Brain Research 1646 (2016) 44–52

Contents lists available at ScienceDirect

Brain Research

journal homepage: www.elsevier.com/locate/brainres

Research report

Language effects in second-language learners: A longitudinal electrophysiological study of spanish classroom learning

Laura Soskey [a], Phillip J. Holcomb [b,c], Katherine J. Midgley [b,c,*]

[a] Department of Psychology, University of Rochester, Rochester, NY, USA
[b] Department of Psychology, San Diego State University, San Diego, CA, USA
[c] Department of Psychology, Tufts University, Medford, MA, USA

ARTICLE INFO

Article history:
Received 30 October 2015
Received in revised form
18 March 2016
Accepted 14 May 2016
Available online 24 May 2016

Keywords:
Language learning
Spanish learning
ERPs
N400

ABSTRACT

How do the neural mechanisms involved in word recognition evolve over the course of word learning in adult learners of a new second language? The current study sought to closely track language effects, which are differences in electrophysiological indices of word processing between one's native and second languages, in beginning university learners over the course of a single semester of learning. Monolingual L1 English-speakers enrolled in introductory Spanish were first trained on a list of 228 Spanish words chosen from the vocabulary to be learned in class. Behavioral data from the training session and the following experimental sessions spaced over the course of the semester showed expected learning effects. In the three laboratory sessions participants read words in three lists (English, Spanish and mixed) while performing a go/no-go lexical decision task in which event-related potentials (ERPs) were recorded. As observed in previous studies there were ERP language effects with larger N400s to native than second language words. Importantly, this difference declined over the course of L2 learning with N400 amplitude increasing for new second language words. These results suggest that even over a single semester of learning that new second language words are rapidly incorporated into the word recognition system and begin to take on lexical and semantic properties similar to native language words. Moreover, the results suggest that electrophysiological measures can be used as sensitive measures for tracking the acquisition of new linguistic knowledge.

1. Introduction

The question of how the mature brain acquires a second language, particularly later in life after a first language has become well established, has received increasing attention in recent years (e.g., Doughty, 2003; Kroll and Tokowicz, 2005). Second language learning, especially in the school setting, presents certain challenges to the language learner and there is evidence that such learning is both more difficult and less likely to result in native-like competence in both language production and comprehension (e.g., Johnson and Newport, 1989; Weber-Fox and Neville, 1996). Here we sought to determine whether previously reported very early changes in the neural response to words in a newly learned language (L2 – McLaughlin et al., 2004) differ from similar neural responses to words in a well-established native (L1) language.

Fundamental to second language learning is acquisition of a new vocabulary (e.g., Nation, 1993). However, most of the work on L2 word processing has focused on advanced classroom learners and behavioral measures of language processing (e.g., Domínguez et al., 2013; Elgort and Piasecki, 2014; Veivo and Jarvikivi, 2013). Comparatively less is known about the timing and sequence of the neuro-cognitive mechanisms that are used by L2 learners during the acquisition of words in their new second language (L2).

Recently a small group of studies has demonstrated the utility of the event-related potential (ERP) technique in examining neuro-cognitive changes underlying word learning in beginning L2 users. In a seminal study, McLaughlin et al. (2004) investigated L2 acquisition at the earliest stages of L2 learning using a paradigm developed by Chwilla et al. (1995). Chwilla et al. had observed that the N400 for the target words in a semantic priming paradigm was largest when the target letter string was a pseudoword, was of intermediate amplitude for target words that followed unrelated primes, and the least negative when a target word was preceded by a semantically related word. In the procedurally similar Mclaughlin et al. study participants were university students who had never studied the target second language (French) prior to

* Correspondence to: SDSU Research Foundation, 6505 Alvarado Rd, Suite 203, San Diego, CA 92180, USA.
E-mail address: kmidgley@sdsu.edu (K.J. Midgley).

http://dx.doi.org/10.1016/j.brainres.2016.05.028
0006-8993/

enrollment in an introductory university course in that language. ERPs and behavioral data were collected during three sessions spaced over the first semester of French learning. McLaughlin et al. found that even in the first session, after an average of 14 h of classroom instruction, learners' ERPs demonstrated changes in that pseudoword target items elicited a larger N400 than unrelated or related prime-target pairs. This is very similar to what Chwilla et al. had previously reported for monolingual language processing. McLaughlin et al. also found that these priming effects became more pronounced across sessions (session 2=63 h of instruction, session 3=138 h), suggesting that the effects were a direct result of increasing L2 proficiency. Importantly, ERP data from the first session suggested significant cognitive learning had occurred before behavioral data showed any difference between the experimental group and the control group. The results from this study suggest that the neural consequences of L2 learning can be measured throughout the earliest phases of word acquisition and that the N400 component is a good measure for assessing such learning.

Another study that addresses early L2 learning in university students was conducted by Stein et al. (2006). Participants were English-speaking students in a German language-immersion exchange program in Switzerland. Experimental sessions were conducted before German (L2) learning occurred, and then about 5 months later after intense L2 instruction. ERPs were collected to individual word items from English (L1), German (L2), and Romansh (an unknown language) on both days. Although Stein et al. did not report amplitude differences between L1 and L2 in the N400 during either session, they did find that the duration of the L2 N400 waveform was reduced at the second (post-learning) session.

Finally, more recently Yum et al. (2014) examined the very initial phases of orthographic and semantic acquisition in monolingual native English speakers exposed to Chinese words under controlled laboratory conditions over 10 sessions. Behavioral performance on tests of L2 word learning showed steady improvement over sessions, and these data were used to separate the learners into those who learned quickly and those that took longer to learn new L2 words. ERPs to new L2 words in the two groups revealed qualitatively distinct learning patterns. While fast learners showed an increase in anterior N400 amplitude with training, slow learners showed increasingly more posteriorly distributed positive-going ERPs with learning. The authors suggested that these different patterns might reflect underlying learning strategies used by the two groups with slow learners relying more on holistic pattern processing of Chinese characters and fast learners using a strategy whereby they kept track of a limited set of character parts and their relationships (much like what happens in alphabetic languages).

A parallel (and larger) literature has looked at ERP changes as a function word learning in general – i.e., without telling participants the newly learned items are in a new L2 and/or using items that plausibly could be new L1 words (e.g., Bakker et al., 2015; Borovsky et al., 2010, 2012; Mestres-Miss,é et al., 2007; Perfetti et al., 2005). One prominent theme in this literature is whether there is a need for a consolidation period after new words are learned in terms of their incorporation into the lexical system. Several studies have suggested that new L1 words can be lexicalized very quickly after just a few learning encounters and that word-like N400 effects can be obtained from these items (e.g., Borovsky et al., 2010; Mestres-Miss,é et al., 2007). However, more recently Bakker et al. (2015) have reported evidence that the nature of the ERP effects obtained immediately after learning are somewhat different from those where a longer period of consolidation (24 h in their case) had taken place. Only after consolidation to new word ERPs show effects of lexicalization. In the current study we took advantage of this finding and used a relatively long consolidation period between testing sessions (see below).

1.1. The present study

McLaughlin et al. (2004) and Yum et al. (2014) both demonstrated the rapidity of changes in early L2 acquisition in the form of increases in N400 amplitude as a function of learning. One inference that can be drawn from these findings is that as learners acquire words in their new L2 these new lexical items rapidly start to be processed in a similar fashion to L1 items. In this way L2 word acquisition might be viewed as being similar to learning new words in L1. However, McLaughlin et al. and Yum et al. did not directly compare processing of new L2 items to established L1 items, so it is not clear from their studies how the changes in the N400 they reported are related to those generated by L1 items. This is important because a recent study by Midgley et al. (2009) demonstrated that comparisons between L1 and L2 words in participants learning a new L2 actually show an attenuation of the N400 in L2 compared to L1 and this L1–L2 difference (what they called "language effects") was smaller in more proficient bilinguals. This suggests that the N400 in addition to being sensitive to early learning might also be a more stable marker of language proficiency. However, Midgley et al. did not do a fine grained comparison of L2 learners over the early course of learning and their sample of learners was on average more advanced than those in McLaughlin et al. or Yum et al. So, it is possible that the changes reported by these latter authors actually reflect a short term modulation of the N400 to new lexical items, and that after more extensive L2 learning takes place, the N400 to L2 words declines somewhat. This could happen because initially L2 words might be processed via strong lexical links to their L1 translations, which in turn are used to access meaning. Large N400s would then be due to the strong mapping between the L1 translations and meaning (e.g., as in the RHM model of Kroll and Stewart, 1994). According to this view N400s to new L2 items should not necessarily differ from those recorded to their L1 translation equivalents, although there might be a delay in the time course of the N400 due to the extra step of L2 to L1 lexical mapping. Alternatively, N400s to L2 items during very early learning (as in McLaughlin et al. and Yum et al.) might be on the continuum reported by Midgley et al. – that is, they might be smaller than L1 N400s, but with increasing proficiency they would continue to grow in amplitude. The present study sought to determine (a) whether, with learning, new L2 items show increasing N400 amplitude; and (b) if N400 amplitude to L2 items during learning is smaller than N400 amplitude to equivalent L1 items.

In order to longitudinally examine changes to the N400 to new L2 words, the current study recruited native English-speakers enrolled in an introductory university Spanish course. All participants were initially naïve L2 learners. The stimuli for the experiment were taken from the course curriculum such that data would reflect learning due to formal classroom instruction. ERPs were collected from all participants to both L1 and L2 words in a go/no-go lexical decision paradigm where go events were pseudowords and no-go events were real words. Learners were tested at three time points over the course of the four-month academic semester. We made the following predictions. First, we predicted that we would replicate findings from Midgley et al. (2009), such that L1 items would elicit larger N400 waveforms than L2 words. Second, because Midgley et al. (2009) compared intermediate and proficient bilinguals and showed that the language effect at the N400 was reduced in the more proficient bilinguals, we predicted that as in McLaughlin et al. and Yum et al., as Spanish proficiency increased across sessions, the N400 to new L2 items would increase resulting in a decline in language effects.

2. Results

2.1. Behavioral data

The post-test scores for experimental sessions 1, 2, and 3 were 62.8% (SD=12.1%), 76.8% (SD=9.7%), and 78.7% (SD=9.3), respectively (see Table 1) which represented a significant improvement in Spanish word recognition across sessions ($F(2,22)=68.67$, $p<0.0001$).

Button-press responses to pseudoword items in the go/no-go paradigm (see Table 2) indicated that participants were able to perform the experimental task. They also demonstrated an effect of LEARNING SESSION with a significant reduction in the real word false alarm rate to new Spanish words across sessions ($F(2,22)=22.26$, $p=.0004$ – see Table 2).

2.2. Visual inspection of ERPs

Plotted in Fig. 1(a) are ERPs at 12 representative electrode sites across the scalp time-locked to all Spanish critical items as a function LEARNING SESSION. Plotted in 1b are the same data at the Pz and FPz sites but also including the ERPs to all English (L1) words in the third recording session. Plotted in Fig. 1(c) are both the L1 and L2 ERPs at FPz and Pz as a function of LEARNING SESSION. As can be seen in all of these plots early on in the waveforms a similar pattern of activity is apparent and includes a central-anterior negativity peaking at about 100 ms (N1) followed by a large positivity peaking between 200 and 300 ms across the scalp (P2). Following the P2 there are clear differences in Fig. 1 (a) for the Spanish items as a function of LEARNING SESSION between 300 and 600 ms. Fig. 1(c) suggests there are smaller differences of LEARNING SESSION for English items in this same time epoch, although there do appear to be later differences (600–800 ms) for these items. Fig. 1(b) shows that the effect of LEARNING SESSION in the 300–600 ms range seen for Spanish items extends into the difference between Spanish and English words.

Fig. 2 shows voltage maps depicting the differences between L1 and L2 (L2 is subtracted from L1) for each LEARNING SESSION. It is clear that in Session 1 there is a robust and widespread difference between L1 and L2 starting at 350 ms, which is depicted by the deep blue color covering the entire scalp. In Session 1 these differences are localized to the central and posterior region of the scalp, but persist until the end of the recording epoch. The voltage maps at 450 ms best visualize the language effect on the N400 component. As the sessions progress, this difference at 450 ms remains robust, but is more focused in the posterior region. It can also be observed that while posterior differences between L1 and L2 persist for the entire recording epoch in Session 1, in Sessions 2 and 3 this posterior difference is attenuated earlier in the recording epoch.

2.3. ERP data analysis

2.3.1. 200–350 ms epoch

In this early epoch there was a main effect of LANGUAGE ($F(1,11)=34.20$, $p<0.0001$) whereby L1 items were more negative-going (mean=2.12 μV) than L2 items (mean=3.40 μV – see Fig. 1 (C)). There was no effects involving the LEARNING SESSION variable.

Table 1
Mean(SD) percentage of correctly translated Spanish items after each ERP session.

	ERP 1 Post-test	ERP 2 Post-test	ERP 3 Post-test
Score	62.77*(12.1)*	76.83*(9.7)*	78.71*(9.3)*

Table 2
Mean (SD) percentage of correct button presses (Hits) to pseudoword items and percentage of false alarms (FAs) to real word items by list type in each session.

	English List		Spanish List	
	Hits	FAs	Hits	FAs
Session 1	95.61 *(6.28)*	0.58 *(1.14)*	80.70 *(16.94)*	12.13 *(7.75)*
Session 2	98.25 *(4.67)*	0.58 *(0.86)*	75.88 *(19.75)*	2.92 *(2.63)*
Session 3	96.05 *(6.78)*	1.02 *(0.90)*	73.68 *(15.28)*	2.78 *(2.04)*

2.3.2. 350–600 ms epoch

This is the epoch, which encompasses the typical time range of the N400, there were again differences between the two languages, (LANGUAGE, $F(1,11)=5.07$, $p=0.046$) which now differed as a function of scalp site (LANGUAGE × DISTRIBUTION, $F(11,121)=6.33$, $p=0.0032$). This interaction reflects the comparatively larger posterior negativity for L1 than L2 words as can be seen in Fig. 1 (b). There was also a main effect of LEARNING SESSION ($F(2,22)=6.09$, $p=0.0086$) as well as a significant interaction between LEARNING SESSION, LANGUAGE and DISTRIBUTION ($F(22,242)=2.29$, $p=0.046$). A visual inspection of the FPz and Pz electrodes in Fig. 1(c) suggests that the ERPs in this epoch differ for L1 and L2 as a function of learning. Follow-up analyses decomposing the above three-way interaction confirmed that the effect of LEARNING SESSION was only significant in L2 ($F(2,22)=4.01$, $p=.038$) and not in L1 ($F(2,22)=1.3$, $p=.29$).

2.3.3. 600–800 ms epoch

In the post-N400 epoch we again observed robust LANGUAGE effects and LEARNING SESSION effects both of which differed as a function of scalp distribution (LANGUAGE × DISTRIBUTION, $F(11,121)=12.15$, $p<.0001$; LEARNING SESSION × DISTRIBUTION, $F(22,242)=2.42$, $p=.033$). L1 words remained more negative then L2 words at posterior sites, while L2 words tended to be somewhat more negative-going than L1 words at anterior sites (see Fig. 1(c)). There were, however, no significant interactions between LEARNING SESSION and LANGUAGE (all Fs < 1.6).

2.4. Time-course analysis

A time-course analysis was performed to provide a more fine-grained look at language effects across learning sessions. The results of these analyses are reported in Table 3 and can be visualized in the voltage maps in Fig. 2. As can be seen there were robust, widely distributed language effects (L1 more negative-going than L2) in Session 1 across the latency ranges starting before the N400 and extending after. However, language effects become more localized in time and more posterior in distribution in the later sessions. And at anterior sites there was even a flip-over in the polarity of the language effect in later epochs (after the traditional N400) with L2 items producing larger negativities than L1 items (see the red anterior area in Session 3 voltage maps in Fig. 2).

3. Discussion

In previous research it has been shown that in mature L2 learners of an intermediate level, L2 items elicit an attenuated N400 component compared to L1 words (Midgley et al., 2009). There is also evidence that the N400 component is sensitive to L2

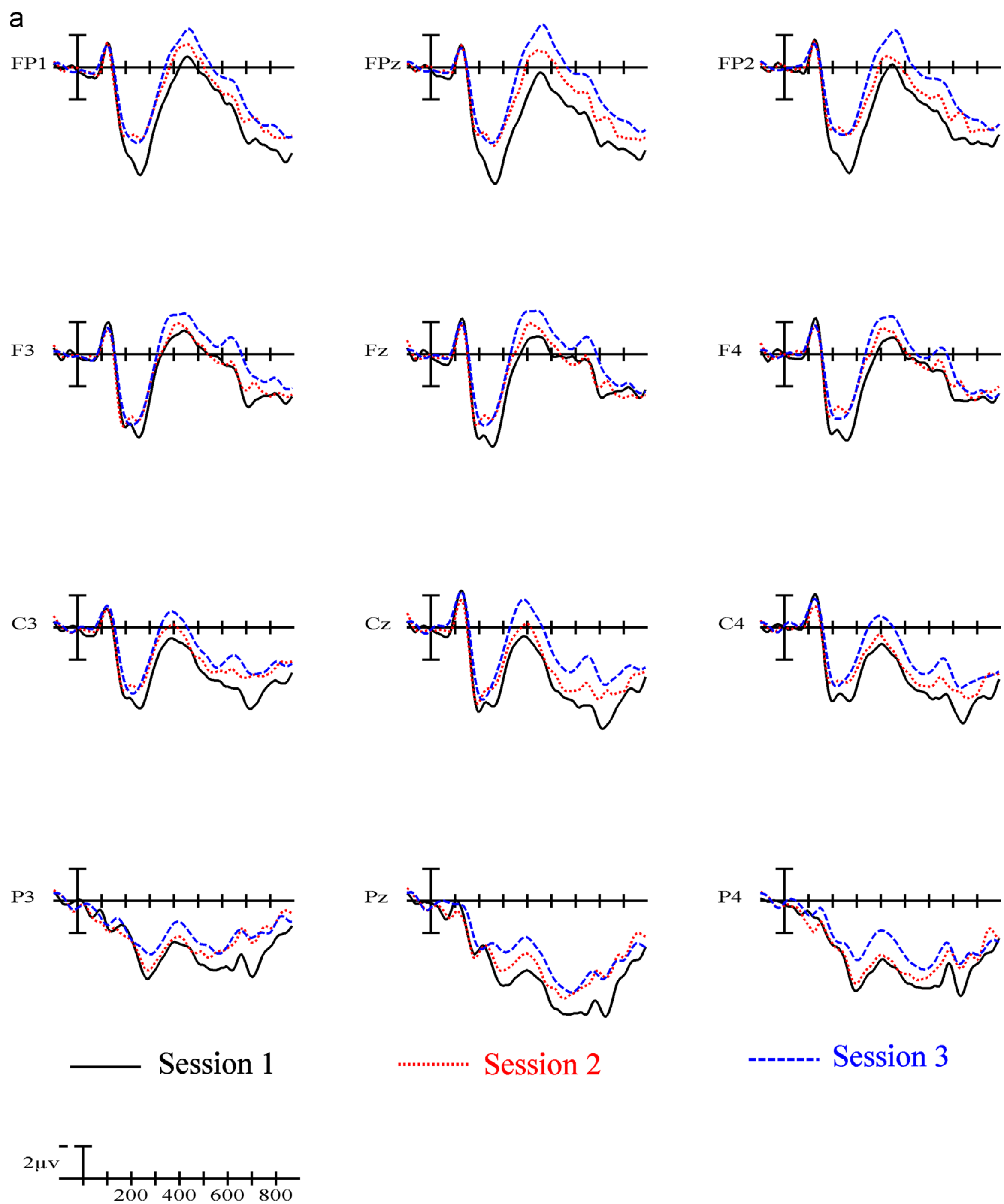

Fig. 1. (a) ERPs to L2 items at 12 electrode sites contrasting the three LEARNING SESSIONS. (b) The same as 1a, but showing only the FPz and Pz sites and including the third session of L1 words. (c) ERPs to L2 (right) and L1 (left) contrasting the three LEARNING SESSIONS at FPz and Pz.

lexicality judgments in the earliest stages of L2 acquisition (McLaughlin et al., 2004; Yum et al., 2014). The current study aimed to expand this body of research by investigating language effects on the N400 by comparing ERPs to L1 and L2 words in early L2 learners. To attain the most realistic data possible, monolingual English-speaking students in an introductory Spanish course were studied in three sessions over the course of five months of learning. In each session, ERPs were collected to L1 and L2 items that were passively read for meaning. Because the N400 component has been understood to reflect lexical/semantic processing of a word (Kutas and Hillyard, 1980, 1984), this was the component of most interest in this experiment. Based on previous studies, we predicted that the N400 to L2 items would be reduced in amplitude in comparison to L1 items. Importantly, we expected that this language effect would diminish across sessions as participants became more proficient in their L2.

We observed a significant effect of language in each session independently, particularly at posterior electrode sites, such that L2 words always elicited an attenuated N400 component compared to L1 words. This finding replicates the results from Midgley et al. (2009) in intermediate L2 learners. The important difference here is that the present study focused on the earliest stage of L2 learning. Though Midgley et al. (2009) showed similar language effects in intermediate learners, the question remained as to

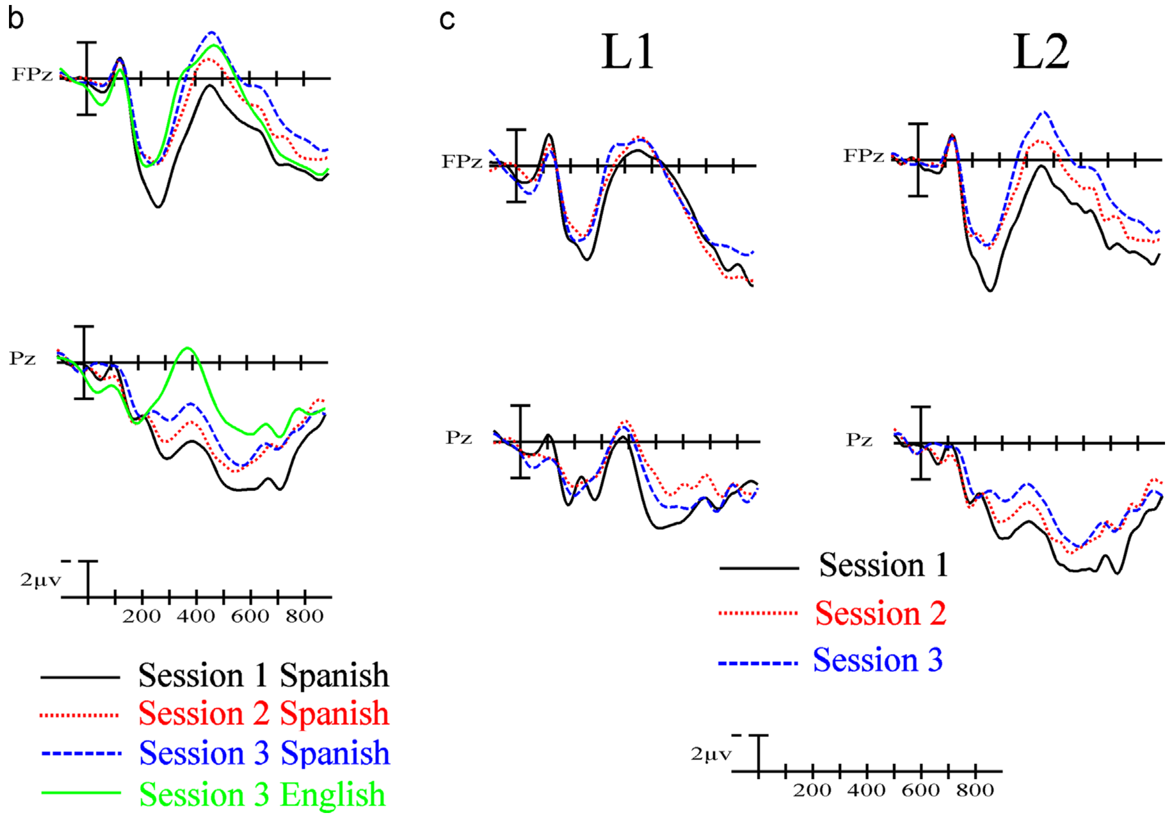

Fig. 1. (*continued*)

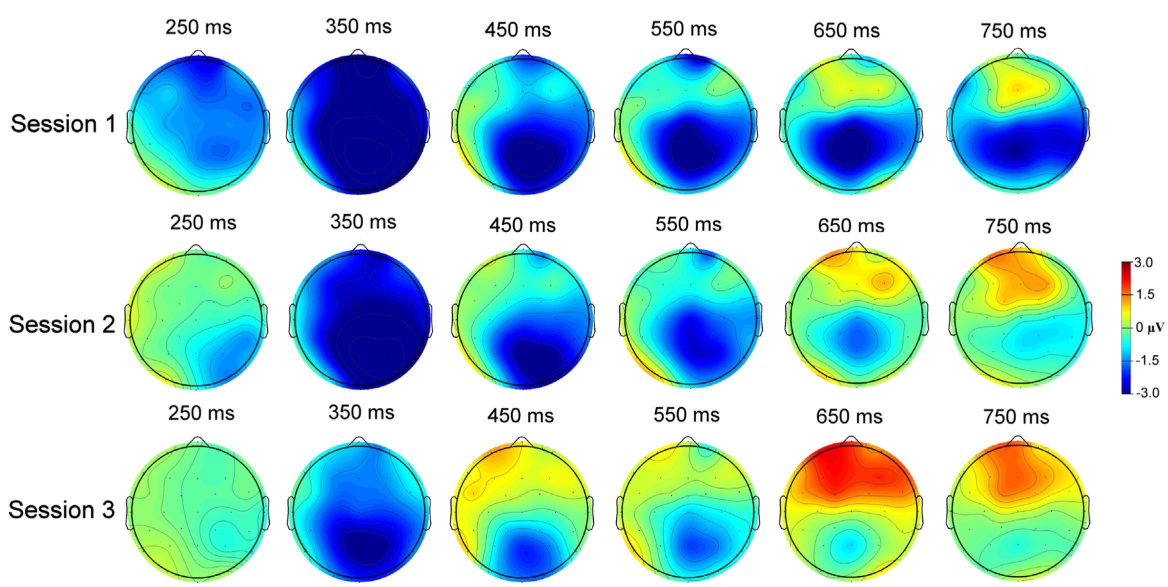

Fig. 2. Voltage maps in 100 ms epochs showing the difference between L2 and L1 ERPs in session 1 (top) session 2 (middle) and session 3 (bottom).

whether this language effect described all stages of L2 learning, or only emerged after people completed the earliest stages. When considering languages with similar alphabets (e.g., English and Spanish), it is plausible that L2 words would be processed similarly to L1 low-frequency words or pseudowords in the initial phase of L2 learning (eliciting a larger N400 than high-frequency L1 words; Kutas and Federmeier, 2000; Chwilla et al., 1995), and then with increasing proficiency begin to demonstrate the observed language effect. However, the results of the present study confirm that the language effect is found to be robust even after only 53 days of L2 learning.

Not only did we find an effect of language in each session, but we found that this effect diminished across sessions, such that the L2 N400 component continued to increase in amplitude from session one to session two and from session two to session three. This effect suggests that the L2 N400 is mediated by language proficiency in language learners, especially when these results are considered in conjunction with the results from Midgley et al. (2009) who found a language effect in intermediate learners, but not in proficient bilinguals. This attenuation of the language effect that we observed also supports the findings by McLaughlin et al. (2004) which reported rapid changes in brain responses to L2

Table 3
Time-course analysis (100 ms intervals) of language effect and interaction of language effect by distribution.

		200–300	300–400	400–500	500–600	600–700	700–800
Session 1	Language effect	**	***	***	***	*	*
	Language effect × distribution				**	*	
Session 2	Language effect		***	**	**		
	Language effect × distribution		*	**	*		
Session 3	Language effect		***				
	Language effect × distribution		**	**	*	**	*

** $p < 0.05$.
*** $p < 0.01$.
* $p < 0.10$.

items after only 14 h of instruction, and then continued differentiation of ERPs to L2 pseudowords and real words across several subsequent sessions. These findings, along with those of the present study suggest that there is significant plasticity, even in the adult language system.

As an explanation for the attenuated L2 N400, or the language effect, Midgley et al. (2009) hypothesized that this difference between L1 and L2 processing was caused by differences in orthographic neighborhood size of the items in L1 compared to those in L2. It can be assumed that L1 items are much more richly interconnected in the lexicon than are L2 words, and in monolingual studies it has been shown that words with larger orthographic neighborhoods elicit higher-amplitude N400s (Holcomb et al., 2002). This could explain why the L2 N400 amplitude in the current study increased over time – as participants become more proficient with their new L2 vocabulary, the L2 words themselves benefited from more overall connectivity which resulted in greater N400s (Midgley et al., 2009; Kounios and Holcomb, 1994).

Our study design included two different types of test lists – mixed (Spanish and English) and pure language lists (English or Spanish items only). Based on an earlier study by Alvarez et al. (2003) in more proficient Spanish learners we predicted that in a mixed list switching from L2 items to L1 items within the list might elicit a larger N400 for L1 items (this is what Alvarez et al. found) and therefore a larger effect of language when contrasting English and Spanish ERPs. In fact, we observed no significant interaction between language effect and list type – N400 language effects did not differ as a function of list type (pure vs. mixed). One possibility for this lack of a list effect is that switch effects in mixed lists at the earliest stage of learning do not resemble the profile of switch effects in the intermediate learners studied by Alvarez et al. (2003).

Though the present study provides solid evidence for the language effect in early L2 learners and the rapidity of L2 acquisition, the study has several limitations. First, the data in current study were collected from only 12 participants. This admittedly small number of learners may miss some of the subtler effects of learning that are larger N study could detect. However, despite having few participants, the results were statistically reliable indicating that this effect of language on the N400 and its attenuation over time are a robust effect.

The present study only followed participants over five months of language learning, and already there were significant increases in N400 amplitude to L2 items and a decline in the difference in N400 between languages. The earlier Midgley et al. (2009) study examined much more proficient, intermediate learners and showed similar significant language effects after an average of 5 years of L2 learning. So, it is curious that similar language effects appear to persist over a long period of seemingly long period of language learning. It is possible that the evolution of the language effect slows after an initial L2 learning phase and may only slowly increase during the intermediate learning phase. Another possibility is that language effects are sensitive to the specific items tested. In the current study we carefully chose items that were part of the Spanish curriculum that our participants were learning. In Midgley et al. (2009) L2 items were chosen from lists that most intermediate L2 learners know. It seems likely that although they were on average more proficient some of the items tested in Midgley et al.'s might have been less familiar resulting in larger language effects. It will be important for future studies to follow participants over several years to investigate the trajectory of N400 language effects.

Another finding worth further exploration was that we found no significant effect of language prior to 300 ms except in Session 3. Midgley et al. (2009) found an early main effect of language prior to 300 ms in their intermediate L2 learners but not in proficient bilinguals. Any effects in this early epoch are believed to be sensitive to orthographic manipulations of stimuli (Holcomb et al., 2002), which is why, unlike the N400 component, these effects may vary with type of language. Different languages vary by types of alphabets used, incorporation of different types of accents, word-length, bigram frequency, and by many other visual characteristics, and early effects may reflect these differences. It would be quite relevant for a future study to investigate the effect of type of language on these early components. For example, a study could compare native English-speakers learning different languages (e.g., Spanish, French, German, or Dutch) whose words have varying bigram frequencies or amounts of orthographically similar words in comparison to English. It would also be interesting to categorize these languages by how orthographically shallow (high letter-to-sound correspondence) or deep (low letter-to-sound correspondence) they are, and compare ERPs with this difference as a factor. Spanish is a much more orthographically shallow language than English, and it would also be valuable to investigate these early language effects in Spanish-speakers in the earliest stage of learning English to see if there are significant differences before 200 ms and when these differences arise. Studies investigating early differences in ERP components in early language learners would serve to elaborate on the full time course of word processing to include pre-lexical orthographic processes, rather than just lexical and semantic processing at the N400.

4. Conclusions

The present study provides the first evidence of a language effect at the typical N400 component in the earliest stage of L2 learning in the adult brain. Our results also suggest that this language effect evolves quickly as a result of increasing L2 proficiency.

We hypothesize that this increase in N400 amplitude to L2 items is the result of increased L2 orthographic connectivity in the mental lexicon due to L2 learning. In addition to the few existing studies that have investigated L2 learning, the findings from the current study suggest that while learning a second language as an adult may seem difficult and slow, the brain's plasticity allows for a quite rapid integration of words from a second language into the mental lexicon.

5. Methods

5.1. Participants

Twelve Tufts undergraduate students (7 female, mean age = 18, SD = 0.70) who were enrolled in Introductory Spanish were paid to participate. All were native English speakers, were right-handed, had normal or corrected-to-normal vision, and had no history of traumatic brain injury. English was reported to be the first language learned by all participants (L1), with no other languages learned before the age of 8. Prior to enrolling in Spanish I at Tufts, 9 of the participants had no formal exposure to Spanish (L2), and 3 participants had learned Spanish for 2-4 years in high school.

Participants reported their abilities to read, speak and comprehend English and Spanish (1 = unable; 7 = expert) as well as how frequently they read or communicated (spoke or wrote) in both languages (1 = rarely; 7 = very frequently). All participants consistently rated their L1 language use and ability as significantly higher than in their L2 ($t(1,11) = 29.11$, $p = 0.001$) (see Table 4).

5.2. Stimuli

The stimuli created for this study were 228 three- to seven-letter non-cognate morphemically simple Spanish words being taught in the Tufts Spanish I curriculum and their translations into English. The average length of the English items was 4.79 letters (SD = 1.11) while the average length of the Spanish items was 5.36 (SD = 1.15). The average written CELEX log frequency of the English items was 2.19 (SD = 0.03) which is equivalent to 155 occurrences per million items.

The probe items consisted of 76 English non-words and 76 Spanish non-words. The non-word items were based on their respective original language and retained phonetic and orthographic properties of that language. The average length of the English non-word items was 4.87 letters (SD = 1.18) while the average length of the Spanish non-word items was 5.17 (SD = 1.22).

Each group of 228 English items and 228 Spanish items were divided into four sub-groups of 57 items each. These English and Spanish sub-groups were then combined to create 4 lists that all included an English block (57 English items, 19 English non-word items), a Spanish block (57 Spanish items, 19 Spanish non-word items), and a Mixed block (57 English items, 19 English non-word items, 57 Spanish items, 19 Spanish non-word items). The word items and non-word items were presented in a random order in each block. The lists were created from the subgroups such that if a participant saw “apple” in the English block they would not see “manzana” in the Spanish block or in the Mixed block, in order to avoid repetition of translation equivalents. These blocks were presented in a counterbalanced manner across participants, and the participants did not see the blocks in the same order from one session to the next.

Table 4
Mean (SD) of self-reported language ability and use frequency in English (L1) and Spanish (L2).

	L1 – English	L2 – Spanish
Language ability	6.88*(0.25)*	2.08*(0.48)*
Language use frequency	7.00*(0.00)*	1.52*(0.86)*

5.3. Procedure

5.3.1. Training session

The first of four total sessions was a vocabulary training program where participants were exposed to and tested on all of the Spanish words that were used in the study. The training session took place in a room with 0–2 other participants in the same room, each at his or her personal computer station. The participants worked silently and at their own pace. First, participants were given a pre-test where they were asked to provide the English translations for any of the 228 Spanish items they already knew. Next the participants performed a training task that used a quarter of the Spanish items. During the training task participants saw a fixation point followed by a single Spanish item followed by the simultaneous appearance of the same Spanish item on the right and its English translation equivalent on the left of the same screen (Fig. 3(a)). The participants were allowed to study the translation screen as long as they felt they needed to remember the translation, and pressed a key on the keyboard to proceed to the next item.

After being trained on one section of the items, participants were tested on that same group of items in a different order. During testing, participants saw a fixation point, an English word, and then a screen with two Spanish words on it, one being the correct translation of the previous English word. To choose the correct Spanish translation, participants pressed the “1″ on the keyboard if the correct translation was on the left, or a “0″ on the keyboard if the correct item was on the right. The participants were not limited on time to choose between the two Spanish items presented. After the participants chose between the two Spanish items, the correct Spanish translation appeared on the screen before the next English word was shown (Fig. 3(b)).

This training/testing pattern was performed four times during the training session so that all participants were exposed to all Spanish items used in the experiment. The training/testing lists were presented in a counterbalanced order across participants. At the end of the training session participants were administered a post-test which was identical to the pre-test to assess learning.

5.3.2. Experimental sessions

Participants were exposed to three experimental sessions. Sessions one and two averaged 53 days apart, and Sessions two and three averaged 67 days apart. We chose these intervals to roughly correspond to equivalent points in the semester when students would be most available to come to the laboratory for testing. On all three experimental days, participants were seated in a comfortable chair approximately 1.5 m away from a monitor. Stimuli appeared in white Verdana text centered on a black screen. The maximum height and width of the stimuli were such that no saccades would be required during reading of the single word stimuli. Each trial consisted of a fixation cross, a blank screen, an item, a blank screen, and then a symbol during which participants were allowed to blink (see Fig. 3(c) for stimuli durations and example of trial) followed by a blank screen. Other than during the blink screen, participants were asked to remain still and relaxed.

The participants completed a go/no-go lexical decision task, in which they were asked to press one button on a game pad as quickly as possible whenever they detected a non-word (25% of trials). No response was required to other stimuli (i.e., critical items). We and others have successfully used this task in previous ERP language studies where the goal is to minimize response

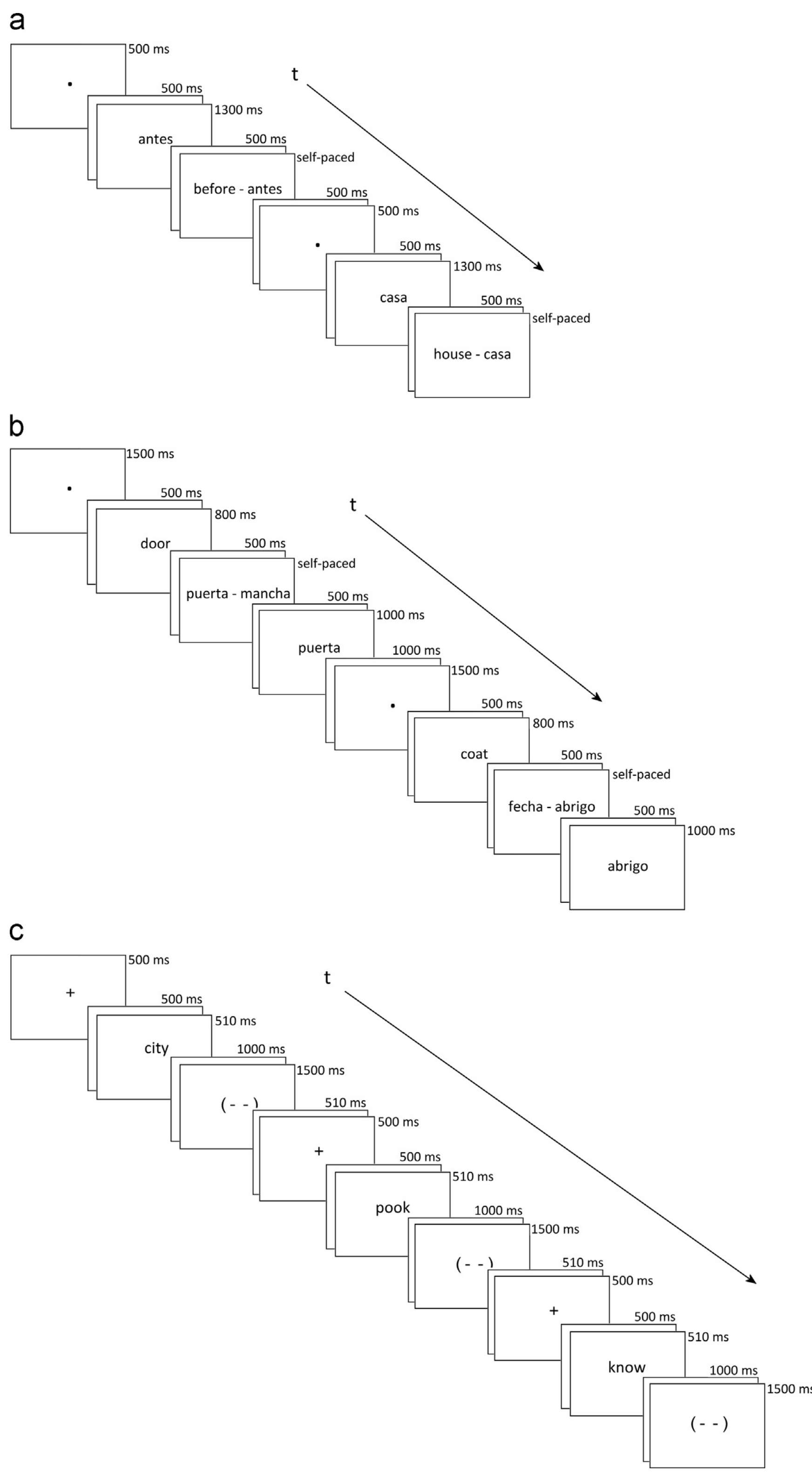

Fig. 3. (a) A typical vocabulary training trial. The duration of each presentation is noted to the left. (b) A typical vocabulary testing trial. The duration of each presentation is noted to the left. (c) A typical experimental trial. The duration of each presentation is noted to the left.

based processes in the ERPs of interest (e.g., Yum et al., 2014). Before the experimental lists, participants were shown a practice list to familiarize them with the task. After the experimental section, participants were asked to provide English translations of all 228 Spanish words used to create the items for the study in order to test for learning and comprehension (see Fig. 3(c) for a schematic of the task).

5.4. EEG recording procedure

During each experimental session participants were tested in a darkened sound attenuated room while seated in a comfortable

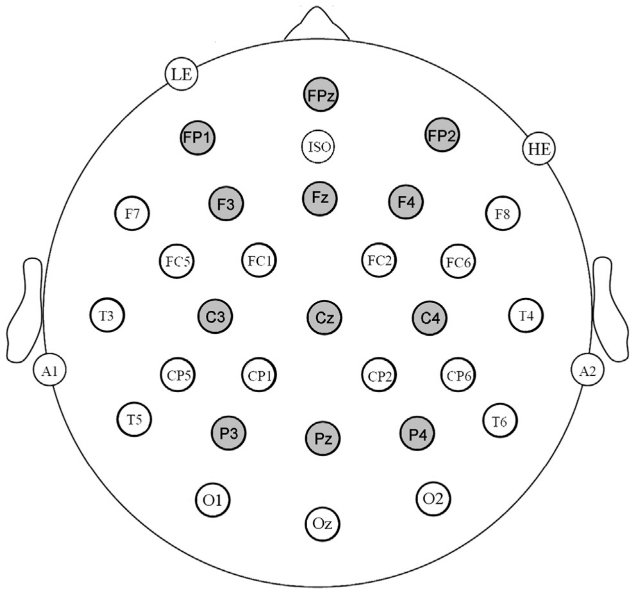

Fig. 4. Electrode montage showing the locations of 29 scalp electrodes. The 12 electrodes highlighted in grey represent electrode sites used in data analysis (FP1, FPz, FP2, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4).

chair. The electroencephalogram (EEG) was then recorded from 29 tin scalp electrodes which were embedded in an elastic cap (Electro-Cap International, see Fig. 4). To monitor blinks and eye movements, additional free electrodes were attached below the left eye (LE) and to the right of the right eye (VE). All electrodes were referenced to an electrode placed on the left mastoid bone (A1), and continuous recording from an electrode on the right mastoid bone was used to monitor and account for differential mastoid activity. All 29 head electrode impedances were below 5 kΩ, eye electrodes were below 10 kΩ, and mastoid electrodes were less than 2 kΩ. The EEG was amplified using an SA Bioamplifier at a bandpass of 0.01–40 Hz, and the EEG was measured at a rate of 200 Hz throughout the experiment.

5.5. Data analysis

The analysis of ERPs was conducted using an average of the mean amplitude values which were low-pass filtered at 15 Hz (identical analyses using a 30 Hz low pass filter produced the same pattern reported in the Section 2). A representative subset of 12 electrodes was chosen for analysis (see grey sites in Fig. 4) based on previous similar research (e.g., Midgley et al., 2008; Midgley et al., 2009). To capture the activity before, during, and after the typical N400 waveform, mean amplitudes were measured in three epochs (pre-N400 between 200 and 350 ms, N400 between 350 and 600 ms, and post-N400 between 600 and 800 ms) for the three levels of LEARNING SESSION (Session 1 vs Session 2 vs. Session 3), two levels of LANGUAGE (L1 vs. L2), two levels of LISTTYPE (Pure and Mixed), and 12 levels of DISTRIBUTION (12 electrodes; see Fig. 4). Repeated measures analyses of variance (ANOVAs) were performed separately on the data in each of these epochs. A Geisser and Greenhouse (1959) correction was applied to all repeated measures with more than one degree of freedom in the numerator. Significant interactions involving the LANGUAGE AND LEARNING SESSION variables were followed up with separate ANOVAs looking at LEARNING SESSION separately for the two languages. Language effects were further explored using separate time-course analysis on 100 ms-epochs between 200 and 800 ms for each LEARNING SESSION. Only p values that survived FDR correction are included in this table (Groppe et al., 2011).

Acknowledgements

This research was supported by HD25889 to PJH and KJM.

References

Alvarez, R.P., Holcomb, P.J., Grainger, J., 2003. Accessing word meaning in two languages: an event-related brain potential study of beginning bilinguals. Brain Lang. *87* (2), 290–304.
Bakker, I., Takashima, A., van Hell, J.G., Janzen, G., McQueen, J.M., 2015. Tracking lexical consolidation with ERPs: lexical and semantic-priming Effects on N400 and LPC responses to newly-learned words. Neuropsychologia 79, 33–41.
Borovsky, A., Kutas, M., Elman, J.L., 2010. Learning to use words: event related potentials index single-shot contextual word learning. Cognition 116 (2), 289–296.
Borovsky, A., Elman, J.L., Kutas, M., 2012. Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. Lang. Learn. Dev. 8 (3), 278–302.
Chwilla, D.J., Brown, C., Hagoort, P., 1995. The N400 as a function of the level of processing. Psychophysiology *32*, 274–285.
Domínguez, L., Tracy-Ventura, N., Arche, M.J., Mitchell, R., Myles, F., 2013. The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. Biling. Lang. Cogn. 16, 558–577.
Doughty, K., 2003. The Handbook of second Language Acquisition. Blackwell., Malden, MA.
Elgort, I., Piasecki, A.E., 2014. The effect of a bilingual learning mode on the establishment of lexical semantic representations in the L2. Biling. Lang. Cogn. 17, 572–588.
Geisser, S., Greenhouse, S., 1959. On methods in the analysis of profile data. Psychometrika 24, 95–112.
Groppe, D.M., Urbach, T.P., Kutas, M., 2011. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. Psychophysiology 48 (2011), 1711–1725.
Holcomb, P.J., O'Rourke, T., Grainger, J., 2002. An event-related brain potential study of orthographic similarity. J. Cognit. Neurosci. 14, 938–950.
Johnson, J.S., Newport, E.I., 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. Cognit. Psychol. 21, 60–99.
Kounios, J., Holcomb, P.J., 1994. Concreteness effects in semantic processing: event-related brain potential evidence supporting dual-coding theory. J. Exp. Psychol. Learn. Mem. Cogn. 20, 804–823.
Kroll, J.F., Stewart, E., 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. J. Mem. Lang. 33, 149–174.
Kroll, J.F., Tokowicz, N., 2005. Models of bilingual representation and processing: looking back and to the future. In: Kroll, J.F., de Groot, A.M.B. (Eds.), Handbook of Bilingualism: Psycholinguistic approaches. Oxford University Press, New York, NY, pp. 531–553.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427), 203–205.
Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.
Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cognit. Sci. 4, 463–470.
McLaughlin, J., Osterhout, L., Kim, A., 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. Nat. Neurosci. 7 (7), 703–704.
Mestres-Missé, A., Rodriguez-Fornells, A., Münte, T.F., 2007. Watching the brain during meaning acquisition. Cereb. Cortex 17 (8), 1858–1866.
Midgley, K.J., Holcomb, P.J., Grainger, J., 2009. Language effects in second language learners and proficient bilinguals investigated with event-related potentials. J. Neurolinguist. 22 (3), 281–300.
Midgley, K.J., Holcomb, P.J., Walter, J.B., van Heuven, Grainger, J., 2008. An electrophysiological investigation of cross-language effects of orthographic neighborhood. Brain Res. 1246 (3), 123–135.
Perfetti, C., Wlotko, E.W., Hart, L.A., 2005. Word learning and individual differences in word learning reflected in event-related potentials. J. Exp. Psychol. Learn. Mem. Cogn. 31 (6), 1281–1292.
Stein, M., Dierks, T., Brandeis, D., Wirth, M., Strik, W., Koenig, T., 2006. Plasticity in the adult language system: a longitudinal electrophysiological study on second language learning. NeuroImage *33* (2), 774–783.
Veivo, O., Järvikivi, J., 2013. Proficiency modulates early orthographic and phonological processing in L2 spoken word recognition. Biling. Lang. Cogn. 16, 864–883.
Weber-Fox, C.M., Neville, H.J., 1996. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. J. Cognit. Neurosci. 8, 231–256.
Yum, Y., Grainger, J., Midgley, K.J., Holcomb, P.J., 2014. Initial second language vocabulary learning: a longitudinal study using a new electrophysiological assessment technique. Psychophysiology 51, 364–373.