Brain & Language 139 (2014) 84–98

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Early stage second-language learning improves executive control: Evidence from ERP

Margot D. Sullivan [a], Monika Janus [a], Sylvain Moreno [b], Lori Astheimer [c], Ellen Bialystok [a,b,*]

[a] York University, Canada
[b] Rotman Research Institute and University of Toronto, Canada
[c] University of Massachusetts, United States

ARTICLE INFO

Article history:
Accepted 13 October 2014
Available online 6 November 2014

Keywords:
ERP
Second-language training
Executive control
Go–nogo
Sentence judgment
Bilingualism
Neuroplasticity

ABSTRACT

A growing body of research has reported a bilingual advantage in performance on executive control tasks, but it is not known at what point in emerging bilingualism these advantages first appear. The present study investigated the effect of early stage second-language training on executive control. Monolingual English-speaking students were tested on a go–nogo task, sentence judgment task, and verbal fluency, before and after 6 months of Spanish instruction. The training group ($n = 25$) consisted of students enrolled in introductory Spanish and the control group ($n = 30$) consisted of students enrolled in introductory Psychology. After training, the Spanish group showed larger P3 amplitude on the go–nogo task and smaller P600 amplitude on the judgment task, indicating enhanced performance, with no changes for the control group and no differences between groups on behavioral measures. Results are discussed in terms of neural changes underlying executive control after brief second-language learning.

1. Introduction

Lifelong bilingualism has been shown to enhance performance in verbal and nonverbal executive control tasks across the lifespan (review in Bialystok, Craik, Green, & Gollan, 2009). The presumed source of this effect is the well-established finding that both languages are active for bilinguals when one is being used, requiring a control mechanism to focus attention on the target language and avoid intrusions from the competing language (Gollan & Kroll, 2001; Green, 1998). This control mechanism is presumed to be the executive control system used generally to control attention. Evidence from fMRI supports this interpretation by demonstrating overlap in the networks involved in domain-general executive control and language switching for bilinguals (Abutalebi & Green, 2008; Luk, Green, Abutalebi, & Grady, 2012). However, little research has examined the emergence of these changes over time to determine how much bilingual experience is necessary for these advantages to be manifest, and no research has jointly investigated the behavioral and neural outcomes of second-language (L2) learning at very early stages to capture the first evidence of such changes. L2 learners are in a sense "bilinguals in training" and so potentially provide powerful evidence for the origin and possible mechanism of processing differences found with bilingualism.

Osterhout and colleagues (e.g., Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006) have examined the emergence of linguistic processing differences that follow from early L2 learning by recording event-related potentials (ERPs). In one study, McLaughlin, Osterhout, and Kim (2004) showed that after just 14 h of university-level French instruction, students showed significant changes in N400[1] amplitude during a lexical decision task, even though behavioral performance was at chance. These results not only demonstrate changes in the neural response after second-language learning, but also show that these brain-related functional changes associated with experience precede behavioral changes. More recent research by this group (Tanner, McLaughlin, Herschensohn, & Osterhout, 2013) compared native German speakers with English speakers enrolled in college-level German courses on grammaticality judgments in German. Native German speakers and 3rd year L2 learners both showed the classic P600 effect associated with syntactic violations, but subsequent analyses indicated that there was individual variation within a group of 1st year learners. Early learners displayed either N400 or P600 waveforms in response to syntactic violations. Moreover, the

* Corresponding author at: Department of Psychology, York University, 4700 Keele St., Toronto, ON M3J 1P3, Canada.
E-mail address: ellenb@yorku.ca (E. Bialystok).

[1] The N400 effect reliably occurs when processing semantic anomalies and the P600 effect reliably occurs when processing syntactic anomalies (Osterhout & Nicol, 1999).

http://dx.doi.org/10.1016/j.bandl.2014.10.004
0093-934X/

amplitude of the P600 was positively correlated with behavioral performance for all learners but not for native speakers. These results support the notion that linguistic processing is highly malleable in the early stages of L2 training, and that brain functional changes due to training can be stage-like and vary among individuals.

Evidence from two imaging studies has also demonstrated the effect of brief L2 training in adulthood on brain structure. First, adults who underwent an intense three-month language training program in one of three languages (Arabic, Dari, or Russian) at an interpreter academy in Sweden showed increased hippocampal volume and increased cortical thickness of three left hemisphere regions: middle frontal gyrus, inferior frontal gyrus, and superior temporal gyrus (Mårtensson et al., 2012). Second, in a diffusion tensor imaging (DTI) study using monolingual English speakers as participants, Schlegel, Rudelson, and Tse (2012) found that 9 months of Chinese instruction led to improved efficiency of white matter tracts within the left hemisphere connecting language areas, the temporal area of the right hemisphere, and within the tracts crossing the genu of the corpus callosum. Not only do these results demonstrate that brain structural reorganization plays a role in adult L2 learning, but also that the connectivity changes elicited from training parallel the results found in lifelong bilinguals. Luk, Bialystok, Craik, and Grady (2011) showed better maintained white matter structure in tracts across the corpus callosum and connecting frontal to posterior regions in bilinguals than in comparable monolinguals. Similar results were reported by Mohades et al. (2012) in a study comparing monolingual and bilingual 8- to 11-year-olds, with the largest group difference found along the left inferior occipitofrontal fasciculus, a white matter tract associated with semantic processing.

Together, these results from the bilingualism and L2 training literature indicate that functional and structural adaptations to the brain can occur developmentally along a continuum of bilingualism. Investigation of these early changes associated with language learning has largely been confined to language tasks and left hemisphere regions associated with linguistic processing. No study to date has examined the *early* effects of L2 training on the neuronal processes related to executive control, a set of processes associated with bilingualism. There may also be connections in the other direction, namely executive control processes influence language learning (Bartolotti, Marian, Schroeder, & Shook, 2011) and language processing (for review see Ye & Zhou, 2009). Our concern, however, is not with the factors responsible for successful L2 learning (which undoubtedly include better executive control) but rather with the consequences of L2 learning on executive control, given equivalent initial levels of executive control for individuals who do or do not receive the training. The primary challenge is to understand how training in one domain (learning a second language) can lead to changes in another domain (nonverbal executive control). Understanding this process is the aim of the present study.

There is some evidence that brief forms of *within-domain* training can modify executive control networks. Rueda, Rothbart, McCandliss, Saccomanno, and Posner (2005) trained children for 5 days with exercises focused on attention and control and reported that ERP waveforms for the trained children showed an adult-like N2 effect on the child attentional network task (ANT) with no changes found for the untrained children. Similarly, Moreno et al. (2011) demonstrated that brief training produced effects *across* domain. Preschool children who were given 20 days of musical training showed larger P2 amplitude on the go–nogo task, a measure of response inhibition, than did children who received training in visual art. The change in P2 amplitude was positively correlated with improved verbal intelligence scores, supporting the claim for cross-domain influence.

The present study extended this research by investigating whether brief L2 learning is sufficient to produce early modification of neuronal responses to executive control tasks for both within- and across-domain tasks in young adults. Because children are in the process of developing competence across all domains of knowledge, a more stringent test of experience-induced plasticity must be based on evidence from adults. All of the tasks used in the present study have previously demonstrated advantages in fully bilingual adults compared to monolinguals: verbal fluency (letter condition; Luo, Luk, & Bialystok, 2010), metalinguistic task (grammaticality judgment; Moreno, Bialystok, Wodniecka, & Alain, 2010), and nonverbal control (go–nogo; Moreno, Wodniecka, Tays, Alain, & Bialystok, 2014).

Behavioral studies of verbal fluency have shown that bilinguals generally produce fewer words than monolinguals (Bialystok, Craik, & Luk, 2008; Gollan, Montoya, & Werner, 2002; Rosselli et al., 2000). However, if the two groups are matched on vocabulary size, then bilinguals perform equivalently to monolinguals in the semantic condition (e.g., name "animals") but outperform them in the more difficult letter condition (e.g., name words that begin with the letter "F") that requires executive control (Luo et al., 2010). In addition to being more effortful because of the difficult selection criterion, the letter condition is associated with brain regions involved in the executive control network (Grogan, Green, Ali, Crinion, & Price, 2009).

Metalinguistic tasks were the first domain to demonstrate better performance by bilingual children than monolinguals. In one set of studies by Bialystok (1986), children were asked to judge only the grammaticality of simple sentences, but in some conditions, the meaning conflicted with the grammaticality ("Apples grow on noses"), requiring executive control to focus attention on the grammar and inhibit the salient but distracting meaning. Bilingual children were more accurate than monolinguals in judging these anomalous sentences (review in Bialystok, 2001). Moreno et al. (2010) extended this work to adults by administering a similar task while recording ERPs. As with children, both monolinguals and bilinguals performed comparably in the simple conditions, but for the conflict sentences that required executive control, bilinguals showed larger N400 amplitude for semantically anomalous sentences and smaller P600 amplitude for syntactically incorrect sentences that was more bilaterally distributed than was found for monolinguals. This is evidence for a continued bilingual advantage into adulthood for metalinguistic tasks that recruit the executive control system.

Finally, a commonly-used task to investigate nonverbal control is the go–nogo paradigm in which participants press a key if one kind of stimulus appears and refrain from responding if another kind of stimulus is presented. It is a simple task that assesses frontally-mediated inhibitory control and has well-established ERP signatures (Pfefferbaum, Ford, Weller, & Kopell, 1985). Nogo trials show greater amplitude for both negative (N2) and positive (P3) waves compared to go trials reflecting the requirement for greater attentional resources in those nogo trials (Lavric, Pizzagalli, & Forstmeier, 2004). In two studies, bilinguals showed larger N2 amplitude than monolinguals indicating better control on nogo trials and a more efficient executive control system (Fernandez, Tartar, Padron, & Acosta, 2013; Moreno et al., 2014).

The studies showing bilingual advantages in verbal fluency, metalinguistic decisions, and nonverbal inhibition are essentially correlational in that participants were obviously not randomly assigned to groups. All the bilinguals in those studies were fluent and practiced in the use of two languages over a long period of time. What is not known is how much bilingualism is required for these effects to appear and whether their emergence can be more precisely traced to the use of a second language. The current study investigated the effects of early stage L2 training on brain-related and behavioral changes in executive control. Evidence for such changes would provide support for neuroplasticity in adult language learners and show the early emergence of a neural profile that resembles bilingual processing.

2. Materials and methods

2.1. Participants

Participants were 55 English-speaking monolingual undergraduate university students between the ages of 17 and 32 years old. The training group consisted of students taking first year introductory Spanish and the control group consisted of students recruited from an introductory Psychology course who had never taken Spanish or any other second-language course (see Table 1 for background characteristics). It is important to note that both groups of participants were enrolled at the same university, taking the same overall number of courses, and did not differ on years of education; the only known difference between the groups was whether or not they were receiving L2 training. Students in the two groups were equivalent on all background measures as well as in socioeconomic status as indicated by maternal education.

2.2. L2 training

The introductory Spanish course at York University, Toronto, Canada is described by the department as an intensive full-year course that gives equal attention to listening, speaking, reading, and writing skills. Students are expected to attain the knowledge of core grammatical structures as well as a minimum vocabulary of 2500 Spanish words. In addition, a self-report questionnaire that was designed to assess the extent of L2 training the students received during the 6 months of the course was administered to the Spanish-learners at posttest. Participants in this group provided ratings on the number of hours spent speaking, reading, and writing Spanish over the training period (see Table 2).

2.3. Procedure

Participants were tested at the beginning (September) and end (March) of the academic year. In the first session, participants were given the Language and Social Background Questionnaire (LSBQ; Luk & Bialystok, 2013), Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) to assess English receptive vocabulary, and Kaufman Brief Intelligence Test: Matrices Subtest (K-BIT; Kaufman & Kaufman, 1990) to assess nonverbal fluid intelligence. All three experimental tasks were administered in both sessions. ERPs were recorded during the sentence judgment and go–nogo tasks.

2.4. ERP recording

ERP tasks were administered on a Dell desktop computer with a 19-inch LCD monitor. Participants were seated in a comfortable armchair with the computer screen positioned 50 cm from eye level. The lights in the testing room were kept off during the task. The electroencephalogram (EEG) was continuously recorded using a BioSemi Active-Two amplifier system (BioSemi, Amsterdam, Netherlands) from an array of 64 active Ag–AgCl electrodes located at standard scalp positions (International 10/20 system sites) as well as the left and right mastoids. Data were collected using a DC recording and were digitized at a sampling rate of 512 Hz. During the recording, all electrodes were referenced to the CMS (Common Mode Sense) electrode, with the DRL (Driven Right Leg) electrode serving as the ground. Impedances of the electrodes were kept below 20 kΩ. The EEG system and each step of the set-up procedure were explained to each participant. Each participant completed a series of ocular movements to record the electro-oculogram at the start and end of each ERP testing session.

EEG data were analyzed off-line using Brain Electrical Source Analysis (BESA) software (Version 5.1.8; MEGIS Software, Gmbh). Continuous EEG was segmented into 1000 ms epochs from 200 ms before to 800 ms after target onset (go or nogo stimulus, or critical word), and was baseline corrected to the 200 ms prestimulus interval. During pre-processing, all trials were visually inspected and those containing excessive noise due to ocular or movement artifacts were removed. All participants included in the final analyses had at least 80% of trials in each condition and amplitude thresholds for the signal were adjusted on a participant-by-participant basis to include >80% of the target stimuli in the average.

Based on the electro-oculogram, averages were calculated for each participant for horizontal and vertical eye movement and for blinks. Principle component analysis was conducted on these averages and scalp projections related to these eye movement components were subtracted from EEG data for each participant in order to minimize signals due to each type of eye movement (see Picton et al., 2000). ERPs were then re-referenced to the average of the left and right mastoid and averaged for each individual participant, separately for each task condition and electrode site. A bandpass filter of 0.01–30 Hz was used, and a 60 Hz notch filter was then applied to the data of each participant. All analyses were conducted on correct trials only.

2.5. Tasks

2.5.1. Verbal fluency

Participants were asked to produce as many English words as possible within 60 s according to a criterion. There were three trials for category fluency (clothing, fruits, occupations) and three trials for letter fluency (letters F, A, S). Instructions for letter fluency included the restriction to exclude proper nouns, numbers, and variations of a word already produced. Scores were the number of acceptable items produced in each condition, averaged across the three trials.

2.5.2. Sentence judgment

The sentences were presented in English and were based on materials developed by Osterhout and Nicol (1999) and also used by Moreno et al. (2010) (see Appendix A for the complete list). Frames were manipulated to create sentences that were correct, "A new computer will *last* for many years", syntactically incorrect, "A new computer will lasting for many years", semantically anomalous, "A new computer will paint for many years" or filler sentences that were both incorrect and anomalous. Filler sentences were included to equate the number of acceptable and unacceptable responses but were not analyzed. Correct and semantically anomalous target words were matched in frequency from the

Table 1
Group means (and standard deviations) for background measures by group.

Group	Age	Years of education	Number of female participants	Number of right-handed participants	PPVT	K-BIT
Spanish learners $N = 25$	20.6 (3.0)	13.9 (1.7)	18	22	106.4 (9.5)	97.2 (12.8)
Controls $N = 30$	19.7 (1.7)	13.7 (1.4)	22	29	106.3 (10.9)	95.1 (10.8)

Table 2
Self-report estimates in hours, standard deviations, and range for time spent in Spanish activities during the period of the course for Spanish learners.

Hours per week speaking in class			Hours per week speaking outside of class			Hours per week writing			Hours per week reading		
M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
2.6	1.5	0–5	2.6	2.2	0–8	2.4	2.0	0–8	3.5	3.0	1–15

Kucera and Francis (1967) norms (correct: $M = 96$ times per million, semantically anomalous: $M = 70$ times per million, $p > .2$) and length (correct: $M = 4.94$, semantically anomalous: $M = 4.52$, $p > .3$).

Sentences were presented one word at a time on the computer screen, after which participants made a judgment about whether or not it was grammatically correct, regardless of the meaning. There were 120 sentences, consisting of 30 exemplars of each of the four sentence types. Counterbalancing across lists ensured that only one version of each sentence was present on a given list viewed by each participant. Each trial began with a white fixation cross for 500 ms, after which the sentence appeared one word at a time in a white font against a black background. Each word was shown in the center of the screen for 300 ms, with a blank screen between each word that lasted 450 ms (post-stimulus interval). A 500 ms blank screen interval followed each sentence. Afterwards, the answer prompt screen appeared in which a white fixation cross was shown again in the center of the screen with a lowercase 'c' (correct) in the lower left hand corner of the screen and a lowercase 'i' (incorrect) in the lower right hand corner. The prompt remained on screen until the participant pressed the letter key 'a' on the keyboard for correct or the letter key 'l' for incorrect.

ERP analyses were conducted on mean amplitudes of difference waves (violation condition – control condition) measured during the N400 (300–500 ms) and P600 (600–800 ms) time windows. The central-posterior electrodes of interest for the N400 difference were those in a 2 (anterior–posterior) by 3 (medial–lateral) grouping (C1, Cz, C2, CP1, CPz, CP2), and the electrodes of interest for the P600 difference were in a 2 (anterior–posterior) by 3 (medial–lateral) grouping (CP1, CPz, CP2, P1, Pz, P2). Four-way repeated measures ANOVAs were conducted for each violation condition with group as the between-subjects factor, and session, anterior–posterior electrode position, and medial–lateral electrode position as the within-subject factors. For analyses where the assumption of sphericity was violated, degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity. The Bonferroni correction was applied to all post hoc pairwise comparisons for significant interactions.

2.5.3. Go–nogo

Stimuli included two triangles (pointing upwards or downwards) and two rectangles (vertical or horizontal) that were white or purple and presented on a black background. The variation in shape and orientation was to reduce repetition effects in visual processing. Participants were instructed to press the mouse key when the shape was white and to inhibit their response when the shape was purple. There were 200 trials consisting of 160 go trials (80%) and 40 nogo trials (20%). For each trial, a white fixation cross appeared against the black background for 500 ms, followed by a blank screen that lasted for a variable duration between 0 and 500 ms, after which the stimulus appeared in the center of the screen for 300 ms. A blank screen was presented for 900 ms before the start of the next trial (post-stimulus interval). Therefore, participants had a total of 1200 ms to make a response following the onset of the stimulus. Each participant completed 200 trials consisting of 160 go trials (80%) and 40 no-go trials (20%). Go and no-go trials were presented in randomized order in a single block. A practice block of 20 trials preceded the task. Recordings of ERPs and accuracy rates were made for both go and nogo trials and RTs were recorded for go trials only.

Unlike the mean amplitude analyses of the sentence judgment components (N400/P600), peak amplitude and peak latency were measured during the N2 (200–350 ms) and P3 (350–500 ms) time windows, both because of the characteristics of the waveforms (the N2/P3 complex provides measurable peaks) and to be consistent with prior research (Moreno et al., 2014). The frontal–central electrode sites were arranged in a 2 (anterior–posterior) by 3 (medial–lateral) grouping (F1, Fz, F2, FC1, FCz, FC2). Four-way repeated measures ANOVAs were conducted with group as the between-subjects factor, and session, condition (go vs. nogo), anterior–posterior electrode position, and medial–lateral electrode position as within-subject factors.

Table 3
Mean number of words (and standard deviations) for letter and category fluency conditions by group at both testing sessions.

Group	Fluency condition	Number of words produced at pretest *M* (*SD*)	Number of words produced at posttest *M* (*SD*)
Spanish learners $N = 24$	Letter	13.7 (3.1)	15.1 (3.2)
	Category	16.6 (2.2)	18.4 (3.0)
Controls $N = 30$	Letter	12.5 (3.3)	13.4 (3.5)
	Category	15.8 (2.4)	16.6 (2.1)

3. Results

3.1. Verbal fluency

Data from the verbal fluency task are presented in Table 3. One Spanish-learner dropped the course before the end of term and

Table 4
Mean accuracy scores (and standard deviations) for the three types of sentences in the grammaticality judgment task by group for both testing sessions.

Group	Sentence	Pretest accuracy *M* (*SD*)	Posttest accuracy *M* (*SD*)
Spanish learners $N = 18$	Correct	0.96 (0.04)	0.96 (0.06)
	Semantically Anomalous	0.86 (0.19)	0.92 (0.08)
	Syntactically Incorrect	0.92 (0.08)	0.90 (0.15)
Controls $N = 23$	Correct	0.96 (0.06)	0.95 (0.05)
	Semantically Anomalous	0.88 (0.08)	0.93 (0.06)
	Syntactically Incorrect	0.93 (0.15)	0.92 (0.16)

was removed from the analyses, leaving a total of 54 participants. A three-way repeated measures ANOVA for group, session (pretest, posttest), and condition (letter, category) showed a main effect of group, $F(1,52) = 5.21$, $p < .05$, $\eta^2 = .09$, with more words produced by the Spanish-learners. There were also main effects of session, $F(1,52) = 22.08$, $p < .001$, $\eta^2 = .29$, with more words at posttest, and condition, $F(1,52) = 59.10$, $p < .001$, $\eta^2 = .53$, with more words produced in the category condition. There were no interaction effects, $Fs < 2.5$.

3.2. Sentence judgment

Data from 14 participants (7 Spanish-learners and 7 controls) could not be analyzed due to either poor behavioral performance (<50% accuracy on multiple conditions, $n = 4$), noisy EEG data ($n = 9$), or dropping the course ($n = 1$). The final sample consisted of 18 Spanish-learners (15 females) and 23 controls (18 females).

3.2.1. Accuracy data

Mean accuracy scores (Table 4) for the sentence judgment task were analyzed with a three-way repeated measures ANOVA for group, session, and condition. A significant session by condition interaction was found, $F(1.73,67.36) = 5.37$, $p < .01$, $\eta^2 = .12$, whereby accuracy increased only on semantically anomalous sentences. There was a main effect of condition, $F(1.66,64.54) = 8.19$, $p < .01$, $\eta^2 = .17$, in which accuracy for correct sentences was higher than for semantically anomalous ($p < .001$) and syntactic violations ($p = .069$), with no difference between the latter two. There were no group effects or interactions, $Fs < 1$.

3.2.2. N400: Time window 300–500 ms

The ERP waveforms for the three conditions at pretest and posttest are shown in Fig. 1. All analyses were conducted on the differences in mean amplitude between each violation condition and correct sentences. These data are shown in Fig. 2 for semantic violations and 3 for syntactic violations. Repeated measures ANOVAs by condition revealed a significant interaction of session by anterior–posterior electrode position for semantically anomalous sentences, $F(1,39) = 3.93$, $p = .05$, $\eta^2 = .09$. Specifically, at pretest, anterior electrodes showed a larger mean amplitude difference (-1.84 μV) than posterior electrodes (-1.59 μV), whereas at posttest, posterior electrodes revealed a greater mean amplitude difference (-1.66 μV) than anterior ones (-1.33 μV). There were no other significant main effects or interactions involving session or group, $Fs < 2.5$.

Analysis of the N400 for sentences with syntactic violations revealed a significant three-way interaction of group, session, and medial–lateral electrode position, $F(2,78) = 5.17$, $p < .01$, $\eta^2 = .12$. Post hoc comparisons examining the effect of group indicated that at pretest there was a group difference in amplitude for left-lateralized electrodes (Spanish -1.48 μV vs. control .10 μV, $p < .05$) and at posttest for right-lateralized electrodes (Spanish $-.96$ μV vs. control 1.29 μV, $p < .05$). There was a significant main effect of group, $F(1,39) = 5.21$, $p < .05$, $\eta^2 = .12$, with the Spanish-learners showing an overall negative amplitude difference ($-.87$ μV) and the control group showing a positive amplitude difference (1.08 μV) during this time window. There were no other significant main effects or interactions involving session or group, $Fs < 1$.

3.2.3. P600: Time window 600–800 ms

For the semantic violation condition, the repeated measures ANOVA for the P600 revealed a marginally significant group by medial–lateral electrode position interaction, $F(2,78) = 2.92$, $p = .06$, $\eta^2 = .07$. Post hoc comparisons examining the effect of medial–lateral electrode position, revealed that for the Spanish-learner group the mean amplitude difference for the left side electrodes ($-.57$ μV) differed from the right side electrodes (.19 μV) ($p < .05$). This difference was not seen in the control group ($M_{\text{left}} = -1.55$ μV, $M_{\text{right}} = -1.41$ μV). There were no other significant main effects or interactions involving session or group, $Fs < 2$.

Importantly, analysis of the P600 on the syntactic violation condition showed a significant three-way interaction of group, session, and medial–lateral electrode position, $F(1.94,75.55) = 3.31$, $p < .05$, $\eta^2 = .08$. Post hoc comparisons examining the effect of session revealed a marginally significant decrease in P600 mean amplitude difference ($p = .056$) after L2 training for the Spanish group across right side electrodes CP2 and P2 (see Fig. 3 for ERPs), with no change for the control group. There was a significant group effect, $F(1,39) = 4.02$, $p = .05$, $\eta^2 = .09$, with the control group showing an overall larger mean amplitude difference (7.01 μV) than the Spanish-learner group (4.19 μV). Independent samples t-tests across an average of all 6 electrodes (CP1, CPz, CP2, P1, Pz, and P2) for the syntactic violation condition indicated no significant group difference at pretest, $t(39) = -1.16$, $p = .26$. There were no other significant main effects or interactions involving session or group, $Fs < 2$.

3.3. Go–nogo

Data from 12 participants (4 Spanish-learners and 8 controls) were excluded due to technical difficulties/noisy EEG data ($n = 11$) or having dropped the Spanish course ($n = 1$). The final sample for this task consisted of 21 Spanish-learners (16 female) and 22 controls (15 female).

3.3.1. Go–nogo behavioral data

A two-way repeated measures ANOVA for group and session on each of the go–nogo behavioral measures in Table 5 showed no significant effects or interactions, $Fs < 3.2$.

3.3.2. N2: Time window 200–350 ms

Fig. 4 presents the ERPs recorded at pretest and posttest, and the difference waveforms for the nogo condition minus the go condition are displayed in Fig. 5. N2 peak analyses revealed a main effect of condition, $F(1,41) = 67.02$, $p < .001$, $\eta^2 = .62$, whereby nogo conditions elicited a larger negativity than go conditions, as expected (Lavric et al., 2004; Pfefferbaum et al., 1985). Independent samples t-tests across all 6 frontal–central electrodes for go, $t(41) = -.98$, $p = .33$ and nogo N2 amplitude $t(41) = .28$, $p = .78$, indicated no significant pretest group differences. There was a significant three-way interaction of group by session by anterior–posterior electrode site, $F(1,41) = 4.95$, $p < .05$, $\eta^2 = .11$. Post hoc comparisons examining the effect of anterior–posterior electrode position indicated that at posttest, anterior electrode amplitude ($M_{\text{Spanish\_anterior}} = -2.97$ μV, $M_{\text{control\_anterior}} = -3.51$ μV) differed from posterior electrode amplitude ($M_{\text{Spanish\_posterior}} = -2.34$ μV, $M_{\text{control\_posterior}} = -4.11$ μV) for both groups ($p < .05$). There was a significant interaction of group by anterior–posterior electrode site, $F(1,41) = 6.64$, $p < .05$, $\eta^2 = .14$. Post hoc comparisons indicated that the Spanish-learners had larger anterior (-3.30 μV) than posterior (-2.80 μV) amplitudes ($p < .05$), whereas the control group did not display this pattern ($M_{\text{anterior}} = -3.32$ μV, $M_{\text{posterior}} = -3.64$ μV). There was a marginally significant four-way interaction of group by session by condition by anterior–posterior electrode site, $F(1,41) = 3.86$, $p = .056$, $\eta^2 = .09$. Post hoc comparisons examining the effect of anterior–posterior electrode position indicated that the Spanish learner group showed larger go anterior amplitude compared to posterior amplitude at both pretest ($M_{\text{anterior}} = -2.22$ μV, $M_{\text{posterior}} = -1.30$ μV, $p < .01$) and posttest ($M_{\text{anterior}} = -1.08$ μV, $M_{\text{posterior}} = -0.16$ μV, $p < .001$), whereas the control group showed no such differences at pretest ($M_{\text{anterior}} = -0.86$ μV, $M_{\text{posterior}} = -0.66$ μV) or posttest ($M_{\text{anterior}} = -1.40$ μV,

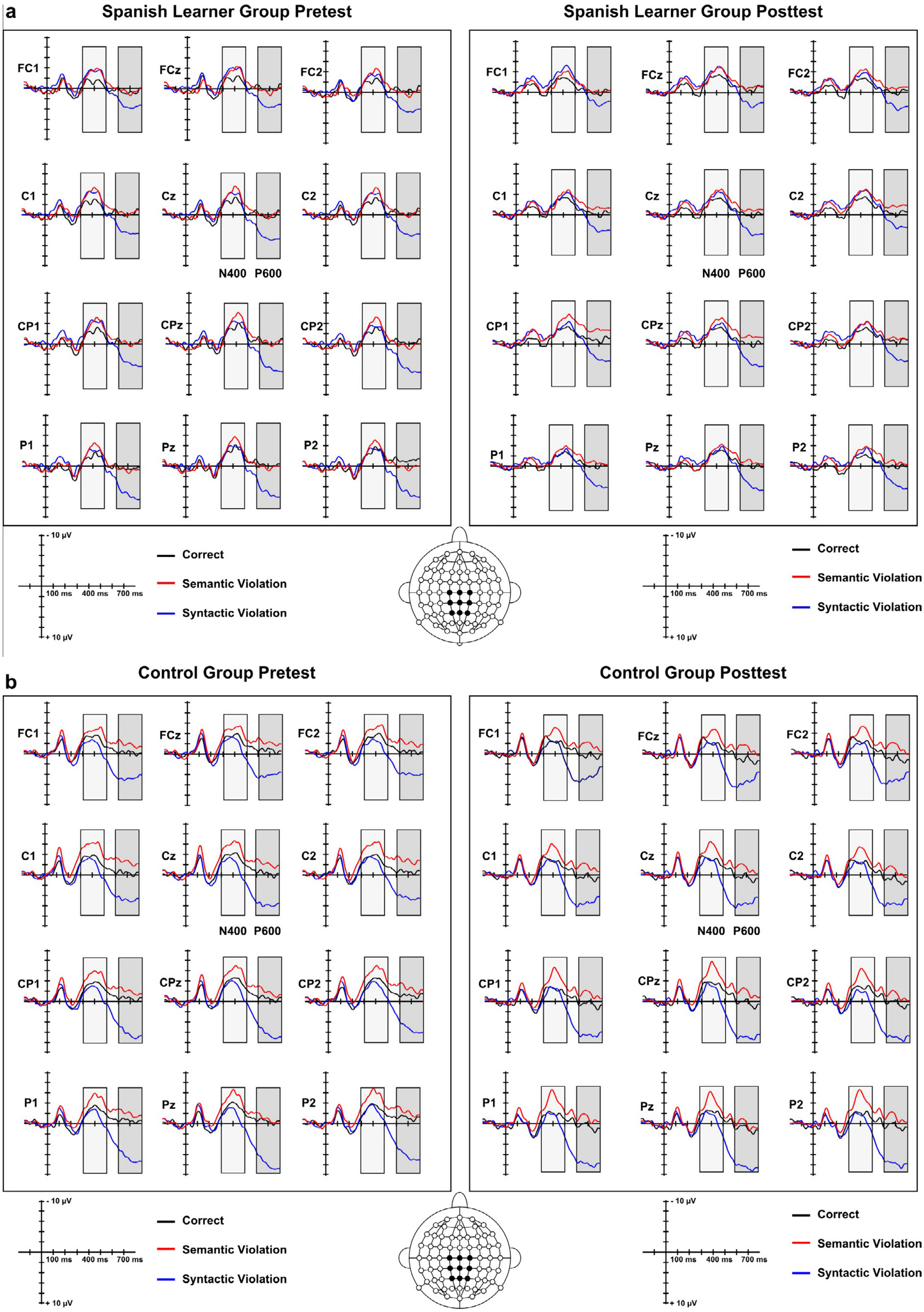

Fig. 1. Event-related potentials (ERPs) recorded during the sentence judgment task at pretest and posttest for (a) the Spanish learner group and (b) the control group.

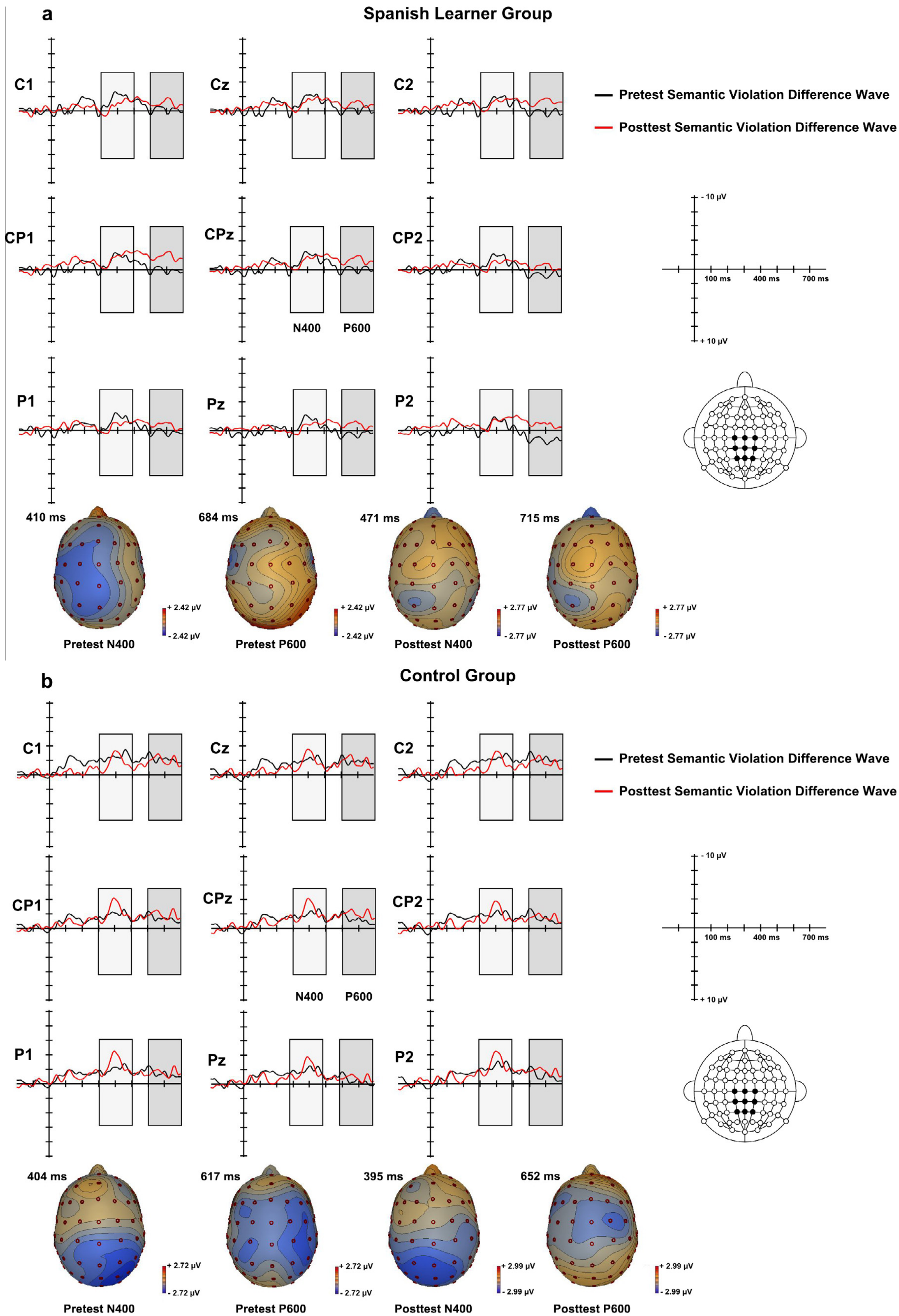

Fig. 2. Waveforms showing the difference between the semantic violation and correct conditions for the sentence judgment task at pretest and posttest for (a) the Spanish learner group and (b) the control group. Topography maps represent the peak activity for each group, at pretest and posttest, during both the N400 and P600 time windows.

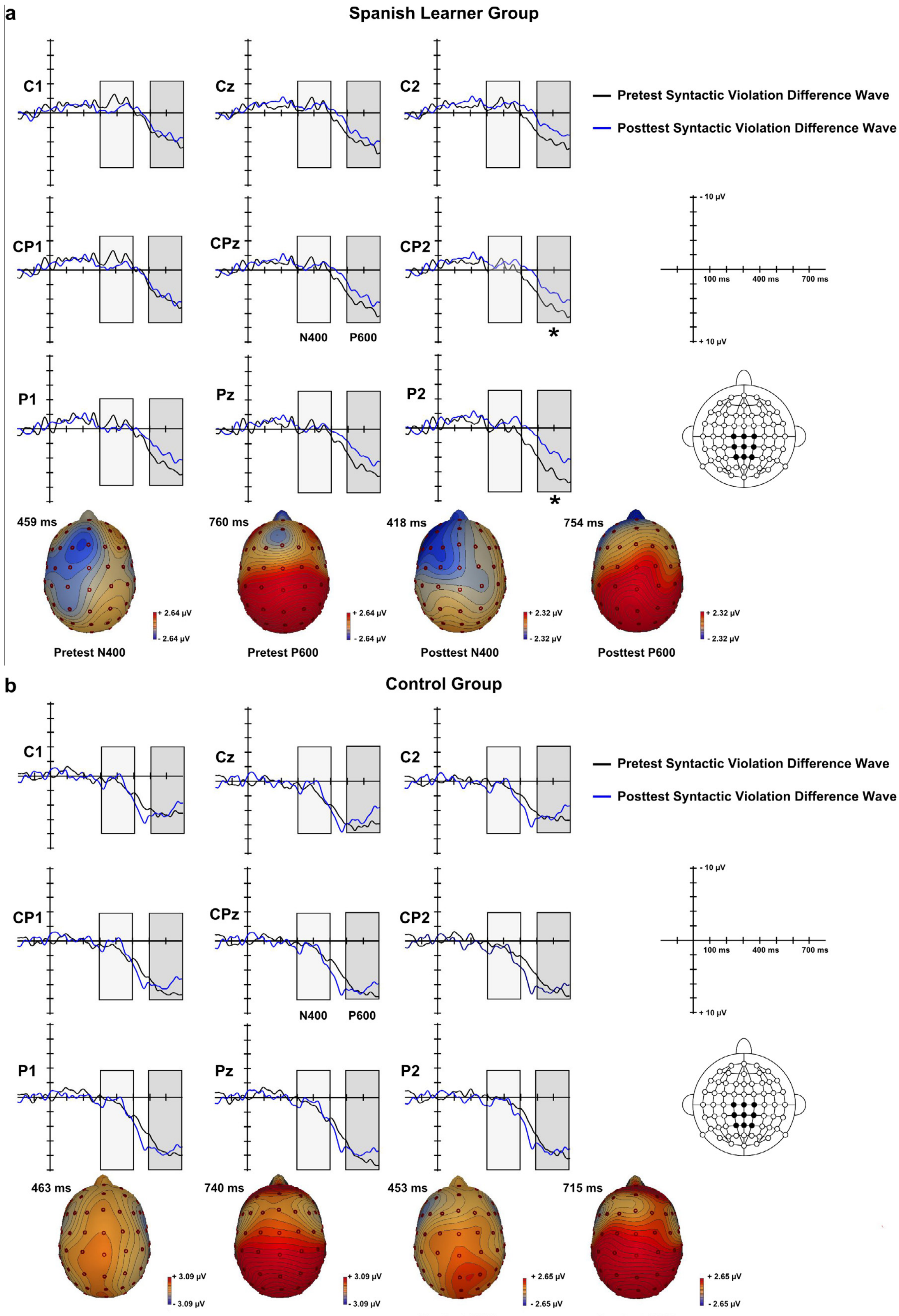

Fig. 3. Waveforms showing the difference between the syntactic violation condition and correct conditions for the sentence judgment task at pretest and posttest for (a) the Spanish learner group and (b) the control group. The asterisk denotes a significant decrease in the mean P600 amplitude difference for the Spanish learner group after L2 training ($p < .05$). Topography maps represent the peak activity for each group, at pretest and posttest, during both the N400 and P600 time windows.

Table 5
Group means (and standard deviations) for behavioral performance on the go–nogo task.

Group	Go–nogo measure	Pretest M (SD)	Posttest M (SD)
Spanish learners $N = 21$	RT (ms)	357 (69)	356 (64)
	Go % Correct	96.3 (10.7)	96.6 (8.0)
	False Alarms/40	5.0 (3.5)	4.0 (3.2)
	Nogo Accuracy Cost[a]	0.07 (0.23)	0.06 (0.12)
	d-prime[a]	3.55 (0.86)	3.76 (0.89)
Controls $N = 22$	RT (ms)	339 (32)	339 (27)
	Go % Correct	99.3 (1.0)	99.1 (1.6)
	False Alarms/40	3.5 (3.3)	4.0 (3.1)
	Nogo Accuracy Cost[a]	0.09 (0.09)	0.09 (0.07)
	d-prime[a]	4.01 (0.61)	3.90 (0.65)

[a] Nogo accuracy cost was calculated as the difference between the proportion of correct go trials minus the proportion of correct nogo trials divided by the proportion of correct go trials (see Lahat, Todd, Mahy, Lau, & Zelazo, 2010). The discriminability index (i.e., *d*-prime or d') is based on signal detection theory and is calculated using the formula $d' = z\,(\text{hits}) - z\,(\text{false alarms})$ (Schulz et al., 2007).

$M_{posterior} = -1.57\ \mu V$). At posttest the control group did show larger nogo posterior amplitude ($-6.66\ \mu V$) compared to anterior amplitude ($-5.63\ \mu V$) ($p < .01$), whereas the Spanish learners did not ($M_{anterior} = -4.86\ \mu V$, $M_{posterior} = -4.51\ \mu V$). There were no other significant main effects or interactions involving session or group, *Fs* < 2.6.

For N2 latency, there was a significant main effect of session, $F(1,41) = 9.34$, $p < .01$, $\eta^2 = .14$, showing overall reduced latency at posttest. There was a marginally significant group effect, $F(1,41) = 3.53$, $p = .067$, $\eta^2 = .08$, with slower latency for the Spanish-learner group (287 ms) than the control group (275 ms). To investigate whether this may be due to group N2 latency differences at pretest, an independent samples *t*-test was performed on the averaged go and nogo latencies across the 6 electrodes (F1, Fz, F2, FC1, FCz, and FC2) at pretest. There was a significant group difference for pretest nogo latency, $t(41) = 7.20$, $p < .001$, with the Spanish-learner group having longer latency ($M_{Spanish} = 306$ ms, $SD = 16.65$) than the control group ($M_{control} = 271$ ms, $SD = 14.89$). There was a significant group by session interaction, $F(1,41) = 14.37$, $p < .001$, $\eta^2 = .22$. Post hoc comparisons indicated that the Spanish-learner group had significantly shorter latency after L2 training ($p < .001$) but there was no latency change for the control group (see Fig. 4 for ERPs). There was a significant interaction of session by condition, $F(1,41) = 4.70$, $p < .05$, $\eta^2 = .10$, which indicated that nogo latency ($M_{pretest} = 289$ ms, $M_{posttest} = 270$ ms) was significantly reduced from pre- to posttest ($p < .001$), whereas go latency did not change across sessions ($M_{pretest} = 287$ ms, $M_{posttest} = 280$ ms). There were no other significant main effects or interactions involving session or group, *Fs* < 3.

3.3.3. P3 time window 350–500 ms

Analyses of P3 peak amplitude revealed a significant interaction of group by session, $F(1,41) = 4.51$, $p < .05$, $\eta^2 = .10$. Post hoc comparisons revealed a significant increase in P3 amplitude after L2 training across all six frontal–central electrodes: F1, Fz, F2, FC1, FCz, and FC2 ($p < .05$), with no such changes observed in the control group (see Fig. 4 for ERPs). Independent samples *t*-tests across all 6 frontal–central electrodes for go, $t(41) = -1.46$, $p = .15$ and nogo P3 amplitude $t(41) = -1.44$, $p = .16$, indicated no significant pretest group differences. There was a main effect of condition, $F(1,41) = 71.11$, $p < .001$, $\eta^2 = .63$, whereby nogo conditions elicited a larger positivity than go conditions, as expected. Finally, there was a session by anterior–posterior electrode interaction, $F(1,41) = 4.81$, $p < .05$, $\eta^2 = .10$, which indicated that at both pretest ($M_{anterior} = 5.44\ \mu V$, $M_{posterior} = 7.85\ \mu V$) and posttest, posterior electrodes showed larger amplitude than anterior electrodes ($M_{anterior} = 7.08\ \mu V$, $M_{posterior} = 8.71\ \mu V$) ($p < .001$). There were no other significant main effects or interactions involving session or group, *Fs* < 2.5.

Although the P3 amplitudes for nogo trials in the pretest were not significantly different across the two groups, they appear to be visually different (see Fig. 4). Therefore, to confirm the effect showing an increase in amplitude for the Spanish group while controlling for pretest amplitude, we selected participants in the two groups with equivalent pretest P3 amplitude to compare their results at posttest. Matched pairs were selected from Spanish and control groups that had a difference in pretest nogo P3 amplitude of <0.80 μV (range: 0.01–0.61), resulting in 13 pairs of participants. An analysis of this subset analysis resulted in a group by session interaction, $F(1,24) = 3.99$, $p = .06$, $\eta^2 = .14$ (see Table 6), and post hoc comparisons showed a significant increase in P3 amplitude after L2 training across all six frontal–central electrodes for the Spanish group: F1, Fz, F2, FC1, FCz, and FC2 ($p < .05$), with no such changes observed in the control group.

There was a significant four-way interaction for group, session, condition, and medial–lateral electrode position for P3 peak latency, $F(2,81.86) = 3.64$, $p < .05$, $\eta^2 = .08$. Post hoc comparisons examining the effect of session indicated that for the Spanish-learner group, nogo latency was significantly shorter after training at left-frontal ($p < .01$), midline ($p < .05$), and right-frontal sites ($p < .05$), with no change in P3 latency measures for the control group (see Fig. 4 for ERPs). A significant three-way interaction of group by session by condition, $F(1,41) = 4.23$, $p < .05$, $\eta^2 = .09$, also indicated that P3 nogo latency was significantly reduced ($p < .05$) after training for the Spanish-learner group, with no change for the control group. An independent samples *t*-test for the averaged nogo P3 latency across all 6 frontal–central electrodes again showed pre-test differences in latency, $t(41) = 2.07$, $p < .05$ between the Spanish learners ($M_{Spanish} = 412$ ms, $SD = 31.63$) and the control group ($M_{control} = 392$, $SD = 32.53$), with the Spanish learners having a longer nogo P3 latency at pretest. There was a significant three-way interaction of group by anterior–posterior by medial–lateral electrode position, $F(2,82) = 4.73$, $p < .05$, $\eta^2 = .10$. Post hoc comparisons examining the effect of medial–lateral electrode position indicated that for the control group, the latency of posterior electrodes FC1 (397 ms) and FC2 (391 ms) were significantly different ($p < .05$), whereas the Spanish-learners did not show this pattern, with FC1 (408 ms) and FC2 (408 ms) being equivalent. There were no other significant main effects or interactions involving session or group, *Fs* < 3.5.

3.3.4. ERP-grade correlations

At posttest, the Spanish-learners were asked to provide a self-report of their progress in the course and report their expected grade. An averaged difference score (posttest P3

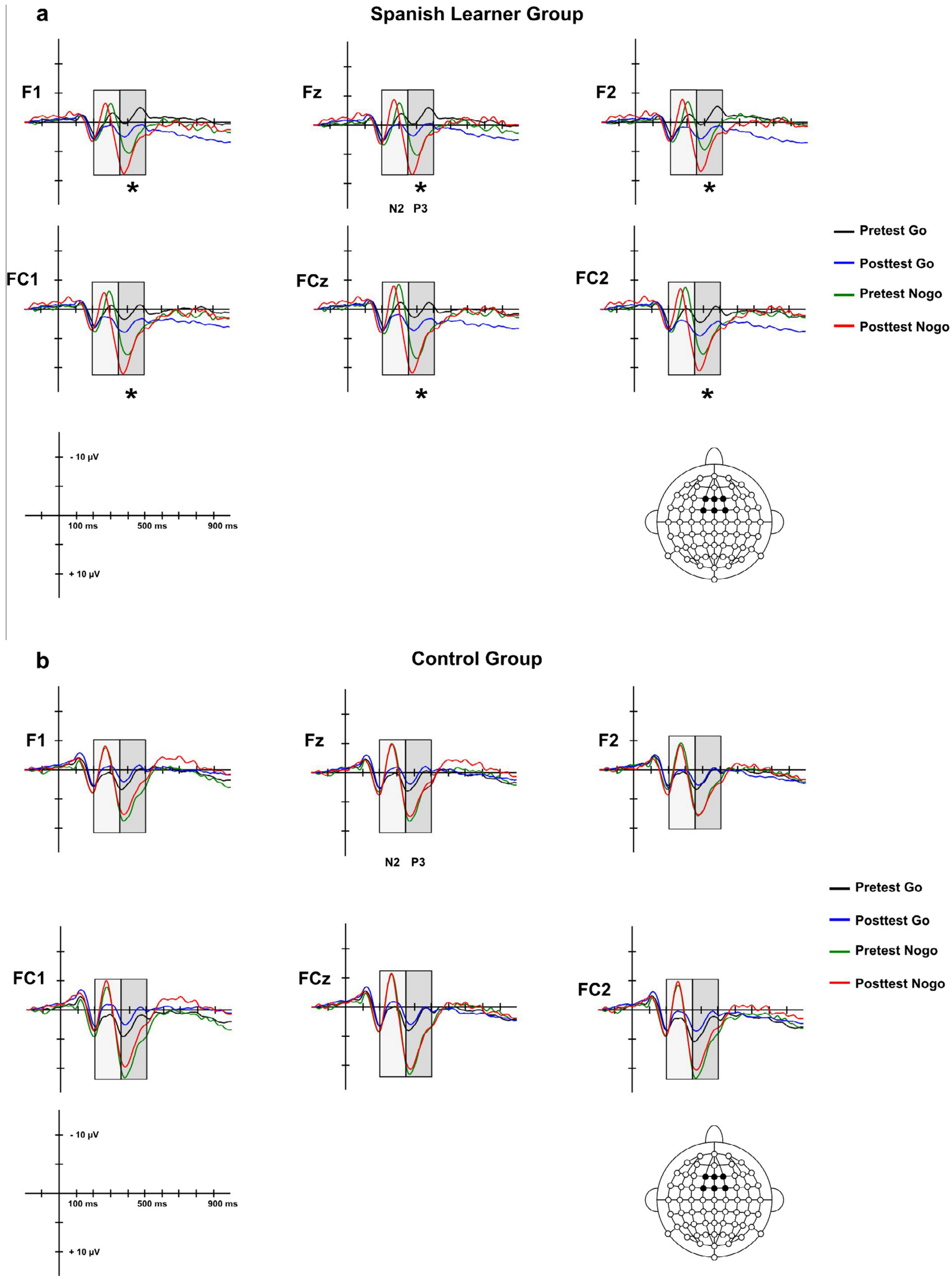

Fig. 4. Event-related potentials (ERPs) elicited during go and nogo trials at pretest and posttest for (a) the Spanish learner group and (b) the control group. Asterisks denote a significant increase of P3 peak amplitude for the Spanish learner group after L2 training ($p < .05$).

amplitude – pretest P3 amplitude) was created separately for go and nogo conditions across the electrodes of interest: F1, Fz, F2, FC1, FCz, FC2. The P3 amplitude differences in each condition were then correlated with the self-reported Spanish grades. Two participants were removed from the analysis because their averaged difference scores were deemed to be outliers (based on >2 *SD*). For the

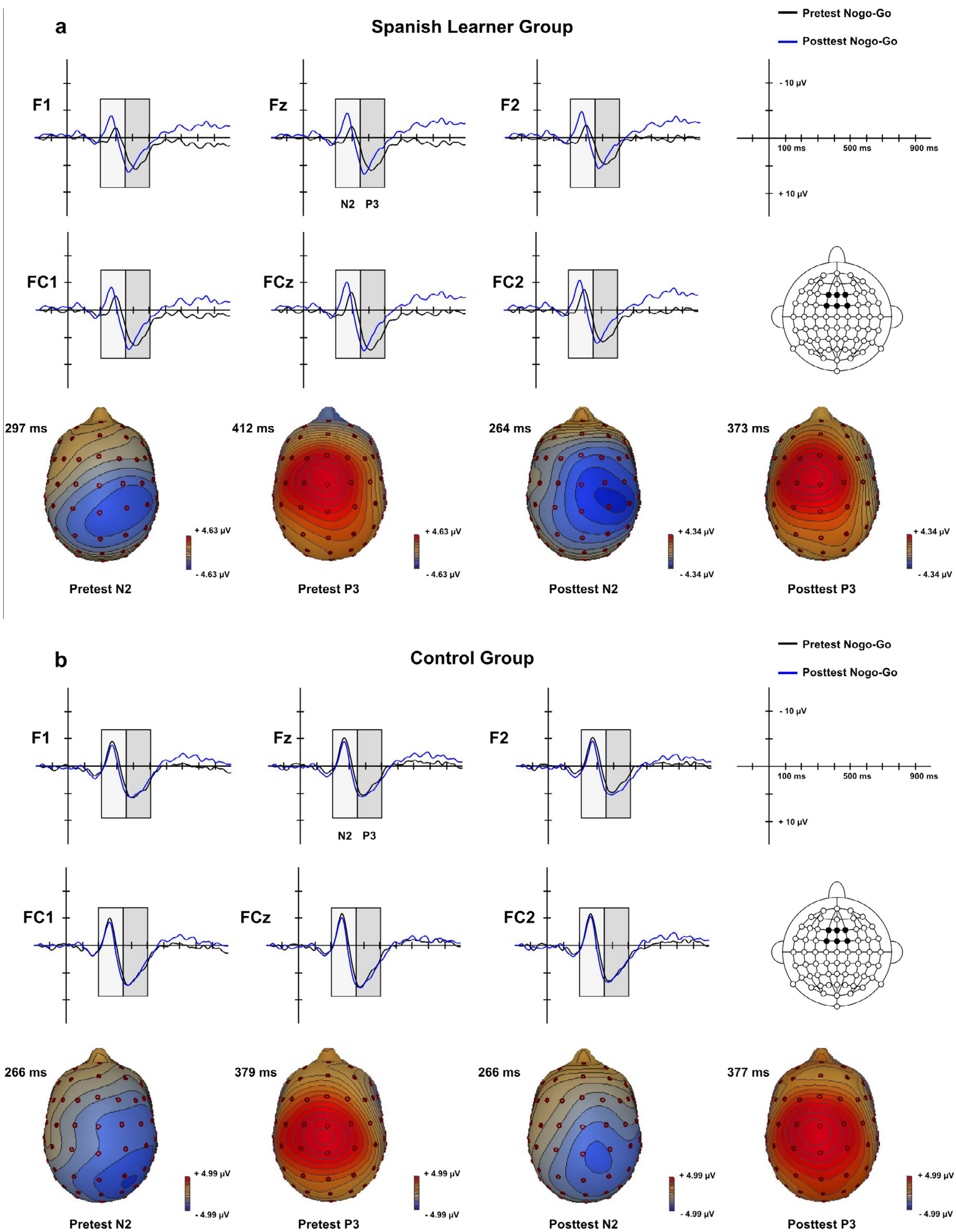

Fig. 5. Waveforms showing the difference between go and nogo trials at pretest and posttest for (a) the Spanish learner group and (b) the control group. Topography maps represent the peak activity for each group, at pretest and posttest, during both the N2 and P3 time windows.

remaining 19 Spanish-learners, there was a significant correlation between the averaged nogo difference and Spanish grades, $r = .46$, $p = .047$ (see Fig. 6); higher self-reported grades were associated with a larger change in P3 peak amplitude after L2 training. There was no correlation between the averaged go difference score and self-reported grades. For the sentence judgment task, there was no

Table 6
Subset analysis means (and standard deviations) for Group matched pretest nogo P3 peak amplitude.

Group	Go–nogo condition	Pretest P3 amplitude *M* (*SD*)	Posttest P3 amplitude *M* (*SD*)
Spanish learners $N = 13$	Go	3.31 (4.56)	6.95 (4.46)
	Nogo	10.92 (6.89)	12.73 (6.60)
Controls $N = 13$	Go	5.31 (4.56)	3.91 (4.46)
	Nogo	10.96 (6.60)	10.95 (6.60)

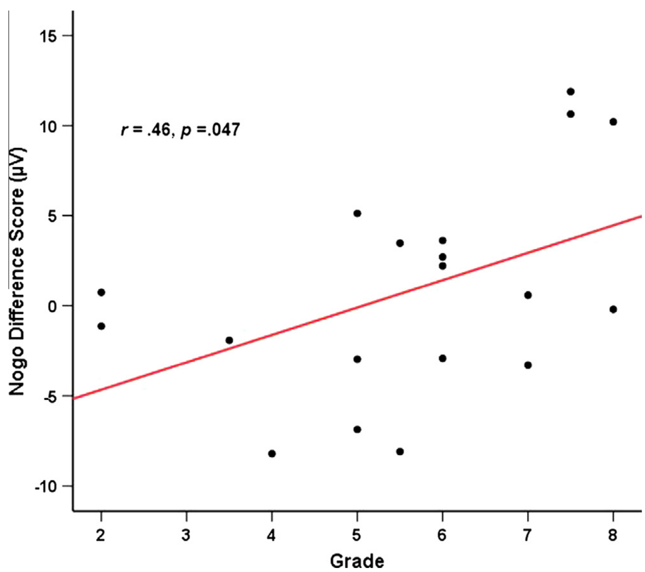

Fig. 6. Pearson correlation between self-reported Spanish grades and the averaged nogo difference score (posttest nogo P3 amplitude – pretest nogo P3 amplitude) across the electrodes of interest: F1, Fz, F2, FC1, FCz, and FC2.

relationship between the averaged P600 difference scores (posttest – pretest) across electrodes CP2 and P2 for the syntactic violation condition and self-reported Spanish grades, although the P600 had shown a reduction in amplitude after L2 training.

4. Discussion

Participants in the Spanish group had spent 6 months in an introductory language class and were far from bilingual, but after this relatively brief L2 training they demonstrated significant modulations in the electrophysiological signal during verbal and nonverbal conflict tasks that were similar to those found for bilinguals. As in the earlier study by McLaughlin et al. (2004), these electrophysiological changes occurred in the absence of behavioral changes. For the control group, there were no electrophysiological changes in any measure in the post-test. Participants in the control group were students studying at the same university and undertaking the same level of course work as the training group, but they were not enrolled in a second-language course. Thus, these participants were not an active control group and did not receive another specific type of training or any incentive for their participation. Although only one group received the relevant intervention, we consider the comparison valid because of the similarity of university environment and background characteristics for the two groups. The present study is the first to extend the effects of early language training on linguistic outcomes (cf., Osterhout et al., 2006) to nonverbal executive control. Our interpretation is that brief L2 training produces cross-domain, neuroplastic effects and advances our understanding of the mechanism behind the bilingual advantage in executive control. Because the L2 learners in the present study are not fluent bilinguals, these results showing emerging changes with language learning is evidence that brain functional changes occur along a continuum of bilingualism experience.

The Spanish learners revealed a significant processing change on the sentence judgment task after training, showing decreased P600 mean amplitude for the syntactic violation condition. Importantly, this reduced P600 effect is similar to the pattern displayed by bilinguals and occurred across two electrode sites (CP2 and P2) within the ROI used in the Moreno et al. (2010) study. Smaller P600 mean amplitude indicates less effortful processing on this task (Hahne & Friederici, 1999, 2001), and observing this effect after only 6 months of L2 training provides crucial information for our primary question of how short term language instruction influences executive control in L2 learners. The larger N400 amplitude reported by Moreno et al. (2010) for bilinguals processing semantically anomalous sentences was not found in the current study. Consistent with the interpretation that the larger N400 relates to the processing and subsequent inhibition of irrelevant linguistic stimuli, monolingual speakers of English as well as our L2 learners may not recruit the same inhibitory processes as bilinguals when performing grammaticality judgments in English. In spite of this difference, the changes in ERP waveforms from the pretest to posttest sessions for the P600 component found in the Spanish learners shows movement towards the pattern found for full bilinguals.

None of the results for the behavioral data showed significant differences between groups. It is important to note, however, that both the judgment and fluency tasks were administered in the participants' L1, making the modification in the ERP response in the judgment task more dramatic. In other words, brief Spanish training modified participants' processing of English. For the judgment task, most scores were close to ceiling, but accuracy in the most difficult condition, the anomalous sentences, improved at posttest for both groups, possibly reflecting a practice effect. Moreno et al. (2010) also found accuracy to be equivalent for monolinguals and bilinguals in this task. Similarly, both groups improved at posttest on both verbal fluency tasks with no difference between groups. Linck, Kroll, and Sunderman (2009) found that students learning Spanish in a study-abroad context showed reduced performance on a verbal fluency task in English, their L1. The present study found no effect of L2 learning on verbal fluency in English, but the level of Spanish obtained by these classroom learners after 6 months in an English-speaking environment was considerably less than that of the study-abroad students after the same time. Nonetheless, the ERP data are consistent with modifications on L1 processing as a consequence of L2 learning (Kroll, Dussias, Bogulski, & Valdes-Kroff, 2012).

The most striking result is the finding that brief L2-training led to processing changes on a nonverbal executive control task. The Spanish learners showed changes in the P3 peak amplitude and in the latency of the P3 and N2 components, whereas no changes were found in the control group. However, independent samples *t*-tests showed pretest group differences in N2 and P3, so this result needs to be interpreted with caution. Some pre-existing group differences in electrophysiology may arise because this study did not involve pretest matching and random assignment to conditions. The change in P3 amplitude was found only for L2-learners and not the control group and supports our argument that short-term L2 training modifies the executive control system. These results are different from those found for full bilinguals in which bilingual participants showed a larger N2 amplitude than monolinguals (Fernandez et al., 2013; Moreno et al., 2014), but such discrepancies are not surprising given the massive differences in experience. It is possible that the P3 component is more amenable to early training effects than the N2 component, which requires more experience to be modified. Importantly, the increase in P3 peak amplitude is consistent with training-induced strengthening of the neural network involved in response inhibition. In line with this interpretation, go–nogo studies using populations with compromised executive control, including children with attention deficit hyperactivity disorder (Fallgatter et al., 2004) and adults with Parkinson's disease (Bokura, Yamaguchi, & Kobayashi, 2005), show smaller nogo P3 amplitude than their respective control groups. Supporting the relation between L2 learning and these outcomes, the increase in nogo P3 peak amplitude was correlated with self-reported Spanish grades.

5. Conclusions

The pretest/posttest design used in the present study provides strong evidence linking the difference between groups to language experience and not to some extraneous factor. At pre-test, participants in both groups were similar on background and task measures, but at post-test, only the Spanish learners showed changes in task processing, and only in the predicted tasks and in the predicted directions. In the majority of research on bilingualism, monolinguals and bilinguals are compared at one point in time, leaving the possibility of an undetected group difference. Thus, the present results give insight into the mechanism underlying advantages in executive control and support the interpretation in those studies that bilingualism is responsible for those differences.

As in previous research, significant effects were found in the neural evidence with no behavioral differences between groups. This pattern has been reported for bilinguals performing metalinguistic (Moreno et al., 2010) and cognitive tasks (Luk, Anderson, Craik, Grady, & Bialystok, 2010) and for L2 learners on a lexical decision task (Osterhout et al., 2006). Not surprisingly, behavioral and neural evidence document different aspects of performance, but importantly, these forms of evidence also appear on a different timetable. As training modified processing of both a verbal and nonverbal executive control task, this validates that even at the early stages of L2 learning executive control is modified in a domain-general manner.

The present results demonstrate early plasticity from experience in adult participants. The remarkable finding is that only 6 months of Spanish instruction has the potential to begin reorganizing the 20-year old monolingual brain to resemble bilingual processing when performing verbal and nonverbal conflict tasks. This is powerful evidence for the effect of experience on functional brain responses before the outcomes of that experience can be detected through behavioral measures, and signifies the benefits of lifelong second-language learning.

Acknowledgments

This research was supported by Grant A2559 from the Natural Sciences and Engineering Research Council, Canada (NSERC) to Ellen Bialystok. We thank Tashua Case, Ashley Chung, Zehra Kamani, Mike Rakoczy, and Kalinka Timmer for their assistance with the study, as well as all of our participants.

Appendix A

Experimental sentences used in the current study, developed by Osterhout and Nicol (1999). For each item the correct, syntactically anomalous, and semantically anomalous version is depicted.

The cats won't *eat/eating/bake* the food that Mary gives them.
The astronomer's argument might *prove/proving/shout* that there are three canals on the moon.
In case of a break-in, the alarm system will *warn/warning/swear* that there is an intruder.
The new species of orchid will *grow/growing/sing* in tropical regions.
This expensive ointment will *cure/curing/loathe* all known forms of skin disease.
This old electric blender doesn't *crush/crushing/own* ice cubes anymore.
This exotic spice may *add/adding/seek* the oriental flavour that John enjoys.
The new fighter plane can *fly/flying/walk* faster than anyone had expected.
The boxes in the attic may still *hold/holding/find* many old photographs and souvenirs.
This test of reasoning might *fail/failing/hate* to discriminate among students.
The puppy seems to *like/liking/call* to sleep a lot during the day.
The cowboy always gives his horse a chance to *drink/drinking/fish* from the stream.
Billy bumped his bicycle, causing it to *fall/falling/sneeze* into the street.
The therapist hoped that the new drug would *calm/calming/clean* the patient who was so anxious.
The new software package will *print/printing/glue* very elaborate pictures.
The publisher hoped that the textbook would *draw/drawing/hear* students with a variety of interests.
William thought that he would *fit/fitting/dig* right in with the crowd at the reception.
Mary knew that the food at the hotel would *cost/costing/fight* too much.
The hikers noticed that the boulder seemed *to rest/resting/live* precariously on the mountain.
At the aquarium, there are otters that *swim/swimming/fly* and do tricks for the crowds.
Every day at three, the newspapers should *land/landing/dance* on the porch out front.
The tree in the backyard can't *sprout/sprouting/sell* new buds in this weather.
Most physicians believe that the new drugs can *prevent/preventing/study* many forms of disease.
The composer agreed that his music should *enchant/enchanting/question* the public.
The repairman thinks that the leaky tub might *bother/bothering/ask* the tenants downstairs.
This rare herb can *heal/healing/count* the pains in your back.
The farmhouse is so old that it *scares/scaring/writes* the neighbours.

The teacher said our report must not *last/lasting/cry* for more than ten minutes.
In the nation's landfills, chemicals of different sorts may *mix/mixing/hope* to create lethal substances.
The simulated accident might *frighten/frightening/ignore* the children enough so that they will wear their bike helmets.
The plumber said that the leaking water might *seep/seeping/speak* out from behind the refrigerator.
The fingerprints on the gun could *prove/proving/judge* that the defendant is innocent.
The beavers in the pond sometimes *chew/chewing/melt* the garden hose.
The fancy French clock doesn't *tell/telling/ask* the time during power failures.
Critics say that the rap songs might *tend/tending/learn* to lead young people astray.
The new brand of toothpaste could *help/helping/beg* to provide protection against disease.
Those small spiders would often *spin/spinning/burn* beautiful webs.
The pacifier we bought in Japan will *soothe/soothing/drop* the cranky baby.
The skyscraper being built by the city would *block/blocking/send* out the sunlight.
These grapevines don't *grow/growing/jog* well in sandy regions.
At the end of the day, the dog always *waits/waiting/peaks* in the driveway.
We hoped that the news of the award would *cheer/cheering/wash* up the depressed student.
The circus elephants get on their hind legs and *stand/standing/chirp*, which impresses the audience.
Simple vegetable oil is used to *fry/frying/plow* the vegetables.
The black widow spider likes to *hide/hiding/sigh* in dark places.
Fountain pens shouldn't be used to *sketch/sketching/dust* since they were designed only for writing.
It was hard to get the infant to *smile/smiling/vote* for the photographer.
The defendant's account of the incident didn't *match/matching/paste* the one given by the codefendant.
The movers didn't think that the piano would *weigh/weighing/cough* as much as it did.
Susan was worried that her kitten would *scratch/scratching/lift* the young child.
The peregrine falcon chicks always *chirp/chirping/staple* until the father brings food.
The general admits that the missile might *explode/exploding/call* before leaving the area.
The assistant was told that the alibi would *prevent/preventing/consider* an indictment.
My grandfather's habit of chasing geese might *shock/shocking/stab* you, but he's quite normal otherwise.
People hope that the sculpture will *inspire/inspiring/invent* new forms of artistic expression.
So many bugs live in the garden, they must *eat/eating/buy* a head of lettuce every minute.
The booklet says that the contraceptive will *fail/failing/complain* if used too sparingly.
My brother bet that this spider could *climb/climbing/type* faster than you could.
The baby's teeth are so sore that the pacifier might *hurt/hurting/cheat* her too much.
The supervisor's report found that the factory should *train/training/hug* workers more thoroughly.
One kangaroo at the San Diego Zoo would sometimes *sit/sitting/write* all day.
The portrait of Uncle Henry doesn't *look/looking/sing* like him.
The new heater in the maid's room should *dry/drying/find* the laundry.
The strawberry beds might *tempt/tempting/sneeze* rabbits and other animals.
The colours in the sweater should not *fade/fading/walk* when the sweater is washed.
The new chemical additive may *tend/tending/desire* to lower the freezing point of water.
The sea lions can *bask/basking/edit* on the beach all day.
The security camera at the bank will now *take/taking/trip* photographs of everyone.
The bull that escaped could *smash/smashing/send* the wooden fence around the meadow.
The award winning play will *run/running/leap* for several more months.
The couple's newborn baby *sneezes/sneezing/types* so much that they took her to the doctor.
The hiker used his last match to *start/starting/tie* the fire.
Where the road *forks/forking/believes*, we couldn't figure out which way to go.
Bob's rubber raft *hit/hitting/loves* a rock, which punctured it.
After Jane's accident, she found it difficult to *drive/driving/boil* for several months.
The pet cats will soon *eat/eating/describe* their evening meal.
The raging bull will *charge/charging/whistle* at the man.
The new romance novel should *sell/selling/leak* in every store this year.
Alison used a hammer to *break/breaking/kiss* the small lock.
Betsy went out to the orchard to *pick/picking/melt* apples for a pie.
The powerful magnet will *pull/pulling/learn* defective parts from the assembly line.
The lever on the basement wall does not *shut/shutting/lift* off the power supply.
The new detergent is supposed to *clean/cleaning/burn* the floors with ease.
The newly planted grass will *grow/growing/swim* quite a bit during the next year.
The hidden door will *open/opening/cook* when the secret code is spoken.
A new computer will *last/lasting/paint* for many years.
The local beers in Seattle will *satisfy/satisfying/trip* every beer drinker.
The red ants in Arizona will *bite/biting/wash* you if you are not careful.
The new crop of corn should *feed/feeding/scrape* everyone in the state.
The noisy ducks will soon *fly/flying/skip* away from the lake.

References

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes, 23*, 557–582. http://dx.doi.org/10.1080/01690960801920602.

Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology, 2*(324), 1–10. http://dx.doi.org/10.3389/fpsyg.2011.00324.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 57*, 489–510. <http://www.jstor.org/stable/1130604>.

Bialystok, E. (2001). How does experience change cognition? Evaluating the evidence. *British Journal of Psychology, 102*, 303–305. http://dx.doi.org/10.1111/j.2044-8295.2011.02008.x.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*, 89–129. http://dx.doi.org/10.1177/1529100610387084.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older adult bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 859–873. http://dx.doi.org/10.1037/0278-7393.34.4.859.

Bokura, H., Yamaguchi, S., & Kobayashi, S. (2005). Event-related potentials for response inhibition in Parkinson's disease. *Neuropsychologia, 43*, 967–975. http://dx.doi.org/10.1016/j.neuropsychologia.2004.08.010.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test-III*. Circle Pines, MN: American Guidance Service.

Fallgatter, A. J., Ehlis, A. C., Seifert, J., Strik, W. K., Scheuerpflug, P., Zillessen, K. E., et al. (2004). Altered response control and anterior cingulate function in attention-deficit/hyperactivity boys. *Clinical Neuropsychology, 115*, 973–981. http://dx.doi.org/10.1016/j.clinph.2003.11.036.

Fernandez, M., Tartar, J. L., Padron, D., & Acosta, J. (2013). Neurophysiological marker of inhibition distinguishes language groups on a non-linguistic executive function test. *Brain and Cognition, 83*, 330–336.

Gollan, T. H., & Kroll, J. F. (2001). The cognitive neuropsychology of bilingualism. In B. Rapp (Ed.), *What deficits reveal about the human mind/brain: A handbook of cognitive neuropsychology* (pp. 321–345). Philadelphia: Psychology Press.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology, 16*(4), 562–576. http://dx.doi.org/10.1037/0894-4105.16.4.562.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81. http://dx.doi.org/10.1017/S1366728998000133.

Grogan, A., Green, D. W., Ali, N., Crinion, J. T., & Price, C. (2009). Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cerebral Cortex, 19*, 2690–2698. http://dx.doi.org/10.1093/cercor/bhp023.

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. http://dx.doi.org/10.1162/089892999563328.

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. http://dx.doi.org/10.1017/S1366728901000232.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman brief intelligence test—Manual*. Circle Pines, MN: American Guidance Service.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes-Kroff, J. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.). *The psychology of learning and motivation* (Vol. 56, pp. 229–262). San Diego: Academic Press. http://dx.doi.org/10.1016/B978-0-12-394393-4.00007-8.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lahat, A., Todd, R. M., Mahy, C. E. V., Lau, K., & Zelazo, P. D. (2010). Neurophysiological correlates of executive function: A comparison of European-Canadian and Chinese-Canadian 5-year-old children. *Frontiers in Human Neuroscience, 3*, 1–10. http://dx.doi.org/10.3389/neuro.09.072.2009.

Lavric, A., Pizzagalli, D. A., & Forstmeier, S. (2004). When 'go' and 'nogo' are equally frequent: ERP components and cortical tomography. *European Journal of Neuroscience, 20*, 2483–2488. http://dx.doi.org/10.1111/j.1460-9568.2004.03683.x.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language. *Psychological Science, 20*, 1507–1515. http://dx.doi.org/10.1111/j.1467-9280.2009.02480.x.

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition, 74*, 347–357.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*, 605–621.

Luk, G., Bialystok, E., Craik, F. I. M., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of Neuroscience, 31*, 16808–16813. http://dx.doi.org/10.1523/JNEUROSCI.4563-11.2011.

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. L. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis on functional neuroimaging studies. *Language and Cognitive Processes, 27*, 1479–1488. http://dx.doi.org/10.1080/01690965.2011.613209.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*, 29–41. http://dx.doi.org/10.1016/j.cognition.2009.08.014.

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., et al. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage, 63*, 240–244. http://dx.doi.org/10.1016/j.neuroimage.2012.06.043.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language learning: Minimal instruction produces rapid change. *Nature Neuroscience, 7*, 703–704. http://dx.doi.org/10.1038/nn1264.

Mohades, S. G., Struys, E., Van Schuerbeek, P., Mondt, K., Van De Craen, P., & Luypaert, R. (2012). DTI reveals structural differences in white matter tracts between bilingual and monolingual children. *Brain Research, 1435*, 72–80. http://dx.doi.org/10.1016/j.brainres.2011.12.005.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science, 22*, 1425–1433. http://dx.doi.org/10.1177/0956797611416999.

Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of Neurolinguistics, 23*, 564–579. http://dx.doi.org/10.1016/j.jneuroling.2010.05.002.

Moreno, S., Wodniecka, Z., Tays, W., Alain, C., & Bialystok, E. (2014). Inhibitory control in bilinguals and musicians: ERP evidence for experience-specific effects. *PLoS One, 9*(4), e94169. http://dx.doi.org/10.1371/journal.pone.0094169.

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A means for exploring the neurocognition of second language processing. *Language Learning, 56*, 199–230. http://dx.doi.org/10.1111/j.1467-9922.2006.00361.x.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain response to syntactic and semantic anomalies. *Language and Cognitive Processes, 14*, 283–317. http://dx.doi.org/10.1080/016909699386310.

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography and Clinical Neuropsychology, 60*, 423–434. http://dx.doi.org/10.1016/0013-4694(85)91017-X.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, J. R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*, 127–152. http://dx.doi.org/10.1111/1469-8986.3720127.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., et al. . *Applied Neuropsychology, 7*, 17–24. http://dx.doi.org/10.1207/S15324826AN0701_3.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences, 102*, 14931–14936. http://dx.doi.org/10.1073/pnas.0506897102.

Schlegel, A. A., Rudelson, J. J., & Tse, P. U. (2012). White matter structure changes as adults learn a second language. *Journal of Cognitive Neuroscience, 24*, 1664–1670. http://dx.doi.org/10.1162/jocn_a_00240.

Schulz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., & Halperin, J. M. (2007). Does the emotional Go–nogo task really measure behavioral inhibition? Convergence with measures on a non-emotional analog. *Archives of Clinical Neuropsychology, 22*, 151–160. http://dx.doi.org/10.1016/j.acn.2006.12.001.

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition, 16*, 367–382. http://dx.doi.org/10.1017/S1366728912000302.

Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience and Biobehavioral Reviews, 33*, 1168–1177. http://dx.doi.org/10.1016/j.neubiorev.2009.03.003.