

INITIAL SECOND LANGUAGE VOCABULARY LEARNING:
A LONGITUDINAL ERP STUDY

A thesis submitted by

Yen Na Yum

In partial fulfillment of the requirements

for the degree of

Master of Science

in

Psychology

TUFTS UNIVERSITY

May, 2012.

ADVISER:

Phillip Holcomb

Abstract

Studying lexical learning across writing systems in beginning learners is one way to provide insight into how second language (L2) learners approach an entirely unfamiliar orthography and to determine whether previous language learning effects are generalizable to languages that have minimal overlap. This study examined initial L2 orthographic and semantic acquisition in monolingual native English speakers who learned a fixed set of Chinese words under controlled laboratory conditions. Participants were tracked to obtain longitudinal behavioral and event-related potentials (ERP) data in 10 sessions of L2 vocabulary learning. Verbal backward translation and semantic categorization were assessed. Behavioral data supported that participants gradually acquired the L2 stimuli items and several ERP components showed changes during the course of the study. Compared to L1 items, an enhanced P2 component was seen to L2 items in semantic categorization. A small but growing N400 component and a late anterior effect in L2 items were seen to increase at later sessions. These effects appeared to depend on individual differences and the degree of successful learning, as participants who had better behavioral results showed distinct patterns of ERP activation compared to those who performed less well. Complete learners showed a negative shift over the anterior region of the brain while incomplete learners showed a posterior positive effect, both of which were right-lateralized. The observed differences could relate to differential processing styles or strategy use. Interestingly, a shift in latency of the N400 component was seen to semantic categorization in English (L1 items), suggesting that L1 could be influenced at a very early point in L2 learning.

Although being bilingual has many advantages, learning a second language (L2) in adulthood seems like a daunting task to many people. Until relatively recently, it was believed that L2 learning after a certain critical period would result in poorer outcome, with the ultimate L2 attainment never reaching the natural fluency of a first language. This was perhaps based on the false notion that neurons, and by extent brain networks, stop growing in adulthood. However, mounting evidence shows that the brain undergoes structural changes specifically related to learning well into adulthood (May, 2011). This plasticity in the adult brain prompts many questions related to neural changes resulting from learning an L2 in adulthood: What is the time course of these changes and can we identify neural signatures of L2 learning? Do changes in language processing differ for L1 and L2? Are changes dependent on features of the languages such as their writing systems?

Vocabulary knowledge forms the basis of spoken communication and reading comprehension from the earliest point of L2 learning. Both in beginning and advanced bilingual speakers, the fundamental role of vocabulary in the fluent use of L2 is well documented (e.g. Nation, 1993; 2001). A number of studies have examined novel word learning in adulthood in first language, where participants learned a set of low frequency vocabulary items or pseudowords in which phonotactic rules of their native language were preserved (Frishkoff, Perfetti, and Collins-Thompson, 2010; Mestres-Misse, Rodriguez-Fornells, & Munte, 2007; Perfetti, Wlotko, & Hart, 2005). These studies found that novel words in which meaning acquisition was achieved showed behavioral and neuronal changes approaching that of real L1 words. Integration of the new vocabulary items into the existing lexicon is relatively easy, because the orthographic and phonological systems were familiar to the learners. In contrast, any second language learning involved building up new orthographic, phonotactic, or grammatical

rules. Effects of these learning processes and rule abstraction would be especially prominent in the initial phase of L2 learning.

An important factor in L2 research is the proficiency or fluency of the L2 learner. Not surprisingly, experience with a foreign language changes how the words in that language are perceived and processed as well as how they are used. From initially being unfamiliar with the language to achieving native-like proficiency, the L2 learner should display changes in brain activities in response to L2 vocabulary items. A recent study by Midgley, Holcomb, and Grainger (2009) compared how words in L1 and L2 were processed by bilinguals of different proficiency levels, measured by event-related potentials (ERP). A few ERP components were found to differentiate less proficient and more proficient bilinguals of English and French. In particular, by 150ms after stimulus onset, the peak amplitude was more negative for L1 than L2 words at posterior sites on the scalp. This difference was found in beginning L2 learners but not in more proficient bilingual speakers. Additionally, the N400 component was smaller to L2 words than L1 words in less proficient learners, but virtually the same in more experienced bilinguals. Neuroanatomical data have also supported the role of L2 proficiency in bilingual lexical processing (Abutalebi, Cappa, & Perani, 2001; Perani & Abutalebi, 2005; Stein, Federspiel, Koenig, et al., 2009). The activation patterns for L1 and L2 words were more similar in bilinguals with higher L2 proficiency. In particular, activation in the inferior frontal gyrus seen to L2 words was reduced when participants became more proficient in L2, suggesting more automaticity and less reliance on L1 translation (Stein et al., 2009).

Many of the bilingual studies conducted to date used English and Romance languages that share an alphabet. Although significant differences were seen, the close relationship between L1 and L2 made it difficult to isolate the unique contribution of various processes during L2

word learning. To tackle this problem, some researchers have chosen to create an artificial language or use pseudowords to better manipulate word features such as familiarity and phonological similarity (e.g. De Groot and Keijzer, 2000; Finkbeiner & Nicol, 2003). However, motivation to learn the new vocabularies might be compromised in these experiments. In a longitudinal study, using natural languages and recruiting participants who have an interest in learning this language should improve attention and learning.

Natural languages belong to different writing systems that vary in their mappings of graphic forms to linguistic forms. Different writing systems have been shown to involve different reading networks in the brain, as revealed by neuroimaging studies (Bolger, Perfetti, & Schneider, 2005; Perfetti, Liu, Fiez, Nelson, Bolger, & Tan, 2007). Recent work by Perfetti and colleagues has suggested that languages which use different writing systems (English and Chinese in their studies) are learned as an L2 in a different manner than are more related languages (Nelson, Liu, Fiez, & Perfetti, 2009; Perfetti and Liu, 2005; Perfetti et al., 2007). In their System Accommodation Hypothesis, they used the Piagetian concepts of assimilation and accommodation to explain these differences. According to this account, assimilation occurs when the word recognition system of L1 can be used to process words in the writing system of L2 without modification, whereas accommodation implies an adaptation of the L1 word recognition system to process L2 words. Perfetti et al. (2007) presented evidence for an asymmetry in the mechanisms involved in learning L2 words, with English L1 readers developing new structures in order to read in Chinese (accommodation), whereas Chinese L1 readers using pre-existing reading system to read English stimuli (assimilation). Their results showed a pattern of accommodation in both typical classroom learners and a group of native English speakers with no prior knowledge of Chinese trained in the laboratory (Liu, Dunlop, Fiez,

& Perfetti, 2007).

One goal of this study was to determine which of the second language effects and mechanisms previously reported in Romance languages such as French also apply to languages that do not share the same type of writing system and/or alphabet. Chinese was chosen because of its minimal overlap in both its written and spoken form with English. Chinese has been argued to be a language of high contrast to English (see Wang, Perfetti, & Liu, 2003 for discussion). It is a morphosyllabic language-- each character corresponds to a single syllable and a morpheme. Chinese is also commonly said to be "logographic," where visual word forms have a direct route to meaning with little or no contribution from phonology. Although some characters contain sub-lexical units known as radicals that can carry phonological information, the grapheme to phoneme mapping is inconsistent and unpredictable.

Since a reliable system of grapheme to phoneme conversion does not exist, it has been proposed that phonological decomposition in Chinese word processing was not used in lexical identification. In addition, phonological processing can only be completed when whole-word orthographic processing has been completed (Chen, Flores d'Arcais, & Cheung, 1995; Tan, Laird, Li, & Fox, 2005). Some have proposed that learning to read Chinese would involve a more direct mapping of whole word characters onto sound and meaning (Liu, Wang, & Perfetti, 2007). This does not mean phonological information was not learned in initial word acquisition, but phonological information could be ambiguous because of the large number of homophones in Chinese. For these reasons, orthography should be more reliable than phonology for L2 Chinese learners who are engaging in visual word identification. This is in contrast to alphabetic and syllabic systems where phonological information can be very helpful in lexical processing (Grainger & Holcomb, 2009).

Experimental evidence has shown that adult native English speakers can correctly differentiate incorrect orthographic structure from correct Chinese characters within a few months of college-level classroom learning (Wang, Liu, & Perfetti, 2003). Liu, Perfetti, and Wang (2006) conducted an ERP study using a delayed naming task in college students learning to read Chinese. They found that Chinese words produced a larger P200/N200 component compared to English words in the first-term learners. This was interpreted as increased visual processing (occipital N200) and general lexical processing (P200). By the second term of Chinese learning, the occipital effects had diminished but the more frontal effects were still evident. The dissociation of the two components was also taken to indicate that lexical identification of Chinese characters did not proceed in parallel. Instead, whole word identification could be broken down into identification of lexical constituents (Perfetti et al., 2007). Although the N400 component might have been expected to change, they found that results were too variable and inconclusive on this component. Together, these studies showed that learners were sensitive to the lexical properties of the Chinese items that they encountered during classroom learning.

Second language (L2) learners show a number of differences in language processing compared to monolinguals, which might be attributed to acquiring an additional language that requires different orthographic, phonological, and semantic processes. In some cases, changes in processing are not limited to the L2, but also extend to L1 processing. The mutual influence between L1 and L2 in bilinguals and second language learners is one that has received increasing attention in research. One of the main findings that has emerged from this line of study is that cognate status influences L2 vocabulary learning (e.g. De Groot & Keijzer, 2000). Cognates are words that are orthographically or phonologically similar or identical across languages (e.g.,

table in English and *table* in French). Cognates are easier to learn and recall compared to non-cognates (De Groot & Keijzer, 2000; Sánchez-Casas, Davis, & García-Albea, 1992; Sánchez-Casas & García-Albea, 2005). The advantage in learning L2 items that are cognates with L1 words is easy to understand, since the L1 lexical entry can easily cue the new L2 word, orthographically or phonologically. Interestingly, Van Hell & Dijkstra (2002) showed L2 cognates gave L1 items a processing advantage even when it is in an exclusively L1 context. Additionally, interference effects were observed when cognates do not share semantic content. A recent study by Lagrou, Hartsuiker, and Duyck (2011) showed L1-L2 interference for an auditory lexical decision task. English-Dutch bilinguals were given English words with interlingual homophones (e.g., *lief* and *leaf*) that are unrelated in meaning, and the reaction time to these homophones was slowed as compared to non-homophones.

These findings were predicated on the similarity, if not overlap, of lexical items in the L1 and L2 of the bilinguals. So although the effects were robust in the many alphabetic languages that were studied, a further question was whether these kinds of similarity effects could be replicated in languages that are less similar. A recent study (Thierry & Wu, 2004) examined this question by showing Chinese-English late bilinguals word pairs in their L2 (English). The word pairs were either related or unrelated semantically in English, but half also had a hidden manipulation of form repetition but no semantic relatedness when translated into Chinese. Results showed that native English controls had no sensitivity to the hidden manipulation, but the native Chinese speakers had longer reaction times, higher error rates, and larger N400 ERP shifts. Since Chinese and English provided a good contrast in terms of minimal overlap of lexical features, the study's results strongly implied that late bilinguals' L1 and L2 interact and create interference for each other even when the two languages were dissimilar.

Although studies have shown that there are neural changes associated with proficiency level in L2, there remain several important gaps in our understanding of how L2 vocabulary items are learned and processed. Specifically, because these studies used cross-sectional designs, it is not clear at which point in L2 learning the observed effects emerge and what the trajectories of these changes are. However, we know that there are rapid changes in orthographic and semantic learning in the beginning phase of L2 lexical learning. As a study by McLaughlin and colleagues showed, the N400 component is sensitive to the level of exposure to a new L2 very early in the course of L2 learning (McLaughlin, Osterhout, & Kim, 2004). Even before learners were able to categorize words they had been exposed to in language class, their ERPs reliably differentiated these learned items from L2 pseudowords in the N400 component. At slightly more advanced points in L2 acquisition, newly learned items showed evidence of having acquired semantic representations, producing significant N400 attenuation in a semantic priming paradigm. Thus, early changes in L2 vocabulary processing could be detected in ERPs prior to indications in similar behavioral testing results. This is one reason to be concentrating on the initial period of L2 acquisition, and the temporal sensitivity of ERP makes it ideal for detecting the timing of the sequence of processes involved in L2 word processing.

Recent studies have shown that adults were highly sensitive to certain linguistic information in L2, showing measurable changes after hours or even minutes of learning (McLaughlin, Osterhout, & Kim, 2004; Gullberg, Roberts, Dimroth, Veroude, & Indefrey, 2010). Methodologically, however, it is particularly challenging to study learners in the very initial stages of L2 word acquisition because most experimental tasks require a threshold of L2 word knowledge beyond that of the newest L2 learners in a typical classroom environment. Consequently, most of the previous research in L2 acquisition has examined L2 learners that

were relatively far along in the process of becoming bilingual, making it difficult to isolate the precise changes in the neural processes engaged in initial L2 vocabulary acquisition. Other studies have compared participants with different proficiency levels in cross-sectional designs. One way to study initial L2 learning without the variability typically found in the classroom environment would be to train L2-naïve participants in the laboratory. The laboratory lexical training approach has been used extensively in the past to investigate the processes underlying L2 acquisition (see De Groot & Van Hell, 2005 for review). Most laboratory studies examining lexical learning used a small set of novel items that participants could learn within a few visits, so although a controlled environment is achieved, the scope of word learning is more restricted. The stimuli become highly familiar or even overlearned to the learners, which might not resemble a typical language learning situation. To address this issue, we conducted a pilot study that expanded the scope of the learning, tracking initially naïve L2 learners who acquired an L2 vocabulary in the laboratory using a longitudinal design. Monolingual participants were taught a small set of Chinese vocabulary items consisting of 84 single characters. Learning was achieved by pairing these items with their translation equivalents in the participants L1. Along with L1-L2 word association training, a forced-choice translation recognition test was also used. Training took place over four two-hour laboratory sessions, and ERPs were recorded to all to-be-learned words during training and testing. The goals of this pilot study were relatively modest-- the primary interest was in determining if initially L2 naïve learners could be taught a rudimentary vocabulary in a new language in relatively few sessions. Further, it was of interest whether the ERPs patterns produced by the new vocabulary items were comparable to those observed in university learners of L2. Behavioral results showed improvements over the four sessions and changes in ERP data were obtained. Plotted in Figure 1 are the ERPs at the O2 electrode site for

the L2 learners in a forced-choice recognition test (time-locked to the Chinese items). As can be seen, there is a systematic decrease at this site in an early posterior positivity over the course of learning. It seemed plausible that this effect might be related to the language effects reported by Midgley et al. (2009), where an early posterior positivity was larger for L2 items in less competent learners and was not present in more proficient users. The data from this pilot laboratory study suggested that laboratory learners in this paradigm were on a similar trajectory of learning as that of classroom learners. This study provides crucial preliminary data on the feasibility of using this approach to study a number of thorny issues in L2 acquisition and forms the foundation of the proposed study.

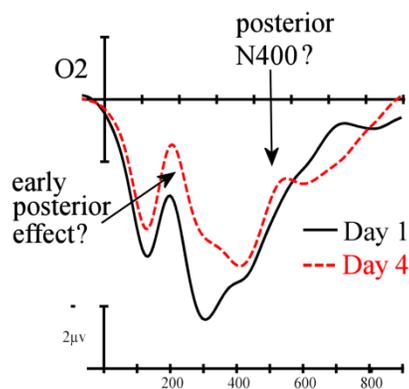


Figure 1 Lexical learning effects in laboratory-trained L2 learners (n = 12).

Current Study

The current study investigated the mental and neural processes involved in L2 vocabulary acquisition during the earliest points of learning under rigorously controlled learning conditions. Rather than exclusively studying students in university second language courses, which was generally a population with high variability in terms of language background, this study recruited participants who had not had prior exposure to the to-be-learned language (or any related

language). Participants had ten training sessions in the lab, of which four had concurrent ERP recordings. Training tasks were designed to focus on different linguistic information so participants could quickly acquire a visual and semantic familiarity to the new L2. Exposure to the new L2 should be linear over the ten sessions because the initially non-Chinese speaking participants were brought into the lab at regular intervals and the stimuli were carefully counterbalanced. As this experiment tested the same participants before and after learning, data from different sessions should show learning effects. Due to the larger scale of the study, it was also possible to separate participants into two groups according to their behavioral performance, which would allow us to investigate individual differences in L2 lexical learning.

The first ERP recording session coincided with the initial exposure to new L2 words on day 1, and served as a baseline to compare to subsequent ERP/behavioral results on ensuing ERP sessions (4th, 7th, and 10th days). ERPs were time locked to each word presented in the tasks. Although all of the tasks showed learning-related effects, for our results and discussions, we focused on two assessment tasks-- Backward Translation and Semantic Categorization Assessment Task (SCAT). This was because these two paradigms provided direct measures of a deeper level of learning and were more similar to other tasks found in the literature.

In Backward Translation Task, L2 words were presented as probes for participants to verbally translate into L1. It was predicted that over the sessions, translation accuracy would improve. In addition, L2 words would show early N/P200 changes reflecting a familiarity with the new word forms. An increase in the N400 was predicted, reflecting growing semantic connections. In the SCAT, L2 behavioral performance was expected to improve and ERPs elicited by L2 words were predicted to show early changes and an increased N400 similar to the Backward Translation Task. Based on previous studies and the pilot data, we also expected that

language effects would be present in the learners, with differences between L1 and L2 both in early components and the N400. These language effects were expected to change with time, because as participants gained more L2 experience, language processing should be easier in L2.

Method

Participants

Participants were 28 native English speakers (13 females; mean age = 19.6, SD = 1.6) who had never learned or had extensive exposure to Chinese. They were also not fluent in any other L2, but up to 6 years of classroom exposure to an alphabetic language was permitted. Participants were recruited by using advertisements that sought "people who are interested in starting to learn Chinese but have never done so." Extensive language histories were collected using a standard questionnaire. All participants were right-handed university students with normal or corrected-to-normal visual acuity and a normal neurological profile.

Apparatus & Data Analysis

For ERP recordings, participants sat in a comfortable chair in a sound attenuating room while 32 channels of EEG data (all electrodes < 5KOhm) were digitized continuously (SA Instruments amplifiers, 6db cutoffs .01 and 40Hz, 12 bit a/d 250 Hz sampling rate). Stimuli were presented as white letters/characters on a black background on a 19 inch monitor (visual angle < 3 degrees vertical and horizontal). Participant responses were collected with a multi-button response box. Average ERPs were formed from artifact-rejected single trials of EEG for each assessment task. In the case where behavioral measures were collected in tandem with the EEG, only trials with correct responses were included in average ERPs, unless otherwise noted. Mean amplitude and time course data were calculated from the ERPs. Data were analyzed with repeated measures and mixed design analysis of variance (ANOVA) which included two factors of scalp distribution as

well as variables of interest. The topographic variables (ANTERIOR-POSTERIOR: pre-frontal vs. frontal vs. central vs. parietal vs. occipital and HEMISPHERE: left vs. midline vs. right) and SESSION (Sessions 1 vs. 4 vs. 7 vs. 10) were within-subject factors while LEARNING-SUCCESS (Complete vs. Incomplete) was a between-subject factor. In the set of analysis examining language effects, the variable LANGUAGE (Chinese vs. English) was added to the overall ANOVA. The Geisser-Greenhouse correction was applied for all repeated-measures ANOVA with more than one degree of freedom to account for nonsphericity (Geisser & Greenhouse, 1959).

Stimuli

A total of 200 words including nouns, verbs, and adjectives were selected from vocabulary words taught in 1st semester Chinese language classes at Tufts University. All L2 items consisted of either one or two Chinese characters, and visual complexity was controlled where possible. Stimuli included another 100 distracter words in both L1 and L2 that participants did not learn. The distracter items were controlled for word length and lexical frequency. All L1 items consisted of common English words, with word lengths between 2 and 11 letters and mean frequency of 512.61 occurrences per million. All to-be-learned L2 items and the L1 items associated with them were presented with equal frequency in each session to eliminate possible presentation frequency effects.

Procedure

Participants went through ten automated sessions with the primary goal of learning a total of 200 words in the new L2. Three tasks were used for training and two tasks for assessment. All of the tasks were designed to be amenable to ERP data collection and to allow the tracking of performance changes, but only results from the two assessment tasks were reported in this paper.

ERP recordings were conducted in addition to vocabulary instruction and behavioral testing on four of these lab visits (Sessions 1, 4, 7, and 10); on the other 6 visits only L2 instruction and behavioral testing were conducted and ERP data were not collected.

Training

The first task was a go/no-go n-back word familiarity task. The 200 to-be-learned L2 words were mixed with 100 not-to-be-learned L2 words and presented one-at-a-time visually (500ms L2 word on, 1500 ms inter-trial-interval-- ITI). Participants were instructed to press a button whenever they detect a repeated item (n-back task with 10% repetitions between 1 and 3 trials back). This task focused learners' attention on the visual word form and allowed an initial familiarization with the lexical form of the new L2 words. In addition, this task provided a baseline response before training, to serve as comparison to other measures of learning on subsequent test days. In pilot testing, we have found that performance on this task was reasonable even during the first session.

The second task was a word-word association training, the primary training procedure. Word association has a rich history in experimental psychology and has been demonstrated to be an effective vocabulary training technique in previous studies of L2 learning (see De Groot & Van Hell, 2005). By pairing each to-be-learned word with its L1 translation, learners could associate meaning and other information with the new word. In our study, each of the 200 to-be-learned L2 word was paired with its L1 translation in L1-L2 association (500 ms L1 word, 500 ms blank, 1000ms L2 word, and 2500 ms ITI). Participants did not have to respond to the words, but were encouraged to pay attention. Blinking was allowed during ITI. In our pilot study, this task clearly worked since participants were able to eventually report the translation of more than 70% of the newly learned words.

The third training task was a translation recognition task. Pairs of words were presented one after another (800 ms probe, 500 ms blank, 800 ms target, wait for response, 800ms feedback, and 2500ms ITI). Participants were instructed to indicate by button press whether the second presented word was the correct translation of the first one. It was a forced-choice response, so learners either pressed the "yes" or the "no" button. The word pairs were presented in two blocks. Block 1 consisted of half of the pairs and these were tested in the learned direction of translation (L1 probe-L2 translation). Block 2 consisted of the other half and the word pairs were tested in the reversed direction (L2 probe-L1 translation). Performance on this task provided behavioral assessment as well as additional training for learners, since the correct translation was shown at the end of each trial as feedback.

Assessment

The first assessment task was a **L2 to L1 backward translation task**. In each trial, participants saw a learned L2 word, and their task was to verbally produce the L1 translation. They could take as much time as needed and could choose to pass if they did not know the translation. Regardless of their answers, the correct L1 translation was presented after the response as feedback (800 ms L2 word, wait for response, 500ms L1 feedback, 2500ms ITI). This production task required learners to produce a specific L1 translation rather than just to recognize it, and thus provided a direct behavioral measure of the number of words that the learners had internalized in the new L2 vocabulary.

The second assessment task was the go/no-go **semantic categorization assessment task (SCAT)**. The learners were presented with L2 words one-at-a-time visually in block one, while the L1 translation equivalents were presented in block 2 (800 ms on, 2500 ms ITI). Participants were asked to press a button to occasional probe words in a specific semantic category (body

parts or food and drinks - 17% of trials) -- ERPs were recorded to L1 and L2 non-probe "critical" words. Since the two languages were tested separately in pure blocks, participants performed the same task in L1 and L2 and translation was not required, which allowed L1 and L2 words to be directly compared for language effects.

Results

Overall Behavioral Results

To further inspect how learning affected the neural processing of newly-acquired L2 words and L1 items, we divided all learners into those who did well behaviorally (complete learners) and those who did less well in the ten sessions (incomplete learners). The participants were separated into the two groups using their overall behavioral performance in all learning sessions and all assessment tasks. A two-sample t-test showed significant differences in the group means in all the sessions, $t(1,26) > 3.59$, $p < .01$ (see Table 1 for group means). This grouping of participants was used in the contrasts of LEARNING-SUCCESS in both Backward Translation and SCAT.

Table 1. Mean behavioral performance in all tasks for complete (n = 14) vs. incomplete learners (n = 14) across learning sessions.

Session Number		Session 1	Session 4	Session 7	Session 10	Mean of All 10 Sessions
Mean	Complete Learners	32.2% (5.1%)	70.1% (6.5%)	80.8% (5.0%)	85.55% (5.5%)	69.4% (3.1%)
	Incomplete Learners	24.6% (6.0%)	44.8% (8.2%)	54.2% (11.9%)	59.1% (14.9%)	46.7% (8.9%)

Backward Translation L2 Items

The ERP waveforms time-locked to the L2 probe items in the Backward Translation Task were plotted in Figure 2a. Visual inspection of the figure showed a large positive deflection starting around 150ms post-stimulus onset. This was followed by a negative component that peaked between 300 and 400ms and continued over most of the anterior sites until 600ms. At posterior sites, there was first a positive-going component followed by several smaller peaks between 200 and 500ms. Differences between sessions started to occur at 200ms post-stimulus onset and continued into later epochs. Visual inspection of the contrast between complete and incomplete learners revealed that complete learners showed a widely distributed anterior negative shift in ERP amplitude from 300 to 700ms with learning, while incomplete learners did not show anterior changes but showed a posterior positive shift (see Figure 2b).

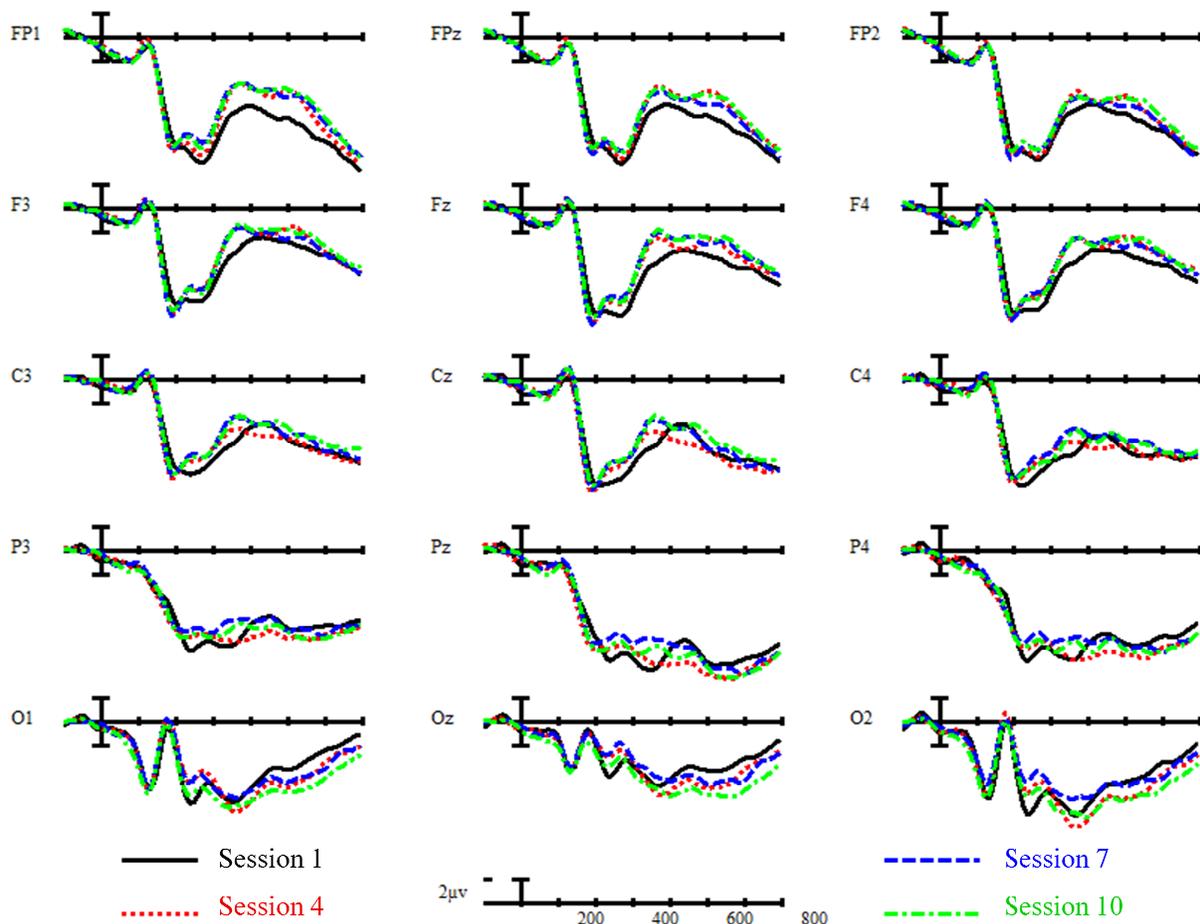


Figure 2a ERPs of Chinese (L2) probe items in the backward translation task (n=28).

Four time windows were tested for statistical differences— 200-300ms, 300-400ms, 400-500ms, and 500-700ms. The only significant result was found in the 400-500ms epoch.

400-500ms. In this epoch, an interaction of LEARNING-SUCCESS x SESSION was found, $F(3,78) = 2.86, p = .0499$. The complete learners showed an increase in the negative component as the sessions went by, but the incomplete learning group became more positive in general.

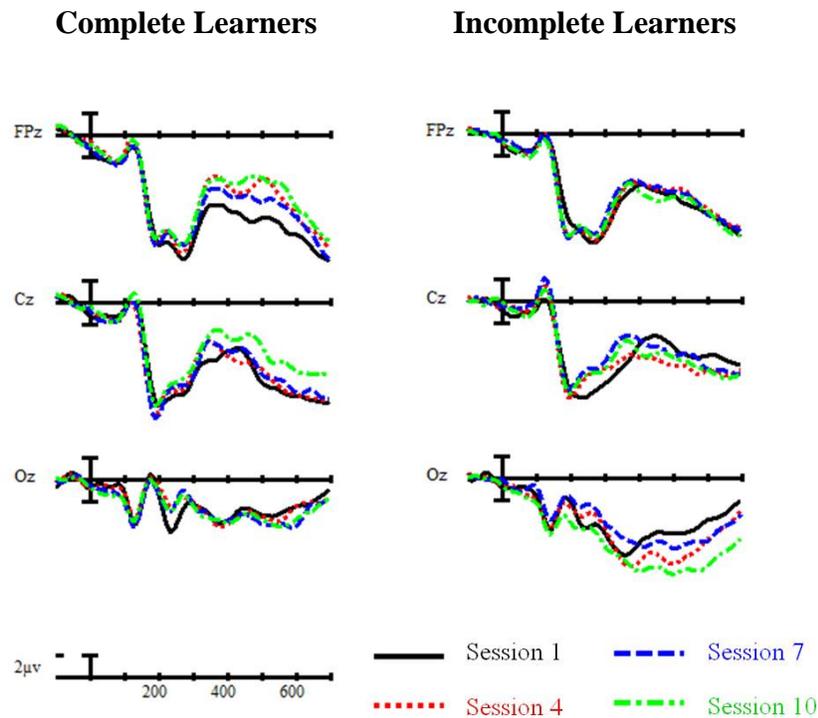


Figure 2b ERPs of L2 items in complete (left) and incomplete (right) learners in the Backward Translation Task.

Complete Learners

Incomplete Learners

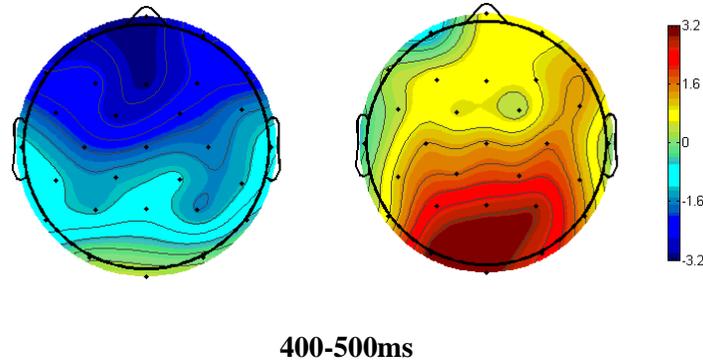


Figure 2c Voltage maps calculated from differences waves of Session 10 minus Session 1 at 400-500ms.

Behavioral Data

In general, behavioral results were consistent with predictions. The mean percentage of correct answers was 67.9% (SD = 14.6) for complete learners and 45.4% (SD = 13.5) for incomplete learners. A SESSION (all 10 sessions) x LEARNING-SUCCESS (complete vs. incomplete) between-subject factorial ANOVA revealed a main effect of SESSION, $F(9,260) = 38.6, p < .001$, indicating that both groups of participants performed significantly better with time. A main effect of LEARNING-SUCCESS was also found, showing that complete learners did better than incomplete learners in Backward Translation, $F(1,260) = 91.3, p < .001$. The interaction of the two factors was significant, $F(9,260) = 2.2, p = .025$, indicating that the complete learner group made greater improvements than the incomplete learner group over time. Table 2 showed a breakdown of the behavioral performance data for the two groups in all ten sessions.

Table 2 Percentage hits in Backward Translation Task across all sessions.

Session Number	1	2	3	4	5	6	7	8	9	10
Complete Learners:	13.7%	31.4%	52.9%	65.0%	74.5%	82.3%	85.4%	88.2%	92.7%	92.9%
Mean % (SD)	(5.4)	(10.5)	(9.4)	(16.7)	(16.5)	(17.6)	(18.3)	(20.4)	(14.1)	(16.7)
Incomplete Learners:	8.3%	14.7%	30.7%	34.9%	46.5%	53.5%	59.6%	65.9%	69.2%	70.9%
Mean % (SD)	(4.7)	(10.5)	(11.5)	(10.4)	(12.6)	(16.2)	(17.0)	(16.1)	(17.9)	(18.2)

SCAT L2 Items

The ERP waveforms for the L2 critical items of the four recordings of the SCAT were plotted in Figure 5a. As can be seen, the anterior and central parts of the scalp showed a negative shift with learning, while posterior sites showed a positive shift. The differences occurred mainly between session one and subsequent sessions, with divergences starting as early as 200ms and continuing until 700ms.

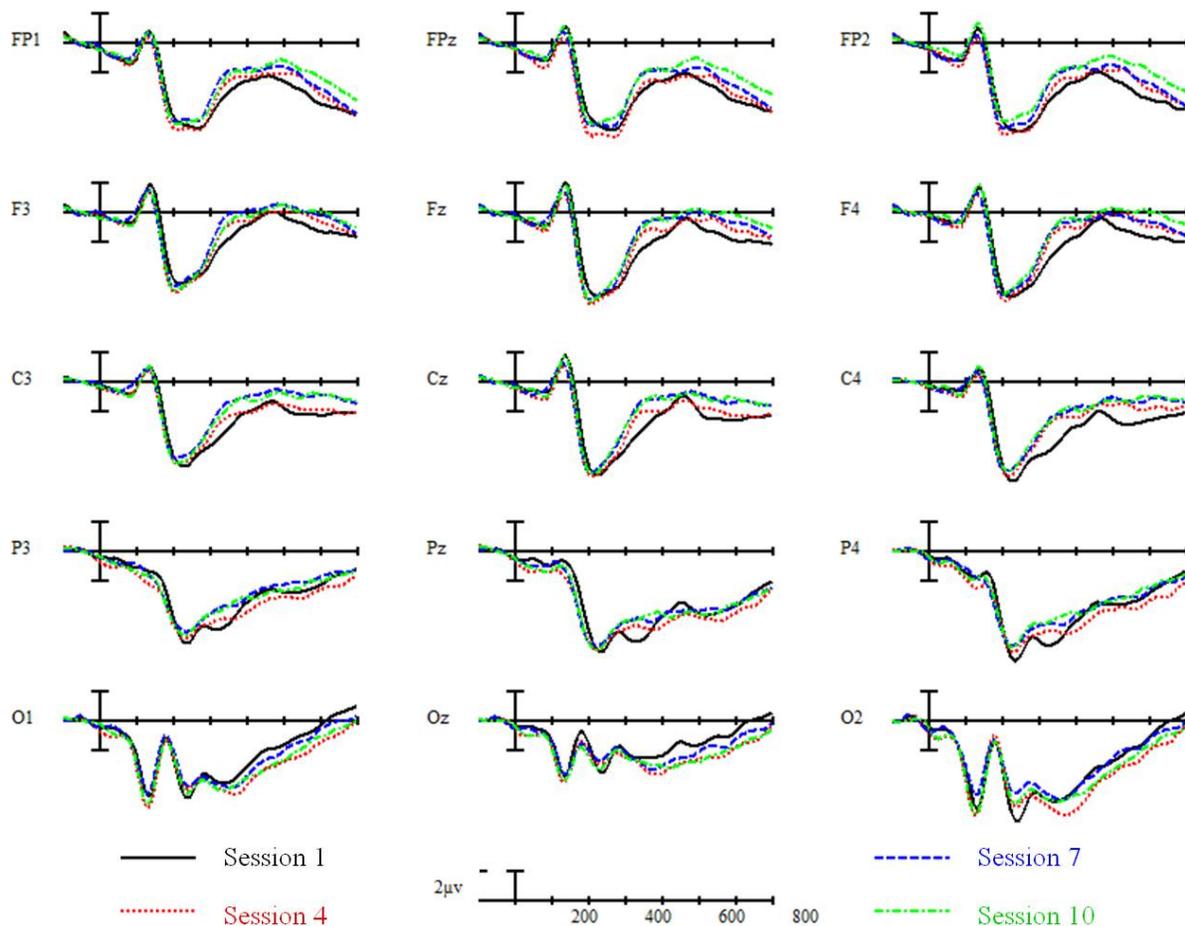
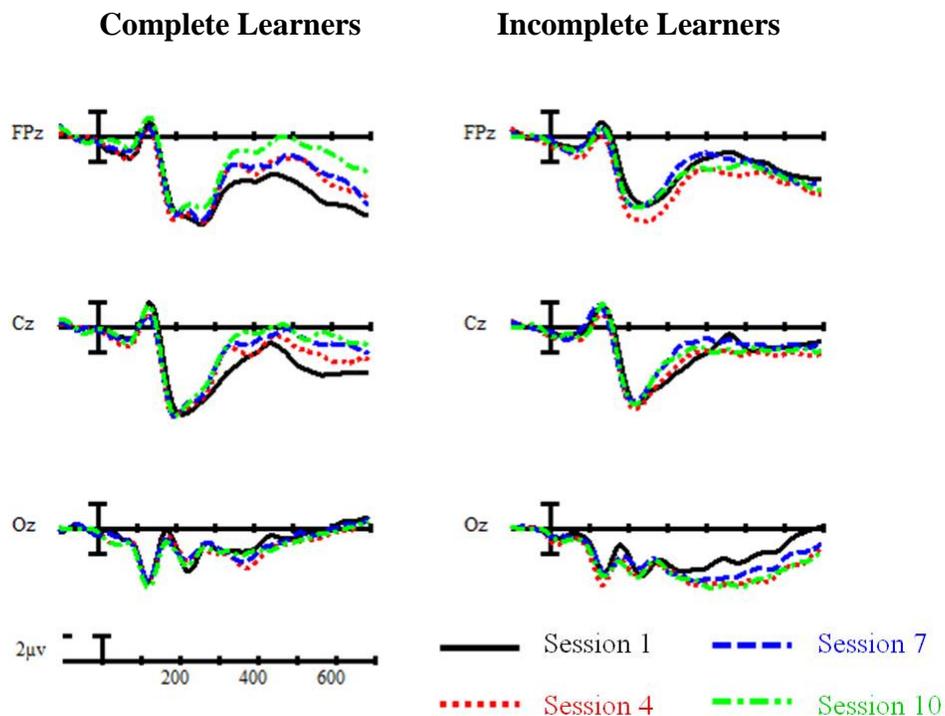


Figure 3a ERPs of Chinese (L2) critical items in SCAT (n=28).

Also plotted were the ERPs for the two groups separated by behavioral performance (see

Figure 3b). More subtle differences between these groups were seen. First, it could be noted that the changes over learning sessions in the waveforms of the complete learners were more orderly than those of the incomplete learners. For the complete learners, changes were seen starting at 200ms post-stimulus onset, with a slight reduction of a widely-distributed positive-going component. At 300-500ms, a negative deflection was pronounced at anterior and central sites, and it increased in negativity over the learning periods. Posterior sites exhibited little or no difference. While the progressions were less consistent for this group at frontal and central sites, in general a small negative shift was observed between 200-450ms. However, the pattern for incomplete learners at posterior sites was more evident. A sustained positive shift became enhanced with learning starting at around 200ms, and showed a larger difference at right posterior sites compared to left posterior sites. Voltage maps for the contrast between the two learning groups were shown in Figure 3c.



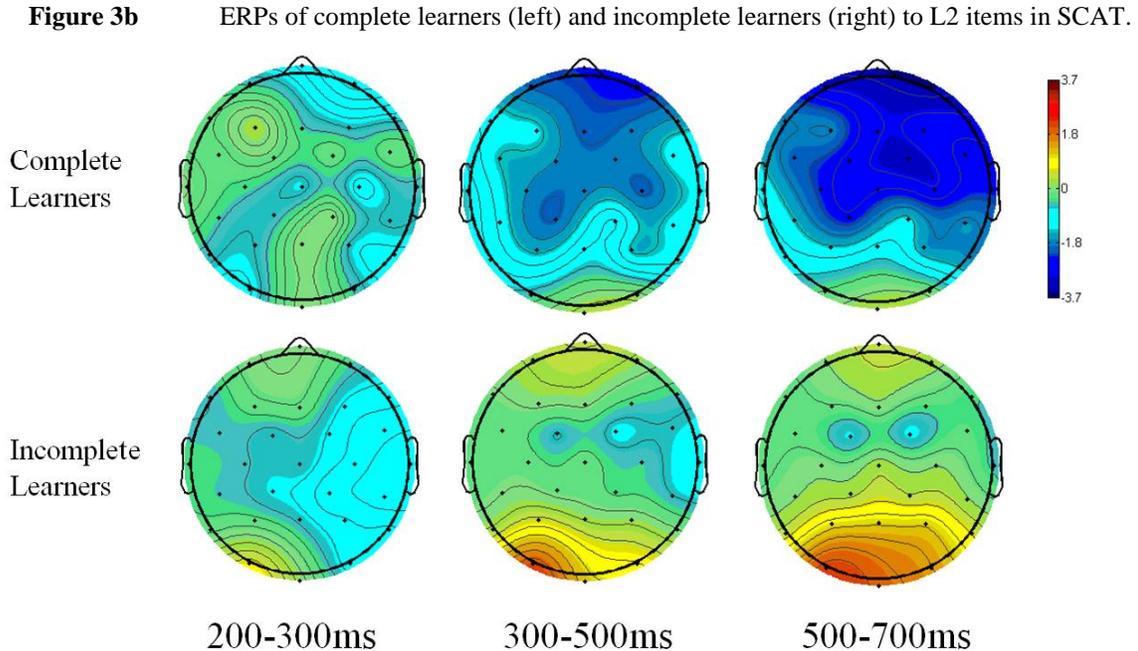


Figure 3c Voltage maps of L2 items in SCAT for complete (top row) and incomplete (bottom row) learners, calculated from differences waves of Session 10 minus Session 1.

200-300ms. A significant interaction between SESSION and HEMISPHERE was observed in this window, with the right hemisphere showing a larger difference than the left hemisphere and the midline sites as the sessions progressed, $F(6,156) = 3.6, p = .008$. In addition, a SESSION x LEARNING-SUCCESS x ANTERIOR-POSTERIOR x HEMISPHERE 4-way interaction was found, $F(24,624) = 2.5, p < .001$. A negative shift at right occipital sites and a positive shift at left occipital sites were seen in the incomplete group, as compared to the complete group which did not exhibit much difference between sessions at these sites.

300-500ms. The four-way interaction of SESSION x LEARNING-SUCCESS x ANTERIOR-POSTERIOR x HEMISPHERE was significant, $F(24,624) = 1.9, p = .006$. The difference was driven by the negative shift at anterior and central sites for complete learners and the positive shift at posterior sites for incomplete learners. This general pattern was right-lateralized for both groups,

as indicated by the interaction with the HEMISPHERE factor.

500-700ms. In this window, SESSION interacted with the ANTERIOR-POSTERIOR factor, with a negative shift at anterior sites and a positive shift at occipital sites, $F(12,312) = 2.9$, $p = .042$. A SESSION x LEARNING-SUCCESS interaction was also found, with complete learners showing more negative waves than incomplete learners, $F(3,78) = 3.8$, $p = .014$. Additionally, this interaction was modulated by the two topographic variables (SESSION x LEARNING-SUCCESS x ANTERIOR-POSTERIOR x HEMISPHERE, $F(24,624) = 1.9$, $p = .007$). The pattern of effects in this window was similar to that of the 300-500ms window, where again complete learners showed a negative shift over the anterior region of the brain and incomplete learners showed a posterior positive effect, both of which were right-lateralized.

Behavioral Data

All participants improved in behavioral performance in SCAT over time, as predicted. A SESSION (Sessions 1, 4, 7, and 10) x LEARNING-SUCCESS (complete vs. incomplete) between-subject ANOVA showed that the main effects of SESSION and LEARNING-SUCCESS were both marginally significant, $F(3,104) = 8.8$, $p = .054$ and $F(1, 104) = 6.5$, $p = .084$ respectively. Follow-up analyses revealed that there were significant improvement between Sessions 1, 4, and 7, but not between Sessions 7 and 10. The complete and incomplete learning groups performed on the same level at Session 1, but the complete learning group performed significantly better for the remaining sessions. The interaction between SESSION and LEARNING-SUCCESS was significant, $F(3, 104) = 4.7$, $p = .004$, indicating that complete learners performed significantly better than incomplete learners at later sessions. For the breakdown of performance by sessions, see Table 3.

Table 3 Mean percentage hits for the SCAT (L2 items) by sessions.

Session Number	1	4	7	10	Mean
Complete Learners:	24.3%	61.0%	81.4%	86.7%	63.3%
Mean % (SD)	(3.2)	(5.4)	(4.8)	(5.3)	(4.7)
Incomplete Learners:	27.4%	37.9%	54.5%	59.8%	44.9%
Mean % (SD)	(3.8)	(5.5)	(6.5)	(6.7)	(5.6)

SCAT L1 Items

The ERP waveforms for the L1 critical items were plotted in Figure 4a. Visual inspection of the waves revealed a small negative component post-onset followed by a positive component at 200ms. A negative-going wave starting at around 200ms peaked near 350ms and the negativity increased by session. It appeared consistently across most of the scalp except for occipital sites, where a positive shift was observed instead. At 500-700ms, there was a positive deflection, and its amplitude was enhanced with time at anterior sites and occipital sites. There were no observable differences between complete and incomplete learners for this analysis.

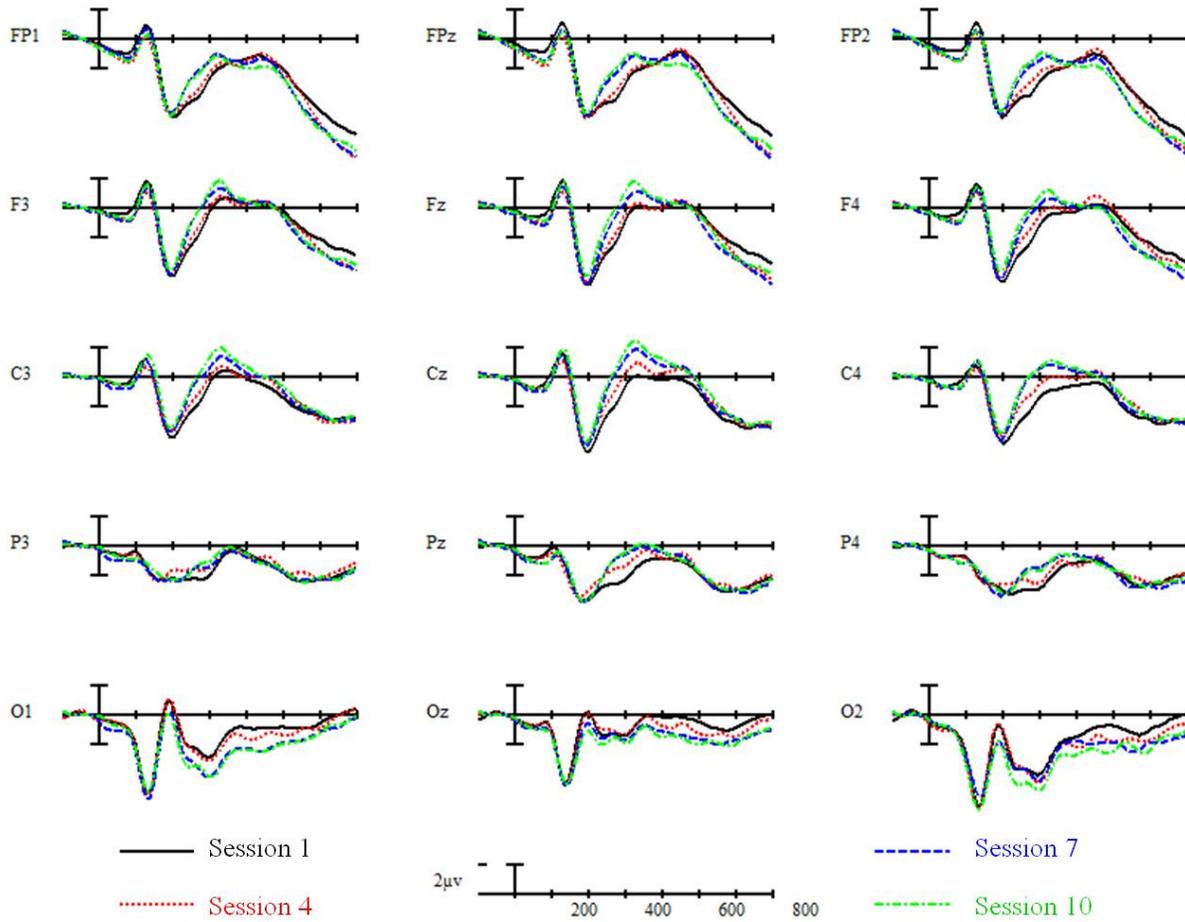


Figure 4a ERPs of English (L1) critical items in SCAT (n=28).

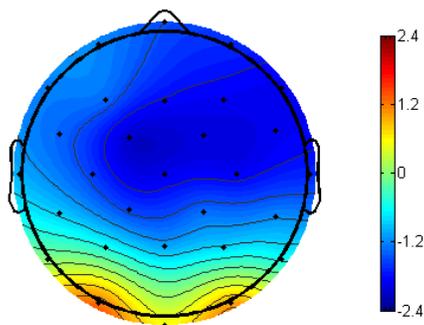


Figure 4b Voltage map of L1 items in SCAT for all learners from 200-350ms, computed from difference waves of Session 10 minus Session 1.

200-350ms. A main effect of SESSION was found in this time window, showing that the

waves were becoming more negative with time, $F(3,78) = 3.1, p = .042$. SESSION interacted with HEMISPHERE, $F(12,312) = 4.5, p = .002$, indicating that the effect was stronger at anterior and central sites than posterior sites. The SESSION x HEMISPHERE interaction was significant as well, showing that the effect was right-lateralized, $F(6,156) = 3.4, p < .001$. See Figure 4b for the voltage map of this effect.

350-500ms. The SESSION x HEMISPHERE interaction was marginally significant in this window, $F(12,312) = 2.2, p = .082$. Although the negative shift over central sites and positive shift over occipital sites continued, the effect dissipated at other sites.

500-700ms. No significant result was found in this window.

Behavioral Data

The participants performed at an average of 92.7% hit accuracy in the English version of the SCAT over the 4 ERP sessions. Complete and incomplete learners did not perform at a different level ($p > .05$ in all cases), although LEARNING-SUCCESS was trending towards significance ($F(1, 104) = 6.2, p = .089$). Examination of the data suggested that the complete learners were performing better than incomplete learners in later sessions (see Table 4 for details of the behavioral results).

Table 4 Behavioral data for the SCAT (English block) by session.

Session Number	1	4	7	10	Mean
Complete Learners:	93.8%	96.0%	96.0%	92.9%	94.6%
Mean % (SD)	(2.0)	(1.3)	(1.4)	(3.3)	(2.0)
Incomplete Learners:	93.8%	92.1%	91.9%	85.2%	90.8%
Mean % (SD)	(2.3)	(2.4)	(2.3)	(5.5)	(3.1)

SCAT Language Effects

The contrasts between the ERPs generated by L1 (English) and L2 (Chinese) critical items in Sessions 1 and 10 are plotted in Figure 5a and Figure 5b respectively. Voltage maps calculated by subtracting ERPs of L2 items from L1 items were presented in Figure 5c. Visual inspection of Figure 4a revealed that the waveforms of L1 and L2 started diverging from 200ms post-stimulus onset. The first positively-going wave peaking near 200ms (P2) was larger for L2 items than L1 items. The difference in P2 was especially prominent at posterior sites. In the next negative deflection in the 300 to 500ms window (N400), L2 had later onset and smaller peak amplitude compared to L1. After 500ms, L2 items generally evoked slightly more negative waves than L1 items in the anterior regions. Examination of Figure 5c showed that the LANGUAGE effects were similar between Sessions 1 and 10. The main difference appeared to occur at the late anterior effect between 500 and 700ms, where L2 was much more negative than L1 at Session 10 compared to Session 1.

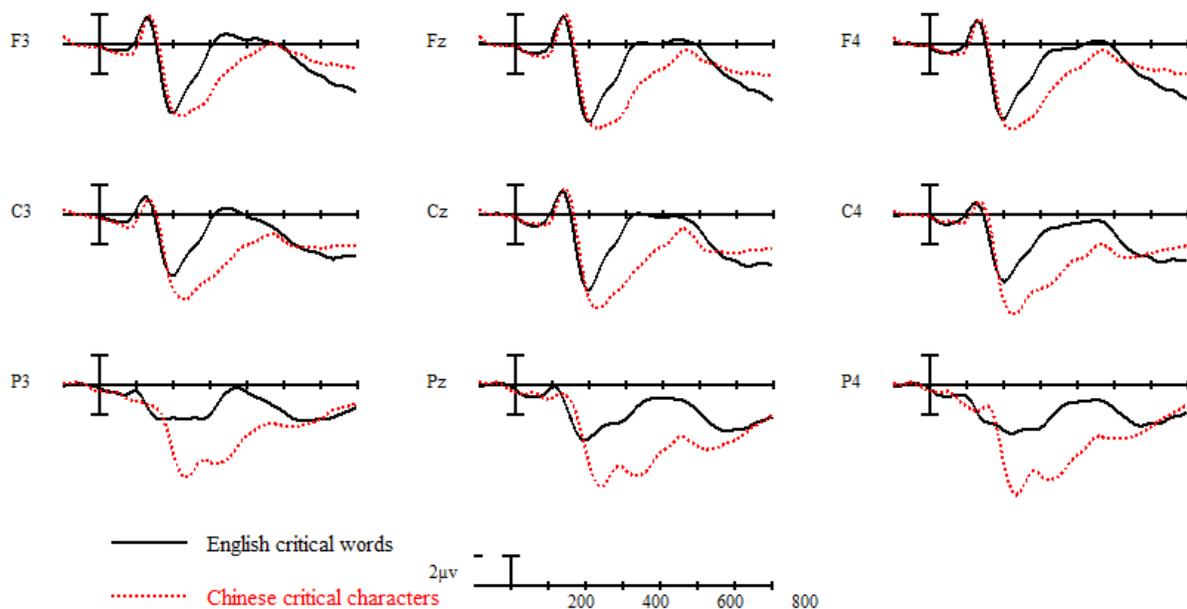


Figure 5a Language effects in SCAT in Session 1 for all learners (n=28).

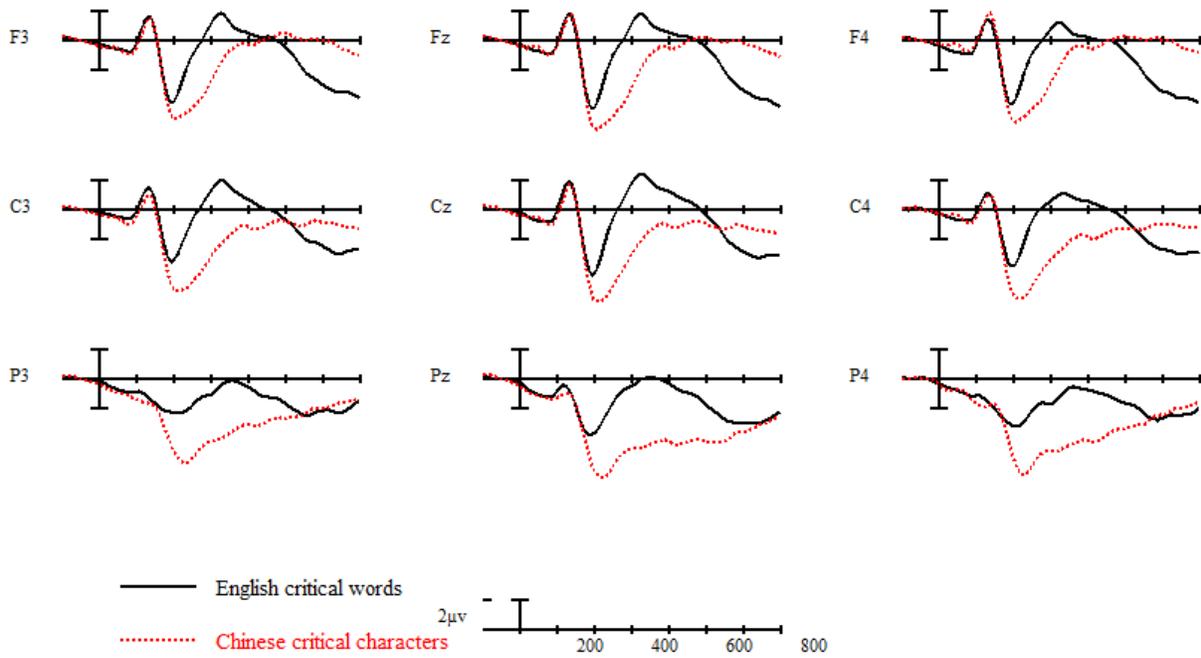


Figure 5b Language effects in SCAT in Session 10 for all learners (n=28).

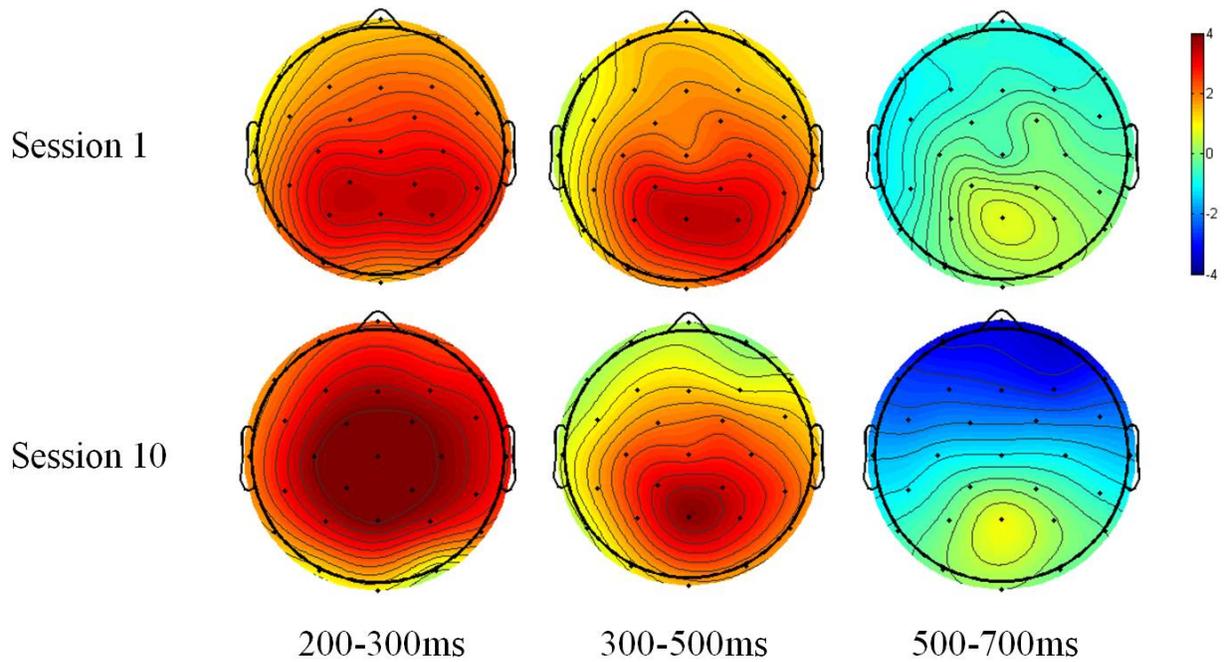


Figure 5c Voltage maps of the language effect in SCAT in Session 1 (top row) and 10 (bottom row) computed by subtracting ERPs of L1 items from ERPs of L2 items.

200-300ms. In this epoch, a main effect of LANGUAGE was found ($F(1, 26) = 8.2, p = .008$), with L1 items being more negative than L2 items. The factor of LANGUAGE also interacted with scalp sites (LANGUAGE x ANTERIOR-POSTERIOR, $F(4,104) = 34.8, p < .001$; and LANGUAGE x ANTERIOR-POSTERIOR x HEMISPHERE, $F(8, 208) = 3.9, p < .001$). The interaction of LANGUAGE x ANTERIOR-POSTERIOR suggested that the difference between L1 and L2 was larger at posterior sites than anterior sites. The 3-way interaction of LANGUAGE x ANTERIOR-POSTERIOR x HEMISPHERE further showed that the difference was most evident at right posterior sites. The SESSION factor interacted with LANGUAGE and scalp sites as well (SESSION x LANGUAGE x ANTERIOR-POSTERIOR, $F(12, 312) = 5.2, p < .001$; and SESSION x LANGUAGE x HEMISPHERE, $F(6, 156) = 3.0, p = .009$). The two interactions pointed to stronger difference at anterior and central sites but diminished difference at occipital sites, as well as a reduced right-lateralization.

300-500ms. A main effect of LANGUAGE was observed ($F(1,26) = 36.71, p < .001$) with L1 items being more negative than L2 items. This was modulated by scalp sites (LANGUAGE x ANTERIOR-POSTERIOR, $F(4,104) = 27.09, p < .001$; LANGUAGE x HEMISPHERE, $F(2,52) = 10.6, p < .001$; and LANGUAGE x ANTERIOR-POSTERIOR x HEMISPHERE, $F(8, 208) = 8.5, p < .001$). The difference between L1 and L2 waves was most pronounced at posterior midline sites. A significant SESSION x LANGUAGE x HEMISPHERE interaction was found, $F(6, 156) = 2.5, p = .026$. Examination of Figure 4c showed that the language effect was initially more right-lateralized, but it became less so over the sessions.

500-700ms. A main effect of LANGUAGE was found ($F(1,26) = 8.2, p = .008$) with L1 items being more positive than L2 items in this epoch. There were interactions of LANGUAGE and scalp

distribution (LANGUAGE x ANTERIOR-POSTERIOR, $F(4,104) = 20.9, p < .001$; and LANGUAGE x ANTERIOR-POSTERIOR x HEMISPHERE, $F(8, 208) = 3.6, p < .001$). The LANGUAGE x ANTERIOR-POSTERIOR interaction indicated that L1 items were more positive-going than L2 items only at the front of the head but were still slightly more negative at the back. Additionally, the 3-way interaction of LANGUAGE x ANTERIOR-POSTERIOR x HEMISPHERE showed that the residual posterior negativity was focused on the midline sites. The SESSION factor interacted with LANGUAGE and ANTERIOR-POSTERIOR, $F(12,312) = 2.5, p = .005$, indicating that the anterior difference between L1 and L2 items became stronger over the sessions, as can be seen in Figure 5c.

Discussion

L2 Items

In this short longitudinal study of L2 vocabulary learning, participants learned a set of 200 vocabularies within ten lab visits. Results from this experiment showed that there were systematic improvements in task performance in both translation and categorization of L2 items over a short period of laboratory learning. The behavioral data suggested that L2 learning of visually distinct symbols could occur rapidly in a controlled environment in initially L2-naive participants. This was in accordance with our predictions and some previous findings. We further examined whether there were concurrent neural changes that could be the neural markers of L2 lexical learning. A number of ERP patterns emerged over the four recording sessions, which could be related to both familiarity with the stimuli and learning effects.

Early posterior effects similar to that observed in the pilot study (see p.11) were seen in both tasks. However, it was modulated by LEARNING-SUCCESS, where only incomplete learners

exhibited changes at right occipital sites. For incomplete learners, an enhanced occipital component occurred in the 4th and 7th sessions, but by the 10th session the difference had diminished. This suggested that this process was helpful in the initial stage of learning these L2 items, and with sufficient proficiency with these items, it was no longer necessary. The complete learners did not show this pattern of results at all, perhaps indicating that they had used a different strategy from the beginning. This effect suggested a change in the very early stage of visual learning, before successful L2 word acquisition that eventually all learners reached. Vaid and Frenck-Mestre (2002) reported a processing bias for the right hemisphere in a group of French-English bilinguals when identifying orthographically marked L2 words. They interpreted this as perceptual processing in L2 visual word identification. The current effect might also be related to perceptual processing of sub-lexical units. Since Chinese items could be broken down into smaller parts called radicals that convey information, it seemed likely that L2 learners would engage in this kind of visual analysis early in the course of processing. However, this process was not necessary for word identification because holistic recognition was also possible given the repeated exposure to the same words. Indeed, coarse-grained whole-word processing might be more efficient for task performance because task demands did not require retrieval of orthographic information, which might explain this pattern of results.

In the N400 epoch, there was a significant interaction between sessions in complete and incomplete learner groups for both the Backward Translation Task and the SCAT. As predicted, the peak amplitude of the N400 at central and posterior sites became more negative with time in complete learners. Successful L2 vocabulary learning would strengthen the L2 items and forge links to both other L2 items and the corresponding L1 items. We interpreted the stronger N400 response as increased semantic processing due to activation of these semantic networks. In

incomplete learners, the N400 had the same peak amplitude over sessions, but the peak latency shifted about 100ms earlier so that it peaked at around 350ms in Session 10 as compared to 450ms in Session 1. An earlier N400 could imply that the learners were taking a shorter time to recognize whether they were able to translate the L2 items. While the behavioral data clearly indicated that learning had taken place, the amplitude of the N400 did not change significantly in incomplete learners.

The pattern of an anterior negative shift in complete learners and a posterior positive shift in incomplete learners at 300-500ms was consistent across tasks. This suggested that these components reflected general L2 word processing mechanisms, but these processes were very different between the groups. Participants were generally matched in education and reading levels, however, individual differences might not be apparent in L1 because of their high expertise in it. These different learning trajectories in learner groups probably reflected several types of individual differences in second language learning. The ability to quickly encode the visual and semantic information of the L2 items is related to working memory capacity, which is known to vary among individuals and can affect language comprehension (Daneman & Carpenter, 1980). We took care that the L1-L2 paired associations were presented at each evenly-spaced experimental session, so the stimuli should have been re-encoded the same number of times across a similar time span. However, the capacity to retain information could also differ for each learner so that a participant who had a longer retention interval would be at an advantage in task performance. These effects have been observed in laboratory word-training experiments that lasted more than a week (Healy & McNamara, 1996). Due to the relative difficulty in recognizing newly-acquired L2 vocabulary, it was likely that working memory capacity and long-term memory retention interval added to the difference observed between

complete learners and incomplete learners. The current study did not independently measure these cognitive abilities, but future investigations should include these as possible covariates.

Recent evidence showed that being bilingual confers advantage in learning novel words with unfamiliar phonological features (Kaushanskaya & Marian, 2009). To minimize prior experience of learning a new orthography, potential participants who were early bilinguals, fluent in their second language, or had learned non-alphabetic languages were excluded from participation. However, some participants had classroom L2 learning experience in other alphabetic languages and some participants were monolingual. Task differences may also have contributed to the observed differences. In the Backward Translation Task, the participants needed to identify the L2 word to retrieve the exact lexical entry to respond and verbally translate. In the SCAT, participants could respond correctly as long as they could identify whether the target word belonged to the specific category. This discrepancy in task demands affected the behavioral performance, but did not appear to have pronounced effects on the waveforms of these L2 words.

L1 Items

For the L1 critical items in SCAT, a significant difference was found in the 200-350ms epoch, with an increasing negativity over time at most sites. The difference was stronger in anterior and central regions and right-lateralized. This unexpected longitudinal change in the negative component could reflect a shift in latency and amplitude in the N400 component due to repeated exposure to the target stimuli. Our prediction was that with time, learners would gain a higher familiarity of the items, and something akin to a frequency effect might emerge. However, that usually manifest as a reduction in the N400 component, which was the opposite of what was found. In fact, this increase in negativity in the L1 items was similar to the change in L2 items in

the same semantic categorization task, except this effect occurred earlier in the time course. One possible interpretation was that L2 word acquisition had affected how their L1 translations were processed. In building up semantic associations for L2 words, the same semantic links were augmented in the L1 words. This could create a larger semantic network for these L1 words, thus resulting in more activation and enhancing the N400 component. If this was the case, it would offer strong support for the interactive nature of languages in a bilingual because this task was purely in L1 and did not require translation or other bilingual processes.

No difference was observed between complete and incomplete learning groups for this early effect, suggesting that the change was not correlated with success of lexical learning, but perhaps reflected more automatic processes. It had been shown in previous research that L1 and L2 lexical interference effects exist in non-cognates and languages that belonged to different writing systems. A recent study (Takahashi, Suzuki, Shibata, Fukumitsu, Gyoba, et al., 2010) showed a similar latency shift in the N400 in Japanese kindergartners who were exposed to English as an L2 compared to those who were monolingual. Although the populations and the tasks are very different, it seems likely that there are mechanisms that change L1 processing without high L2 proficiency. A proper control experiment with a group of monolingual participants exposed to the same stimuli without paired learning of L1 and L2 words would confirm whether the L1 effect was in fact caused by novel L2 vocabulary learning. Another recent study by Thierry and Wu (2004) showed interference effects in pure L2 lists in late bilingual learners. Their manipulation was more subtle because participants were not aware that the study had a bilingual component at test. Compared to that found in Thierry and Wu's (2004) study, this effect was seen in speakers that had much less experience in the L2. This could have important implications for the interconnectedness of languages in a bilingual, showing that

languages interact at a noticeable level even at the very beginning of L2 learning. However, in the current study, the L1 stimuli were shown repeatedly in relation to the L2 translations, so strategies and explicit memory processes might have played a role in moderating this change in L1.

Language Effects

Unlike the Backward Translation task, where the L1 and L2 items were processed in different contexts, the blocked design of the SCAT allowed a direct comparison of L1 and L2 in the same language users. Significant differences between the languages were found quite early in word processing, beginning at around 200ms post-target onset. The P2 component (P200/N200 in Liu et al.'s (2006) study) was larger in L2 items than L1 items. Unlike the results from Liu et al. (2006), the P2 component in anterior region was stronger in later sessions. Given that their sample had one year of L2 learning and ours had much less, this could reflect the development of lexical processing as shown by the P2 component. In contrast, the occipital negativity in this window was reduced in later sessions, which perhaps indicated less visual processing resulting from experience with the L2 visual wordform.

L1 was significantly more negative than L2 from 300ms to around 500ms at all sites but the effect was strongest in central and posterior regions. The N400 result replicated the findings of Midgley et al. (2009) where less proficient English-French and French-English bilingual speakers both showed similar effects. The current study extended their results by showing converging evidence with Chinese as an L2 of a different writing system. Since the two languages are dissimilar, one might expect very different language processes in these initial learners. The current effects were particularly notable because they occurred at a very early time point in L2 vocabulary learning (~16hrs), without formal instructions. Additionally, we found

that another language effect at this initial period of learning in a late anterior effect occurring around 500-700ms that increased in strength in later sessions. This effect might indicate development of a later-onset N400-like component, since L2 elicit a later N400 in general.

Conclusions

In sum, a group of initially naive native English speakers learned Chinese characters through 10 automated lab sessions. A number of ERP components measured during the course of the study showed changes in response to learning. An early posterior positive effect was found for the newly acquired L2 words, likely indicating increased visual processing. Later anterior negativity was evoked by L2 items in both tasks, especially in complete learners. A general pattern emerged between learners of varying degrees of learning success, with an anterior negative shift in complete learners and a posterior positive shift in incomplete learners. The difference could be related to strategy use or individual differences that affect language learning ability. Even in a controlled environment, there appeared to be many factors that could affect learning performance at this early stage of learning vocabularies in a new L2. Compared to L1 words, L2 words evoked an enhanced P2 component and a smaller N400. A late anterior effect around 500-700ms was also observed at later sessions. These language effects were as predicted and showed a similar pattern to L2 learners of other alphabetic languages, suggesting that these learning effects could be extended to the learning of a different writing system. Finally, it was found that processing of L1 items changed from the by learning the semantic equivalents in a new L2 that belonged to a different writing system and did not share orthographic or phonological features. This suggested that languages that a person knows are not stationary or isolated but start to interact with each other with minimal contact.

There are convincing behavioral and ERP changes in the current laboratory learning study showing changes in initial L2 vocabulary acquisition. However, the current study only employed a set of vocabularies in the visual modality, whereas natural language learning relies heavily on phonological and interactive processes. It will be an interesting question to determine whether laboratory learners could take advantage of additional phonological information to cement their L2 knowledge. Questions remain also on whether learners in the laboratory learning paradigm acquired L2 words in the same way as those in a traditional classroom. There are speculated differences due to the method of obtaining L2 information, but how these might affect initially naïve learners need further empirical testing. As individual differences appear to emerge in this relatively homogeneous group, cognitive and motivational factors that influence L2 vocabulary learning should be identified and explored as well. Language learning and usage are both highly dynamic processes that change with experience even in adulthood; further studies in how the changes occur will benefit the understanding of bilingualism and L2 learning.

References

- Abutalebi, J., Cappa, S. F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, *4*, 179–190.
- Baddeley, A. D., Gathercole, S. & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158-173.
- Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping*, *25*, 92–104.
- Chen, H-C., Flores d'Arcais, G. B., & Cheung, S-L. (1995). Orthographic and phonological activation in recognizing Chinese characters. *Psychological Research*, *58*. 144-153.
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, *19*, 450-466.
- De Groot, A. M. B. & Keijer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign language vocabulary learning and forgetting. *Language Learning*, *50*, 1-56.
- De Groot, A. M. B. & Van Hell, J. G. (2005). The learning of foreign language vocabulary. In J.F. Kroll & A.M.B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9-29). New York, NY, USA.
- Finkbeiner, M. & Nicol, J. (2003). Semantic category effects in second language word learning. *Applied Psycholinguistics*, *24*, 369-383.
- Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2010). Lexical quality in the brain: ERP evidence for robust word learning from context. *Developmental Neuropsychology*, *35*, 376-403.

- Geisser, S., & Greenhouse, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95-112.
- Grainger, J. & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistic Compass*, 3, 128-156.
- Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult language learning after minimal exposure to an unknown natural language. *Language Learning*, 60, 5-24.
- Kaushanskaya, M. & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, 16, 705-710.
- Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 952-965.
- Liu, Y., Wang, M., & Perfetti, C. A. (2007). Threshold-style processing of Chinese characters for adult second-language learners. *Memory & Cognition*, 35, 471-480.
- Liu, Y., Dunlop, S., Fiez, J., & Perfetti, C. A. (2007). Evidence for neural accommodation to a writing system following learning. *Human Brain Mapping*, 28, 1223-1234.
- Liu, Y., Perfetti, C. A., & Wang, M. (2006). Visual analysis and lexical access of Chinese characters by Chinese as second language readers. *Linguistics and Language*, 7, 637-657.
- May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends in Cognitive Sciences*, 15, 475-482.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word

- learning: Minimal instruction produces rapid change. *Nature Neuroscience*, 7, 703-704.
- Healy, A. F., & McNamara, D. S. (1996). Verbal learning and memory: Does the modal model still work? *Annual Review of Psychology*, 47, 143-172.
- Mestres-Missé, A., Rodriguez-Fornells, A., & Munte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, 17, 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. *Journal of Neurolinguistics*, 22, 281-300.
- Nation, I. S. P. (1993). Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon*. (pp. 115-134). Amsterdam: Benjamins.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nelson, J. R., Liu, Y., Fiez, J., & Perfetti, C.A. (2009). Assimilation and accommodation patterns in ventral occipitotemporal cortex in learning a second writing system. *Human Brain Mapping*, 30, 810-820.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15, 202-206.
- Perfetti, C. A. & Liu, Y. (2005). Orthography to phonology and meaning: Comparisons across and within writing systems. *Reading and Writing*, 18, 193-210.
- Perfetti, C. A., Liu, Y., Fiez, J., Nelson, J., Bolger, D. J., & Tan, L-H. (2007). Reading in two writing systems: Accommodation and assimilation in the brain's reading network. *Bilingualism: Language and Cognition*, 10, 131-146.

- Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1281-1292.
- Sagara, N. & Alba, M. (2006). The Key Is in the Keyword: L2 Vocabulary Learning Methods with Beginning Learners of Spanish. *The Modern Language Journal*, *90*, 228-243.
- Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology*, *4*, 293–310.
- Sánchez-Casas, R. M. & García-Albea, J. E. (2005). The representation of cognates and noncognates words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? In J. F. Kroll and A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226–250). New York: Oxford University Press.
- Stein, M., Federspiel, A., Koenig, T., Wirth, M., Lehmann, C., Wieest, R., Strik, W., Brandeis, D., & Dierks, T. (2009). Reduced frontal activation with increasing L2 proficiency. *Neuropsychologia*, *27*, 2712-2720.
- Takahashi, J., Suzuki, Y., Shibata, H., Fukumitsu, Y., Gyoba, J., Hagiwara, H., & Koizumi, M. (2010). Effects of non-native language exposure on the semantic processing of native language in preschool children. *Neuroscience Research*, *69*, 246-251.
- Tan, L-H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping*, *25*, 83-91.

- Thierry, G. & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *NeuroReport*, 15, 1555-1558.
- Vaid, J. & Frenck-Mestre, C. (2002). Do orthographic cues aid language recognition? A laterality study with French-English bilinguals. *Brain and Language*, 82, 47-53.
- Van Hell, J. G. & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9, 780-789.
- Wang, M., Perfetti, C. A., & Liu, Y. (2003). Alphabetic readers quickly acquire orthographic structure in learning to read Chinese. *Scientific Studies of Reading*, 7, 183-208.