Effects of visual complexity and sublexical information in the occipitotemporal cortex in the reading of Chinese phonograms: A single-trial analysis with MEG

Chun-Hsien Hsu a,⁎, Chia-Ying Lee b, Alec Marantz c,d

a Institute of Neuroscience, National Yang-Ming University, Taiwan
b Institute of Linguistics, Academia Sinica, Taiwan
c Department of Linguistics, New York University, USA
d Department of Psychology, New York University, USA

ABSTRACT

We employ a linear mixed-effects model to estimate the effects of visual form and the linguistic properties of Chinese characters on M100 and M170 MEG responses from single-trial data of Chinese and English speakers in a Chinese lexical decision task. Cortically constrained minimum-norm estimation is used to compute the activation of M100 and M170 responses in functionally defined regions of interest. Both Chinese and English participants’ M100 responses tend to increase in response to characters with a high number of strokes. English participants’ M170 responses show a posterior distribution and only reflect the effect of the visual complexity of characters. On the other hand, the Chinese participants’ left hemisphere M170 is increased when reading characters with high number of strokes, and their right hemisphere M170 is increased when reading characters with small combinatorial of semantic radicals. Our results suggest that expertise with words and the decomposition of word forms underlies processing in the left and right occipitotemporal regions in the reading of Chinese characters by Chinese speakers.

1. Introduction

Skilled readers show remarkable efficiency in visual word recognition. There has been a growing body of research concerning the manner in which readers extract visual features, word forms, and lexical and meaning-related information during the early stages of visual word recognition. With excellent temporal resolution, electrophysiological methods have proved most appropriate for addressing this kind of question. For example, in studies using event-related potentials (ERPs), Sereno, Rayner, and Posner (1998) have reported lexicality and frequency effects in the N1 ERP component. Hauk et al. (2006) have reported a lexicality effect around 200 ms and an earlier interaction between lexicality and the orthographic typicality around 160 ms. Similar inferences have been drawn from magnetoencephalography (MEG) studies that have demonstrated differentiated early MEG components reflecting processes at the form level or orthographic level in visual word recognition (Solomyak & Marantz, 2009; Tarkiainen, Cornelissen, & Salmelin, 2002; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999). Tarkiainen and colleagues have demonstrated that the Type I activity, also called the M100 response, is found around 100 ms in or near the primary visual cortex (V1) and is sensitive to the low-level analysis of visual features. The differentiation between letter and symbol strings was picked up by a later component, Type II activity, around 150 ms (greater for letter than symbol strings) in the inferior-temporal cortex, reflecting object-level processing (Tarkiainen et al., 1999, 2002). In other work, a robust difference between orthographic and non-orthographic stimuli has been reported in the M170 or the N170 responses (Bentin, Mouchetant-Rostaing, Giard, Eckallier, & Penrier, 1999; Solomyak & Marantz, 2009; Tarkiainen et al., 1999; Zweig & Pykkänen, 2009). Studies have suggested that this response originates from the left occipitotemporal region, the so-called visual word form area (VWFA), and that its function is specific to processing orthographic stimuli (Cohen et al., 2000; McCandliss, Cohen, & Dehaene, 2003).

Although a pattern seems to be emerging from these studies, indicating that the earliest electrophysiological response to visual word recognition in the brain peaks around 100 ms after stimulus onset and reflects the analysis of the surface feature of a visual word, and a subsequent component at around 150–200 ms reflects lexicality or the identification of word forms, the findings are still inconsistent across studies. Controversy still surrounds the nature of the representations accessed during the early processing of a word. This inconsistency could be attributable to the fact that most studies only look at one or two linguistic properties, such as lexicality, lexical frequency, or word length, in a factorial design,
in order to determine the effect of a particular variable on brain responses. However, most psycholinguistic variables, such as frequency or word length, are continuous in nature, and many co-vary in complicated ways. By dichotomizing the continuous variables into categories (e.g., high versus low frequency) and trying to match other variables (e.g., word length), factorial designs may result in substantial loss of statistical power and require selection of atypical stimuli (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Moreover, high correlation between the confounding and manipulated variables can be problematic to their interpretation.

One possible way to overcome these problems is to use regression analysis in order to find the best-fit relationship between brain activities and a set of predicting variables. This approach has only very recently been applied to the field of human electrophysiology (Dambacher, Kliegl, Hofmann, & Jacobs, 2006; Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006; Hauk et al., 2006; Solomyak & Marantz, 2009, 2010). For example, Hauk et al. (2006) applied it to examine how 10 psycholinguistic features of 300 stimulus words modulate the processing of visual word recognition in different time courses at distinct brain regions. They found that variables associated with the surface structure of a word (such as word length and orthographic typicality) showed effects on the left inferior-temporal cortex within the first 100 ms after stimulus onset, and that lexical frequency showed its effect slightly later at 110 ms, with a semantic variable modulating a widely distributed cortical network shortly after 160 ms, simultaneously with lexicality.

A series of studies conducted by Solomyak and Marantz (2009, 2010) also employed single-trail correlational analysis, this time with MEG in the source space, to investigate whether lexical access occurs within 200 ms after perceiving a word. They first found that M170 is sensitive to form properties (such as bigram, trigram, and lexical frequency), but not to a targeted lexical property (heteronym frequency ratio: relative frequency of the different meanings of a heteronym) and concluded that true lexical properties did not affect the processing until after 300 ms, while earlier activation of M170 is primarily modulated by orthographic form. Their later study showed that the orthographic form features exert an effect on earlier stages in processing around 130 ms, and then the M170 is sensitive to the conditional probability between affixes and stems of morphologically complex words (Solomyak & Marantz, 2009), which reflects morphological decomposition in earlier stages of reading based on the visual word forms of stems and affixes.

Chinese is characterized as an ideographic writing system that presents the highest contrast to alphabetic systems such as that used for English. The character forms the basic unit of the Chinese writing system. Each character is composed of basic strokes, and these strokes are then combined to form components called radicals. Approximately 80% of the Chinese characters are phonograms that consist of a semantic radical (usually on the left) and a phonetic radical (usually on the right). The semantic radical is related to the meaning of a given character, while the phonetic radical is related to its pronunciation. In addition, the phonetic radicals are usually stand-alone characters with their own pronunciations and meanings, which do not contribute to the meaning of the character. There is increasing evidence indicating that reading a complex character involves the processing of its radicals (Ding, Peng, & Taft, 2004; Feldman & Siok, 1997, 1999; Hsu, Tsai, Lee, & Tzeng, 2009; Lee, Tsai, Huang, Hung, & Tzeng, 2006; Lee, Tsai, Su, Tzeng, & Hung, 2005; Lee et al., 2007). Lee et al. (2006) conducted a priming experiment and found a reduced N400 activation at 50 ms and 100 ms stimulus onset asynchronies (SOA) when a target (e.g., 雨/yu3/ (rain) was semantically related to the phonetic radical (e.g., 風/feng1/ (wind) embedded in the prime (e.g., 棟/feng1/). Furthermore, such a priming effect was absent at 500 ms SOA. Since the reduction in the N400 amplitude indexes the associative semantic relation in prime-target pairs (Kutas & Federmeier, 2000), the result suggests that phonograms are decomposed at the earlier stage of visual word recognition, and that the semantic value of the phonetic radical is temporarily accessed. The evidence for automatic sublexical semantic activation of Chinese phonetic radicals supports the hypothesis that sublexical decomposition is purely structural and arises in the early stage of visual word recognition (Ding et al., 2004; Perfetti, Liu, & Tan, 2005).

However, it remains unclear what kind of properties of the characters contribute to their decomposition into radicals. The hypothesis of a VWFA suggests that this region is sensitive to informative sublexical units (Dehaene, Cohen, Sigman, & Vinckier, 2005), e.g., open-bigrams in English. For Chinese characters, Feldman and Siok (1997) have introduced the variable "combinability of radicals," defined as the number of phonograms that share the same phonetic or semantic radical, as the appropriate measurement for the component frequency of phonograms. Several studies have demonstrated faster response latency in reading characters with large combinability compared with reading characters with small combinability in a lexical decision task and a semantic judgment task (Chen & Weekes, 2004; Feldman & Siok, 1997, 1999; Hsiao, Shillcock, & Lavidor, 2006, 2007). In fact, studies have demonstrated that a variety of sublexical properties can be found to modulate the early stages of lexical processing. Hsu et al. (2009) have demonstrated that large phonetic combinability characters elicited greater negativity at N170 than small phonetic combinability characters. Furthermore, this effect was further modulated by phonetic consistency, which indicates whether the pronunciation of a phonogram agrees with phonograms containing the same phonetic radical. The combinability effect on the N170 was only found in the reading of high-consistency characters (Hsu et al., 2009). Meanwhile, reading high-consistency characters produced greater negativity of the N170 than reading low-consistency characters in the bilateral occipitotemporal electrodes. Several fMRI studies on reading Chinese have also demonstrated effects of frequency, consistency, and lexicality in the bilateral occipitotemporal area (Kuo et al., 2003; Lee, Huang, Kuo, & Tzeng, 2010; Lee et al., 2004). All this evidence suggests that the N170 originating in bilateral occipitotemporal region may index early decompostion processing in reading Chinese characters.

It is noteworthy that most studies of alphabetic writing systems suggest that the left lateralized occipitotemporal regions play important roles in earlier processes of visual word recognition, compared to a bilateral activation for Chinese characters and Japanese Kanji (Bolger, Perfetti, & Schneider, 2005; Nakamura, Dehaene, Jobert, Bihan, & Kouider, 2005; Tan, Laird, Li, & Fox, 2005). Nakamura et al.’s (2005) study found increased activation in the anterior and posterior fusiform gyrus regions in both hemispheres of Japanese participants as they read Kanji and Kana, respectively. A study involving Koreans who were educated in both Chinese and written English reported a left-lateralized N170 effect for both English and Korean words but a bilateral N170 effect for Chinese characters and pictures (Kim, Yoon, & Park, 2004). Such cross-linguistic differences provide support for the phonological mapping hypothesis of the left-lateralized N170 effect (Maurer & McCandliss, 2007). That is, the left lateralization of the reading-related N170 effect may reflect the spelling-to-sound mapping consistency of alphabetic writing systems.

In addition, an fMRI study has shown that after training with Chinese characters’ pronunciations and meaning, native English speakers’ fusiform activation is stronger in the right hemisphere than in the left hemisphere in response to Chinese characters (Liu, Dunlap, Fiez, & Perfetti, 2007). Furthermore, native English
speakers who have learned Chinese characters show a stronger N170 activation on the right posterior scalp than on the left posterior scalp (Liu, Perfetti, & Wang, 2006), which is consistent with the assumption that the amplitude of N170 activation is associated with the language experience of visual word recognition (McCandliss, Posner, & Givón, 1997; Wong, Gauthier, Woroch, Debuse, & Curran, 2005). These results suggest that expertise of Chinese character recognition affects processing in the RH fusiform gyrus regions.

Meanwhile, although radical combinability, consistency, frequency, and lexicality have been identified as modulating N170 responses, our corpus analysis across 3967 phonograms (selected from the Academia Sinica Balanced Corpus; Huang & Chen, 1998) shows that the number of strokes is significantly correlated to the frequency of the character ($r = -1.139, p < .001$), phonetic combinability ($r = -.140, p < .001$) and semantic combinability ($r = -.190, p < .001$). To overcome the colinearity among these variables, the single-trial regression analysis which has been used in lexical-semantic processing. We hope to evaluate the contributions from various features of a character reflecting different aspects of Chinese character recognition in modulating early and later activations from the bilateral character regions. It is thus important to investigate whether the activation of bilateral occipitotemporal regions in later time windows would show the effects of lexical-semantic processing.

Two particular issues are addressed in the present study. First, we hope to evaluate the contributions from various features of a character reflecting different aspects of Chinese character recognition in modulating early and later activations from the bilateral occipitotemporal regions. For example, would the decomposition of characters, which may be indexed by radical combinability, be performed primarily by the left hemisphere, as the morphological decomposition in the reading of English words? To this end, the influences of the number of strokes, combinability, character frequency, and two variables which index the semantic-level properties will be estimated by applying the linear mixed-effects model (Bates & Sarkar, 2007) to the distributed source data from MEG and ERP studies will be adopted (Hauk et al., 2006; Solomyak & Marantz, 2009, 2010). In addition, there are also studies suggesting that the fusiform gyrus regions contribute to processes of conceptual knowledge (Mechelli, Sartori, Orlandi, & Price, 2005; Wheatley, Weisberg, Beauchamp, & Martin, 2005). It is thus important to investigate whether the activation of bilateral occipitotemporal regions in later time windows would show the effects of lexical-semantic processing.

Second, we are going to further distinguish the effect of the form level from the linguistic level of character recognition on neural responses by comparing responses from groups with different features of a character reflecting different aspects of Chinese character recognition in modulating early and later activations from the bilateral occipitotemporal regions. For example, would the decomposition of characters, which may be indexed by radical combinability, be performed primarily by the left hemisphere, as the morphological decomposition in the reading of English words? To this end, the influences of the number of strokes, combinability, character frequency, and two variables which index the semantic-level properties will be estimated by applying the linear mixed-effects model (Bates & Sarkar, 2007) to the distributed source data from MEG and ERP studies will be adopted (Hauk et al., 2006; Solomyak & Marantz, 2009, 2010). In addition, there are also studies suggesting that the fusiform gyrus regions contribute to processes of conceptual knowledge (Mechelli, Sartori, Orlandi, & Price, 2005; Wheatley, Weisberg, Beauchamp, & Martin, 2005). It is thus important to investigate whether the activation of bilateral occipitotemporal regions in later time windows would show the effects of lexical-semantic processing.

2. Methods

2.1. Design and stimuli

Four hundred Chinese phonograms were selected as target characters from the Academia Sinica Balanced Corpus (Huang & Chen, 1998). The corpus is based on more than five million words (approximately 10 million characters) culled from textbooks, newspapers, works of literature, popular works of fiction, and transcripts. Target characters were selected using the following procedures. First, following Feldman and Siok’s (1997) definition, 3697 phonograms were identified for calculating the combinability of phonetic and semantic radicals. Then, 2711 phonograms which could be words on their own were chosen. The target characters were selected to form a similar distribution in the number of strokes around the norm. Finally, the target characters were coded according to number of strokes, character frequency, phonetic combinability, semantic combinability, semantic ambiguity, and their noun-to-verb ratio.

Previous studies have used subjective ratings as measurements for the semantic ambiguity of Chinese characters (Perfetti & Tan, 1998; Tan et al., 2000), and this method was adopted here as follows: 5640 characters from the Academia Sinica Balanced Corpus were randomly separated into 15 subsets. Each subset included 376 characters. Two-hundred and seventy native Chinese college students participated in the rating task, and each participant received one subset. Thus, every 15 participants would complete a corpus, and each character had an average rating value from 18 participants. Participants were also asked to give each character a value from one (precise meaning) to five (vague meaning) according to the diversity of the meanings when the character was being used in a word. They were presented with characters one by one in a random order on a screen. In each trial, one character was presented in the center of the screen and remained there until the participant made a response by pressing a number key (1–5). For instructions, they were asked to generate two-character words by using the character in as many words as possible, no matter the position of the target character in a two-character word, and to estimate the number of different meanings corresponding to the target character. The rating 1 thus represents that the meaning of the character is explicit and that only one meaning is available (e.g. 糖/tan2/y, sweet) or that the character is a “bind-” character (Taft, 2006), which is always bound with the other character even though there is no explicit meaning when it stands alone (e.g. 碎/le4/ is always used with 腐/se4/ as 垃圾/le4/ which means garbage). The rating 2 indicates that there are two or three different meanings available for the character (e.g. 程/cheng2/y, rules or a journey). The rating 3 indicates that the character has three or four different meanings (e.g. 火/hua1/y, flower, spend, colorful, blurred, pattern, trick, peanut, etc.). The rating 4 indicates that the meaning of the character is fairly diverse, and there are four to five different meanings available for the character (e.g. 風/feng1/y, wind, style, landscape, and disease). The rating 5 indicates that the meaning of the character is highly diverse and can have more than five different meanings (e.g. 花/hua1/y, flower, spend, colorful, blurred, pattern, trick, peanut, etc.).

Table 1 illustrates the mean of each variable of interest and their cross-correlation coefficients. The number of strokes was negatively correlated with character frequency, phonetic combinability, semantic combinability and semantic ambiguity, while semantic ambiguity was positively correlated with character frequency and negatively associated with the number of strokes and semantic combinability. None of the other variables were significantly correlated. To balance the response during the lexical decision, four hundred items were created by combining two
radicals that do not co-occur in Chinese (pseudo-characters) or by combining two radicals and changing one stroke (non-characters).

2.2. Experimental procedure

The participants were ten native Chinese speakers and five native English speakers. The participants' approval for the study was obtained from the Institutional Review Boards at NYU, and all participants were paid for their participation. Experimental stimuli were projected onto a screen above the participants' heads while they lay in the magnetically shielded room that houses the MEG. The continuous MEG data were then epoched with 100 ms pre-stimulus intervals and 900 ms post-stimulus intervals, and baseline corrected using the pre-stimulus data. Trials with amplitude or not. MEG data were recorded continuously by a 157-channel axial gradiometer whole-head MEG system (Kanazawa Institute of Technology, Kanazawa, Japan) throughout the task. A band-pass filter (DC to 100 Hz) was applied during the recording, which was acquired with a sampling frequency of 1 kHz.

Prior to MEG acquisition, each participant's head shape was digitized, and head position indicator coils were used to localize the position of the participant's head inside the MEG helmet. The head-shape digitization and head position indicator locations were later used to co-register the MEG coordinate system to that of each participant's structural MR images.

2.3. Analysis methods

2.3.1. MEG preprocessing and minimum-norm estimation

In off-line processing, MEG data were first noise reduced using the time-shift PCA algorithm (de Cheveigné & Simon, 2007). Jousmäki and Hari (1996) have suggested that the R peak of the electrocardiogram, which has maximum amplitude above 100 fT, is usually superimposed in the MEG signal over the temporo-occipital scalp. Because these artifacts may have a considerable effect on the time-shift PCA algorithm (de Cheveigné & Simon, 2007).

The continuous MEG data were then epoched with 100 ms pre-stimulus intervals and 900 ms post-stimulus intervals, and baseline corrected using the pre-stimulus data. Trials with amplitude variations larger than 1.5 pT were excluded from averaging and subsequent linear mixed-effects model analyses. Each participant's MEG data was then averaged and low-pass filtered at 40 Hz. The participants' structural MR images were processed in FreeSurfer (Cortechs Labs, La Jolla, CA and MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging, Charleston, MA) to create a cortical reconstruction of each brain. The MNE toolbox (MGH/ HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging, Charleston, MA) was then used to calculate a cortically constrained minimum-norm solution for each participant's MEG data with 5124 sources on each participant's cortical surface. The boundary-element model method was used to compute the forward solution, which estimates the resulting magnetic field at each MEG sensor by the activity at each of the 5124 sources. This forward solution was then employed to create the inverse solution, which identified the spatio-temporal distribution of activity over sources that best account for each participant's average MEG data. Only components of activation that were in the direction normal to the cortical surface were retained in the minimum-norm solution, and the resulting minimum-norm estimates were converted into a dynamic statistical parameter map (dSPM), which measures the noise-normalized activation at each source to avoid some of the inaccuracies of standard minimum-norm calculations (Dale et al., 2000).

2.3.2. Regions of interests (ROIs)

Each participant's cortical surface was normalized onto a standard brain supported by FreeSurfer, and then all participants' dSPM solutions were averaged for use in defining regions and time windows of interest. The averaged dSPM solutions of both hemispheres showed a clear pattern of early activation, which consisted of an earlier (~100 ms) negative (current flowing inward) peak in the posterior occipital region, and a slightly later (150–200 ms) positive (current flowing outward) peak in the occipitotemporal fusiform gyrus area. The earlier activation was defined as the M100 and the later activation as the M170. ROIs were created on the regions which displayed the pattern of activity described above in both LH and RH. Although the pattern of M170 activation was fairly consistent across participants, the location of the M100 region seemed to vary between participants. Therefore, the M100 region was defined on individual participant's activation after normalizing their cortical surface onto the standard brain. Then, an inverse solution was computed over each participant's raw MEG data over each identified ROI to compute the trial-by-trial minimum-norm estimation at each time point. The resulting minimum-norm estimates were converted into dSPM values for use in subsequent linear mixed-effects model analyses.

M100 and M170 analyses were generally conducted on signed activity, while that the analysis of English participants' LH M100 responses was conducted on absolute activity (for reasons to be discussed below). The dependent variable was the average amplitude over a 50-ms window centered at each participant's peak latency in each ROI. The Chinese participants' M170 analysis was also conducted on the rise, the height reached at the peak, and fall of the peak in 20 ms windows: a “rise” time window spanning 30–10 ms prior to the peak, a “height” time window centered at the

### Table 1

<table>
<thead>
<tr>
<th>Number of strokes</th>
<th>Cross-correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Character frequency</td>
<td>Phonetic combinability</td>
</tr>
<tr>
<td>Semantic ambiguity</td>
<td>Semantic ambiguity</td>
</tr>
<tr>
<td>13.15 (4.24)</td>
<td>2.36 (1.03)</td>
</tr>
<tr>
<td>7.16 (4.11)</td>
<td>86.68 (68.77)</td>
</tr>
<tr>
<td>1.73 (57)</td>
<td>0.23 (1.05)</td>
</tr>
</tbody>
</table>

Means, standard deviations, and cross-correlation coefficients for the dependent variables used in the linear mixed-effects model.

* p < .05.
** p < .01.
*** p < .001.
peak latency, and a “fall” time window spanning 10 through 30 ms after the peak.

In the remaining trials without amplitude variations larger than 1.5 \(t\)1, trials of real characters were included in subsequent linear mixed-effects model analyses. Chinese participants’ mean number of trials was 359 (ranging from 330 to 396), and English participants’ mean number of trials was 327 (ranging from 287 to 373)\(^1\). The averaged amplitude was analyzed by the linear mixed-effects model with participants and items as crossed random effects, separately for each MEG response and each participant group by employing the lmer program of the lme4 package (Baayen, 2008; Bates & Sarkar, 2007). For fixed effects, we included trial number to minimize variance attributable to practice or fatigue effects. Finally, character frequency was transformed logarithmically. The estimated coefficient \(b\), standard error, and \(t\) value for each variable was reported. The \(p\) values were obtained through Markov Chain Monte Carlo (MCMC) sampling supported by the pvals.fnc program of the languageR package (Baayen, Davidson, & Bates, 2008). These packages are included in the R system for statistical computing (Ver. 2.9.1; R Development Core Team, 2009).

### 2.3.3. Statistical methods for testing effects in VWFA over later time windows

To find the time windows of interest for possible later effects in the VWFA, the same M170 ROIs were used. The activation at each time point over 300–800 ms was normalized within participants, and then all participants’ data were concatenated. Correlations were then computed point by point by across all participants’ normalized values to create a correlation wave of each independent variable of interest over time. The correction procedure was applied to the correlation waves based on Maris and Oostenveld’s (2007) method as follows: A new statistic, \(\sum r\), was defined by summing up all correlation coefficients within a temporal cluster of consecutive significant effects in the same direction (at the raw \(p < 0.05\) significance level). In addition to the original \(\sum r\) values, 10,000 permutations of the independent variables were used to compute the correlation wave as it had been in the original analysis. For each randomly produced wave, the \(\sum r\) statistic was computed for each of its temporal clusters, after which the highest absolute value of \(\sum r\) was taken as the statistic for that permutation. The percentage of values higher in absolute value than the original statistic was taken to be the Monte Carlo \(p\) value for the original cluster, which indicated the statistical significance of the \(\sum r\) statistic for each temporal cluster. Then, the linear mixed-effects model analyses were conducted by fitting the mean amplitude within the significant time cluster of both LH and RH VWFA, with participants and items as crossed random effect.

### 3. Results

#### 3.1. Behavioral results

The mean response times and accuracies of lexical decision tasks are shown in Table 2. Two linear mixed-effects models were used to estimate the group effect and the Chinese participants’ lexicality effect in response time with two random effects, including participants and items. The results showed that the Chinese participants revealed a faster response time than the English participants (\(b = 157, SE = 57, t = 2.74, p < 0.01\)). The Chinese participants’ response time was faster in response to real characters compared to pseudo-characters (\(b = 130, SE = 5.89, t = 22.077, p < 0.001\)) and non-characters (\(b = 26, SE = 7.44, t = 3.501, p < 0.001\)). For error data, the same design of mixed-effects models was used with the logistic link function and binomial variance. Correct and incorrect responses were scored as 0 and 1, respectively. The results showed that the Chinese participants made fewer errors than the English participants (\(b = 2.32, SE = .15, z = 15.28, p < 0.001\)). With regard to the lexicality effect, Chinese participants made fewer errors in response to real characters compared to pseudo-characters (\(b = 1.479, SE = .125, z = 11.834, p < 0.001\)) and made more errors in response to real characters compared to non-characters (\(b = .952, SE = .244, z = −3.904, p < 0.001\)).

#### 3.2. M100 responses

Fig. 1 illustrates the ROIs of the M100 response and the estimated amplitude (dSPM) value from one representative Chinese participant and one representative English participant. For Chinese participants, the mean M100 peak latency was 89 ms (ranging from 79 to 110 ms) in the left hemisphere and 79 ms (ranging from 61 to 101 ms) in the right. For English speaking participants, the mean M100 peak latency was 98 ms (ranging from 71 to 123 ms) in the left hemisphere and 89 ms (ranging from 62 to 103 ms) in the right. One English speaking participant did not show the typical pattern of M100 activity for the left hemisphere and showed a positive peak at 103 ms in the relevant ROI. Table 3 shows the results of M100 analyses. For Chinese participants, the M100 response in the LH and the RH was negatively correlated with the number of strokes (LH: \(b = −.0028, SE = .0012, t = −2.35, p < 0.05\); RH: \(b = −.0045, SE = .0013, t = −3.57, p < 0.01\)). The RH M100 response was also negatively correlated with phonetic combinability (\(b = −.0034, SE = .0013, t = −2.68, p < 0.01\)). As the M100 response was based on negative activity, these results suggest that large visual complexity increased the amplitude of LH M100 responses, and that large phonetic combinability increased the amplitude of RH M100 responses (Fig. 3a–c). English participants’ left hemisphere M100 responses which were measured by absolute values, were also positively correlated with the number of strokes (\(b = .003, SE = .0015, t = 2.04, p < 0.05\) (Fig. 4a), which indicated that large visual complexity increased the amplitude of their LH M100 responses.

### Table 2

Means and standard errors for response times and accuracies in lexical decision task.

<table>
<thead>
<tr>
<th></th>
<th>Chinese participants</th>
<th>English participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Real characters</td>
<td>Pesudo-characters</td>
</tr>
<tr>
<td>Response times (ms)</td>
<td>650(30)*</td>
<td>785(44)</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>93(1)</td>
<td>81(3)</td>
</tr>
</tbody>
</table>

\* Values in parentheses are standard error.
In addition, Chinese participants’ LH and RH M100 activation were increased by the trial number (LH: $b = .0001, SE = .000, t = -3.29, p < .001$; RH: $b = -.0001, SE = .000, t = -4.06, p < .001$), and this did not have significant effect on English participants’ LH nor RH M100 activation (LH: $b = .000, SE = .000, t = -0.89, p = .39$; RH: $b = .000, SE = .000, t = -0.04, p = .92$).

### 3.3. M170 responses

Fig. 2 illustrates the ROIs of the M170 response and the mean amplitude (dSPM) of both Chinese and English participants. For Chinese participants, the mean M170 latency was 165 ms (ranging from 144 to 206 ms) in the left hemisphere and 180 ms (with a range of 151–248 ms) in the right. For English participants, the mean M170 latency was 179 ms (ranging from 146 to 210 ms) in the left hemisphere and 165 ms (with a range of 147–186 ms) in the right. Table 4 shows the results of M170 analyses. For Chinese participants, the LH M170 amplitude was positively correlated with the number of strokes ($b = .0032, SE = .0012, t = 2.56, p < .05$). In addition, the semantic combinability had a negative coefficient in the 20 ms time window centered at the RH M170 responses ($b = -0.0002, SE = 0.0012, t = 2.56, p < .05$). As the analyses were of positive activity, the correlation suggests that visual complexity increased the LH M170 amplitude responses, and that low

---

Table 3

<table>
<thead>
<tr>
<th>Variables</th>
<th>Chinese participants – LH M100; $R^2 = .20$</th>
<th>English participants – LH M100; $R^2 = .48$</th>
<th>Chinese participants – RH M100; $R^2 = .14$</th>
<th>English participants – RH M100; $R^2 = .25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>.0709</td>
<td>.2</td>
<td>.003</td>
<td>.1197</td>
</tr>
<tr>
<td></td>
<td>.0548</td>
<td>.1154</td>
<td>.0445</td>
<td>.0817</td>
</tr>
<tr>
<td></td>
<td>1.29</td>
<td>1.33</td>
<td>1.07</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>.965</td>
<td>−1.197</td>
<td>.000</td>
<td>−0.117</td>
</tr>
<tr>
<td></td>
<td>−.07</td>
<td>−1.17</td>
<td>.000</td>
<td>−1.47</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.0817</td>
<td>.000</td>
<td>.177</td>
</tr>
<tr>
<td></td>
<td>.0445</td>
<td>.0817</td>
<td>.000</td>
<td>.916</td>
</tr>
<tr>
<td>Trial number</td>
<td>-.0001</td>
<td>&lt;.001***</td>
<td>.004</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>−4.06</td>
<td>&lt;.001***</td>
<td>−.04</td>
<td>.916</td>
</tr>
<tr>
<td>Number of strokes</td>
<td>−.0045</td>
<td>&lt;.001***</td>
<td>.000</td>
<td>−.0028</td>
</tr>
<tr>
<td></td>
<td>.0013</td>
<td>.000</td>
<td>.000</td>
<td>−.0028</td>
</tr>
<tr>
<td></td>
<td>−3.57</td>
<td>−1.61</td>
<td>.009</td>
<td>−.161</td>
</tr>
<tr>
<td>Frequency</td>
<td>.0121</td>
<td>.16</td>
<td>.052</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>.0089</td>
<td>.16</td>
<td>.009</td>
<td>.542</td>
</tr>
<tr>
<td></td>
<td>1.36</td>
<td>.58</td>
<td>.50</td>
<td>.542</td>
</tr>
<tr>
<td>Phonetic combinability</td>
<td>−.0034</td>
<td>.000</td>
<td>.000</td>
<td>.51</td>
</tr>
<tr>
<td></td>
<td>.0013</td>
<td>.000</td>
<td>.000</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>−2.68</td>
<td>−.04</td>
<td>.017</td>
<td>.45</td>
</tr>
<tr>
<td>Semantic combinability</td>
<td>−.0001</td>
<td>.017</td>
<td>.017</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>.0001</td>
<td>.017</td>
<td>.017</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>−.79</td>
<td>.71</td>
<td>.71</td>
<td>.45</td>
</tr>
<tr>
<td>Semantic ambiguity</td>
<td>−.0038</td>
<td>.017</td>
<td>.017</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>.014</td>
<td>.71</td>
<td>.71</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>−.27</td>
<td>.45</td>
<td>.45</td>
<td>.45</td>
</tr>
<tr>
<td>NV ratio</td>
<td>−.0063</td>
<td>.71</td>
<td>.71</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>.0049</td>
<td>.45</td>
<td>.45</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>−1.27</td>
<td>−.066</td>
<td>.0164</td>
<td>−.4</td>
</tr>
<tr>
<td></td>
<td>.174</td>
<td>.86</td>
<td>.86</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>.057</td>
<td>.385</td>
<td>.385</td>
<td>.385</td>
</tr>
<tr>
<td></td>
<td>.0066</td>
<td>.661</td>
<td>.661</td>
<td>.661</td>
</tr>
<tr>
<td></td>
<td>.0164</td>
<td>.4</td>
<td>.4</td>
<td>.4</td>
</tr>
<tr>
<td></td>
<td>.86</td>
<td>.385</td>
<td>.385</td>
<td>.385</td>
</tr>
<tr>
<td></td>
<td>.385</td>
<td>.661</td>
<td>.661</td>
<td>.661</td>
</tr>
</tbody>
</table>

*p < .05.

**p < .01.

***p < .001.
semantic combinability increased the amplitude of RH M170 responses (Fig. 3d and e). The English participants’ LH M170 responses were also positively correlated with the number of strokes \((b = .0047, SE = .002, t = 2.29, p < .05)\) (Fig. 4b). Finally, the trial number did not have a significant effect on Chinese participants’ LH nor RH M170 activation (LH: \(b = .000, SE = .000, t = -1.29, p = .78\); RH: \(b = .000, SE = .000, t = .65, p = .51\)), and the English participants’ LH and RH M170 activation were decreased by the trial number (LH: \(b = -.0001, SE = .000, t = -1.97, p < .05\); RH: \(b = -.0001, SE = .000, t = -2.6, p < .05\)).
The current findings about the M170 response help elucidate the role of semantic ambiguity in the processing of Chinese characters. The RH VWFA regions were negatively correlated with semantic ambiguity, indicating that high semantic ambiguity words elicit more negative M170 activity in reading characters with larger phonetic combinability, which seems consistent with Solomyak and Marantz's (2009) finding that high bigram frequency words elicit more negative M100 activity. Although a bigram frequency effect in the opposite direction was also observed, the results suggest that V1 processes of feature integration are also sensitive to inter-feature probabilities of adjacent stroke patterns.

Table 4

<table>
<thead>
<tr>
<th>Variables</th>
<th>Chinese participants – LH M170; $R^2$ = .45</th>
<th>$p$-values</th>
<th>Chinese participants – RH M170; $R^2$ = .36</th>
<th>$p$-values</th>
<th>English participants – LH M170; $R^2$ = .21</th>
<th>$p$-values</th>
<th>English participants – RH M170; $R^2$ = .10</th>
<th>$p$-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>.368</td>
<td>.092</td>
<td>4</td>
<td>.000</td>
<td>.187</td>
<td>.0964</td>
<td>1.94</td>
<td>.081</td>
</tr>
<tr>
<td>Trial number</td>
<td>.000</td>
<td>.000</td>
<td>–29</td>
<td>.872</td>
<td>–.0001</td>
<td>.000</td>
<td>–1.97</td>
<td>.047</td>
</tr>
<tr>
<td>Number of strokes</td>
<td>.0032</td>
<td>.0012</td>
<td>2.56</td>
<td>.012</td>
<td>.0047</td>
<td>.002</td>
<td>2.29</td>
<td>.022</td>
</tr>
<tr>
<td>Frequency</td>
<td>.0053</td>
<td>.0086</td>
<td>.62</td>
<td>.53</td>
<td>–.0098</td>
<td>.0106</td>
<td>–.93</td>
<td>.351</td>
</tr>
<tr>
<td>Phonetic combinability</td>
<td>–.0002</td>
<td>.0012</td>
<td>–.19</td>
<td>.853</td>
<td>–.003</td>
<td>.002</td>
<td>–1.49</td>
<td>.139</td>
</tr>
<tr>
<td>Semantic combinability</td>
<td>–.0001</td>
<td>.0001</td>
<td>1.81</td>
<td>.074</td>
<td>.000</td>
<td>.0001</td>
<td>.01</td>
<td>.994</td>
</tr>
<tr>
<td>Semantic ambiguity</td>
<td>–.0004</td>
<td>.0136</td>
<td>–.03</td>
<td>.97</td>
<td>.0078</td>
<td>.0192</td>
<td>.41</td>
<td>.694</td>
</tr>
<tr>
<td>Noun-to-verb ratio</td>
<td>.0059</td>
<td>.0048</td>
<td>1.24</td>
<td>.217</td>
<td>–.0017</td>
<td>.0077</td>
<td>–.22</td>
<td>.834</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Intercept)</td>
<td>.3973</td>
<td>.0872</td>
<td>4.56</td>
<td>.000</td>
<td>.2698</td>
<td>.0573</td>
<td>4.71</td>
<td>.004</td>
</tr>
<tr>
<td>Trial number</td>
<td>.000</td>
<td>.000</td>
<td>.65</td>
<td>.513</td>
<td>–.0001</td>
<td>.000</td>
<td>–2.6</td>
<td>.011</td>
</tr>
<tr>
<td>Number of strokes</td>
<td>.0004</td>
<td>.0014</td>
<td>3</td>
<td>.757</td>
<td>–.0003</td>
<td>.0018</td>
<td>–.18</td>
<td>.87</td>
</tr>
<tr>
<td>Frequency</td>
<td>–.0088</td>
<td>.0095</td>
<td>–.92</td>
<td>.366</td>
<td>.0024</td>
<td>.0092</td>
<td>.26</td>
<td>.786</td>
</tr>
<tr>
<td>Phonetic combinability</td>
<td>–.0013</td>
<td>.0013</td>
<td>–.93</td>
<td>.367</td>
<td>.000</td>
<td>.0018</td>
<td>–.02</td>
<td>.99</td>
</tr>
<tr>
<td>Semantic combinability</td>
<td>–.0001</td>
<td>.0001</td>
<td>1.69</td>
<td>.097</td>
<td>–.0001</td>
<td>.0001</td>
<td>–.95</td>
<td>.339</td>
</tr>
<tr>
<td>Semantic ambiguity</td>
<td>–.0053</td>
<td>.0149</td>
<td>–.35</td>
<td>.705</td>
<td>–.0098</td>
<td>.0168</td>
<td>–.59</td>
<td>.544</td>
</tr>
<tr>
<td>Noun-to-verb ratio</td>
<td>.0029</td>
<td>.0052</td>
<td>.55</td>
<td>.587</td>
<td>.0013</td>
<td>.0068</td>
<td>.19</td>
<td>.865</td>
</tr>
</tbody>
</table>

3.4. Lexical effects in VWFA over later time windows

The LH VWFA analysis revealed an effect of semantic ambiguity throughout the 514–666 ms ranges ($\sum r = 7.523$ for 153 time-points, $p < .01$). The RH VWFA analysis revealed the effects of character frequency throughout the 604–800 ms ranges ($\sum r = 13.014$ for 197 time-points, $p < .001$), and the effects of semantic ambiguity throughout the 646–798 ms ranges ($\sum r = 7.471$ for 153 time-points, $p < .01$). None of the other variables had a significant effect on LH or RH VWFA activity.

The same linear mixed-effects model analyses were conducted by fitting the mean amplitude within each significant time cluster of the two ROIs (LH VWFA: 514–666 ms; RH VWFA: 646–800 ms) with participants and items as crossed random effects. The activity of LH VWFA regions was negatively correlated with semantic ambiguity ($b = –.06$, $SE = .017$, $t = –3.53$, $p < .001$). The activity of the RH VWFA regions was negatively correlated with character frequency ($b = –.05$, $SE = .015$, $t = –3.42$, $p < .001$) and showed no significant effect of semantic ambiguity ($b = .005$, $SE = .02$, $t = .24$, $p = .79$).

4. Discussion

Both Chinese and English speakers reading Chinese characters showed increased M100 activity in response to characters with higher numbers of strokes, reflecting a general impact of visual complexity on processing in the V1 region. That is, M100 activity is enhanced by complex visual stimuli regardless of language experience. Interestingly, trial number also triggered an increase in Chinese speakers’ M100 activity, but this was not the case with the English speakers’. Since ERP studies have demonstrated an enhancement activation of the V1 region around 80–100 ms that indexes the expectancy of visual stimuli (Luck et al., 1994), the effect of the trial number in Chinese speakers’ M100 activity may reflect an expectation of encountering the stimuli. On the other hand, English speakers responded randomly to the stimuli; therefore trial number did not have any effect on their M100 activities, reflecting no build up of expectation about the nature of the presented stimuli.

Chinese speakers’ RH M100 responses also showed a larger activity in reading characters with larger phonetic combinability, which seems consistent with Solomyak and Marantz’s (2009) finding that high bigram frequency words elicit more negative M100 activity. Although a bigram frequency effect in the opposite direction has also been demonstrated around the 80–100 ms latency (Hauk, Davis, et al., 2006; Hauk et al., 2006), these results suggest that V1 processes of feature integration are also sensitive to inter-feature probabilities of adjacent stroke patterns.

The current findings about the M170 response help elucidate how Chinese speakers’ left and right occipitotemporal regions contribute to the reading of characters (Bolger et al., 2005; Tan et al.,...
Cottrell, & Tarr, 2003). Hence, we need a reconciliation of the visual processes of orthographic forms (Dehaene et al., 2005; Vinckier form analysis, and the anterior area of the VWFA reflects higher le-
sassumption that the posterior area of the VWFA reflects visual
For example, Fig. 2 shows that the English speakers' LH and RH meaningless stimuli not encountered in their language experience.
M170 activity in this experiment reflects a response to LH M170 activation in reading characters with a high number of
visually homogeneous categories (Gauthier, Skudlarski, Gore, & temporal regions are different in the early stages of the reading of
speakers' LH and RH M170 activation respectively, which suggests that
the processes underlying activation in left and right occipito-
temporal regions are different in the early stages of Chinese character recognition. While the effect of semantic-level variables is absent in the M100 and M170 responses in this experiment, the activation of LH and RH fusiform gyrus regions shows the effects of semantic ambiguity and character frequency, respectively, in later time windows. These results are consistent with the assumption that the fusiform gyrus contributes to processes related to conceptual knowledge or semantic networks (Mechelli et al., 2005; Wheateley et al., 2005). Most important, since the form-meaning relationship in Chinese is usually context dependent, high-frequency characters usually have multiple meanings (Perfetti & Tan, 1998; Tan et al., 2000). The linear mixed-effects model analysis present here can separate the effects of character frequency (here observed in the RH) and semantic ambiguity (here observed in the LH). These results further suggest that while processing visually presented Chinese phonograms, the left fusiform gyrus region may contribute to accessing meanings related to the phonograms, while familiarity may account for the activation in the right fusiform gyrus region in the later time windows.

5. Conclusion

Overall, we have demonstrated that Chinese speakers' bilateral occipitotemporal regions are modulated by two different properties of orthographic form during the early stages of character recognition, while English speakers of Chinese characters only show an effect of visual complexity in the M100 responses and in the posterior M170 responses. These results suggest that Chinese speakers use different types of orthographic features in early stages of character recognition. One is the conditional probability between radicals that reflect the decomposition of characters. This conclusion could account for findings in other studies of ERPs and eye movements that radical-level information influences earlier stages of character recognition (Hsiao, Shillcock, & Lee, 2007; Hsu et al., 2009; Lee et al., 2007; Tsai, Lee, Tzeng, & Yen, 2004). Furthermore, it implies that properties of the abstract units of sub-word representations could contribute to the decomposition of words, as studies of English have showed effects of morphological properties on the M170 response (Solomyak & Marantz, 2010; Zweig & Pyllkkänen, 2009). Current models of Chinese character recognition also assume radicals in the input level, and they may account for the effect of radical combinability (Perfetti et al., 2005; Taft, 2006; Yang, McCandliss, Shu, & Zevin, 2008). On the other hand, it is unclear how current models could predict an effect
of the number of strokes in earlier stages of processing. In particu-
lar, Yang, McCandliss, Shu, and Zevin’s (2009) connectionist model
considers the variation in the number of strokes and uses a variety
of codes to represent the hierarchical structure of Chinese ortho-
graphy. Their model may provide a basis for better understanding
of the interaction of different orthographic units.

Acknowledgments
This material is based upon work supported by the National Sci-
ence Foundation under Grant No. 25-91551-F6539. Any opinions,
findings, and conclusions or recommendations expressed in this
material are those of the author(s) and do not necessarily reflect
the views of the National Science Foundation. This work was com-
pleted when the first author was a visiting scholar at New York
University under support from Taiwan’s National Science Council
(NSC-096-2917-I-010-106). This work was also supported by the
Academia Sinica, Taiwan (AS-99-TP-AC1). The authors would like
to thank two anonymous reviewers for their comments on this pa-
rer, and Gwyneth Lewis for help with testing participants and data
analyses. We are grateful to Colin Wilson for suggestions on using
the linear mixed-effects model, and to Jonathan Simon for discus-
sions on using the TSPCA algorithms.

Appendix A. Supplementary material
Supplementary data associated with this article can be found, in

References
Cambridge: Cambridge University Press.
crossed random effects for subjects and items. Journal of Memory and Language,
59(4), 380–412.
Visual word recognition of single-syllable words. Journal of Experimental Psycho-
logy: General, 133(2), 283–316.
approach to blind separation and blind deconvolution. Neural Computation, 7(6),
1129–1159.
ERP manifestations of processing printed words at different psycholinguistic
levels: Time course and scalp distribution. Journal of Cognitive Neuroscience,
revised: universal structures plus writing system variation. Human Brain
Mapping, 25(1), 92–104.
characterization and character decision. Chinese Journal of Psychology, 46,
179–195.
Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Hennaf, M.-
characterization of an initial stage of reading in normal subjects and posterior
(2000). Dynamic statistical parametric mapping combining fMRI and MEG for
predictability effects on event-related potentials during reading. Brain Research,
1084(1), 89–103.
Neuroscience Methods, 162(2), 297–305.
single-trial EEG dynamics including independent component analysis. Journal of
Neuroscience Methods, 134(1), 9–21.
Ding, G., Peng, D., & Taft, M. (2004). The nature of the mental representation of
radicals in Chinese: A primer study. Journal of Experimental Psychology. Learning,
Memory, and Cognition, 30(2), 530–539.


