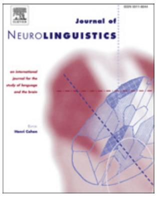
Journal of Neurolinguistics 26 (2013) 440–461

Contents lists available at SciVerse ScienceDirect

Journal of Neurolinguistics

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



Early stage visual-orthographic processes predict long-term retention of word form and meaning: A visual encoding training study

Fan Cao [a,*], Ben Rickles [b], Marianne Vu [b], Ziheng Zhu [b],
Derek Ho Lung Chan [b], Lindsay N. Harris [b], Joseph Stafura [b],
Yi Xu [c], Charles A. Perfetti [b,**]

[a] *Division of Psychology, School of Humanities and Social Sciences, Nanyang Technological University, 14 Nanyang Drive, Room 03-45, Singapore 637332, Singapore*
[b] *Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, USA*
[c] *Department of East Asian Languages and Literatures, University of Pittsburgh, Pittsburgh, PA, USA*

ARTICLE INFO

Article history:
Received 3 September 2012
Received in revised form 11 January 2013
Accepted 20 January 2013

Keywords:
Writing
Visual chunking
Reading
Learning
Chinese
ERP

ABSTRACT

Adult learners of Chinese learned new characters through writing, visual chunking or reading-only. Following training, ERPs were recorded during character recognition tasks, first shortly after the training and then three months later. We hypothesized that the character training effects would be seen in ERP components associated with word recognition and episodic memory. Results confirmed a larger N170 for visual chunking training than other training and a larger P600 for learned characters than novel characters. Another result was a training effect on the amplitude of the P100, which was greater following writing training than other training, suggesting that writing training temporarily lead to increased visual attention to the orthographic forms. Furthermore, P100 amplitude at the first post-test was positively correlated with character recall 3 months later. Thus the marker of early visual attention (P100) was predictive of retention of orthographic knowledge acquired in training.

---

* Corresponding author. Tel.: +65 6592 1825.
** Corresponding author.
E-mail addresses: fcao@ntu.edu.sg (F. Cao), perfetti@pitt.edu (C.A. Perfetti).

0911-6044/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jneuroling.2013.01.003

1. Introduction

The study of learning to read allows observations on the basic and higher level visual processes that allow arbitrary visual forms to become orthographic representations that can support connections to linguistic information. The study of adults learning a new writing system affords a special opportunity for such observations, because the novelty and complexity of visual forms of the new system present a challenge to the learner during the attempts to associate form with meaning.

Chinese presents a distinctive challenge for a learner with an alphabetic first language. Chinese contrasts with English in both the spatial layout of visual forms and the mapping of these forms to meaning and pronunciation. At the script level, Chinese character forms are complex visual-spatial configurations that consist of any or all of eight basic strokes interwoven in patterns to form component radical(s) in a two-dimensional square. At the mapping level, there is no grapheme-phoneme-correspondence in Chinese. The character maps to a single syllable morpheme, thus allowing a direct connection from orthography to meaning as well from orthography to syllable level phonology. Because a given syllable corresponds to many different meanings (producing many homophones), character-phonology mappings are more diffuse than character-meaning mapping (Perfetti, Liu, & Tan, 2005). Accordingly, meaning processes during reading are very dependent on a high quality representation of the orthographic form of the character.

There are several ways to support the acquisition of orthographic form in a nonalphabetic writing system. One is writing, a process that both directs attention to form and allows the possibility of a multimodal (sensory motor plus visual) form representation. Another is visual chunking, an approach that explicitly teaches how to decompose a character into chunks that are functional orthographic units repeatedly appear in different character. In a previous study, the authors (Cao et al., 2012) found that English learners of Chinese benefited from writing characters as they were first introduced to them. Their subsequent recognition of characters learned through writing was enhanced and produced brain activation patterns more similar to native Chinese speakers, compared with instruction that emphasized phonology. This finding is consistent with previous studies on native Chinese adults and children that orthography plays an especially important role in reading Chinese (Peng, Guo, & Zhang, 1985; Song, Zhang, & Shu, 1995).

Indeed writing is a widely-used strategy in Chinese first language literacy acquisition (Wu, Li, & Anderson, 1999), as children practice writing characters as they are introduced into the curriculum. Moreover, the enhancing effect of writing on reading may be general across writing systems, including English (Longcamp, Zerbato-Poudou, & Velay, 2005) and Japanese (Naka, 1998). Most relevant to our study, a recent study of adult learners of Chinese found that a handwriting characters during learning produced greater accuracy in subsequent lexical decision task and a semantic task than did pinyin-typing and reading-only conditions (Guan, Liu, Chan, Ye, & Perfetti, 2011). (Pinyin, an alphabetic system that indicates pronunciation of a character, provided an advantage on phonological tasks.) Neuroimaging studies suggest that writing has an effect on the neural substrate of orthographic processing. In English letter learning, more activation in the left fusiform gyrus is observed following writing compared with reading-only (James & Atwood, 2009). In Chinese, Cao et al. (2012) found greater activation in bilateral fusiform gyrus during character recognition following character-writing compared with pinyin writing.

The effect of writing may extend beyond the visual areas that support orthographic processing to a larger reading network that combines motor and visual-spatial memories. Neuroimaging studies have found that visual perception of single letters activated pre-motor cortex and pre-central gyrus, suggesting an interaction between perception and action (Cao et al., 2012; James & Gauthier, 2006; Longcamp, Anton, Roth, & Velay, 2003). Because motor memories can last for a very long period of time (Shadmehr & Holcomb, 1997), it is possible that the neural effects of well practiced writing include a long-lasting motor representation that serves recognition. In the case of Chinese characters, such a motor representation would center on the sequence of strokes, which follow specific ordering rules. Indeed, behavioral (Flores d’Arcais, 1994) and ERP (Parkinson, Dyson, & Khurana, 2010) evidence suggests that recognition of Chinese characters is affected by information about stroke sequencing.

In contrast with the broad involvement that writing has on motor systems, visual chunking are encoding procedures that direct attention specifically to visual forms that could be functional in orthographic recognition processes. Such encoding procedures would reduce the complexity of visual forms and provide smaller orthographic units for encoding. For example, visual-chunking (Shu, Chen, Anderson, Wu, & Xuan, 2003) organizes strokes into chunks, defined as the basic stroke sequences used to compose radicals and characters. There are 560 basic chunks (stroke sequences) in Chinese, and 118 of them cover 80% of characters (Zhang, 1990). Although some of these chunks make a radical, many are less than a radical. On average, a character contains 10.15 strokes character, whereas 98% characters are composed of 5 or fewer chunks, with 2.1 strokes per chunk. Thus, encoding visual chunks as part of character learning should dramatically reduce visual memory load.

Evidence suggests that native Chinese adults decompose characters into basic chunks in visual word recognition, especially when the characters are low frequency (Zhu & Taft, 1994). There is reason to think that this intermediate unit of the chunk is optimal for learning, relative to the radical and the stroke. Radicals provide large units—a typical character has two radicals—that have significant stroke complexity—5.2 strokes per radical in average. In contrast, the stroke provides a minimal unit that allows no structure to reduce visual memory load. Although there have been suggestions that instruction using basic chunks may be effective for Chinese L1 and L2 learners (Cui, 1997; Wan, 1999; Zhang, 1990), Chinese instruction for native speakers does not explicitly teach basic chunks. Studies of native Chinese children suggest that the ability to perceive characters in terms of major chunks and subchunks is acquired gradually over the early elementary school years and is correlated with vocabulary knowledge, reading comprehension, and teacher's rating of reading level (Anderson et al., 2004; Pak et al., 2005). Our study is the first to examine whether explicit visual chunking facilitates orthographic recognition in Chinese L2 learners.

Comparing writing and visual chunking, the evidence is that both can help to establish a high quality representation of orthography through attention to visual form. However, the locus and function of the visual attention may be different. Writing encourages attention to simple strokes executed in fixed temporal-spatial sequences. Visual chunking encourages attention to structures within the character that can function during recognition. Orthographic enhancement effects may result from both, but the underlying mechanisms may differ and may be detectable during the measurement of cortical activity during recognition.

Accordingly, we compared the learning outcomes of visual chunking and writing instruction, compared with reading-only on English adult learners of Chinese, using ERPs as well as behavioral indicators. If the effect of these enhanced methods is on orthographic recognition, then we may expect to find an effect on an early orthographic recognition component, such as the N170. This early negativity has been widely found in alphabetic (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Maurer, Zevin, & McCandliss, 2008; Zhang et al., 2011) word recognition and is sensitive to character-likeness and frequency in Chinese recognition (Cao, Li, Zhao, Lin, & Weng, 2011; Cao & Zhang, 2011; Zhang et al., 2011). Because we experimentally controlled the learning episodes, we also tested whether we would find ERP evidence for these learning episodes in the form of a P600. The P600, a marker for recognition memory, distinguishing presented from unpresented items (e.g., Rugg & Nagy, 1989). The P600 also distinguishes old from new items when the task only incidentally involves recognition memory (e.g., lexical decision, Curran, 1999). Moreover, and most relevant for our study, the P600 has been observed in English word learning experiments that measured in the same session as the learning (Perfetti, Wlotko, & Hart, 2005). Our question is whether it can also be observed over a longer interval following learning. We tested the long-term retention of any learning effects by having participants perform reading tasks while ERPs were recorded both shortly after instruction and 3 months later. This delayed test allows observations not only about long-term retention but also about the predictive value of measures taken at the first time point for long-term retention measured at the second time point.

2. Methods

2.1. Participants

Thirty undergraduate students enrolled in introductory Chinese at the University of Pittsburgh participated in the study. These students took a pre-test of 60 characters they had learned in class (Pre-test 1), in which they were asked to write down the pinyin, tone and meaning of a given character. Participants were assigned to one of three groups such that the groups were equated on average pre-test performance. The three groups were reading only, character-writing, or visual chunking. All participants went through training sessions and the first ERP session right after training, and twenty of them returned for a retention test and a second ERP session. After elimination of participants due to artifact, there were 7 participants in the reading-only group, 5 in the writing group and 6 in the visual chunking group who completed the whole study including all tasks at both post-test and retention test. However, we included as many participants as possible for a specific task at a specific time. Table 1 presents the demographic information of participants who completed the whole study and the performance on Pre-test 1 in each group. Statistics suggest no significant differences between groups on age, accuracy on tone production, pinyin production and meaning production in the pre-test 1 ($F(2,15) = 0.689$, $p = .517$; $F(2,15) = 0.679$, $p = .522$; $F(2,15) = 1.515$, $p = .252$; $F(2,15) = 0.813$, $p = .462$ respectively). According to an informal interview, all participants met the following criteria: (1) monolingual English speaker, (2) not from a Chinese heritage family, (3) right-handed, (4) free of neurological disease or psychiatric disorders, (5) no Attention Deficit Hyperactivity Disorder (ADHD), and (6) no learning disability. Participants continued to have Chinese classes during the training and testing period, and there was no control of Participants behavior with regard to Chinese learning out of the training sessions.

2.2. Procedures

2.2.1. Behavioral training and testing

One hundred and twenty characters chosen from the students' Chinese textbook that had not been taught in class at the time of the experiment were evenly assigned to one of two conditions: characters to be learned or novel. The characters were matched on variables of spatial structure (left-right, up-down, and simple), number of strokes, and frequency of the English translation (CELEX, Baayen, Piepenbrock, & Gulikers, 1995) across conditions. The chunks used to compose the characters were defined according to Hu (2004). In average, there were 2.9 chunks in each character, with the range of 1–5. In total, 102 different chunks were used, 34 of which occurred in more than one character. Another pre-test (Pre-test 2) was given before learning started to make sure that none of the 120 characters were known.

The training for each character was divided into three 800-ms segments (Fig. 1). For the first 800 ms, participants saw a character in the center of a computer screen; for the second 800 ms, they saw that character's pinyin (alphabetic spelling) and listened to a recording of its pronunciation; for the final 800 ms, they saw the English translation of the character. This sequence was followed by a 15-s period, during which participants were asked, according to their experimental condition, to view the character (reading only group), to write down the character from memory (writing group), or to watch how the character is composed by several chunks (visual chunking group). The entire sequence was repeated three times in a row for each character. The 60 to-be learned characters were presented to every participant on each of 6 days.

On each day of training, learning was followed by a testing session. The testing session included a lexical decision task, a character-meaning matching task and a character-sound matching task. In the lexical decision task, participants were asked to judge whether the stimulus presented on the screen was a real character or not. There were 60 learned characters and 60 novel characters. Each stimulus

Table 1
Demographic information and accuracy (%) on pre-test 1 of participants in each group.

	Age	Pre-test 1 Meaning	Pre-test 1 Pinyin	Pre-test 1 Tone
Reading	19.1 (0.9)	37.0 (23.8)	39.3 (21.8)	30.7 (20.3)
Writing	19.2 (0.8)	37.0 (27.8)	37.3 (19.6)	26.7 (16.2)
Visual chunking	18.7 (0.8)	53.9 (28.3)	58.6 (27.2)	43.1 (33.8)

Note: numbers in the parentheses are standard deviations.

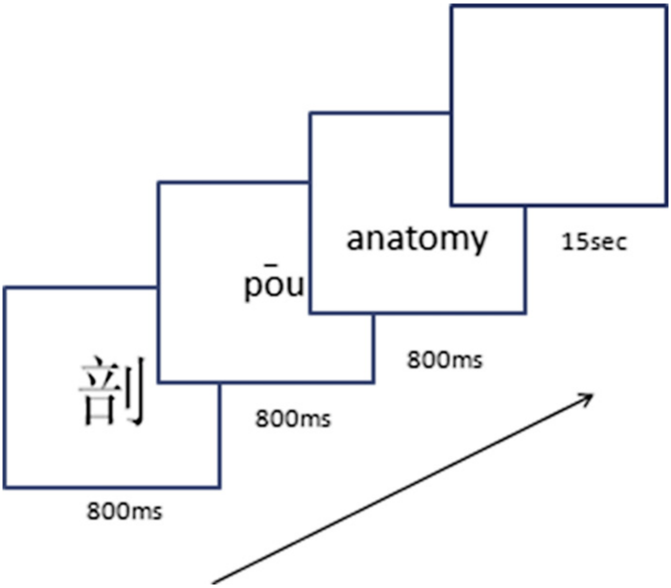


Fig. 1. Indication of the behavioral training procedures.

was presented for 800 ms followed by a blank of 1200 ms. Half of the learned characters and half of the novel characters were converted into false characters by deleting a stroke, adding a stroke or switching the position of two radicals. In the character-meaning matching task, participants were asked to judge whether the English word is the correct meaning of the character. In the character-sound matching task, participants were asked to judge whether the sound is the correct pronunciation of the character they saw. There were 60 learned characters in both tasks. The character was presented for 800 ms followed by a blank of 200 ms. Then an English word was presented for 800 ms in the meaning matching task and in the syllable matching task, a spoken Chinese syllable was presented for 800 ms. A blank of 1200 ms followed the English word or the Chinese sound. In all three tasks, participants were told to press a button with their index finger for a "yes" response and another button with their middle finger for a "no" response. The presentation order of the different types of stimuli was randomized. The same materials were used for 6 training days.

Following learning and testing on the last day of training, participants completed a post-test to assess their proficiency on trained characters. The post-test required participants to provide the pinyin, tone and meaning of each of the 60 characters they had learned. The pre-test and post-test were paper–pen tests, and the learning and testing were computerized using E-Prime. After the post-test on the last day, there was an ERP session in which brain waves were recorded during three tasks: Chinese lexical decision, English lexical decision and character-meaning matching. About 3 months later (range 2.5–3.4 months), participants were asked to return for a retention test and a second ERP session. The retention test was the same as the post-test, and the second ERP session was the same as the first session.

2.2.2. ERP procedures

2.2.2.1. Chinese lexical decision. Stimuli included 60 characters learned in the experiment and 60 novel characters. Half of the learned and half of the novel characters were converted into false characters by adding or deleting a stroke or by switching the position of two radicals. In addition, 60 characters learned in class were included, 30 of which were converted into false characters. ERP data analysis focused on the three types of real characters: characters learned in the experiment (i.e., learned), characters learned in class (i.e., familiar), and novel characters (i.e., novel). Each stimulus was presented for 800 ms, followed by a 400 ms blank. A red cross was presented at the end of the blank for 1000 ms. The participant was asked to press the "yes" or "no" button once they saw the red cross in order to exclude the readiness potential from the ERP epoch. The participant was asked to hold eye-blinking till they saw a signal on the screen which was a picture of an eye. There was an eye-blinking signal every 3 trials.

2.2.2.2. English lexical decision. In this task, 30 monosyllabic English words and 30 monosyllabic English pseudowords were included. The frequency of the English words was 338.818 per million (Baayen et al., 1995) which was matched with the frequency of the English translation of the Chinese characters. The procedure of this task is the same as the Chinese lexical decision task.

2.2.2.3. Character-meaning matching. In this task, stimuli were 60 characters learned in the experiment plus 60 characters learned in class. On each trial, the character was presented for 800 ms, followed by a blank of 200 ms, and then an English word for 800 ms, followed by a blank of 400 ms. For half of the stimuli, the character and the English word matched, and the other half were foils. A red cross was presented at the end of the blank and last for 1000 ms. The participant was asked to press the "yes" or "no" button once they saw the red cross in order to exclude the readiness potential. The participant was asked to hold eye-blinking till they saw a signal on the screen which was a picture of an eye. There was an eye-blinking signal every 2 trials.

2.3. ERP data acquisition and pre-processing

ERPs were recorded using 128 Ag/Cl electrodes in a dense array with impedances kept below 50 kΩ using a Geodesic Hydrocel system (Electrical Geodesics, Inc., Eugene, OR). Data were sampled at 500 Hz, and amplified with a hardware bandpass 0.01–200 Hz filter (time constant ~10 s). Pre-processing was done using NetStation software (Electrical Geodesics, Inc., Eugene, OR). Offline, data were filtered with a 30 Hz lowpass filter, and segmented according to word and task. In Lexical decision tasks segments started 200 ms before and extended 800 ms after the target word (1000 ms total). In the character meaning matching task, two different, 1200-ms epochs were created for each word (prime and target) within the word pairs, in each case starting 200 ms before and extending 1000 ms after word onset. Eye movements and blinks were automatically detected using separate pairs of electrodes, one for eye-blinks and another for eye-movements. Changes in amplitude exceeding ±140 μV were considered blinks and 75 μV were considered eye movements both within a sliding 640 ms window. Channels with variation less than ±0.5 μV within 150 ms or variation more than ±200 μV across the entire epoch were excluded. After automatic artifact detection, channels that were marked for exclusion on over 15% of the trials were excluded from processing. Data from bad (excluded) channels were interpolated using data from surrounding electrodes, re-referenced with respect to all electrodes, and baseline corrected relative to the 200 ms portion of data immediately preceding the stimulus. Subjects with more than 12 channels marked bad (excluded) by automatic artifact rejection were removed from the analysis. On average, data from only 2 subjects from each task at Time 1 and Time 2 needed to be removed due to excessive artifacts. Data for each remaining subject were visually examined for non-neural artifacts.

2.4. ERP data analyses

Based on the hypotheses that writing and visual chunking enhance early orthographic recognition and episodic memory, our analysis focused on N170 and P600. Accordingly, we inspected the grand average wave forms to determine whether effects appeared in the time windows and electrode locations where these components are usually observed. For the N170, 130–190 ms served as the time window over which mean amplitudes were computed within a cluster of electrodes in the left and right occipital and temporal regions that typically show the N170 (electrodes 51, 58, 59, 60, 64, 65, 66, 69, 70 in the left hemisphere and 83, 84, 85, 89, 90, 91, 95, 96, 97 in the right hemisphere) (Fig. 12). For the P600, 450–650 ms served as the time window with mean amplitudes computed within a cluster of parietal area electrodes (55, 54, 79, 61, 62, 78, 67, 72, 77) (Fig. 12). During visual inspection, we also discovered a P100, which appeared to vary according to learning condition. The mean amplitude of P100, which was observed on the left and right occipital and temporal electrodes, was calculated within the N170 clusters for the time window of 80–130 ms.

To analyze the variation of these components, we carried out ANOVAs for lexical decisions and for character meaning matching. For N170, we compared 30 real English words to 30 real Chinese learned characters in the lexical decision task across all subjects in order to establish the sensitivity of N170 to orthography familiarity. Then we examined whether there was a group difference in the amplitude of N170 for Chinese learned, familiar and novel characters in the lexical decision task. This analysis was also done on real characters in each type. For P600 and P100, we compared brain responses to the 60 learned Chinese characters and 60 familiar characters in the character-meaning matching task.

We were also interested in examining whether early visuo-orthographic brain responses (i.e., P100 and N170) that were sensitive to our training at time 1 predict recall on the retention test at time 2 in participants with special visuo-orthographic training versus those without the special training. Therefore, we conducted brain-behavior correlation analyses of the amplitude of N170 at time 1 in both hemispheres with composite recall score at time 2 (i.e., the sum of recall on the pinyin, meaning and tone test) within the visual-chunking group and the other two groups combined separately. We also conducted brain-behavior correlation analyses of the amplitude of P100 at time 1 in both hemispheres with composite recall score at time 2 within the writing group and the other two groups combined separately. We corrected for multiple comparison ($p = .05/8$).

3. Results

3.1. Behavioral results

3.1.1. Performances during the training

Behavioral measures showed that performance improved in all tasks across the 6 days of training. Table 2 and Fig. 2 show the improvement in accuracy and RT of the lexical decision tasks. An ANOVA of group (reading, writing and chunking) by day (Day 1 through 6) by character type (learned and novel) on accuracy and RT found that the learned characters were more accurate and faster than the novel characters ($F(1,23) = 164.363$, $p < .001$ for accuracy and $F(1,23) = 18.213$, $p < .001$ for RT). Performance Improvement is verified by a significant main effect of day ($F(5,115) = 13.224$, $p < .001$ for accuracy and $F(5,115) = 5.688$, $p < .001$ for RT). There was a significant interaction of character type by day for accuracy ($F(5,115) = 10.283$, $p < .001$) but not for RT. A simple effects test found that accuracy for the learned characters increased over days ($F(5,125) = 20.240$, $p < .001$), whereas accuracy for the novel characters did not change significantly over days (Fig. 2). The main effect of group was not significant for either accuracy or reaction time. The interactions with the group factor were not significant either.

The improvement in character-syllable matching accuracy and RT over the 6 days of training is seen in Table 2 and Fig. 2. An ANOVA of group (reading, writing and chunking) by day (Day 1 through 6) on accuracy and RT found a significant main effect of day ($F(5,120) = 60.867$, $p < .001$ for accuracy and $F(5,130) = 18.592$, $p < .001$ for RT). The main effect of group was not significant. There were no significant interactions.

Character-meaning matching accuracy and RT also improved over the course of the 6 days of training (Table 2 and Fig. 2). An ANOVA of group (reading, writing and chunking) by day (Day 1 through 6) on accuracy and RT found a significant main effect of day ($F(5,120) = 50.341$, $p < .001$ for accuracy and $F(5,130) = 29.314$, $p < .001$ for RT). The main effect of group was not significant. There were no significant interactions.

3.1.2. Performances on post-test and retention test

Table 3 presents accuracy in recalling meaning and pronunciation at the post-test and the long-term retention test. Recall for both meaning and pronunciation declined from post-test to long term retention and was higher at both test points for meaning than for pronunciation. These conclusions were confirmed by an ANOVA of group (reading, writing and chunking) by measurement (meaning

Table 2
Means and standard deviations of accuracy and reaction time for each group on each day on the lexical decision, character-syllable matching and character-meaning matching task.

		Day1	Day2	Day3	Day4	Day5	Day6
Lexical decision							
Learned							
Accuracy (%)	Reading	64(18)	75(10)	75(12)	83(6)	84(7)	88(6)
	Writing	67(19)	81(08)	84(08)	84(10)	86(5)	84(9)
	Visual chunking	75(13)	80(12)	82(09)	87(05)	86(6)	88(7)
Reaction time (ms)	Reading	782(110)	811(157)	705(98)	719(137)	714(108)	694(86)
	Writing	842(165)	845(139)	812(147)	767(104)	776(109)	784(131)
	Visual chunking	709(137)	705(182)	672(84)	669(138)	669(130)	664(127)
Novel							
Accuracy (%)	Reading	50(19)	56(11)	54(10)	55(13)	54(15)	56(13)
	Writing	56(14)	60(12)	63(14)	61(9)	61(13)	62(14)
	Visual chunking	61(10)	66(9)	67(8)	67(11)	66(12)	64(11)
Reaction time (ms)	Reading	765(115)	830(179)	732(121)	744(155)	750(134)	742(142)
	Writing	862(192)	873(146)	833(104)	786(110)	825(110)	811(127)
	Visual chunking	755(150)	752(200)	735(127)	719(158)	719(165)	691(120)
Character-syllable matching							
Accuracy (%)	Reading	50(15)	67(16)	75(13)	82(14)	83(15)	91(5)
	Writing	53(9)	67(13)	68(9)	73(10)	81(10)	80(19)
	Visual chunking	50(14)	66(11)	72(11)	80(10)	82(14)	86(11)
Reaction time (ms)	Reading	1228(140)	1179(157)	1125(167)	1080(193)	1077(182)	1051(163)
	Writing	1232(116)	1196(134)	1154(143)	1067(150)	1048(121)	1053(70)
	Visual chunking	1316(139)	1201(131)	1199(167)	1076(173)	1080(139)	1033(126)
Character-meaning matching							
Accuracy (%)	Reading	68(14)	80(15)	85(14)	89(09)	92(14)	97(3)
	Writing	61(19)	74(8)	84(10)	86(12)	90(11)	94(5)
	Visual chunking	61(16)	77(12)	86(10)	89(10)	91(12)	95(5)
Reaction time (ms)	Reading	782(93)	672(117)	615(87)	583(73)	588(96)	557(84)
	Writing	831(127)	752(70)	712(149)	645(135)	621(96)	635(108)
	Visual chunking	829(177)	698(167)	677(149)	593(123)	592(109)	611(121)

production, pinyin production and tone production) by time (post-test and retention test): main effect of time ($F(1,22) = 281.628$, $p < .001$) and of measurement ($F(2,44) = 73.012$, $p < .001$). Accuracy on pinyin and meaning was higher than on tone ($t(29) = 7.013$, $p < .001$; $t(29) = 9.021$, $p < .001$, respectively) and accuracy on meaning was higher than on pinyin ($t(29) = 8.625$, $p < .001$). There was no significant main effect of group or interactions.

3.1.3. Behavioral performances during ERP sessions

Table 4 presents accuracy and reaction time data for each condition on the Chinese lexical decision task during the ERP sessions. An ANOVA of group (reading, writing and chunking) by time (Time 1 and Time 2) and by character type (learned, familiar and novel) for accuracy and reaction time found a significant main effect of character type for accuracy ($F(2,32) = 119.026$, $p < .001$) and RT ($F(2,36) = 6.693$, $p = .003$). There was a significant main effect of test time on reaction time ($F(1,18) = 8.403$, $p = .01$) and a significant interaction of time and character type for accuracy ($F(2,32) = 6.066$, $p = .006$) and RT ($F(2,36) = 7.212$, $p = .002$). The main effect of group was not significant. A simple effect test on accuracy found that at Time 1, accuracy on learned characters was higher than on familiar characters ($t(24) = 5.272$, $p < .001$) and novel characters ($t(24) = -11.511$, $p < .001$). Accuracy on familiar characters was higher than on novel characters ($t(24) = -6.788$, $p < .001$). At Time 2, accuracy on learned and familiar characters was higher than that on novel characters ($t(21) = 9.340$, $p < .001$; $t(21) = 12.287$, $p < .001$, respectively). The difference between the learned and familiar characters was not significant. A simple effect test on reaction time found that at Time 1, reaction time for learned characters was faster than for familiar characters ($t(24) = 4.703$, $p < .001$) and novel characters ($t(24) = 3.240$, $p = .003$). Reaction times for familiar and novel characters were not different. At Time 2, reaction time was not different among the three character types.

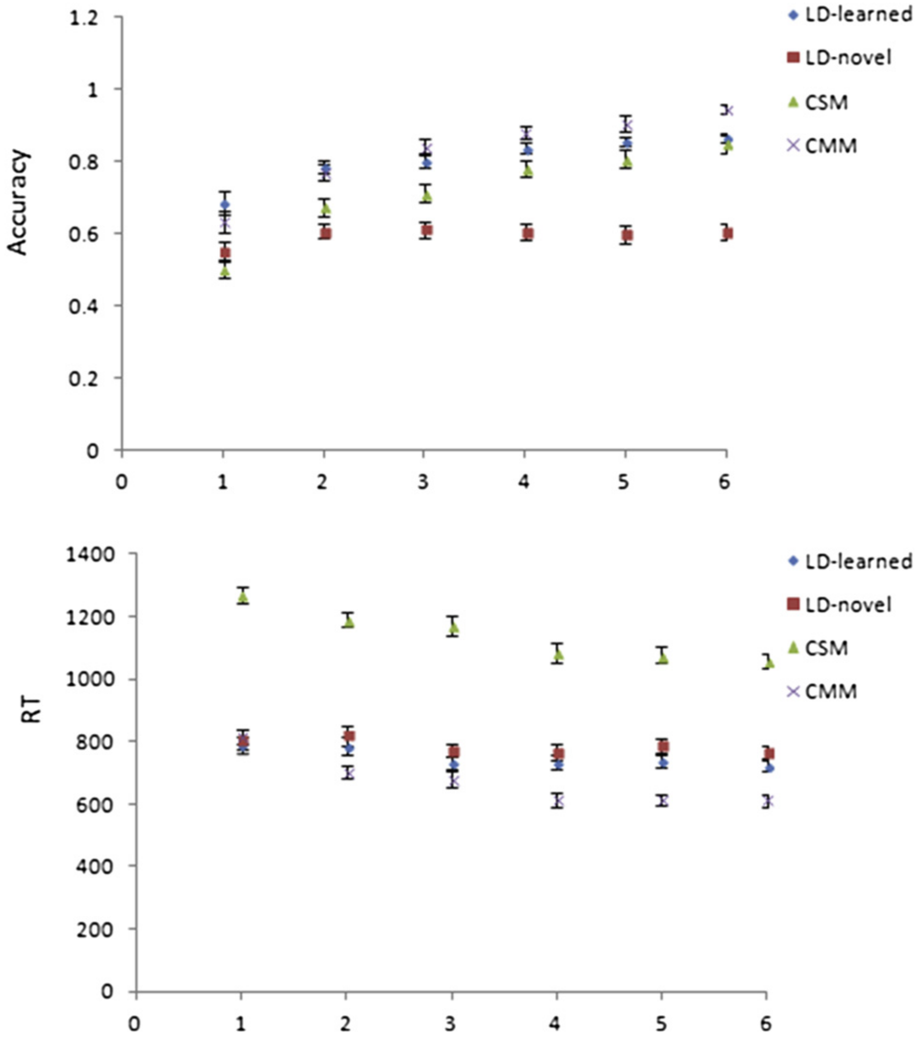


Fig. 2. Increases of accuracy and decreases of RT on the lexical decision task (LD-learned and LD-novel), character-syllable matching task (CSM) and character-meaning matching task (CMM) collapsed across all subjects from day 1 to day 6. Error bars are standard errors.

Table 5 presents accuracy and reaction time data for the English lexical decision task during ERP sessions. ANOVAs of group (reading, writing and chunking) by time (Time 1 and Time 2) for accuracy and reaction time found only a significant main effect of time for RT ($F(1,20) = 10.218$, $p = .005$) with Time 2 faster than Time 1. There were no main effect of group or interaction between group and time.

Table 3
Means and standard deviations of accuracy on the recall test in the post-test and retention test for each group.

Accuracy (%)	Post-test			Retention test		
	Meaning	Pinyin	Tone	Meaning	Pinyin	Tone
Reading	96(8)	85(11)	73(25)	33(11)	19(13)	10(10)
Character-writing	91(11)	75(17)	56(20)	36(13)	24(14)	15(19)
Visual chunking	94(14)	83(24)	69(36)	43(22)	27(22)	19(23)

Table 4
Means and standard deviations of accuracy and reaction time for each group on the Chinese lexical decision task during ERP sessions.

	Time 1			Time 2		
	Learned	Familiar	Novel	Learned	Familiar	Novel
Accuracy (%)						
Reading	83(8)	75(10)	62(11)	80(5)	82(4)	64(5)
writing	84(15)	74(8)	60(19)	83(8)	75(6)	63(10)
Visual chunking	85(11)	77(13)	64(13)	82(8)	86(7)	69(9)
Reaction time (ms)						
Reading	307(80)	343(74)	360(65)	361(75)	354(73)	364(91)
writing	368(88)	407(51)	393(75)	307(127)	301(110)	313(118)
Visual chunking	418(121)	424(112)	393(139)	363(100)	372(97)	377(113)

Table 6 presents accuracy and reaction time data for each condition on the character-meaning matching task during ERP sessions. ANOVAs of group (reading, writing and chunking) by time (Time 1 and Time 2) and by character type (learned and familiar) for accuracy and reaction time found a significant main effect of time ($F(1,17) = 31.892, p < .001$) and character type ($F(1,17) = 4.571, p = .047$) for accuracy. There was a significant interaction between time and character type for accuracy ($F(1,17) = 36.680, p < .001$) and for reaction time ($F(1,19) = 8.899, p = .008$). There was no main effect of group. A simple effect test on accuracy found that the accuracy on learned characters was higher than familiar characters at Time 1 ($t(25) = 2.781$, $p = .01$), and lower than familiar characters at Time 2 ($t(23) = -6.416, p < .001$). A simple effect test on reaction time found that the reaction time on learned characters was faster than that on familiar characters at Time 1 ($t(26) = -2.556$, $p = .017$), and slower than familiar characters at Time 2 ($t(24) = 2.648$, $p = .014$).

3.2. ERP results

As we hypothesized, we found that the N170 was larger for English words than Chinese characters in the left occipito-temporal region in the lexical decision tasks at both Time 1 and Time 2 (Figs. 3 and 4). The visual chunking group had a larger N170 than the other two groups combined in bilateral occipito-temporal areas for learned, familiar and novel characters at Time 1 (Figs. 5 and 6). Similarly as hypothesized, the P600 was larger for learned characters than familiar ones at Time 1 but not Time 2 in the character-meaning matching task (Figs. 7 and 8). We also found that the P100 at bilateral occipito-temporal areas was larger for the writing group than the other two groups combined for learned characters and familiar characters at Time 1 (Figs. 9 and 10). Finally, the amplitude of P100 in the left occipito-temporal area predicted long-term retention on the recall test in the non-writing training participants (i.e., the reading-only group and the visual-chunking group combined) (Fig. 11).

3.2.1. N170 effect

We found that the amplitude of N170 was greater for English words than Chinese characters at the left occipito-temporal cluster at both ERP sessions (Figs. 3 and 4). ANOVA of language (English and

Table 5
Means and standard deviations of accuracy and reaction time for each group on the English lexical decision task during ERP sessions.

	Time 1	Time 2
Accuracy (%)		
Reading	94(7)	97(2)
Writing	90(11)	91(16)
Visual chunking	95(3)	97(2)
Reaction time (ms)		
Reading	319(111)	300(86)
Writing	391(111)	298(83)
Visual chunking	350(80)	308(63)

Table 6
Means and standard deviations of accuracy and reaction time for each group on the character-meaning matching task during ERP sessions.

	Time 1		Time 2	
	Learned	Familiar	Learned	Familiar
Accuracy (%)				
Reading	94(6)	87(9)	75(8)	90(6)
Writing	89(10)	82(9)	70(13)	86(7)
Visual chunking	88(12)	86(11)	71(15)	87(18)
Reaction time (ms)				
Reading	281(75)	318(73)	324(78)	286(70)
Writing	366(64)	394(73)	359(63)	353(53)
Visual chunking	336(69)	335(110)	340(81)	328(70)

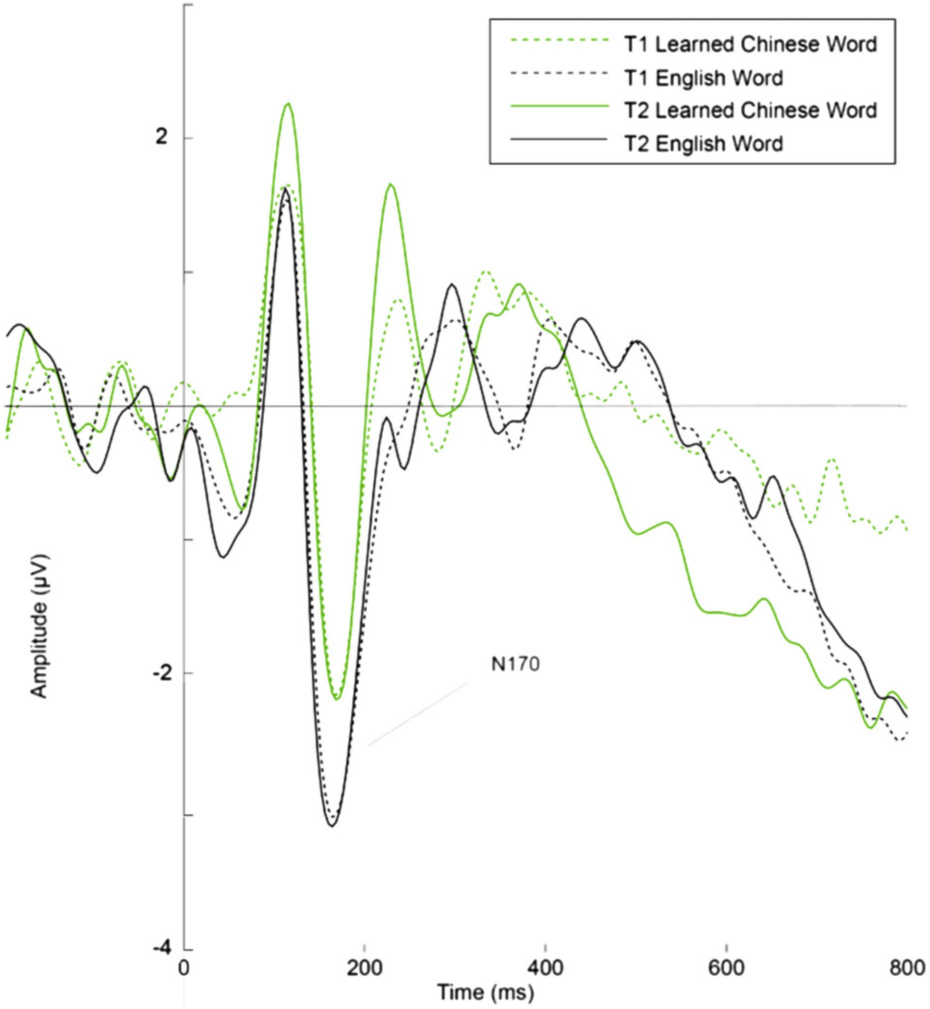


Fig. 3. Differences between English words and Chinese characters at the amplitude of N170. At the left occipito-temporal area, N170 is significantly greater for English words than for Chinese characters at both time 1 and time 2.

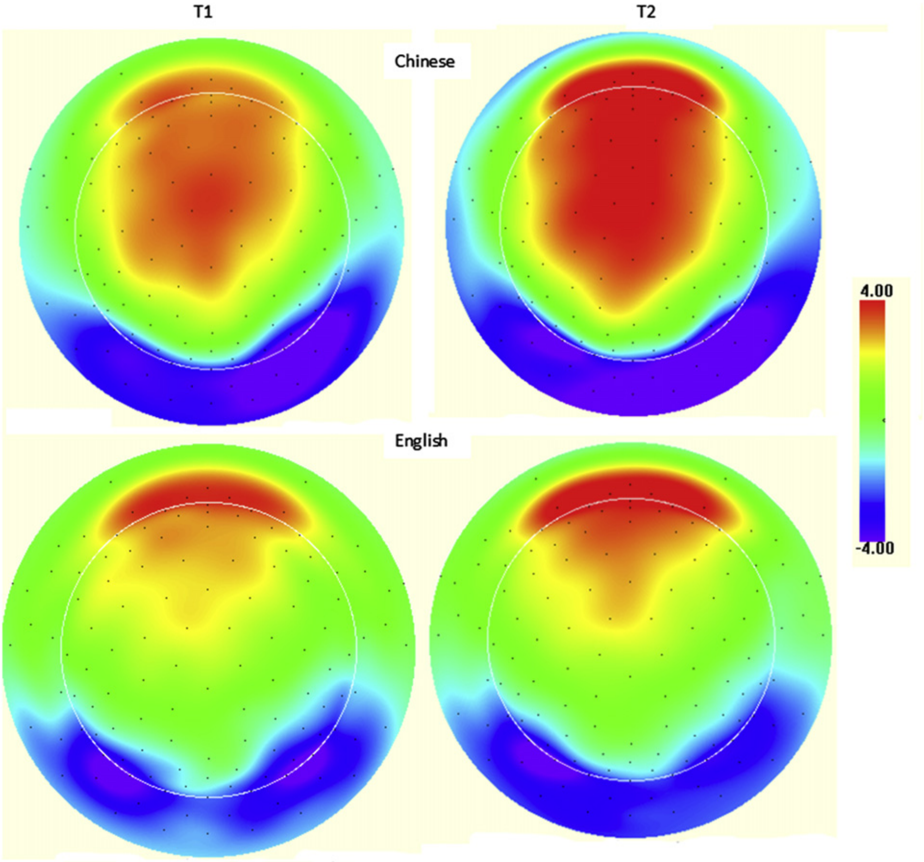


Fig. 4. Topographic maps for English words and Chinese characters in the lexical decision task at the time window of N170. N170 is significantly greater for English words than for Chinese characters at both time 1 and time 2 in the left.

Chinese) by time (first and second session) by hemisphere (left and right) on all subjects ($N = 17$) revealed a significant interaction between language and hemisphere ($F(1,16) = 13.418$, $p = .002$). Simple effect test revealed that there was a significant main effect of language ($F(1,16) = 8.901$, $p = .001$) in the left hemisphere with a greater amplitude for English than Chinese and there was no effect of language in the right hemisphere. There were no other significant main effects or interactions.

For Chinese characters, the N170 amplitude in both hemispheres was greater for the visual chunking group than the other two groups combined for the learned, familiar and novel characters in the lexical decision task at Time 1 (Figs. 5 and 6). ANOVA of hemisphere (left and right) by character type (learned, familiar and novel) by group (chunking ($N = 9$) and the other two groups combined ($N = 13$)) revealed a significant main effect of hemisphere with a higher amplitude in the right hemisphere than in the left hemisphere ($F(1,20) = 15.460$, $p = .001$). The main effect of group was also significant ($F(1,20) = 5.302$, $p = .032$) with a higher amplitude in the chunking group than the other two groups combined. The other main effect and interactions were not significant. Because prior knowledge on pre-test 1 may be correlated with the amplitude of N170, we ran an ANCOVA with the pinyin and meaning recall at pre-test 1 partialled out. We found that N170 was still larger in the chunking group than the other two groups combined ($F(1,18) = 6.001$, $p = .025$).

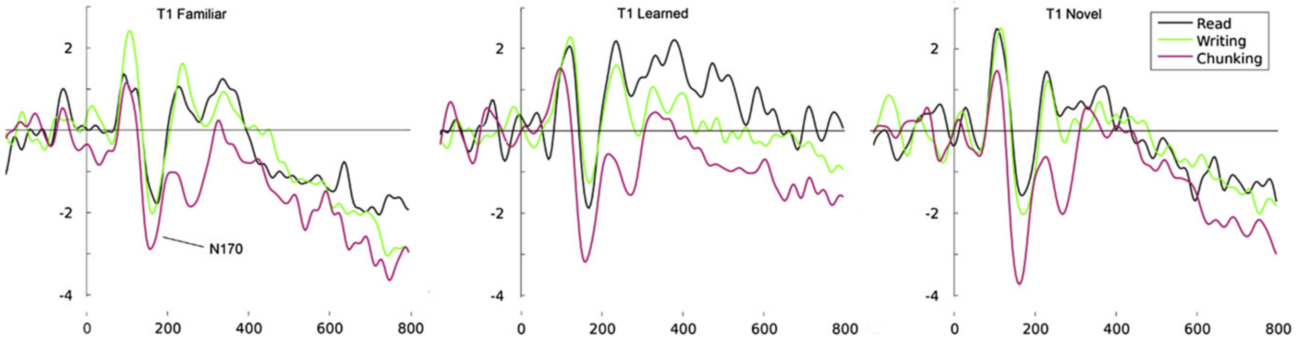


Fig. 5. Group differences at the amplitude of N170. At the left occipito-temporal area, N170 is significantly greater for the visual chunking group than the other groups for learned, familiar and novel characters at time 1.

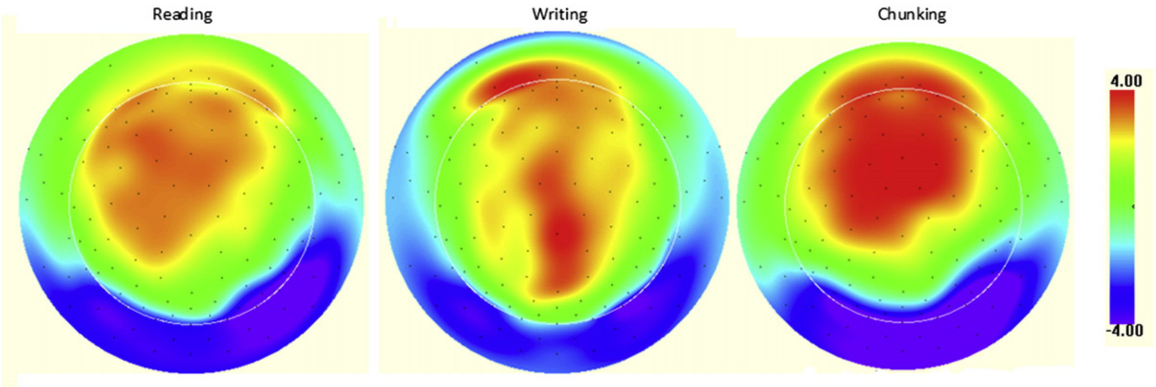


Fig. 6. Topographic maps for each group across all types of characters at the time window of N170. The N170 is greater for the chunking group than the other groups in both hemispheres.

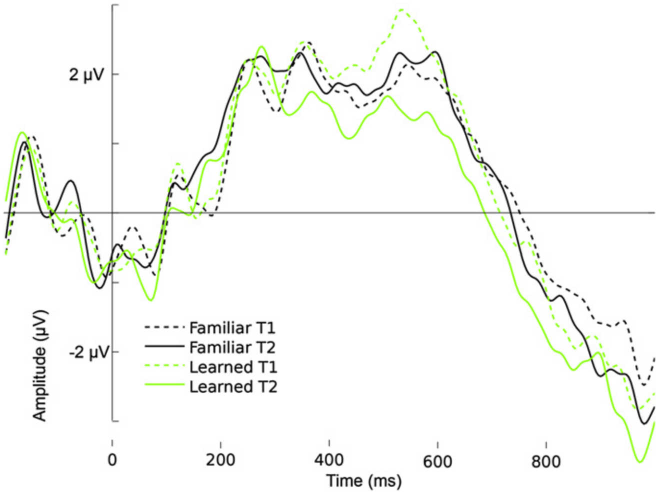


Fig. 7. Differences between learned and familiar characters at the amplitude of P600. At bilateral parietal areas, P600 is significantly greater for the learned characters than the familiar characters at time 1 but not time 2.

3.2.2. *P600 effect*

In the character meaning matching task, the amplitude of P600 during the presentation of the character was greater for learned Chinese characters than familiar Chinese characters at time 1 but not time 2 (Figs. 7 and 8). ANOVA of character type (learned and familiar) by time (Time 1 and Time 2) on all subjects ($N = 18$) revealed a significant interaction between character type and time ($F(1,17) = 10.187$, $p = .005$). Simple effect tests found that the P600 amplitude was significantly larger for learned than familiar characters at time 1 ($F(1,23) = 5.199$, $p = .032$) but not time 2 ($F(1,19) = 2.526$, $p = .128$). The main effects and the other interaction effects were not significant.

3.2.3. *P100 effect*

In the character meaning matching task, during the presentation of the learned character and familiar character, the P100 showed a larger amplitude in the writing group than the other two groups for the Time 1 ERP session (Figs. 9 and 10) at the bilateral occipito-temporal cluster. ANOVA of character type (learned and familiar) by hemisphere (left and right) by group (writing ($N = 8$), and the other two groups combined ($N = 15$)) revealed a significant main effect of group ($F(1,21) = 6.422$, $p = .019$) with a higher amplitude for the writing group than the other two groups combined. The other main effects and interactions were not significant. Because prior knowledge on pre-test 1 may be correlated with the amplitude of P100, we ran an ANCOVA with the pinyin and meaning recall at pre-test 1 partialled out. We found that P100 was still larger in the writing group than the other two groups combined ($F(1,18) = 4.983$, $p = .039$).

In order to further investigate whether the N170 and P100 effect in the bilateral occipito-temporal area was caused by the visual chunking and writing training respectively rather than by different baseline levels in different groups, we included N170 and P100 amplitude of English words in the lexical decision task at time 1 in two ANOVAs of group (chunking, and reading and writing combined for N170; writing, and reading and Chunking combined for P100). We found the group effect was not significant ($F(1,22) = 1.066$, $p = .313$) for N170 and ($F(1,19) = 0.193$, $p = .665$) for P100.

3.2.4. *Brain-behavioral correlation over time*

Finally, we examined the correlations between ERP indicators at time 1 and retention of learning at time 2 to test whether indicators of visual attention and orthographic processing would predict recall. Significant positive correlation was found between the amplitude of P100 in the left occipito-temporal cluster at time 1 and the recall of the character at the retention test at time 2 for the non-

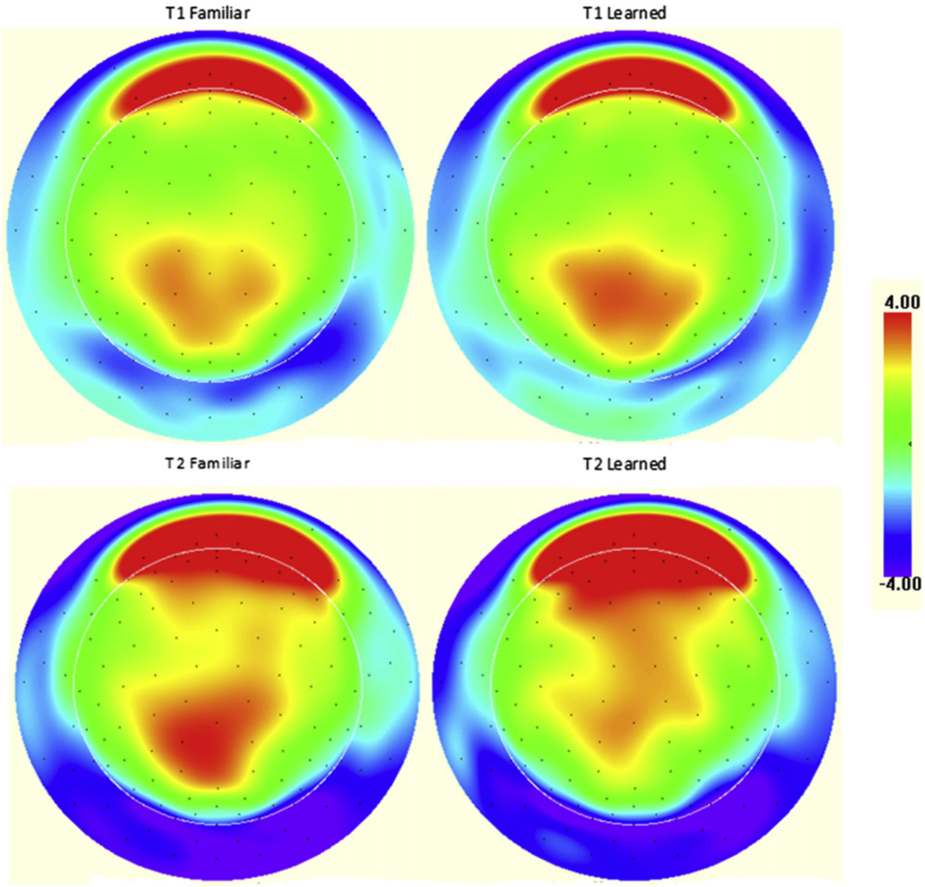


Fig. 8. Topographic maps for familiar and learned characters at time 1 and time 2 for the time window of P600. P600 is greater for learned than familiar characters at time 1 but not time 2.

writing training participants (i.e., the reading-only and the visual chunking group combined) ($N = 11$) ($r = .767$, $p = .006$) at the corrected level for multiple comparison ($p < .05/8$) (Fig. 11). No significant correlations were found for N170, and the correlation within the writing group for P100 was not significant either.

4. Discussion

We have divergent results: The expected effect of orthographic enhancement on character reading did not occur in the behavioral results but did occur in the ERP results. In the behavioral results, accuracy and reaction times did not reliably favor either writing or visual chunking instruction over reading-only. These behavioral results contrast with the findings of previously reported training studies with comparable participant samples and instructional procedures (but using within-subject designs), which showed advantages of writing on subsequent character recognition (Cao et al., 2012; Guan et al., 2011). Although the between-group design of the present study might have worked against statistical power, there is no evidence of this in the data trends, which were mixed in comparisons of the two orthographic enhancement conditions with reading-only. Mean differences in the lexical decision task (RT only, Table 4) and the character meaning task (both accuracy and reaction time, Table 6) slightly favored the reading-only condition at Time 1. During learning (Table 3), both visual chunking

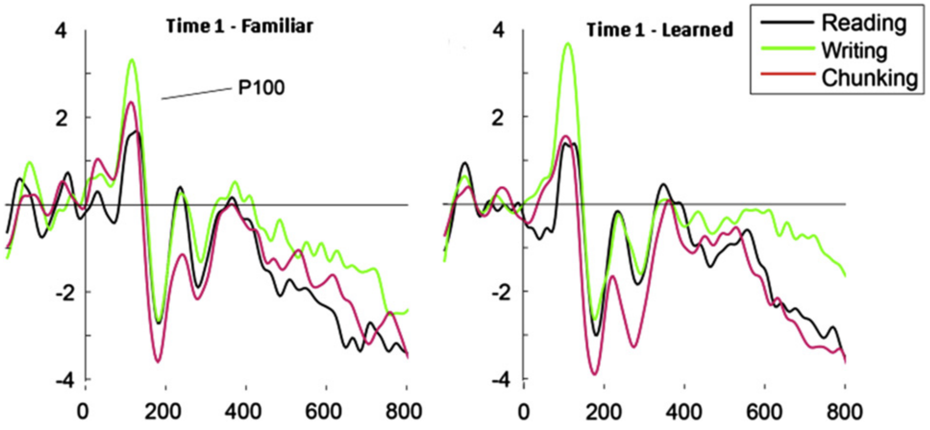


Fig. 9. Group differences at the amplitude of P100. At the left occipito-temporal area, P100 is significantly greater for the writing group than the other groups for learned and familiar characters at time 1.

and writing showed a higher mean accuracy than reading-only through the first three days, but this difference was small and not reliable.

However, the ERPs tell a different story. We expected an N170 to indicate an enhanced orthographic representation, and it seems to have done so. The main N170 effect was due to language: the participants' native language (English) produced a stronger N170 than did their second language, consistent with the results of Liu and Perfetti (2003) who found this pattern for Chinese–English bilinguals (Liu & Perfetti, 2003). Thus we can be confident that the experiment was able to detect an indicator of orthographic processing. As to orthographic enhancement of this effect, the N170 showed a greater amplitude for the visual chunking condition than for reading-only and writing combined, at the first post-test (Time 1). The N170 effect at Time 1 extended to novel as well as familiar characters. With the prior knowledge partialled out, the chunking group was still higher than the other groups for the amplitude of N170 for learned, familiar and novel characters. Taken together, we speculate that visual chunking training fine-tuned the orthographic representation system of characters and that the restructured system can be applied to all characters even novel ones due to the overlap of chunks in different characters. This finding is consistent with a study with artificial orthographies by Yoncheva, Maurer, and McCandliss (2010), who found that the unit size acquired through training (grapheme versus whole word) influenced N170 responses to visual words – greater when training was based on the small unit size (i.e., grapheme) than on the large unit size (whole word). Similarly, visual chunking training draws more attention to the local features of the components of the character than passive viewing and may provide units (comparable to graphemes) in mediating the orthographic enhancement indexed by the N170.

We were also interested in whether we would observe a P600 in response to the presentation of a learned character. This component is observed in recognition memory experiments for stimuli that are "old" (previously presented in the experimental session) (Smith, Stapleton, & Halgren, 1986) and has been taken as an indicator of episodic memory for word learning experiences (Perfetti, Wlotko, et al., 2005). Here, we observed a P600 for learned characters more than familiar characters at the first post-test (Time 1) when the character was presented for a meaning judgment (but not for lexical decision). This is the same situation observed by Perfetti, Wlotko, et al. (2005) for English word learning, i.e., when the participant views a word, knowing that a meaning judgment must be made on a following word. Thus the P600 in this situation may reflect not just a passive recognition process, but a retrieval of the word learning episode. The word learning episode becomes less available over time, as suggested by the disappearance of the P600 at the 3 month delayed test.

A surprising result was the discovery of a P100 effect. Writing training evoked a greater P100 amplitude than either reading-only or visual chunking. Rather than reflecting orthographic processing itself, the P100 is of particular relevance for visual selective attention and working memory (Mangun,

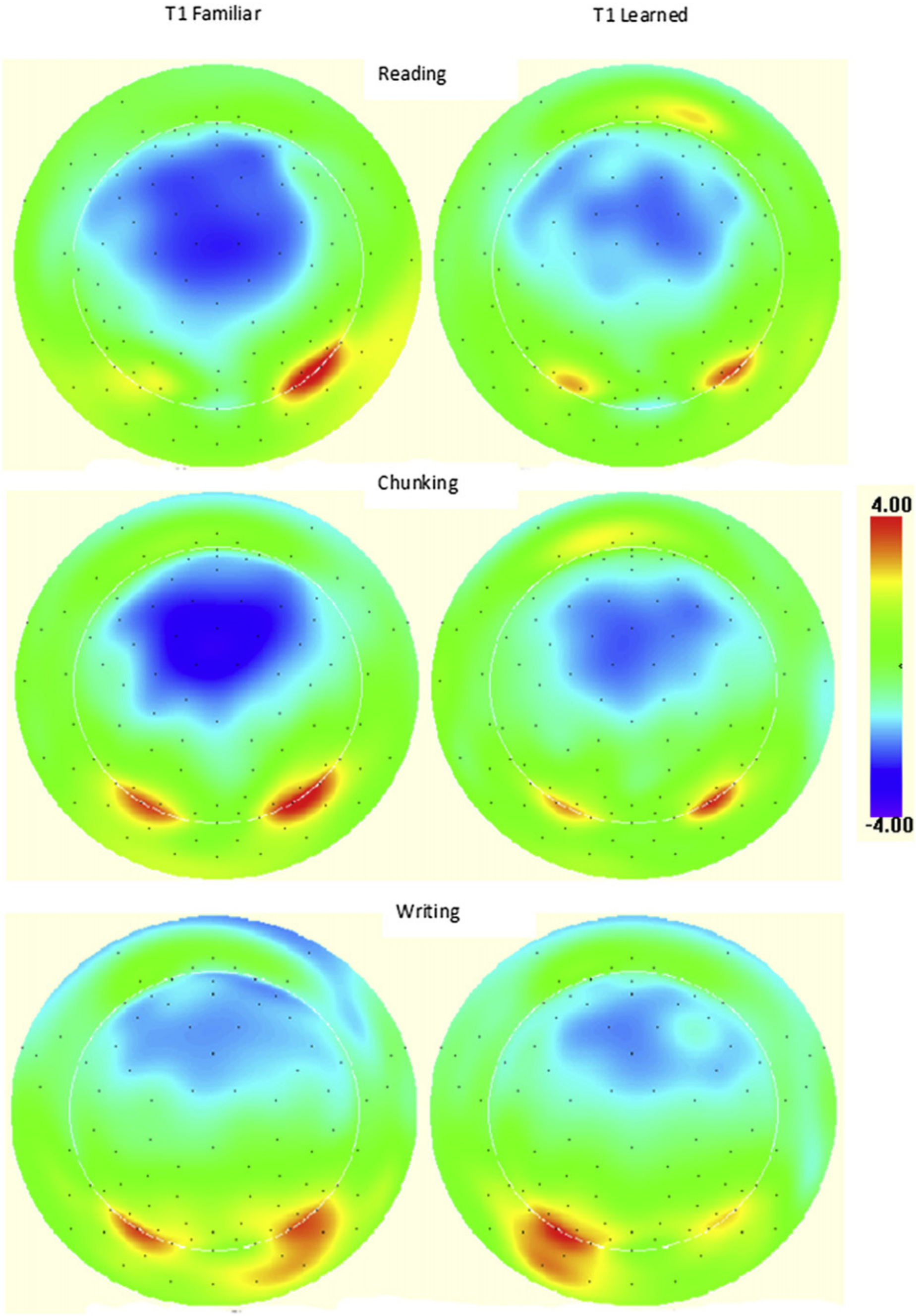


Fig. 10. Topographic maps for familiar and learned characters in each group at the time window of P100. P100 is greater for the writing group than the other groups in both hemispheres.

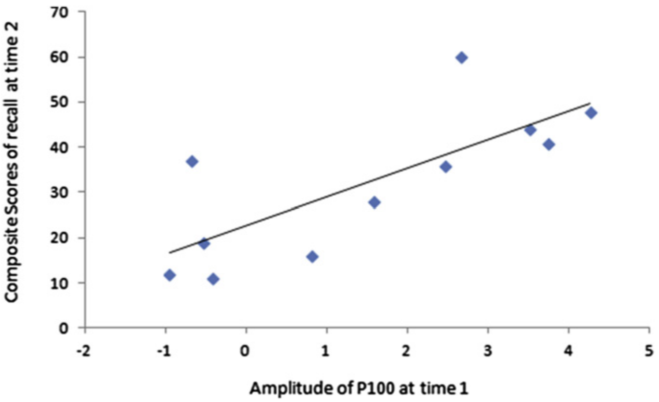


Fig. 11. Brain behavior correlations. The amplitude of P100 at time 1 in the left occipito-temporal area is positively correlated with the composite score of recall at time 2 in the non-writing training participants (the reading-only and visual-chunking group combined).

Buonocore, Girelli, & Jha, 1998; Morgan, Klein, Boehm, Shapiro, & Linden, 2008). It indexes the suppression of irrelevant information (Luck et al., 1994; Martínez et al., 2006), a mechanism that is helpful for efficient encoding. It has been found that increased working memory load leads to increased P100 amplitude in complex visual object processing (Haenschel et al., 2007) and in Chinese character processing (Shimoyama, Nakajima, Ito, & Shibata, 1997). Adults with dyslexia have been found to show reduced amplitude of P100 in a lexical decision task, compared with controls (Dujardin et al., 2011), suggesting a link between poor orthography–phonology connections, which are characteristic of dyslexia, and limited visual attention resources.

When applied to character recognition, a basic question is what level of visual attention is reflected in the P100. One possibility is that the P100 reflects a low level visual sensitivity at the letter or stroke rather than at the whole word level. Indeed, Proverbio and Adorni (2009) found that the P100 was enhanced when attention was drawn to letter-level processing (letter detection), while the N170 was enhanced when the attention was drawn to the lexical level (lexical decision). Extending this observation to the learning situation examined here, it suggests that writing training may lead to processing that includes greater visual attention to the elementary features of the character—sequences of strokes – and thus to greater sensitivity to these features during recognition. One previous fMRI study found increased activation in the superior parietal lobule during Chinese character recognition following writing training (Cao et al., 2012). This study is consistent with our current finding of P100, because the superior parietal lobule is associated with visual-spatial attention (Wojciulik, 1999). The increased visual attention may be due to the motor practice in the writing training. Previous fMRI studies have also found that writing training helps to establish a high-quality orthographic representation in left fusiform gyrus in single letter learning (James, 2010) as well as Chinese character learning (Cao et al., 2012). The current ERP study and precious fMRI studies suggests that writing helps reading presumably because the increased visual attention to the visual features of letters/strokes, the critical component that is promoted by writing through motor practice leads to high-quality representation of orthography. Our results seem to suggest that this visual sensitivity can be generalized to other characters that were not learned through intensive writing training. This possibility would contrast with the results of our previous fMRI study, which found a writing effect at bilateral visuo-orthographic regions and sensori-motor cortex that was specific to the characters that were learned through writing and did not generalize to other characters (Cao et al., 2012). Thus, while writing itself may be character-specific in its *orthographic effect*, the *early visual attention that writing promotes* that can only be identified by high-temporal resolution measure – ERP may become a more generalized procedure.

Our interpretation that the P100 reflects visual learning processes—low level visual attention (P100)—is strengthened by the predictive value of this component for long-term retention. For the non-writing training participants, the P100 effect at time 1 predicted recall of character 3 months later.

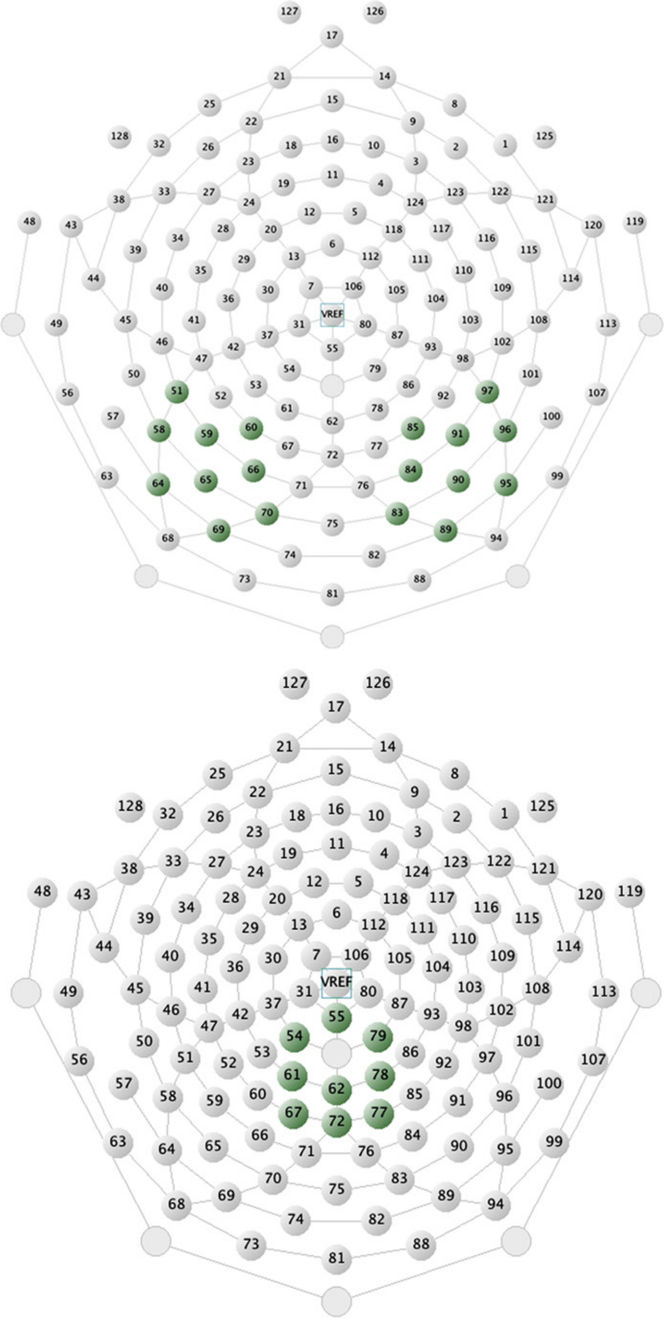


Fig. 12. Montages. The electrode cluster used for P100 and N170 and the cluster used for P600.

This suggests that learners who, in the context of passive viewing (reading only and visual chunking), attend to the visual features of the character and who, thereby, acquire a stronger orthographic representation of the character, may better retain the character and its connection to meaning and pronunciation. Rutman, Clapp, Chadick, and Gazzaley (2010) found that the degree to which participants modulate the early P100 (97–129 ms) event-related potential during selective stimulus encoding significantly correlated with their subsequent recognition for visual stimuli (Rutman et al., 2010). Our results add the possibility that the early stage visual processing indexed by P100 supports the acquisition of a visual representation that can support meaning and sound connections in long-term memory.

The coupling of meaning and sound recall with visual processing aligns with the assumption that a high quality orthographic representation is needed to support lexical identity in Chinese and in reading more generally. Even in alphabetic languages, where phonology is a strong support for reading, sensitivity to orthography is important for reading acquisition. A recent study found that the sensitivity to print as indexed by greater N170 response to words than to symbols in kindergarten predicts reading skills in second grade (Bach, Richardson, Brandeis, Martin, & Brem, in press).

As to the divergence between ERP and behavioral measures, ERP measures are more sensitive to the cortical processes underlying learning and memory than are behavior measures. Indeed, explicit indicators and implicit ERP indicators of language learning sometimes diverge (McLaughlin, Osterhout, & Kim, 2004; Tokowicz & MacWhinney, 2005). For example, in a study of L2 word learning, McLaughlin et al. (2004) found that ERPs showed that learners could discriminate between real words and pseudowords, even when they performed at chance levels in their overt decisions. Although our study found divergence as well, it also suggests that the implicit ERP indicator is predictive of long-term behavioral outcomes.

In summary, our study suggests that writing and visual chunking instruction produce differential orthographic enhancement functions that are detectable by ERP measures. Writing training has advantages for early visual attention and visual chunking training has advantages for orthographic recognition, both of which can be generalized to characters that were not learned through writing or visual chunking. When learners use their own strategies for encoding characters, our results suggest that those who show ERP evidence of early visual attention retain character-meaning-sound connections over a period of three months or more.

References

Anderson, R. C., Li, W., Chen, X., He, Y., Ku, Y.-M., & Shu, H. (2004, August). Visual chunking skills in Chinese character reading and writing. In W. Li (Ed.), *Chinese children's literacy development*, Symposium presented at the 28th International Congress of Psychology, Beijing, China.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *NeuroImage*, in press.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*, 235–260.

Cao, X., Li, S., Zhao, J., Lin, S. e., & Weng, X. (2011). Left-lateralized early neurophysiological response for Chinese characters in young primary school children. *Neuroscience Letters, 492*, 165–169.

Cao, F., Vu, M., Chan, D. H. L., Lawrence, J. M., Harris, L. N., Guan, Q., et al. (2012). Writing affects the brain network of reading in Chinese: a functional magnetic resonance imaging study. *Human Brain Mapping*, . http://dx.doi.org/10.1002/hbm.22017.

Cao, X. H., & Zhang, H. T. (2011). Change in subtle N170 specialization in response to Chinese characters and pseudocharacters. *Perceptual and Motor Skills, 113*(2), 365–376.

Cui, Y. H. (1997). Chinese basic components and Chinese L2 instruction. *Language and Application, 3*.

Curran, T. (1999). The electrophysiology of incidental and intentionalretrieval: ERP old/new effects in lexical decision andrecognition memory. *Neuropsychologia, 37*, 771–785.

Dujardin, T., Etienne, Y., Contentin, C., Bernard, C., Largy, P., Mellier, D., et al. (2011). Behavioral performances in participants with phonological dyslexia and different patterns on the N170 component. *Brain and Cognition, 75*, 91–100.

Flores d'Arcais, G. (1994). Order of strokes writing as a cue for retrieval in reading Chinese characters. *European Journal of Cognitive Psychology, 6*, 337–355.

Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., & Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology, 103*, 509–522.

Haenschel, C., Bittner, R. A., Haertling, F., Rotarska-Jagiela, A., Maurer, K., Singer, W., et al. (2007). Contribution of impaired early-stage visual processing to working memory dysfunction in adolescents with schizophrenia: a study with event-related potentials and functional magnetic resonance imaging. *Archives of General Psychiatry, 64*, 1229–1240.

Hu, J. (2004). GB13000.1 for information processing: components of Chinese characters. In N. L. Committee (Ed.). Chinese Press.

James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science, 13*, 279–288.

James, K. H., & Atwood, T. P. (2009). The role of sensorimotor learning in the perception of letter-like forms: tracking the causes of neural specialization for letters. *Cognitive Neuropsychology, 26*, 91–110.

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory–motor brain network. *Neuropsychologia, 44*, 2937–2949.

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.

Longcamp, M., Anton, J.-L., Roth, M., & Velay, J.-L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *NeuroImage, 19*, 1492–1500.

Longcamp, M., Zerbato-Poudou, M.-T., & Velay, J.-L. (2005). The influence of writing practice on letter recognition in preschool children: a comparison between handwriting and typing. *Acta Psychologica, 119*, 67–79.

Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 887–904.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience, 7*, 703–704.

Mangun, G. R., Buonocore, M. H., Girelli, M., & Jha, A. P. (1998). ERP and fMRI measures of visual spatial selective attention. *Human Brain Mapping, 6*, 383–389.

Martínez, A., Teder-Sälejärvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., et al. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience, 18*, 298–310.

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.

Morgan, H. M., Klein, C., Boehm, S. G., Shapiro, K. L., & Linden, D. E. J. (2008). Working memory load for faces modulates P300, N170, and N250r. *Journal of Cognitive Neuroscience, 20*, 989–1002.

Naka, M. (1998). Repeated writing facilitates children's memory for pseudocharacters and foreign letters. *Memory & Cognition, 26*, 804–809.

Pak, A. K. H., Cheng-Lai, A., Tso, I. F., Hua, S., Wenling, L., & Anderson, R. C. (2005). Visual chunking skills of Hong Kong children. *Reading & Writing, 18*, 437–454.

Parkinson, J., Dyson, B., & Khurana, B. (2010). Line by line: the ERP correlates of stroke order priming in letters. *Experimental Brain Research, 201*, 575–586.

Peng, D., Guo, D. J., & Zhang, S. L. (1985). Access to the lexical information of Chinese characters during semantic judgment. *Acta Psychologica Sinica, 3*, 227–233.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: some implications of research on Chinese for general theories of reading. *Psychological Review, 112*, 43–59.

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.

Proverbio, A. M., & Adorni, R. (2009). C1 and P1 visual responses to words are enhanced by attention to orthographic vs. lexical properties. *Neuroscience Letters, 463*, 228–233.

Rugg, M. D., & Nagy, M. E. (1989). Event-related potentials and recognition memory for words. *Electroencephalography and Clinical Neurophysiology, 72*, 395–406.

Rutman, A. M., Clapp, W. C., Chadick, J. Z., & Gazzaley, A. (2010). Early top–down control of visual processing predicts working memory performance. *Journal of Cognitive Neuroscience, 22*, 1224–1234.

Shadmehr, R., & Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. *Science, 277*, 821–825.

Shimoyama, I., Nakajima, Y., Ito, T., & Shibata, T. (1997). Visual evoked potentials relating to imagery: words for concrete objects versus absolute concepts. *Brain Topography, 9*, 271–274.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: implications for learning to read. *Child Development, 74*, 27–47.

Smith, M. E., Stapleton, J. M., & Halgren, E. (1986). Human medial temporal lobe potentials evoked in memory and language tasks. *Electroencephalography and Clinical Neurophysiology, 63*, 145–159.

Song, H., Zhang, H. C., & Shu, H. (1995). Developmental changes in functions of orthography and phonology in Chinese reading. *Acta Psychologica Sinica, 2*, 139–144.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Studies in Second Language Acquisition, 27*, 173–204.

Wan, Y. X. (1999). Chinese characters categorization and Chinese L2 instruction. *Language Instruction and Research, 4*.

Wojciulik, E. (1999). The generality of parietal involvement in visual attention. *Neuron, 23*(4), 747–764.

Wu, X., Li, W., & Anderson, R. C. (1999). Reading instruction in China. *Journal of Curriculum Studies, 31*, 571.

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.

Zhang, W. X. (1990). *From basic components to structures of Chinese characters*. World Chinese instruction 20.

Zhang, M., Jiang, T., Mei, L., Yang, H., Chen, C., Xue, G., et al. (2011). It's a word: early electrophysiological response to the character likeness of pictographs. *Psychophysiology, 48*, 950–959.

Zhu, X., & Taft, M. (1994). *Complexity effects in Chinese character processing*. Paper presented to the Asian–Australian Workshop on Cognitive Processing of Asian Languages.