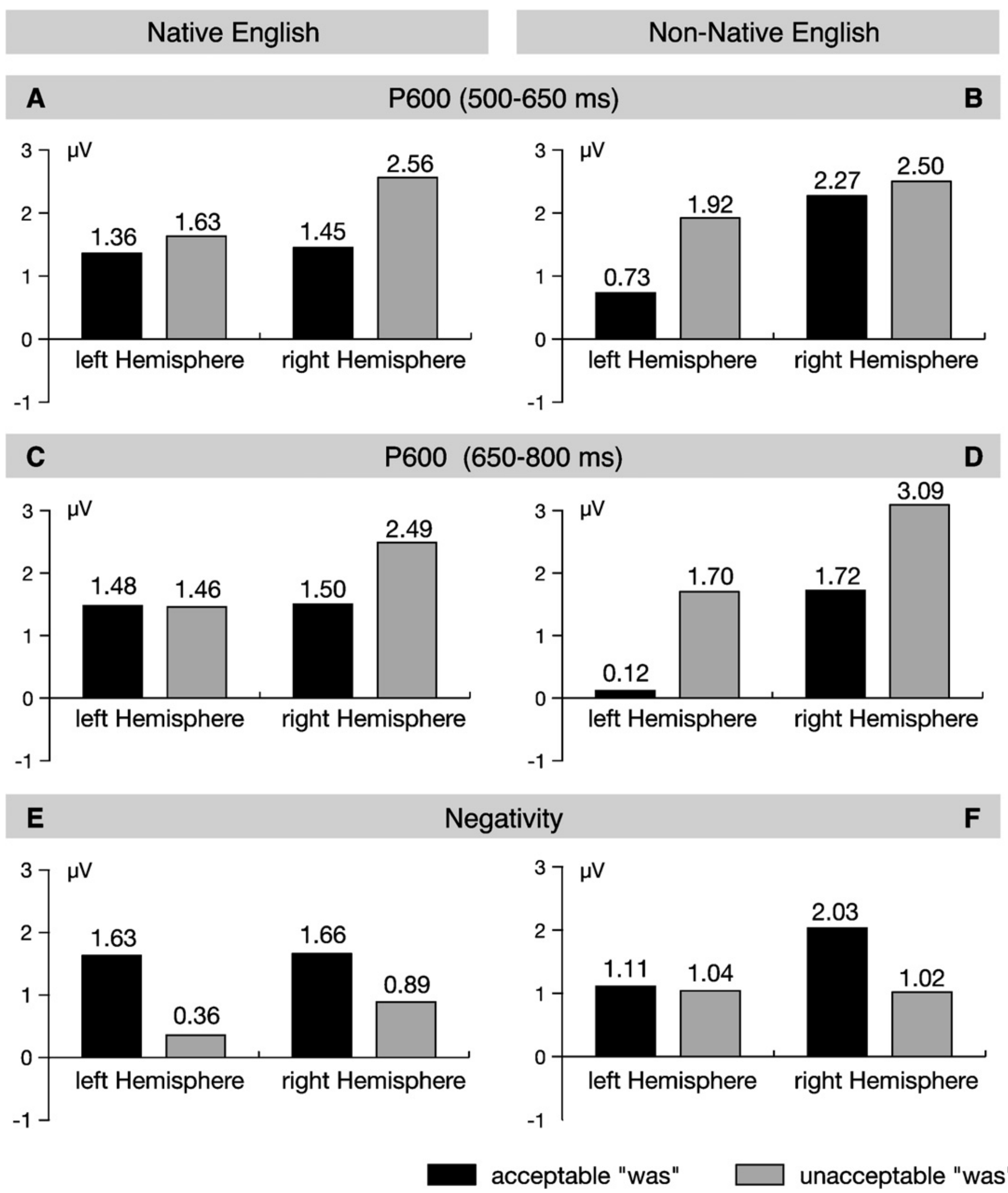

Fig. 3. Mean amplitude values for the P600 and preceding negativity elicited by phrase structure violations in the two measured time-windows. On the left-hand side, mean amplitude values for native readers (A, C, E), and on the right side mean amplitude values for non-native readers (B, D, F) are shown. In black bars the mean amplitude values for acceptable sentences, in gray bars the ones for unacceptable sentences are given.

Analysis of lateral sites did not reveal main effects of Group in the *early* and *late* time-window (all $p > .1$), or Type ($p > .1$) in the *early* time-window. However, a main effect of Type in the *late* time-window ($F(1, 22) = 7.29$, $p < .01$) was significant. Again, unacceptable sentences (2 and 4) showed a more positive-going waveform than that for acceptable sentences (1 and 3) elicited by the critical infinitival marker "*to*" (Fig. 3).

Analysis of the sentence-ending noun "*stock*" in the short sentence analysis resulted in a marginally significant main effect of Type at midline sites ($F(1, 22) = 3.39$, $p < .08$), and a main effect of Type at lateral sites ($F(1, 22) = 4.87$, $p < .04$). Neither main effect interacted with Group (all $p > 1$). Thus, both native and non-native readers of English show a sentence-ending N400 negativity in short sentences (see Fig. 4A and B).

2.2.2.2. The P600 and phrase structure constraints. Brain responses to the auxiliary verb "*was*" in unacceptable sentence types (3) and acceptable sentences (4) are displayed in Fig. 2C and D for native and non-native readers of English, respectively (see also Fig. 1).

Analysis of midline electrodes in the *early* time-window revealed no Group differences ($p > .7$), but of Type ($F(1, 22) = 11.58$, $p < .003$). The auxiliary verb "*was*" in the unacceptable sentences elicited a larger P600 than "*was*" in the acceptable sentences. This effect extended into the *late* time-window ($F(1, 22) = 12.39$, $p < .002$), and was

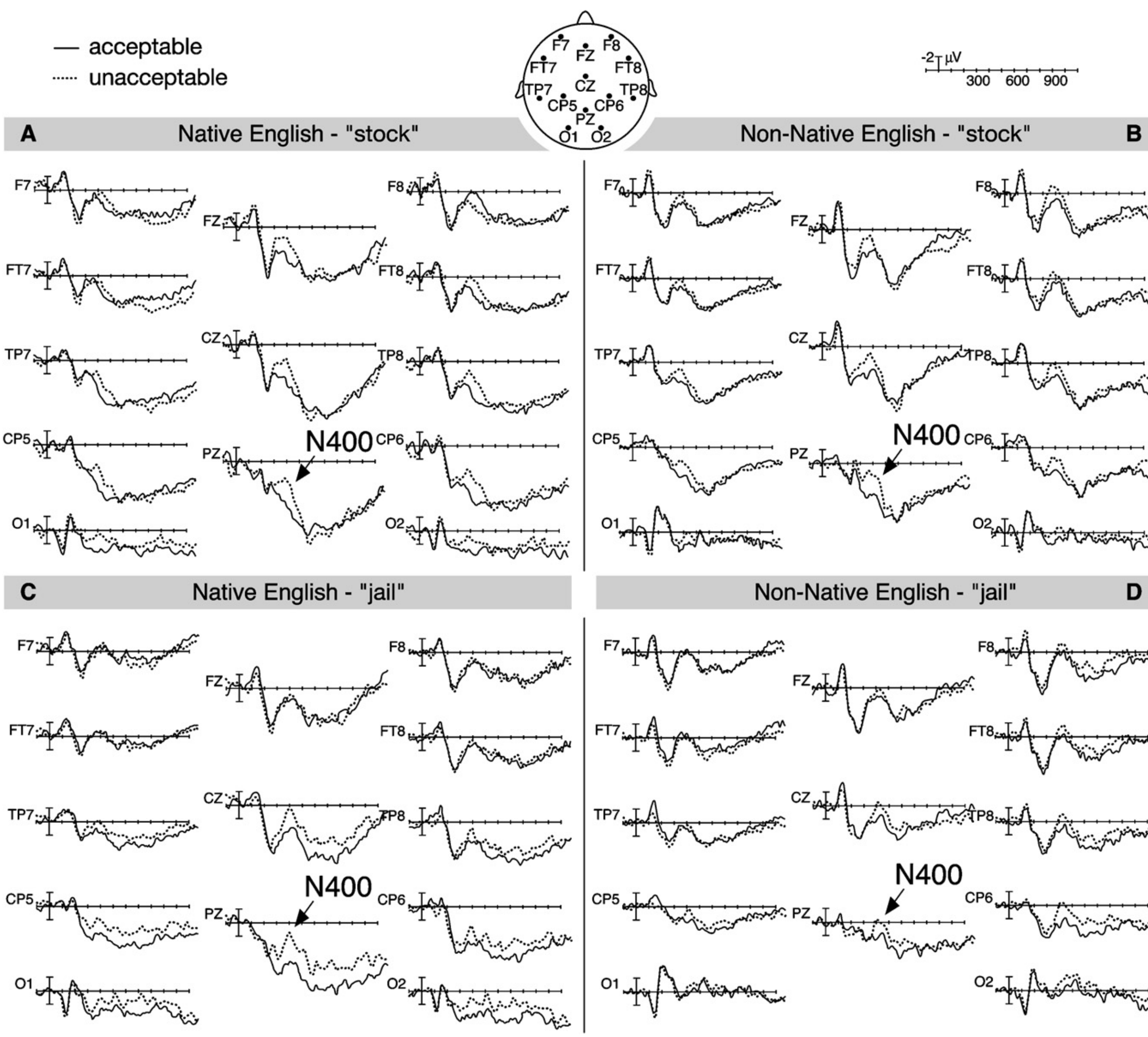

Fig. 4. (A–D): ERPs elicited by the critical sentence-ending words in the temporary syntactic ambiguity condition ("*stock*") for native readers (left-A) and non-native readers (right-B), and in the phrase structure violation condition ("*jail*") for native readers (left-C) and non-native readers (right-D) are given. In black lines we show the brain responses to acceptable, in dotted lines the brain responses to unacceptable critical target words. Waveforms show averages for critical target words from 100 ms prior to stimulus onset up to 1100 ms post-stimulus onset.

Please cite this article in press as: Kotz, S. A. et al., ERPs reveal comparable syntactic sentence processing in native ..., *Acta Psychologica* (2007), doi:10.1016/j.actpsy.2007.10.003

stronger at posterior electrode sites ($F(2, 44) = 5.43$, $p < .02$).

Analysis of lateral sites in the *early* time-window yielded a main effect of Type ($F(1, 22) = 6.77$, $p < .02$) and a four-way interaction of Group, Type, Hemisphere and Electrode Site ($F(4, 88) = 3.16$, $p < .05$). Follow-up analyses by Group did not confirm a three-way interaction between Type, Hemisphere and Electrode Site in either group. However, a marginally significant effect for Type was found in the non-native speakers at left posterior electrode sites ($F(4, 44) = 4.11$, $p < .08$, $w^2 = 0.205$). In the *late* time-window there was a main effect of Type ($F(1, 22) = 9.12$, $p < .006$) and a three-way interaction of Group, Type, and Hemisphere ($F(1, 22) = 4.87$, $p < .04$). Break-down analysis by group revealed a larger right-lateralized P600 effect in native readers of English ($F(1, 11) = 7.27$, $p < .02$, $w^2 = 0.343$), while non-native-readers of English showed a left-lateralized P600 effect ($F(1, 11) = 4.11$, $p < .03$, $w^2 = 0.205$).

Analysis of a negativity preceding the P600 effect did not result in any effect at midline sites (all $p > .1$), but a main effect of Type ($F(1, 22) = 5.57$, $p < .03$) and a three-way interaction of Group, Type, and Hemisphere ($F(1, 22) = 4.05$, $p < .05$). While the negativity preceding the P600 was marginally significant at left hemisphere electrode-sites in native readers ($F(4, 44) = 2.86$, $p < .08$, $w^2 = 0.134$), a significant effect was found for non-native readers at right centro-parietal/occipital electrode-sites ($F(4, 44) = 4.21$, $p < .02$, $w^2 = 0.211$).

Analysis of the sentence-ending negativity comparing the acceptable noun "*jail*" to the unacceptable noun "*jail*" revealed a marginal effect of Group at midline sites ($F(1, 11) = 3.24$, $p < .08$), a main effect of Type ($F(1, 22) = 4.14$, $p < .05$), and an interaction of Type and electrode-site ($F(2, 44) = 10.42$, $p < .001$) confirming that the negativity was largest at Pz. At lateral sites there was a marginal effect of Group ($F(1, 11) = 4.26$, $p < .06$), a marginal effect of Type and electrode-site ($F(4, 88) = 3.23$, $p < .06$), and a significant interaction of Type and hemisphere ($F(2, 88) = 5.01$, $p < .01$) confirming a sentence-ending negativity that was larger over right than left hemisphere electrode-sites and an overall slightly smaller amplitude rise in non-native than native readers of English (see Fig. 4C and D).

3. Discussion

The current experiment set out to investigate the processing of temporary syntactic ambiguity and syntactic phrase structure anomaly in early highly proficient non-native and native readers of English. With this approach we tested the hypothesis whether early highly proficient L2 readers of English parse syntactic knowledge in a native-like manner dependent on factors such as structural similarity between L1 and L2, and structural specificity in L2. Furthermore, the experimental set-up allowed the testing of P600 synergies between syntactic ambiguity and syntactic anomaly in both native and non-native readers of English. Lastly, next to testing L2 compared to L1 syntactic competence we were interested to see whether online ERP measures go hand in hand with behavioural measures.

Offline testing of grammaticality judgment and grammar reception confirm native-like syntactic competence in L2 readers when response time is unlimited. However, behavioural data measured in a delayed acceptability judgment task (as part of the ERP measurement) reveal a different pattern. Overall, non-native readers have lower acceptability rates than native readers. In particular, the judgment of long acceptable sentences is near chance-level (56%), followed by relatively low acceptability rates for short unacceptable sentences (66%). On the other hand, ERP data show that both native and non-native readers of English are sensitive to the detection of syntactic anomaly and temporary syntactic ambiguity even though the distribution and latency of the P600 elicited by syntactic anomaly as well as the negativity preceding the P600 vary as a function of reader type. Furthermore, the P600 elicited by temporary syntactic anomaly is broadly distributed, is delayed in its onset/peak latency, and has a lower amplitude rise than the P600 elicited by phrase structure violations in both groups. Sentence-ending negativities are found in both groups of readers. However, the amplitude rise of the sentence-ending negativity is smaller in non-native than native readers of English when parsing phrase structure information.

In the following we will first address the apparent discrepancy between offline and online behavioural data, and the online behavioural and ERP data. Next, each P600 effect will be discussed with respect to reader type, L1/L2 structural similarity or L2 specificity, and the characteristics of the P600s elicited by the two different syntactic manipulations. In the same context and manner the sentence-ending negativities will be discussed.

3.1. Offline versus online behavioural results

Two differences are apparent when comparing offline and online behavioural data, and ERP measures and online behavioural data. To test non-native readers' knowledge of L2 syntax we used two tests of syntactic competence. In these two tests (Grammaticality Judgment Task by Linebarger et al. (1983), and the Test for Reception of Grammar (T.R.O.G; Bishop, 1982), both L2 and L1 readers of English were tested with no time constraints. The judgment task consisted of a variety of syntactic complexities and anomalies, including verb subcategorization and phrase structure, and the scores are based on a 50% chance level criterion. The T.R.O.G assesses language comprehension via naming and pointing to pictures, and is designed to test the meaning and syntax of words as well as of sentences. While the behavioural results of the T.R.O.G may not be all that surprising as naming and pointing may have been guided by the meaning of words and sentences rather than their syntax, the grammaticality judgment task clearly tested syntactic competence. Even so, both groups of participants judged grammaticality highly correctly.

One explanation for the discrepant offline and online behavioural results could be that participants in the online experiment experienced time pressure which in turn may have affected the acceptability judgment. On the contrary, they did not judge grammaticality under time pressure in the offline tasks. Second, while the proportion of acceptable and unacceptable sentences was equivalent in both the online and offline tasks, the number of syntactic structures tested varied between the tasks. For example, next to subcategorization and phrase structure information, word order, word category, verb agreement, and pronoun reference were tested in the offline grammaticality judgment task. Thus, in comparison to the online acceptability judgment task, the proportion of equivalent syntactic structures was lower in the offline grammaticality judgment task. Lastly, while it is clear that non-native compared to native readers of English accept an unacceptable short sentence (2) as acceptable more often, they also judge acceptable long sentences (4) as unacceptable thus rendering this judgment at chance level. One plausible explanation is that non-native readers do not have trouble detection syntactic anomaly per se, but may encounter increased difficulties when they have to parse a more complex sentence construction such as a sentence in the passive voice. In the current experiment, the long sentences are passive sentence constructions while short sentences are active sentence constructions. As referred to by Rossi and colleagues (2006), Guasti (2004) suggests that the active voice is acquired before the passive voice in L1 syntax acquisition. A similar acquisition dynamic may apply to L2 syntax acquisition. Thus, the use of passive voice rules may be more vulnerable as passive constructions require movement of elements. This is an interesting point in support of the current results and should be followed up in future experiments.

However, the ERP results seem to go hand in hand with the offline behavioural data displaying sensitivity towards L2 syntax. Both non-native and native readers of English elicited a P600 to temporary syntactic ambiguity and phrase structure violations. Dissociations of ERP and behavioural measures have also been discussed in previous L2 syntax research. Weber-Fox and Neville (1996) reported that early and proficient (4–6 years at age of exposure) non-native readers of English (L1 Chinese) show reduced acceptability for unacceptable sentences (phrase structure violations) in comparison to native readers of English, while showing a comparable biphasic pattern of a negativity/P600 in the ERP. Interestingly, in this study non-native participants made more errors judging an unacceptable sentence than an acceptable sentence. However, phrase structure violations were embedded in active rather than passive sentence constructions. Furthermore, data from late L2 learners have also shown that ERP measures dissociate from grammaticality judgment (Osterhout et al., 2000; Tokowicz & MacWhinney, 2005). While participants showed the expected P600 elicited by syntactic violations, behavioural judgment was at chance level (Osterhout et al., 2000) or above chance for sentences constructions that were similar between L1 and L2 or specific to L2 (Tokowicz & MacWhinney, 2005). Both groups of authors have argued that explicit grammaticality judgment may intermix with additional processes such as decision making, response preparation, and motor response. This in turn may render grammaticality or acceptability judgment a more complex and multilayered response than the online detection of syntactic violation and/or syntactic ambiguity as reflected in the ERP. This point is quite relevant in tasks that include a delayed response as the response is moved outside the realm of language-specific processing. Lastly, some authors have argued that the grammaticality judgment task may not be a good measure of L2 linguistic knowledge (McDonald, 2000) or may over-/underestimate L2 syntactic knowledge (Juffs, 2004). In particular, McDonald (2000) pointed out that variables such as syntactic rule types or interpretation of ungrammatical constructions may alter the results of a grammaticality judgment task and renders the task unreliable.

With respect to the current evidence we conclude that there are distinct differences in offline and online behaviour in L2 readers of English as measured by the grammaticality/acceptability judgment tasks. These differences may result from time-constraints as well as from the proportion and type of syntactic rule tested in the respective tasks. Furthermore, dissociations between ERP and behavioural measures go hand in hand with previous evidence (Tokowicz & MacWhinney, 2005; Weber-Fox & Neville, 1996). This applies specifically to the detection of temporary syntactic ambiguity. Furthermore, the results from phrase structure violations reveal that L2 syntactic vulnerability is not restricted to syntactic anomaly, but may also apply to syntactically more complex constructions such as acceptable passive constructions. However, the present data clearly provide further evidence that ERPs and behavioural measures may tap into different processing stages during syntactic processing as previously reported for syntactic processing (Osterhout et al., 2000; Tokowicz & MacWhinney, 2005) and lexical-semantic processing (Kotz, 2001; Kotz & Elston-Guettler, 2004; Kotz & Elston-Guettler, 2007).

3.2. The P600 and temporary syntactic ambiguity

Previous monolingual research on syntactic anomaly detection suggests that the effect size and distribution of the P600 varies as a function of severity of a syntactic anomaly (Coulson et al., 1998; Gunter et al., 1997; Osterhout & Holcomb, 1992). If the syntactic anomaly renders a sentence fully ungrammatical, the effect size of the component should be large. If however, a violation renders a sentence only temporarily syntactically unacceptable the effect size should be small. This hypothesis has been confirmed in studies that investigated syntactic ambiguity and syntactic anomaly detection in parallel (e.g., Kaan & Swaab, 2003). In addition to effect size, it has also been reported that the two syntactic processing types may show topographic differences (e.g., Friederici et al. 2002; Hagoort et al., 1999; Van Berkum et al., 1999).

The P600 elicited by temporary syntactic ambiguity in native and non-native readers is small, is marked by a delayed onset (650–800 ms), and is broadly distributed. These data fit previous results in monolingual English participants (Osterhout & Holcomb, 1992) even though the previously reported P600 latency was longer (500–800 ms) and the component had a fronto-central distribution (Exp.1, but not Exp. 2). However, the fact that both reader groups show the same P600 effect suggests that there are no detectable online processing differences between the groups. These data serve as first online ERP evidence on L2 temporary syntactic ambiguity processing. Furthermore, the data go hand in hand with evidence from eye-movement recordings in highly proficient L2 readers that show native-like syntactic ambiguity resolution (e.g., Frenck-Mestre & Pynte, 1997; Juffs & Harrington, 1996) and are partly in support of a behavioural study by Dussias (2003). In addition, as discussed in Section 1 and by Tokowicz and MacWhinney (2005), L2 sentence structure that does not have an equivalent in L1 may be processed more easily than a structure that may be formed differently in L1 and L2. In the current experiment, temporary syntactic ambiguity in English has no equivalent in Spanish, thus the early highly proficient non-native readers of English may recognize the rule of L2 language-specificity.

In addition, the P600 was followed by a sentence-ending negativity in both reader groups. Osterhout and Holcomb (1992) initially (Exp. 1) offered two interpretations for this negativity. Either it reflects a semantic misfit as has been discussed in the N400 literature (Kutas & Federmeier, 2000) or it is elicited as a result of the unacceptability of a sentence.

In the final discussion of the paper, the authors concluded that the N400-like effect elicited in both short and long sentences can be associated with "the difficulty associated with integrating linguistic material at the semantic-message level, regardless of the cause of interpretative problem" (p. 800) Following this argument, the N400 elicited in both unacceptable sentence types may reflect the participants' difficulty to integrate sentence final words into sentence context when either syntactic ambiguity or syntactic anomaly is embedded in the preceding sentence context. However, both native and non-native readers show the same size N400 effect which supports the conclusion that sentence final integration is comparable between the two groups. This goes hand in hand with data from Kotz (2001) who reported comparable N400 word-level priming effects in early highly proficient L2 speakers of English and monolingual speakers of English. Furthermore, this result is different from the N400 data reported for late learners that often show a delayed N400 and/or N400 amplitude reduction (see for example, Ardal, Donald, Meuter, Muldrew, & Luce, 1990; Weber-Fox & Neville, 1996).

One question that remains open is why the P600 amplitude, peak latency, and distribution diverge from the initial P600 results of Osterhout and Holcomb (1992). Interestingly, while the P600 elicited by temporary syntactic ambiguity had a fronto-central right-shifted distribution in Experiment 1, the effect was more evenly distributed in Experiment 2. As different participants were tested in the two experiments, inter-subject variability may be the cause of such differences. In addition, Kaan and Swaab (2003) clearly showed that the processing of non-preferred syntactic structure is not restricted to a frontal P600. Rather both syntactic anomaly detection and the integration of non-preferred structures (i.e. syntactic ambiguity) co-occur in a posterior P600. However, a posterior P600 elicited by syntactic anomaly was larger and lasted longer than the one elicited by non-preferred structures. These results suggest that posterior positivities may reflect both revision and repair and can be distinguished by amplitude rise and latency. This is confirmed in the present data as well. We extend this interpretation and propose that inter-subject variability may influence the distribution of the P600 elicited by temporary syntactic ambiguity.

Coming back to the initial hypothesis that the size and latency of the P600 may vary as a function of severity of syntactic anomaly, the current results confirm such a hypothesis. As the manipulation of temporary syntactic ambiguity (i.e., verb subcategorization information) did not render a sentence fully syntactically unacceptable, the readers may have recovered more easily. This in turn may have influenced both effect size and duration of the P600. However, we can not exclude the possibility that non-linguistic factors such as working memory capacity, perceivability, and heuristic strategies can alter the underlying processes involved in syntactic ambiguity processing, and may critically influence how a reader resolves ambiguity. Monolingual evidence on syntactic ambiguity resolution points to such a possibility. In these studies effect size differences or even different components as well as distribution differences have been reported for high and low working memory capacity (e.g., Frisch et al., 2002; Mecklinger et al., 1995). It is therefore important to investigate possible linguistic and non-linguistic factors separately in future studies of L2 syntactic ambiguity processing.

3.3. The P600 and phrase structure constraints

Native and non-native readers detected syntactic anomaly that reflects a violation of phrase structure rules in English. In both groups the amplitude rise was larger and the scalp distribution more posterior compared to the P600 elicited by temporary syntactic ambiguity. Also, the latency of the P600 extended across both time-windows thus covering the previously reported duration of a P600 elicited by syntactic anomaly (500–800 ms; Osterhout & Holcomb, 1992). However, the effect was most distinct at left hemisphere leads for non-native readers, while native readers displayed a previously reported right-posterior P600 effect. Thus, while the P600 effect was comparable in both groups, the scalp distribution of the P600 was quite different.

These distributional differences lead to an important question. Do distinct syntactic processes in non-native language comprehension elicit ERP components that tap into different neural sources than the ERP components elicited during native language comprehension? In this case, why would this play a more immanent role in phrase structure rules than in temporary syntactic ambiguity? Both syntactic manipulations render a sentence unacceptable by violating a closed class item such as a preposition (“*to*”) or an auxiliary verb (“*was*”). Furthermore, according to previous L2 evidence (Tokowicz & MacWhinney, 2005) one would expect that if there is L1 and L2 structural similarity, as is the case in the current experiment, L2 syntactic processing should be facilitated in highly proficient non-native readers. To our knowledge only one previously published ERP study provides evidence on early L2 syntactic phrase structure effects (Weber-Fox & Neville, 1996). The authors reported that early bilinguals responded to phrase structure violations (P600) if syntactic rules were acquired before the age of 10. This effect was preceded by an early negativity (N125) and a late negativity (300–500 ms). The negativities varied as a function of age of exposure. An early left anterior negativity (N125) was only found in bilinguals who acquired their second language between the ages of 1 to 3 years, and in late learners (after the age of 10) with a bilateral distribution. A temporo-parietal and left-lateralized negativity (300–500 ms) was reported for early bilinguals (1–3, 4–6, 7–10 years at age of exposure). The authors concluded that changes in ERP asymmetry (as reported for late learners) may be associated with changes in language lateralization and increased involvement of right hemisphere structures.

While such changes in laterality have been reported, most neuroimaging evidence shows that non-identical neural sources in prefrontal cortex vary with language exposure or usage in early bilinguals (e.g., Perani et al., 2003) or vary in the extent of activation in Broca’s area in late proficient bilinguals (Rüschemeyer, Fiebach, Kempe, & Friederici, 2005). In a recent paper Jeong and colleagues (2007) investigated the effect of syntactic similarity during L2 auditory language processing in Korean late trilinguals with equally high proficiency in their L2s (English and Japanese). Comparing activation in syntactically similar languages such as Korean and Japanese, the authors report no activation differences between the two languages in perisylvian language areas while they find activation differences between syntactically dissimilar languages (Korean/ Japanese and English). Latter results nicely demonstrate that next to age of exposure topological differences in L1 and L2 may modify brain activation patterns. However, the current ERP and fMRI evidence is too limited to explain the distributional differences found for phrase structure violations in native and non-native readers of English. Most likely, the distributional differences in the non-native readers result from the negativity preceding the P600 over posterior right hemisphere electrode-sites, while the positivity over left-hemisphere sites was unaffected. In comparison, the right posterior positivity is not modulated by a preceding left-lateralized anterior negativity in the native readers (see also Osterhout & Holcomb, 1992). The true nature of such distributional shifts needs to be further explored by replicating the current results with a larger set of electrodes and a larger number of participants. This should substantiate whether different neural mechanisms underlie syntactic processing of phrase structure rules or not. Lastly, on a note of caution, distributional differences of scalp ERPs do not allow to deduct neural sources as scalp ERPs reflect the summation of potentially numerous neural generators (see for example Van Petten & Luka, 2006). Clearly, additional work is needed to differentiate the effects of early L2 exposure and proficiency, but also structural differences and similarities between languages to fully understand the dynamics of L2 syntactic attainment.

4. Conclusion

To fully understand the human syntactic parsing system it is necessary to look into the multiple structural subtleties that constitute such a system. Here we investigated two syntactic phenomena, namely temporary syntactic ambiguity (i.e., verb subcategorization information) and phrase structure constraints in English. While latter structure was tested with classical syntactic anomaly detection, verb subcategorization information led to temporary syntactic ambiguity. The fact that both deviant syntactic information types elicit a P600 speaks to successful syntactic parsing in native and non-native readers. However, differences in amplitude size, peak latency and distribution indicate that the processing of temporary syntactic ambiguity and syntactic anomaly may be influenced by subject variability as well as potentially different or additional cognitive mechanisms. Furthermore, we tested the boundaries of the human syntactic parsing system by looking into early L2 syntactic attainment. L2 syntactic acquisition in native readers of Spanish with high reading and speaking proficiency in their non-native language English (as evidenced by self-rating scales) showed similar online syntactic parsing to native readers of English, while online behavioural data were different. These results add to previous evidence in early and late L2 learners that ERPs are a sensitive and useful tool to tap into the processing dynamics of L2 syntactic knowledge attainment

Acknowledgements

We wish to thank Ruth De Diego Balaguer for critical input on the Spanish translations and Silke Paulman for edits on a pervious version of the paper.

References

Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language, 39*, 187–205.

Bishop, D. (1982). Test for the reception of grammar (T.R.O.G.). Medical Research Council: Manchester.

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

Dussias, P. (2003). Syntactic ambiguity resolution in second language learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Learning, 25*, 529–557.

Fernàndez, E. (1999). Processing strategies in second language acquisition: Some preliminary results. In E. C. Klein & G. Martohardjono (Eds.), *The development of second language grammars: A generative approach* (pp. 217–240). Amsterdam: John Benjamins.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21*, 175–198.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50A*, 119–148.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*, 78–84.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language, 43*, 476–507.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*, 183–192.

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: L earning, Memory, and Cognition, 22*, 1219–1248.

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31*, 45–63.

Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb-argument structure information in the dynamics of parsing. *Cognition, 91*, 191–219.

Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *NeuroReport, 12*, 3391–3394.

Frisch, S., & Schlesewsky, M. (2005). The resolution of case conflicts: A neurophysiological perspective. *Cognitive Brain Research, 25*, 484–498.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition, 85*, B83–B92.

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

Guasti, M. T. (2004). *Language acquisition: The growth of grammar*. Cambridge, MA: Bradford.

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

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. In S. M. Garnsey (Ed.). *Language and Cognitive Processes. Special Issue: Event-related Brain Potentials in the Study of Language* (vol. 8, pp. 439–483). Hove: Lawrence Erlbaum Associates.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neural architecture of syntactic processing. In C. M. Brown & P. Hagoort (Eds.), *Neurocognition of language*. Oxford: Oxford University Press.

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

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194–205.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 123–141.

Hoover, M. L., & Dwivedi, V. D. (1998). Syntactic processing by skilled bilinguals. *Language Learning, 48*, 1–29.

Huynh, H., & 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.

Jasper, H. H. (1958). Report to the committee on methods of clinical examination in electroencephalography. Appendix: The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371–375.

Jeong, H., Sugiura, M., Sassa, Y., Haji, T., Usui, N., Taira, M., et al. (2007). Effect of syntactic similarity on cortical activation during second language processing: A comparison of English and Japanese among native Korean trilinguals. *Human Brain Mapping, 28*, 194–204.

Juffs, A. (1998). Main verb vs. reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning, 48*, 107–147.

Juffs, A. (2004). Representation, processing, and working memory in second language. *Transactions of the Philological Society, 102*, 199–225.

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning, 46*, 283–326.

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15*, 98–110.

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

Keppel, G. (1982). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Kotz, S. A. (2001). Neurolinguistic evidence for bilingual memory representation: A comparison of reaction times and event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 143–154.

Kotz, S. A., & Elston-Guettler, K. E. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics, 17*, 215–235.

Kotz, S. A., & Elston-Guettler, K. E. (2007). Bilingual semantic memory revisited: ERP and fMRI evidence. In J. Hart & M. Kraut (Eds.), *The Neural Basis of Semantic Memory* (pp. 105–132). Cambridge: Cambridge University Press.

Kučera, H., & Francis, W. N. (1967). *A computational analyis of present day American English*. Providence,R.I.: Brown University.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 1*, 463–470.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13*, 361–392.

McDonald, J. (2000). Grammaticality judgements in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*, 395–423.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience, 7*, 703–704.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition, 23*, 477–494.

Münte, T. F., Heinze, H.-J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. *Neuropsychologia, 36*, 217–226.

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. *Journal of Cognitive Neuroscience, 17*, 1229–1244.

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*, 151–165.

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

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

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes, 8*, 413–437.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*, 739–773.

Osterhout, L., Holcomb, P. J., & Swinney, D. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of experimental Psychology: Learning, Memory and Cognition, 20*, 786–803.

Osterhout, L., McLaughlin, J., Inoue, K., & Loveless, J. (2000). Brain responses in L2 learning. Paper presented at the meeting Bilingualism: From basic research to educational practice, Rovereto, Italy.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., et al. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping, 19*, 170–182.

Rüschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human Brain Mapping, 25*, 266–286.

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

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar. *Studies in Second Language Acquisition, 27*, 173–204.

Van Berkum, J. J., Brown, C. M., & Hagoort, P. (1999). Semantic integration in sentences and discourse: evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657–671.

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language, 97*, 279–293.

Vos, S., Gunter, T. C., Kolk, H. H., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology, 38*, 41–63.

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

Weber-Fox, C. M., & Neville, H. J. (2001). Sensitive periods differentiate processing of open- and closed class words: An ERP study of bilinguals. *Journal of Speech, Language, and Hearing Research, 44*, 1338–1353.

Please cite this article in press as: Kotz, S. A. et al., ERPs reveal comparable syntactic sentence processing in native ..., *Acta Psychologica* (2007), doi:10.1016/j.actpsy.2007.10.003