Research Paper: Native-like Event-related Potentials in Processing the Second Language Syntax: Late Bilinguals

Laleh Esfandiari1, Reza Nilipour2*, Parviz Maftoon1, Vahid Nejati3

1. Department of Foreign Languages, Faculty of Literature, Humanities and Social Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran.
2. Department of Speech Therapy, Faculty of Rehabilitation Sciences, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
3. Department of Psychology, Faculty of Education and Psychology, Shahid Beheshti University, Tehran, Iran.

ABSTRACT

Background: The P600 brain wave reflects syntactic processes in response to different first language (L1) syntactic violations, syntactic repair, structural reanalysis, and specific semantic components. Unlike semantic processing, aspects of the second language (L2) syntactic processing differ from the L1, particularly at lower levels of proficiency. At higher L2 proficiency, syntactic violations are more likely to result in P600, similar to the L1 native speakers.

Objectives: This study aims to assess the effect of proficiency on L2 syntactic processing in late bilinguals and determine whether L1-like cerebral activation patterns will result.

Materials & Methods: In this descriptive quantitative research, the subjects were two groups of Persian-English bilinguals (L1=Persian, L2=English; n=10 high-proficient, n=10 low-proficient; gender=female who started learning English as a Foreign Language (EFL) after the age of 15 through explicit instructions. Within the violation paradigm, Event-related Potentials (ERPs) were collected from the subjects in the neurocognitive lab of Shahid Beheshti University, Iran, in 2019-20. The experimental trials of the ERP task included violated English regular past tense verbs. ERP components were compared with those of the L1 (components closer to P600).

Results: The t-value for P600 peak latency differed significantly only for the Incorrect past tense verb (ICV) condition and only in O2 (P=0.039463, t=2.2205, CI: 0.003112-0.11249, P<0.05) between the two groups (higher in the high proficient group).

Conclusion: P600 for the high-proficient group demonstrated that L2 proficiency was a more determinant factor in L1-like cortical representation of L2 than the age of acquisition and or the type of context.

Keywords: Language; Semantics; Brain waves
Introduction

The emergence of noninvasive electro-physiology, bio-magnetism, metabolism, and hemodynamics measures of brain activities responsible for linguistic and cognitive processing were markers of a paradigm shift in the study of human language and cognition during the past few decades [1]. Examples of these techniques are Electroencephalography (EEG), Magnetoencephalography (MEG), Positron Emission Tomography (PET), and functional Magnetic Resonance Imaging (fMRI). Embedded within the EEG are the neural responses associated with specific sensory, cognitive, and motor events, making it possible to extract these responses from the overall EEG utilizing a simple averaging technique. These specific responses are called Event-Related Potentials (ERPs) to denote electrical potentials associated with specific events [2]. In the first language (L1), different types of processing difficulties elicit different ERP components. Unlike other neuroimaging techniques such as fMRI and MEG, which provide excellent temporal resolution, ERP research (in EEG) has given us many activation patterns known as ERP components useful for detecting effects in online L1 processing (e.g., P600 and N400). These components provide a clear frame of reference for examining the attainment of native language processing in the second language (L2) and artificial language studies [3].

P600 also called Syntactic Positive Shift (SPS), is a positive wave appearing 600 ms after a native speaker processes (morpho) syntactic violation (Figure 1) [4]. P600 has initially been interpreted as reflecting syntactic processes in response to different syntactic violations, such as phrase structures, subcategorization, number, gender, tense, case agreement, and constraints on long-distance dependence. Later, it was interpreted as reflecting syntactic repair and structural reanalysis, or syntactic integration [5]. It was discussed as an ERP component that can reflect both the syntactic and semantic components [6].

Topographically, the precise origin of this process is still unclear because the P600 has so far not been localized using time-sensitive neuroimaging. Few MEG studies have localized the P600 in the middle temporal gyrus and the posterior portion of the temporal cortex, and lesion studies have referred to the basal ganglia as part of the circuit to support the syntax-related P600 processes [7]. P600 is enhanced when a grammatical judgment task is enforced, which directly triggers the processes of repair and reanalysis [8].

Unlike semantic processing, aspects of L2 syntactic processing differ from L1, particularly at lower levels of proficiency [9, 10]. At lower L2 proficiency, Anterior Negativities (ANs) are typically not found; instead, participants show no negativity at all [11] or N400s or N400-like posterior negativities are elicited from them [9]; suggesting the reliance on declarative memory system for L2 syntactic processing. However, ANs have also been found in higher L2 proficiency in some recent studies [12, 13]. The evidence that P600s are generally found in L2, especially at higher proficiency [9, 12, 13], shows reliance on L1 syntactic processing mechanisms due to gained proficiency.

A significant number of behavioral L2 studies have addressed the role of learning context (e.g., study-abroad and at-home university learning) in developing various linguistic skills. In general, results indicate that study-abroad learners improve in measures of fluency and oral skills. The type and amount of the L2 input (formal education vs. implicit immersion into a second language context) are found to influence the acquisition of a second language [14]. In learning L2 explicitly through formal education, the input is restricted, and L2 speakers are exposed to the target language in a structured, usually nonauthentic, and discontinuous way. While in the immersion, for instance, an ESL (English as a second language) condition, the volume of the L2 input is immense and provided from a variety of authentic sources and social settings [14]. Österhout et al. [9] draws on three studies indicating that classroom-based explicit...
L2 instruction can also result in changes in the brain’s electrical activity topographically as well as structurally. Like implicit and immersion types, these changes can occur during the earliest stages of L2 acquisition, and P600 is possible to extract after exposing the explicitly instructed L2 learners to syntactic violations.

Regarding the age of acquisition, behavioral studies on L2 immigrants have shown that most L2 speakers who start acquiring the L2 sometime before puberty demonstrate syntactic performances roughly similar to those of the native speakers, while their performance differs significantly from the native speakers when their onset to learn the L2 is after puberty [15], and they regarded the native-like L2 speakers as exceptions [16]. This interpretation is consistent with the sensitive period in which there is a window of opportunity for learning, and biological changes interfere with the effect of experience at some periods than others. However, recent studies have shown that despite the correlation between proficiency and Age of Acquisition (AoA), some late learners become proficient in L2 with maximum similarity to the native speakers [12].

On the other hand, ERP studies of L2 suggest that besides the effects of the context of learning, AoA, and proficiency level, aspects of L2 language, e.g., syntax vs. semantics, need to be taken into account. For example, although the neurocognition of L2 lexical/semantic processing seems similar to that of L1, the neural processes underlying L2 (morpho) syntax differ from those of the L1 depending on the learner’s level of proficiency (or exposure). At lower levels, L1 brain processes (as indexed by ANs and P600s and late ANs) are not or rarely found. Instead of the automatic structure building in L1, which is indexed by ANs, (morpho) syntax in lower proficiency L2 may depend on lexical/semantic processes initially, as reflected by the N400. In contrast, the presence of ANs and P600s and late ANs at higher proficiency levels suggest that L1-like brain processing is possible for the L2 despite the type and amount of exposure. However, the level of L2 proficiency necessary to achieve native-like brain mechanisms has remained unknown to this day.

Since the 90’s, there is a debate concerning the underlying cognitive processes driving the elicitation of the P600 [13]. Considering the results from a broader study focusing on the effect of L2 proficiency on L2 syntactic as well as semantic processing [17], this study examined the role of L2 proficiency in two groups of late bilinguals while the context of learning and AoA were the least favorable in terms of conditions which are more likely to lead to L1-like activation patterns.

Materials and Methods

The violation paradigm to identify different temporal stages of processing indexes was applied in this study. The present study was descriptive quantitative research using both behavioral and neurocognitive measures. Since sampling was non-random, this study is a non-experimental and descriptive one, and the subjects are being observed in a completely natural and unchanged natural environment. True experiments, although providing analyzable data, often adversely influence the normal behavior of the subjects.

Study subjects

A total of 20 healthy (without serious medical history as reported by themselves) right-handed adult females (Mean±SD age=25.50±5.09 years, age range=19-35 years) self-reported without any neurological or psychiatric pathology were selected as the participants of this study. Sampling was the non-probability purposive in which the subjects were selected non-randomly. The initial population (introduced mainly by their study department at Shahid Beheshti University, Iran since October 2019) who filled out the language proficiency and demographic questionnaires via email [18], did the online Edinburgh handedness inventory, and 51 individuals took the online language proficiency test. Their Mean±SD results of the Edinburgh handedness inventory was 89.40±6.02. After considering the exclusion criteria, i.e., undesirable proficiency level (14 subjects), left-handedness (2 subjects), knowing other languages...
than English and Persian (4 subjects), undesirable age of learning onset (2 subjects), different type of instruction in L2 (2 subjects), and different duration of instruction in L2 (4 subjects), 28 subjects have remained. Of whom 8 (4 high proficient and 4 pre-intermediate) were selected randomly to take the pilot testing of the task (stimuli) both behaviorally and the ERP trials. They were finally removed from the study due to the possible learning and or noticing effect. The ERP results gained from the final 20 subjects were analyzed.

Based on the scores gained in the online Oxford Placement Test (OPT), the subjects were assigned into two groups of 14: the Pre-Intermediate (PI) and the high proficient Advanced level (AD) groups. According to the standard criteria for scoring the online OPT, the participants gaining scores between 99 and 120 were regarded as highly proficient or advance equal to the Common European Framework of Reference for languages (CEFR) C2, and participants with scores between 60 and 79 were assigned in the lower- or pre-/intermediate level (equal to CEFR B2). The rest were excluded. In the selection process of subjects with a lower level of proficiency, those who gained scores around 60-70 (lower band) were selected because we intended to achieve a more accurate difference in performing the tasks in our study (Table 1).

A number of variables were controlled through the demographic questionnaire. The age of acquisition, i.e., the age of onset for participants in both groups, was over 15 (around just after puberty). The pre-intermediate group had the language experience of fewer than five years, and the advanced group had the experience of learning English for more than five and less than 10 years. The educational level of all participants was either BA or MA in fields of study as psychology, linguistics, engineering, neuroscience as well as the English language. They all learned English explicitly through English language institutes or university courses. Therefore, they were regarded as “compound bilinguals” [19] who all knew only two languages, and their mother tongue was Persian. They either volunteered or accepted to participate for monetary rewards.

**Study stimuli**

The original ERP task data were adopted from Newman et al. [20] and consisted of 240 simple declarative English sentences, 40 in each of the six conditions containing correct and violated regular past forms, phrase structure rules, final-word semantic. This study is drawing on the regular past tense sentences from the syntactic violation part of the original study, which includes 80 English declarative sentences in two conditions (experimental: 40 sentences with violated regular past tense verbs, control: 40 sentences with correct regular past tense verbs) (Table 2).

Stimuli were counterbalanced across subjects so that each subject saw either the control or the anomalous version of a given sentence. The control condition (CV) included 40 sentences (control) with the regular past tense verb; all sentences were devised with similar structures, all starting with an expression of past time, e.g., “Yesterday…” or “last week…” or any other past tense marker adverb, followed by a subject pronoun (I, he, or, she), a verb, and a post-verbal argument. For each sentence, a corresponding experimental sentence was created with a violated past tense form by replacing the past tense inflected form with its stem (unmarked) form (e.g., Last week he fail the exam.) (ICV). The visual ERP task presentation of each sentence was initiated by the subject’s pressing one of the two response buttons on the mouse, after which the outline of a box (7×3 degrees of visual angle) would appear in the center of the screen. After a variable (random) delay of 300-1100 ms to attenuate ERP effects associated with the expectation of forthcoming stimuli, the first word of the following sentence appeared. The trial presentation duration was 300 ms (the Stimulus Onset Asynchrony (SOA) of 500 ms). Following the onset of the final word of each sentence, the box outline remained on the screen for a random period of 300 to1100 ms. The words of the sentence were presented one at a time; each word was displayed for 300 ms following a 200-ms delay until the next word. After the box disappeared from the screen, the correctness or incorrectness of the sentence was to be shown via a right-click or left-click, respectively.

**Data collection**

Continuous ERP data were recorded from each participant using a 32 tin electrode with 19 active electrodes. Electrode positions were specified by the International 10–20 system (FP1, FP2, FPz, Fz, F3, F4, FCz, C3, Cz, C4, CPz, P3, T3, T4, T5, T6, Pz, P4, O1, O2, left/right auricles). Impedances were lowered to <5 kΩ. EEG was amplified using an online band pass filter of 208 Hz. Trials containing blinks, eye movements, or excessive noise were identified off-line (using a maximum peak-to-peak amplitude threshold tailored to each participant’s data) and were regarded as noise and finally removed.

The recording duration was 25-30 min using an EEG amplifier (Mitsar) and the WinEEG Software at the psychology lab of Shahid Beheshti University, Department
of Psychology, Tehran, Iran. The stimuli were presented visually on a computer screen at the size of 18.5” and 1369×768 pixels quality using Psytask 1.52. Stimuli were presented, and data were recorded via a Dual-Core Pentium computer under Windows 7 Ultimate.

Results

A two-sample t test was used to compare the ERP latency and amplitude of the performance of the two groups on each condition. ERPs to target words in each condition were elicited across the P600 time window (500-700 ms), i.e., 100 ms before and after the target word.

The between-group comparison (Figure 2 and Table 3) showed the distribution of t value for the comparison of P600 (peak latency and amplitude, respectively) between the two groups across the conditions. Results showed that the t value for P600 peak latency differed significantly only for the Incorrect past tense verb (ICV) condition.

Table 1. Descriptive statistics for the online Oxford placement test proficiency results

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>N</th>
<th>Range</th>
<th>Min</th>
<th>Max</th>
<th>Mean±SD</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High proficient</td>
<td>10</td>
<td>20</td>
<td>100</td>
<td>120</td>
<td>110.70±6.237</td>
<td>38.900</td>
</tr>
<tr>
<td>Low proficient</td>
<td>10</td>
<td>11</td>
<td>60</td>
<td>71</td>
<td>65.80±4.050</td>
<td>16.400</td>
</tr>
</tbody>
</table>

Table 2. The experimental and control tasks of the stimuli

<table>
<thead>
<tr>
<th>Condition Type</th>
<th>Label</th>
<th>Type of Violation</th>
<th>No. of Sentences</th>
<th>Sample Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>ICV</td>
<td>Incorrect regular simple past tense verbs in English with –ed inflections</td>
<td>40</td>
<td>Last week, he fail the exam*</td>
</tr>
<tr>
<td>Control</td>
<td>CV</td>
<td>Correct regular simple past tense verbs in English with –ed inflections</td>
<td>40</td>
<td>Yesterday he seemed to be happy.</td>
</tr>
</tbody>
</table>

*ICV: Incorrect regular past verb, CV: Correct regular past verb; **Sentences containing a violated syntactic form.

Table 3. Between-group comparison: P600 latency

<table>
<thead>
<tr>
<th>Condition*</th>
<th>Channel</th>
<th>P</th>
<th>t</th>
<th>Confidence Interval (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>T6</td>
<td>0.023561</td>
<td>2.4737</td>
<td>0.0087095 0.10689</td>
</tr>
<tr>
<td>ICV</td>
<td>O2</td>
<td>0.039463</td>
<td>2.2305</td>
<td>0.003112 0.11249</td>
</tr>
</tbody>
</table>

*ICV: Incorrect regular past verb, CV: Correct regular past verb.

Figure 2. Distribution of t value for the P600 peak latency: Comparing the high proficient Advanced level (AD) and Pre-Intermediate (PI) group

The more the color goes red, the larger the P600 peak latency for the AD group as compared with the PI group. Conversely, blue areas are the regions where the p600 peak latency is greater for the PI group (CV: Correct regular past tense verb; ICV: Incorrect regular past tense verb); * Denotes P<0.05.
and only in O2 (P=0.039463, t=2.2205, CI: 0.003112-0.11249, P<0.05) between the two groups (higher in the advanced group). For correct sentences, the P600 peak amplitude was smaller for the AD group than the PI.

Within-group comparisons showed the difference between the ERP data elicited by the experimental and control conditions related to the processing of regular English past tense verbs (ICV and CV) by the AD subjects. Within the time window of 100-200 ms, a negative component appeared in the frontal regions (Fp1, Fp2, Fp3, F7, F4, and Fz) upon processing the incorrect English regular past tense verbs in high-proficient EFL learners. In addition, the grand average results for the AD group showed P600 in Pz, P3, and O1 during the processing of the incorrect regular past verb.

**Figure 3.** Grand average of negative ERP waves in the frontal electrodes within the time interval of 100-200 ms and P600 in occipital and parietal regions for the control and experimental conditions while processing the violated and correct English regular past tense verbs in high-proficient EFL learners

---

**ICV= Incorrect regular English past tense verbs**

---

**CV= Correct regular English past tense verbs**

---

**Figure 3** shows the difference between the ERP data elicited by the experimental and control conditions related to the processing of regular English past tense verbs (ICV and CV) by the AD subjects. Within the time window of 100-200 ms, a negative component appeared in the frontal regions (Fp1, Fp2, Fp3, F7, F4, and Fz) upon processing the incorrect English regular past tense verbs in high-proficient EFL learners. In addition, the grand average results for the AD group showed P600 in Pz, P3, and O1 during the processing of the incorrect regular past verb.

**Figure 4** shows the significant difference in N400 amplitude for the incorrect regular past conditions only in O1 for the PI subjects. P600 amplitude did not differ significantly between the two conditions for these subjects; however, not significantly, the P600 latency in F7 was bigger for the PI as compared with the AD subjects.

**Discussion**

The statistically significant difference between the ERP components (peak latency and peak amplitude) elicited from the processing of English regular past tense verbs in pre-intermediate and high proficient adult L2 learners was investigated. Our analysis showed that the latency of the P600 component for incorrect regular past tense sentences in the occipital region (O2) was significantly higher in the high proficient advance group than in the lower proficient pre-intermediate one. It can also be deduced that the P600 amplitude was smaller for the AD than the PI group when the correct
sentences were shown to subjects, i.e., it seems that the PI subjects did expect to observe a conceptual error (semantic violation) because they did not have sufficient English control over correct sentences. However, the statistical analysis does not show a significant difference between the AD and PI groups in the P600 amplitude while processing sentences with violations, not necessarily meaning the absence of the P600 component.

Upon processing the incorrect English regular past tense verbs in high-proficient EFL learners, a harmful component appeared in the frontal regions (Fp1, Fp2, Fp3, F7, F4, and Fz), which could address Newman et al. ’s [21] findings. They indicated that syntactic violations provided greater activations, compared to semantic anomalies in several regions of the superior frontal gyrus, in both hemispheres.

The grand average results showed P600 amplitude differences in Pz, P3, and O1 during the processing of the incorrect regular past verb for the AD subjects, and it did not differ significantly between the correct and incorrect regular past tense conditions of the PI group. During the processing of the incorrect regular past verbs, the P600 latency in F7 was bigger for the PI compared with the AD subjects, which could reflect the time the PI subjects needed for retrieval of the elements. Longer P600 latency for PI subjects could be explained in light of the discussions of the P600 by Friederici (2002) [22] and Hagoort (2003) [23] in that if the P600 reflects the creation or destruction of syntactic relations, then the latency of the P600 should reflect the time needed for retrieval of the elements that participate in those relations, whereas the duration and amplitude of the P600 should be a function of the structure building processes themselves. Then, they predicted that different structural and lexical manipulations should impact the P600 differently. Manipulations that impact retrieval processes should change the latency of the P600, whereas manipulations that impact the number and type of syntactic relations attempted should change the amplitude and or duration of the P600.

The fundamental difference between the acquisition of L1 and L2 in late L2 learners addressing the maturational constraints of L2 capability has been argued by some scholars who reported that much greater cerebral plasticity than previously assumed for late multilingual [24, 25]. Findings of this study were in favor of the latest results in that the subjects with higher proficiency level used the same neural mechanisms as L1 despite their late AoA, type of instruction (explicit), and limited exposure to L2 input; however, this could be due to the nature of the ERP task. Processing different L2 structures seem to be differentially sensitive to AoA and proficiency effects. For example, AoA has a greater impact on phonological and some grammatical processes, while the proficiency level seems to affect more the lexical-semantic and control processes [26].

Conclusion

Findings of this study, along with several similar ERP studies, have shown that adult second language (L2) learners have the increasing opportunity to gain more native first language (L1)-like syntactic processing as their proficiency in L2 increases, regardless of the type of instruction and age of L2 acquisition. The practical implications of ERP studies like the present one for language instruction are that most neuroimaging and ERP studies involve instructed L2 learners. Similarly, our study signifies the benefits of explicit instruction for adult learners in grammatical learning, which can take place in extremely short periods, within a classroom instruction with insufficient and inefficient input compared with the L1, learners could still show native-like processing, even though at a slower rate for some language constructions. However, further studies should be carried out with Iranian children of much earlier ages of onset for learning L2 explicitly. The effect of explicit instruction as mediated with proficiency at lower ages could add to the findings of this study.

Ethical Considerations

Compliance with ethical guidelines

All ethical principles were considered in this study. All participants signed a standard international consent form based on the World Medical Association (WMA) Declaration of Helsinki for Medical Studies (2018).

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors contributions

Conceptualization and supervision: Reza Nilipour; Methodology: Reza Nilipour and Laleh Esfandiari; Investigation, writing – original draft, and writing – review & editing, data analysis: All authors; Data collection: Laleh Esfandiari and Vahid Nejati; Funding acquisition and Resources: Laleh Esfandiari.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

We would like to thank Dr. Reza Khosrowabadi at the Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, for his support and advice, especially the statistical analyses and the data collection. We would also like to thank Dr. Mahdieh Esmaeili from Iran Research Center on Aging, the University of Social Welfare and Rehabilitation Sciences, Tehran, for her help in the EEG lab and task design.

References


[25] Steinhauser K. Event-related Potentials (ERPs) in second language research: A brief introduction to the technique, a selected review, and an invitation to reconsider critical peri-