

An ERP study on initial second language vocabulary learning

YEN NA YUM,^a KATHERINE J. MIDGLEY,^a PHILLIP J. HOLCOMB,^a AND JONATHAN GRAINGER^b

^aPsychology Department, Tufts University, Medford, Massachusetts, USA

^bLaboratoire de Psychologie Cognitive, CNRS & University of Provence, Marseille, France

Abstract

This study examined the very initial phases of orthographic and semantic acquisition in monolingual native English speakers learning Chinese words under controlled laboratory conditions. Participants engaged in 10 sessions of vocabulary learning, four of which were used to obtain ERPs. Performance in behavioral tests improved over sessions, and these data were used to define fast and slow learners. Most important is that ERPs in the two groups of learners revealed qualitatively distinct learning patterns. Only fast learners showed a left-lateralized increase in N170 amplitude with training. Furthermore, only fast learners showed an increased N400 amplitude with training, with a distinct anterior distribution. Slow learners, on the other hand, showed a posterior positive effect, with increasingly positive-going waveforms in occipital sites as training progressed. Possible mechanisms underlying these qualitative differences are discussed.

Descriptors: Second language learning, Vocabulary acquisition, ERPs, Chinese characters

Vocabulary knowledge forms the basis of spoken and written language comprehension from the earliest point of second language (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 a first language (L1), where participants learned a set of low frequency vocabulary items or pseudowords in which the phonotactic rules of their native language were preserved (Batterink & Neville, 2011; Borovsky, Kutas, & Elman, 2010; Frishkoff, Perfetti, & Collins-Thompson, 2010; Mestres-Missé, Rodríguez-Fornells, & Münte, 2007). These studies found that novel words for which some meaning had been acquired showed behavioral and neuronal changes approaching that of familiar L1 words.

However, integration of new vocabulary items into an existing lexicon is not the same as learning words in a new language. The former is relatively easy, because the orthographic and phonological systems used to process new words are familiar to the learners. In contrast, any second language learning will necessarily involve building up new orthographic, phonotactic, and grammatical rules. A number of studies have obtained longitudinal data in order to observe changes in L2 lexical processing that occur with increasing L2 proficiency. One study followed English native speakers learning German, before and after 5 months of immersion in a German-speaking environment (Stein et al., 2006). Following immersion, German words elicited shorter latency N400 and P600 (late positive component [LPC] in their paper) components compared with

the latencies elicited prior to immersion. Amplitude differences in these event-related potentials (ERP) components were not observed.

More directly relevant to the present work is a study by McLaughlin and colleagues that concentrated on the very initial period of L2 learning (McLaughlin, Osterhout, & Kim, 2004). McLaughlin et al. showed that the N400 component is sensitive to the level of exposure to a new L2 very early in L2 learning. In a group of native English speakers learning French as L2, the N400 reliably differentiated learned L2 items from L2-based pseudowords after only 14 h of classroom instruction, and in the absence of any significant differences in behavior. After 63 h of instruction, newly learned items showed N400 attenuation when primed semantically, implying the activation of semantic representations. The results of this study suggest that major changes are already occurring during the very first stages of L2 learning, and that it is therefore important to understand exactly what these early changes involve. Moreover, the ERP changes were seen prior to changes in behavioral measures, thus suggesting that this is a more sensitive technique for studying effects that emerge early during L2 word learning.

Many of the studies of L2 vocabulary acquisition conducted to date involved an L1 and an L2 that share an alphabet. In these cases, L2 learners are likely to use their prior knowledge of the alphabet and the L1 vocabulary to support L2 word learning. Some evidence of this comes from studies of the cognate effect. Cognates, which are words that have the same meaning and similar orthographic or phonological representations across languages, are easier to learn and recall than noncognates (De Groot & Keijzer, 2000). But even noncognate words will be influenced by the overall similarity between two languages. One strategy for limiting such possible influences is to use language pairs that are visually and phonologically distinct, such as English and certain Asian

This research was supported by NIH grants number HD25889, HD043251, and ERC grant number 230313.

Address correspondence to: Yen Na Yum, Room 804A, 8/F Meng Wah Complex, University of Hong Kong, Pokfulam, Hong Kong. E-mail: yumyenna@gmail.com

languages. In their longitudinal study, Ojima, Nakamura, Matsuba-Kurita, Hoshino, and Hagiwara (2011) used this approach by examining Japanese children learning English over a several-year period and found evidence of a developmental progression of increasingly nativelike N400 effects as a function of L2 proficiency.

A number of other studies have investigated the learning of Chinese characters by native English speakers (Liu, Dunlap, Fiez, & Perfetti, 2007; Liu, Perfetti, & Wang, 2006; Liu, Wang, & Perfetti, 2007; Perfetti et al., 2007; Wang, Perfetti, & Liu, 2003). This research points to a rather rapid adaptation to the unfamiliar script with the early formation of structural representations of Chinese characters being followed by the gradual strengthening of their associations with phonology and meaning. For example, Liu, Wang, & Perfetti (2007) examined priming effects in first-term L2 Chinese learners in a word naming task with single character words and a 500-ms stimulus onset asynchrony (SOA; primes presented for 500 ms and immediately replaced by targets until verbal response). They found facilitation from orthographically related primes but no phonological or semantic priming. By the end of the second term, however, the effect of orthographic priming diminished while phonological and semantic priming both facilitated naming. Furthermore, Wang et al. (2003) showed that adult native English speakers with a few months of college-level classroom learning of Chinese were sensitive to orthographic structure, responding differently to real characters, characters with real radicals in illegal positions, and characters with fake radicals. Finally, an ERP study by Liu et al. (2006) provided evidence for effects attributed to visual and lexical processing after one term of tuition in Chinese, with the visual effects diminishing by the end of the second term.

All of the above studies involved classroom learners tested after at least one term of learning. Perhaps the clearest evidence to date for very rapid learning of Chinese characters was provided by Liu, Dunlap et al. (2007) in a laboratory study where native English speakers with no prior knowledge of Chinese were taught the pronunciation and/or the meaning of 60 characters. Twenty-three out of 29 participants showed successful learning and reached 92% accuracy in naming characters and making semantic category judgments after receiving about 6 h of training over a period of 3 days. Most important is that, after such training, these participants showed a pattern of bilateral activation in fMRI in occipital and fusiform areas that is characteristic of skilled Chinese readers, and therefore a sign that these participants had very rapidly adapted to the specificities of Chinese orthography.

Current Study

The central goal of the current study was to provide a more detailed analysis of the mechanisms involved in the very first stages of L2 vocabulary acquisition when L2 and L1 use very different writing systems. Chinese was chosen as the L2 for native English speakers, because it is a language of high contrast to English with minimal overlap in both written and spoken forms. Data on L2 vocabulary acquisition were obtained during the earliest points of learning under rigorously controlled learning conditions. Rather than studying students in university second language courses, which is generally a population with high variability in terms of language background as well as varying degrees of involvement in the learning process, in this study we recruited participants who had not had prior exposure to the to-be-learned language (or any related language), and we used a uniform laboratory teaching regimen that

precisely controls exposure to L2 across learners. Participants received 10 training sessions in the lab, which was meant to roughly correspond to the level of L2 vocabulary exposure during the first semester of a language course. Four of the sessions had concurrent ERP recordings. Training tasks were designed to focus on different linguistic information so participants could quickly acquire both visual and semantic familiarity to the new L2. Exposure to the new L2 was linear over the 10 sessions because the initially non-Chinese-speaking participants viewed carefully counterbalanced stimuli at regular intervals. The larger scale of the study allowed us to categorize participants according to their behavioral performance, thus enabling an investigation of individual differences in L2 lexical learning.

In the present study, we first focus on the N170 ERP component as an index of orthographic processing skill. The amplitude and lateralization of the N170 in response to visual words has been found to reflect expertise in script-specific processing. Specifically, patterns of lateralization for the N170 differ for alphabetic and syllabic scripts, compared to logographic scripts. Alphabetic stimuli systematically elicit left-lateralized N170 responses in skilled readers (Maurer, Brandeis, & McCandliss, 2005; Rossion, Joyce, Cottrell, & Tarr, 2003), whereas reports for N170 responses for logographic stimuli such as Chinese characters have been mixed, including findings of left lateralization (Lin et al., 2011; Maurer, Zevin, & McCandliss, 2008) and bilateral or right-lateralized distribution (Lee et al., 2007; Liu & Perfetti, 2003; Yum, Holcomb, & Grainger, 2011). Left lateralization of the N170 has been argued to reflect phonological recoding of orthographic units (Maurer et al., 2008; Lee et al., 2007), in addition to perceptual expertise alone (Lin et al., 2011). On the other hand, bilateral N170 has been reported for both L2 readers of alphabetic and nonalphabetic languages (Kim, Yoon, & Park, 2004; Maurer et al., 2008; Proverbio, Cok, & Zani, 2002). It is currently unclear how L2 proficiency and tuning of script-specific processing might affect the emergence and lateralization of the N170 component in L2 learners.

In order to examine processing at the semantic level, in the present study we also focused on N400 changes in learners. Much of the literature on the N400 has examined the influence of contextual manipulations such as the difference between related and unrelated words in a priming task or, as originally reported, in sentences with semantically anomalous ending words (Kutas & Hillyard, 1980). However, it is clear that in most language contexts all content words generate N400s as can be seen for words at the beginning of a sentence prior to contextual buildup (e.g., Van Petten & Kutas, 1991) and to single words without a priming context (e.g., Midgley, Holcomb, & Grainger, 2009). The N400 elicited by single words has been argued to reflect the mapping of lexical form onto semantics (Holcomb & Grainger, 2006). Here, we focus on the N400 that results when a task requires semantic assessment of a word and that word is not supported by context, which is typically the condition that produces the large negativity in N400 priming studies. In bilingual studies, it has been shown that single L2 words elicit smaller N400 than L1 words, and L2 words also generate larger N400 in high proficiency L2 users compared to intermediate L2 learners with lower L2 proficiency (Midgley et al., 2009). Thus, N400 amplitude might be used as an index of L2 proficiency in learners.

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). Several tasks, described below,

were used to expose, train, and assess learners in a vocabulary of 200 L2 items. ERPs were recorded in a subsequent semantic categorization task designed to assess participants' passive reading of their new L2 vocabulary. On the basis of prior research focusing on the earliest stages of L2 vocabulary acquisition in alphabetic and logographic languages, we expect to see rapid modulation of N400 amplitude as a function of exposure to L2 words (McLaughlin et al., 2004), accompanied by ERP evidence for developing sensitivity to the structural configurations of the newly learned Chinese words (Liu, Wang, & Perfetti., 2007), possibly in the form of changes in the N170 component (Maurer et al., 2008). Our study will specifically focus on how individual differences in learning speed might differentially affect the acquisition of form and meaning representations of a new vocabulary in an unfamiliar script. To do so, we will compare behavioral and ERP changes over learning in two groups of participants that are defined, according to behavioral assessment, as fast or slow learners.

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 were permitted. These participants were recruited by using advertisements that sought "people who are interested in starting to learn Chinese but have never done so." All participants were right-handed university students with normal or corrected-to-normal visual acuity and a normal neurological profile.

Apparatus and Data Analysis

For ERP recordings, participants sat in a comfortable chair in a sound-attenuating room while 32 channels of electroencephalogram (EEG) data (all electrodes < 5 KOhm) were digitized continuously (SA Instruments amplifiers, 6 db cutoffs .01 and 40 Hz, 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° vertical and horizontal). Participant responses were collected with a multibutton response box. Average ERPs were formed from artifact-rejected single trials of EEG, and only trials with correct responses were included in average ERPs.

Two approaches to analyzing the resulting averaged ERPs were taken. In keeping with the norm in studies of the N170, in one set of analyses ERPs were referenced to the average of the 29 scalp sites (i.e., average reference—Joyce & Rossion, 2005; Yum et al., 2011). The resulting ERP data were measured by calculating mean amplitudes between 160–210 ms. Conversely, in keeping with the norm of studies focusing on the N400 component, in a second set of analyses the ERP data were referenced to the average of the two mastoid electrodes (mastoid reference—Yum et al., 2011). The resulting mastoid reference ERP data were measured by calculating the mean amplitudes between 300–500 ms. Data were analyzed with repeated measures and mixed design analysis of variance (ANOVA). Both sets of analyses included two within-subject factors of scalp distribution: anterior-posterior (prefrontal vs. frontal vs. temporal-parietal vs. occipital for average reference and prefrontal vs. frontal vs. central vs. parietal vs. occipital for mastoid reference) and laterality (left vs. right for average reference and left vs. midline vs. right for mastoid reference). Also

included in the analyses were the within-subject factor of session (Sessions 1 vs. 4 vs. 7 vs. 10) and the between-subject factor of learning rate (fast vs. slow). The Greenhouse-Geisser 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 was selected from vocabulary words taught in first-semester Chinese language classes at Tufts University. The L2 items consisted of either one (75%) or two (25%) Chinese characters written in simplified Chinese script. One-character items were also classified for their structural complexity by doing a median split based on the number of strokes they contained (low stroke number: mean = 5.1, range = 2–7; high stroke number: mean = 10.0, range = 8–13). All L1 items consisted of common English words, with mean word length of 5.0 letters (range: 2–11) and mean frequency of 415 occurrences per million (range: 2–11,437) according to the Celex English database (<http://celex.mpi.nl/>). The mean concreteness of the items was 4.59 on a scale of 1 to 7 (range: 1.95–6.53), as assessed by the MRC psycholinguistics database, but only 80% of the words were in the database (Wilson, 1988). Stimuli included 100 filler words that participants did not learn in both L1 and L2; these appeared in the N-back and translation recognition task. The distracter items were controlled for word length and lexical frequency. 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 10 automated sessions with the primary goal of learning a total of 200 words in the new L2. Four tasks were used for training and assessment of learning on all 10 sessions. On four of these lab visits (days 1, 4, 7, and 10), in addition to vocabulary instruction and behavioral testing, ERPs were collected with an additional task. On the other six visits, only L2 instruction and behavioral testing were conducted, and ERP data were not collected.

Training and behavioral assessment. The first task was a go/no-go N-back word task used to familiarize the learners with the visual word forms to be trained. The 200 to-be-learned L2 words were mixed with 100 not-to-be learned L2 words and presented one at a time visually (500 ms L2 word on, 1,500 ms intertrial 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 orthography 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 found that performance on this task was reasonable even during the first session.

The second task was 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 words was paired with its L1 translation in L1-L2 association (500 ms L1 word, 500 ms blank, 1,000 ms L2 word, and 2,500 ms ITI). Participants did not have to respond to the words, but were encouraged to pay attention. Blinking was allowed during the ITI. In a 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, 800 ms feedback, and 2,500 ms ITI). Participants were instructed to indicate by button press whether the second presented word was the correct translation of the first one (50% were correct). 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.

The fourth training task was an L2 to L1 backward translation task (i.e., provide the English translation of the Chinese word). On each trial, participants saw a learned L2 word and were asked 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, 500 ms L1 feedback, 2,500 ms 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. This task was the principal task used for behavioral assessment of L2 learning and the analysis of individual differences in learning rate.

ERP assessment task. The task participants performed on the four ERP recording sessions was a go/no-go semantic categorization task. The learners were presented with L2 words one at a time visually (800 ms on, 2,500 ms ITI) and were asked to press a button to occasional probe words in a specific semantic category (in the first half of the block, they were asked to respond to body parts, and in the second half they had to press in response to food and drinks—15% of trials). This implies that on all critical nonprobe trials participants simply had to read the stimulus word for meaning. In a second block using this same task, learners were presented with the L1 translation equivalents of the L2 items and again had to detect occasional probe words (in L1). ERPs recorded to nonprobe critical words in the L2 Chinese block were subsequently analyzed. For the N400 analysis, Session 1 data included all nonprobe items, but only items that were successfully identified in the backward translation task were analyzed for Sessions 4, 7, and 10.

Results

To inspect individual differences in learning, we divided learners into two groups, fast and slow learners, based on their behavioral performance. Behavioral performance on translation recognition, verbal backward translation, and semantic categorization were

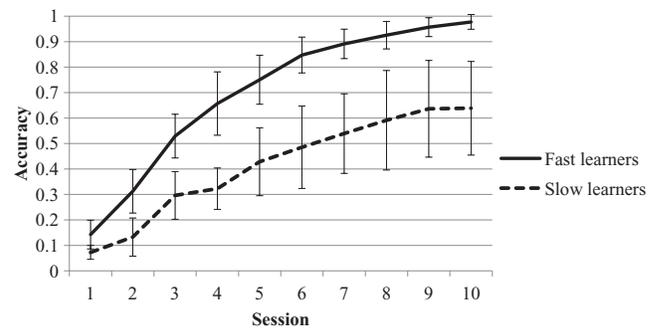


Figure 1. Mean behavioral performance in the backward translation task for fast ($n = 14$) versus slow learners ($n = 14$) across the course of learning (Sessions 1–10). Error bars are standard errors.

highly correlated, but performance on the backward translation task showed the most variability in the data over the sessions. Furthermore, compared to the recognition task, backward translation performance was less likely to be affected by guessing in the early sessions. So, the average of the overall performance (10 data points) in backward translation was used in a median split to determine the fast and slow groups. Other methods of separating learners into groups were explored, and the results were very similar to the separation based on overall performance. Analyses were also performed to examine the influence of visual complexity on behavioral effects of learning. We compared the two groups’ performance to single character words that differ in terms of the number of strokes forming the character, and performance to words differing in terms of the number of characters in the word. For the ERP data, learning effects on orthographic and semantic processing were assessed with N170 and N400 analyses, and this was done separately for the fast and slow learning groups.

Behavioral Results

The mean performance in backward translation across all 10 sessions is shown in Figure 1. A two-sample t test showed significant differences in group means in all the sessions, $t(1,26) > 3.59$, $p < .01$. This grouping of participants was used in the contrasts of learning rate in the ERP analyses as well.

Analyses of character complexity and word length. A repeated measures ANOVA was used to test how character complexity might affect backward translation. Results showed a significant triple interaction between session (Sessions 1 through 10), stroke number (high vs. low) and learning rate (fast vs. slow), $F(9,468) = 7.12$, $p < .001$. This triple interaction remained significant when the analysis was performed on Sessions 1–7 only and when the data were transformed through arcsine transformation, in order to reduce the influence of ceiling effects in the fast participants. Overall, both groups identified characters with a greater number of strokes less accurately. However, in fast learners this influence of stroke number got smaller as learning progressed (Session \times Stroke Number interaction: $F(9,234) = 16.6$, $p < .001$), whereas with the slow learners, the effects did not change (Session \times Stroke Number: $F(9,234) = 2.12$, *n.s.*; Figure 2).

A similar analysis was done to examine the effects of word length, contrasting performance to the single- and the two-character words. As in the complexity analysis, there was a triple interaction between session (Sessions 1 through 10), learning rate

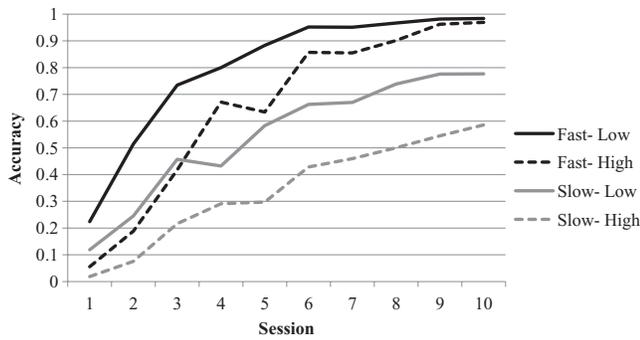


Figure 2. Accuracy in the backward translation task for single-character words with high or low stroke number, for fast and slow learners across the course of learning (Sessions 1–10).

(fast vs. slow), and word length (single vs. double character), $F(9,468) = 3.48, p = .009$. Again, fast learners were less affected by word length as learning progressed (Session \times Word Length interaction: $F(9,234) = 8.96, p < .001$), whereas slow learners showed the opposite pattern (Session \times Word Length: $F(9,234) = 3.83, p = .023$). In other words, the difference between the fast and slow learners was exaggerated in the learning of two-character words compared with single-character words (see Figure 3).

Semantic categorization. All participants improved in behavioral performance in semantic categorization over time (see Figure 4). A Session (Sessions 1, 4, 7, and 10) \times Learning Rate (fast vs. slow) two-factor ANOVA on the hit rate data showed marginally significant main effects of session and learning rate, $F(3,104) = 8.77, p = .054$ and $F(1,104) = 6.50, p = .084$, respectively. The interaction between session and learning rate was significant, $F(3,104) = 4.72, p = .004$, indicating that fast learners performed significantly better than slow learners on later sessions.

ERP Results

N170 epoch analysis (average reference)

160–210 ms. A significant main effect of session showed that later sessions were significantly more negative than the first session, $F(3,78) = 3.79, p = .020$. Additionally, an interaction of Learning Rate \times Session \times Anterior-Posterior \times Laterality was

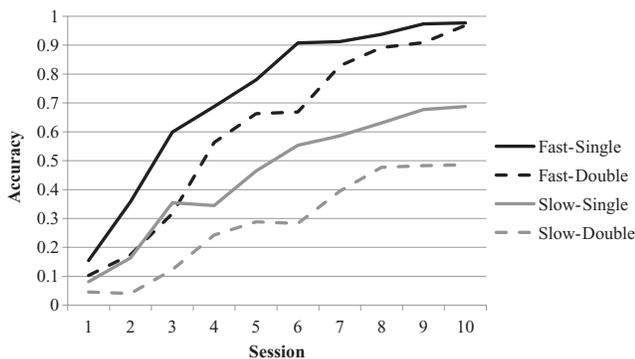


Figure 3. Accuracy in the backward translation task for single-character words compared with double-character words, for fast and slow learners across the course of learning (Sessions 1–10).

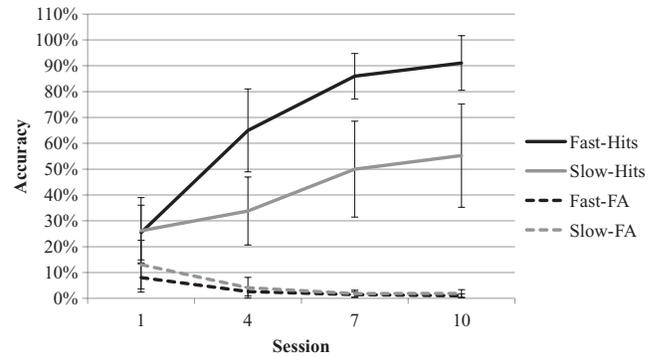


Figure 4. Percentages of hits and false alarms on the semantic categorization task for fast and slow learners on ERP sessions.

found, $F(9,234) = 4.10, p < .001$ (see Figure 5). To untangle the four-way interaction, we ran a series of follow-up analyses at the sites where the N170 is typically most evident (T5/T6 and O1/O2) separately for the two learning groups. Fast learners showed a significant learning effect at the left hemisphere sites (T5/O1), with a main effect of session, $F(3,39) = 3.76, p = .031$. As can be seen in Figure 5 (left side), fast learners produced larger N170 effects in the fourth compared to the first ERP session, and with additional sessions this effect tended to become more left lateralized. There was no evidence for changes across sessions at the corresponding sites in the right hemisphere (T6/O2). Slow learners revealed a different pattern, which included a lack of effects at left hemisphere sites, but a significant main effect of session at T6, $F(3,39) = 3.21, p = .048$. Examination of Figure 5 (right side) showed an inconsistent pattern of effects with an initially larger N170-like right-lateralized effect in ERP in Sessions 4 and 7 but a reversal in Session 10.

N400 epoch analysis (mastoid reference). The ERP waveforms time-locked to the L2 items are plotted in Figure 6. Visual inspection of Figure 6 suggested that differences between sessions began to occur at about 200 ms poststimulus onset and continued into later epochs. Fast learners showed a widely distributed anterior negative shift in ERP amplitude from 300 to 500 ms with learning, while slow learners showed a posterior positive shift.

300–500 ms. The two groups of learners showed the same ERP amplitudes at the first session, but fast learners had significantly more negative waves than slow learners after the first session, and the difference interacted with distributional variables (Learning Rate \times Session \times Anterior-Posterior \times Laterality: $F(24,624) = 1.77, p = .014$). In a follow-up analysis separating the learning groups, fast learners showed significant learning effects at central sites (C3, Cz, C4: $F(3,39) = 3.61, p = .035$). Slow learners showed learning effects with a positive polarity at occipital sites (O1, Oz, O2: $F(3,39) = 3.76, p = .037$).

Summary of ERP results. 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. On the basis of performance in a backward translation task across different testing sessions, participants were categorized as fast versus slow learners, and these two groups showed a different pattern of

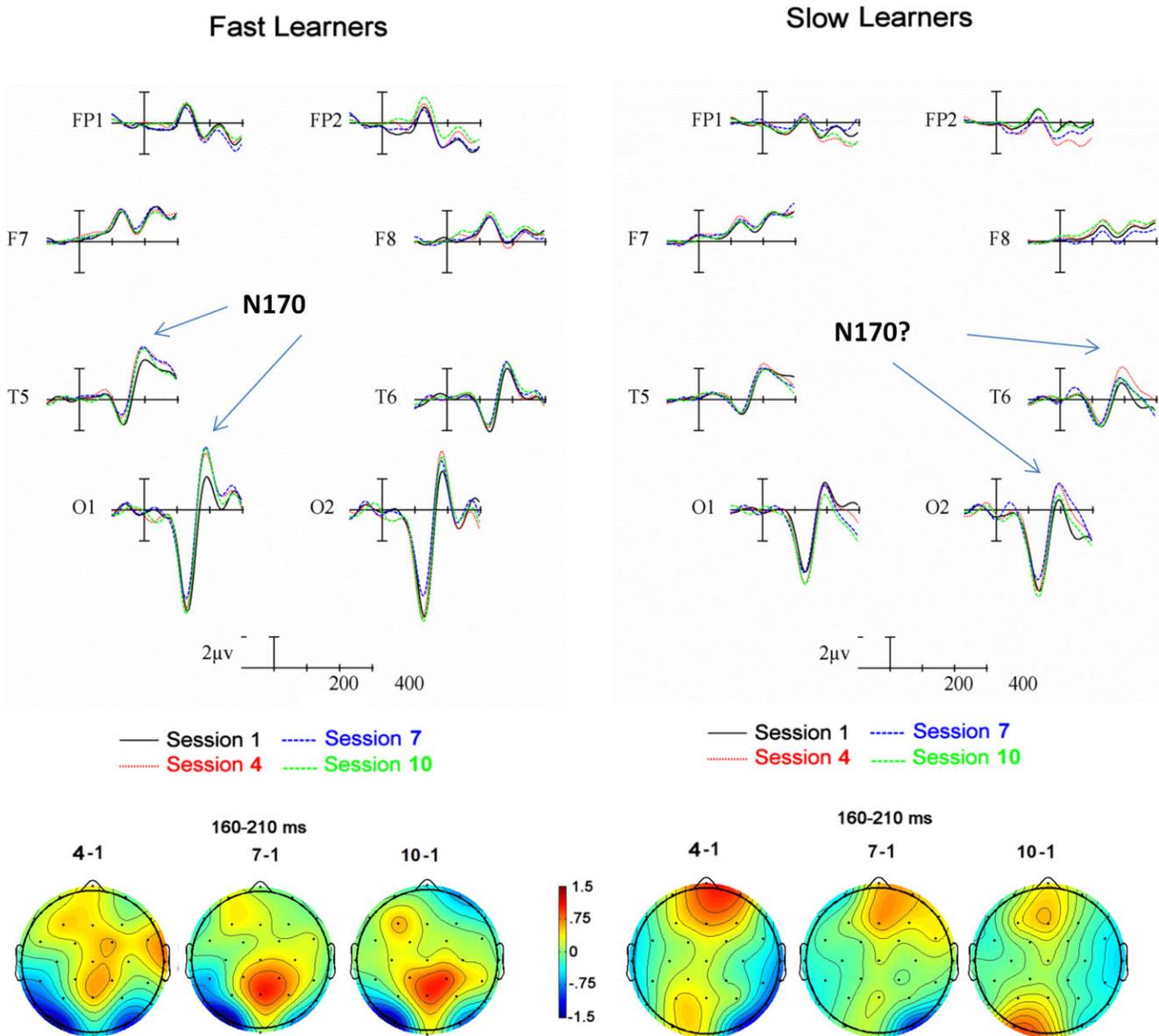


Figure 5. ERPs and voltage maps of L2 items in fast (left) and slow (right) learners obtained with an average reference. These maps were calculated from difference waves obtained by subtracting the voltages in Session 1 from the corresponding values in Sessions 4, 7, and 10 between 160–210 ms poststimulus onset.

learning effects on ERP components. An analysis of N170 effects using an average reference revealed a large increase in N170 amplitude with session that was left lateralized for the fast learners. Slow learners, on the other hand, showed a right-lateralized increase in N170 amplitude across the first sessions that diminished and even reversed with more learning. This contrast in ERP learning effects for the fast and slow learners was even more striking in the N400 analyses performed with a mastoid reference. Fast learners showed a widely distributed anterior negative shift in ERP amplitude from 300 to 700 ms with learning, while slow learners showed a posterior positive shift.

Discussion

In this study of L2 vocabulary learning where a group of English native speakers learned printed Chinese words, participants

showed systematic improvements in performance in a backward translation task over a short period of laboratory learning. The behavioral data suggest that L2 learning of Chinese characters by participants with no prior experience with the Chinese language can occur very rapidly, in line with previous findings (e.g., Liu, Wang, & Perfetti, 2007). 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 4 recording sessions that were interspersed across the 10 training sessions, revealing systematic changes in the ERPs as a function of learning. Most important is that the behavioral data were used to separate participants into a group of fast and slow learners, and these two groups showed qualitatively different patterns of ERP changes with training.

In the analysis of the N170 ERP component, for fast learners we found a rapid left-lateralized increase in N170 amplitude after the

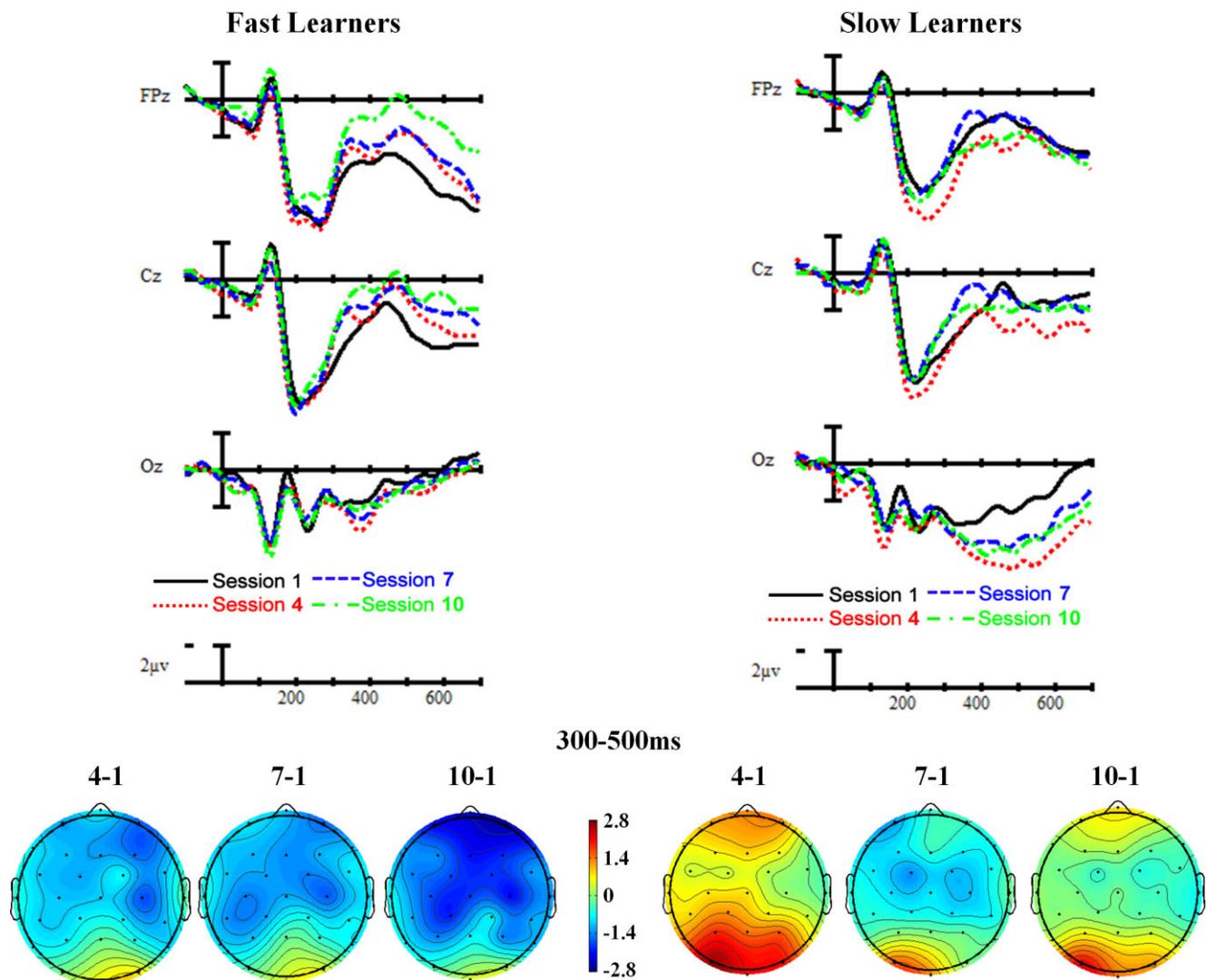


Figure 6. ERPs and voltage maps of L2 items in fast (left) and slow (right) learners obtained with a mastoid reference. The maps were calculated from differences waves of Sessions 4, 7, and 10 minus Session 1 between 300–500 ms poststimulus onset.

first set of four training sessions that remained stable across the following sessions. Slow learners, on the other hand, showed a distinctly right-lateralized increase in N170 amplitude, which again mostly arose during the first four training sessions. In an analysis of later effects (300–500 ms time window), fast learners showed increasingly negative-going waveforms as learning progressed, and these effects had a frontocentral distribution. Again, slow learners showed a very different pattern of learning effects, with a decreased negativity between the first and fourth session that reversed with further learning and had a posterior distribution.

The qualitatively different pattern of learning effects for fast and slow learners, revealed in the ERPs, might well reflect fundamentally different mechanisms used by these two groups of participants to learn Chinese words. It is important to note that the slow learners had achieved the same level of accuracy in the backward translation task by Session 10 as the fast learners had achieved by Session 4, yet the ERPs in the corresponding sessions are completely different (compare the voltage maps for 4-1 fast and 10-1 slow in Figures 5 and 6). The slow learners are indeed learning

Chinese words, but they would appear to be doing so in a different way compared with the fast learners. What could be the nature of these qualitative differences?

One possibility that we tentatively propose is that the fast learners are rapidly developing some form of structural representation of the Chinese words, and they then associate these structural representations with meaning much more efficiently than do the slow learners. Here, the term structural refers to a representation composed of object parts and the spatial relations between these parts. According to this explanation, the fast learners would have detected the presence of certain parts of words that they saw repeatedly in different words, and would use such word parts to create structural representations of each individual word. This would be analogous to using letters to create whole-word orthographic representations in their native language. The slow learners, on the other hand, would be using more holistic representations of the Chinese words, and learning to associate such holistic representations with meaning. Without attempting to define exactly what the nature of this holistic processing might be (see Piepers & Robbins,

2012, for a recent review with respect to face processing), one possibility is that the fast learners might be able to group smaller features into larger subcharacter components, whereas the slow learners would be trying to keep track of the smaller features and the spatial relations between them, as defined with respect to the whole word.

Alternatively, rather than reflecting two different modes of processing, the different pattern of ERP learning effects seen with the fast and slow learners might be related to differences in the relative involvement of the coordinated functions of extracting information about parts and information about their spatial relations. In support of this possibility, Liu and Perfetti (2003) found that, in skilled Chinese readers, a posterior negativity for Chinese characters peaked first in the left hemisphere and then in the right, switching between 150–200 ms poststimulus onset. The authors suggested that this pattern of results reflected the role of left occipital sites in supporting orthographic recognition of radicals or other subcharacter structures, and the role of right occipital sites for the processing of larger spatial relations. Consistent with this interpretation, fast learners in our study might be showing a left-lateralized N170 that reflects enhanced orthographic recognition of subcharacter structure. On the other hand, Maurer, Blau, Yoncheva, and McCandliss (2010) reported a right-lateralized increase in the N170 following training for a novel artificial script in adults using a visual one-back task. A left-lateralized N170 was observed, however, when using a different task that required retrieval of phonological information (Yoncheva, Blau, Maurer, & McCandliss, 2010). An initial script familiarity might therefore be driving the right-lateralized effect, while a shift to a left-lateralized N170 might reflect more developed lexical links to phonological or semantic representation of the novel words.

Returning to our interpretation based on different modes of processing in fast and slow learners, one key advantage of using structural representations compared with holistic representations is that not all the available information need be processed in order to identify a character. In other words, certain specific combinations of parts might suffice in order to identify a character from among the relatively limited number of learned characters so far. This explanation suggests that the fast learners should be less affected by the total number of features (strokes) in a character and less affected by word length (single vs. two-character words), and the results of the behavioral data on length and complexity provide some support that this might be the case. Overall, more complex words (words with many strokes vs. words with few strokes) were harder to learn than simpler words. Most important, however, is that while the influence of character complexity diminished with learning in the fast learners, it tended to increase in the slow learners. Furthermore, the negative impact of word length on learning increased as learning progressed in the slow learners, while the opposite pattern was seen with the fast learners. This is in line with the suggestion that the slow learners are using more information contained in the words than the fast learners are in order to recognize these words. As discussed above, this is what one would expect if these slow learners were using more holistic representations, compared with structural representations.

How might these hypothesized different learning strategies have affected the ERPs generated by the fast and slow learners? The left-lateralized N170 learning effect, seen in the fast learners, would reflect the use of structural representations by this group. This would then enable a more efficient association of form representations with semantics via the L1 translation equivalents, as reflected in the change in N400 amplitude with learning in the

fast learners. Slow learners, on the other hand, would be struggling to memorize holistic representations of characters in order to associate these with the L1 translation equivalents, and it could be this qualitatively different processing that is reflected in the posterior positive shift in the N400 time window seen with these learners.

A study by Hsiao and Cottrell (2009) that compared processing of Chinese characters by expert Chinese readers and naïve Chinese readers provides some support for our hypothesized distinction between holistic and structural processing. This study showed evidence for use of holistic representations when processing Chinese characters in the naïve group of participants, whereas the expert Chinese readers showed no evidence for use of holistic representations. The authors concluded that the Chinese readers had a better awareness of the components of characters, and that this type of structural information is not immediately available to novices. We would argue that, although the majority of our participants may have begun to process Chinese characters holistically, like the novices in the Hsiao and Cottrell (2009) study, many of them switched to more analytical processing following the first phases of word learning. Within the theoretical framework of holistic versus analytic processing, it could therefore be argued that our fast learners quickly adapted to using a more efficient analytic processing strategy, whereas the slow learners may have continued to use a less efficient holistic processing strategy. Finally, it is interesting to note that an informal comparison with results from an independent study testing expert Chinese readers (Yum et al., 2011) suggests that the pattern of word-based ERPs of expert Chinese readers more closely resemble those recorded in the final learning sessions of the current sample of fast compared to slow learners. It will be important in future studies to explicitly compare different learning groups with more proficient readers in the same experimental context to more carefully test this possibility.

The change in amplitude in the N400 window for fast learners was frontally distributed, which differed from the canonical distribution for N400-priming effects. The FN400 ERP component, which does have a frontal distribution and also occurs at this time range, is considered to be a distinct effect from the N400 (Bridger, Bader, Kriukova, Unger, & Mecklinger, 2012; Curran, 1999, 2000; Rugg & Curran, 2007). However, the FN400 is usually reported for items that engage memory processes, and recently presented stimuli generate a reduced FN400 compared to novel stimuli. The effect that we observed (i.e., an increase in negativity with learning) is the opposite of what might be predicted for an FN400 effect. Similarly, repetition typically produces a reduction of the N400 (Rugg, 1990; Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991). Thus, it is unlikely that our results reflect simple effects of repetition or familiarity. In L1 contexts, larger N400 amplitudes generally occur when a word is less expected or known, presumably in part because there are strong orthographic and lexical representations supporting word processing. This might not be the case in a new L2 where the network of representations supporting word processing is weaker and less developed. Consistent with this view are previous findings indicating that the N400 increases as proficiency increases (Midgley et al., 2009; Ojima et al., 2011). So, for fast learners, we suggest that knowledge/network construction is progressing faster towards an L1-like state—and hence a more L1-like N400 pattern.

The present study has shown the viability of using within-laboratory training combined with behavioral testing and ERP recordings to plot the very first phases of L2 vocabulary learning. Native English speakers with no prior knowledge of Chinese

learned to associate Chinese characters with their translation equivalents in English. In this context, learners have to master a new writing system, and L2 words have no form overlap whatsoever with L1 words. Participants were divided into fast and slow learners on the basis of their performance on a backward translation task (i.e., provide the L1 translation of an L2 word),

and these two groups showed strikingly different changes in the ERPs generated by L2 words as learning progressed. These different ERP learning patterns are tentatively interpreted as reflecting different learning strategies, with the fast learners using more structural representations and the slow learners using more holistic representations.

References

- Batterink, L., & Neville, H. (2011). Implicit and explicit mechanism of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*, 3181–3196. doi: 10.1162/jocn_a_00013
- Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, *116*, 289–296. doi: 10.1016/j.cognition.2010.05.004
- Bridger, E. K., Bader, R., Kriukova, O., Unger, K., & Mecklinger, A. (2012). The FN400 is functionally distinct from the N400. *NeuroImage*, *63*, 1334–1342. doi: 10.1016/j.neuroimage.2012.07.047
- Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia*, *37*, 771–785. doi: 10.1016/S0028-3932(98)00133-X
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*, 923–938. doi: 10.3758/BF03209340
- De Groot, A. M. B., & Keijzer, 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. doi: 10.1111/0023-8333.00110
- 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: Oxford University Press.
- 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. doi: 10.1080/87565641.2010.480915
- Geisser, S., & Greenhouse, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112. doi: 10.1007/BF02289823
- Holcomb, P. J., & Grainger, J. (2006). On the time-course of visual word recognition: An ERP investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*, 1631–1643. doi: 10.1162/jocn.2006.18.10.1631
- Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychological Science*, *20*, 455–463. doi: 10.1111/j.1467-9280.2009.02315.x
- Joyce, C. A., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, *116*, 2613–2631. doi: 10.1016/j.clinph.2005.07.005
- Kim, K. H., Yoon, H. W., & Park, H. W. (2004). Spatiotemporal brain activation pattern during word/picture perception by native Koreans. *NeuroReport*, *15*, 1099–1103. doi: 10.1097/00001756-200405190-00003
- Kutas, M., & Hillyard, S. A. (1980). Event-brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*, 99–116. doi: 10.1016/0301-0511(80)90046-0
- Lee, C. Y., Tsai, J. L., Chan, W. H., Hsu, C. H., Hung, D. L., & Tzeng, O. J. L. (2007). Temporal dynamics of the consistency effect in reading Chinese: An event-related potentials study. *NeuroReport*, *18*, 47–51. doi: 10.1097/WNR.0b013e328010d4e4
- Lin, S. E., Chen, H. C., Zhao, J., Li, S., He, S., & Weng, X. C. (2011). Left-lateralized N170 response to unpronounceable pseudo but not false Chinese characters—The key role of orthography. *Neuroscience*, *190*, 200–206. doi: 10.1016/j.neuroscience.2011.05.071
- Liu, Y., Dunlap, S., Fiez, J., & Perfetti, C. A. (2007). Evidence for neural accommodation to a writing system following learning. *Human Brain Mapping*, *28*, 1223–1234. doi: 10.1002/hbm.20356
- Liu, Y., & Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: An ERP study of Chinese bilinguals. *Human Brain Mapping*, *18*, 167–175. doi: 10.1002/hbm.10090
- Liu, Y., Perfetti, C. A., & Wang, M. (2006). Visual analysis and lexical access of Chinese characters by Chinese as second language readers. *Language and Linguistics*, *7*, 637–657.
- Liu, Y., Wang, M., & Perfetti, C. A. (2007). Threshold-style processing of Chinese characters for adult second-language learners. *Memory and Cognition*, *35*, 471–480. doi: 10.3758/BF03193287
- Maurer, U., Blau, V. C., Yoncheva, Y. N., & McCandliss, B. D. (2010). Development of visual expertise for reading: Rapid emergence of visual familiarity for an artificial script. *Developmental Neuropsychology*, *35*, 404–422. doi: 10.1080/87565641.2010.480916
- Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, *1*, 13. doi: 10.1186/1744-9081-1-13
- Maurer, U., Zevin, J. D., & McCandliss, B. D. (2008). Left-lateralized N170 effects of visual expertise in reading: Evidence from Japanese syllabic and logographic scripts. *Journal of Cognitive Neuroscience*, *20*, 1878–1891. doi: 10.1162/jocn.2008.20125
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*, 703–704. doi: 10.1038/nn1264
- Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, *17*, 1858–1866. doi: 10.1093/cercor/bhl094
- 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. doi: 10.1016/j.jneuroling.2008.08.001
- Nation, I. S. P. (1993). Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon*. (pp. 115–134). Amsterdam, The Netherlands: Benjamins.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139524759
- Ojima, S., Nakamura, N., Matsuba-Kurita, H., Hoshino, T., & Hagiwara, H. (2011). Neural correlates of foreign-language learning in childhood: A 3-year longitudinal ERP study. *Journal of Cognitive Neuroscience*, *23*, 183–199. doi: 10.1162/jocn.2010.21425
- 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. doi: 10.1017/S1366728907002891
- Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms “holistic,” “configural,” and “relational” in the face perception literature. *Frontiers in Psychology*. doi: 10.3389/fpsyg.2012.00559
- Proverbio, A. M., Čok, B., & Zani, A. (2002). Electrophysiological measures of language processing in bilinguals. *Journal of Cognitive Neuroscience*, *14*, 994–1017. doi: 10.1162/089892902320474463
- Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, words, and object processing in the visual cortex. *NeuroImage*, *20*, 1609–1624. doi: 10.1016/j.neuroimage.2003.07.010
- Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, *18*, 367–379. doi: 10.3758/BF03197126
- Rugg, M., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251–257. doi: 10.1016/j.tics.2007.04.004
- 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*, 774–783. doi: 10.1016/j.neuroimage.2006.07.008

- Van Petten, C., & Kutas, M. (1991). *Electrophysiological evidence for the flexibility of lexical processing* (pp. 129–174). Amsterdam, The Netherlands: North-Holland–Oxford. doi: 10.1016/S0166-4115(08)61532-0
- Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience*, 3, 131–150. doi: 10.1162/jocn.1991.3.2.131
- 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. doi: 10.1207/S1532799XSSR0702_4
- Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, 20, 6–11. doi: 10.3758/BF03202594
- Yoncheva, Y. N., Blau, V. C., Maurer, U., & McCandliss, B. D. (2010). Attentional focus during learning impacts N170 ERP responses to an artificial script. *Developmental Neuropsychology*, 35, 423–445. doi: 10.1080/87565641.2010.480918
- Yum, Y. N., Holcomb, P. J., & Grainger, J. (2011). Words and pictures: An electrophysiological investigation of domain specific processing in native Chinese and English speakers. *Neuropsychologia*, 49, 1910–1922. doi: 10.1016/j.neuropsychologia.2011.03.018

(RECEIVED February 8, 2013; ACCEPTED November 19, 2013)