Neuroscience Letters 503 (2011) 147–151

Contents lists available at SciVerse ScienceDirect

Neuroscience Letters

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

An event-related potential study on cross-modal conversion of Chinese characters

Qiuyan Wu [a,b], Yuan Deng [a,*]

[a] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 4A Datun Road, Chaoyang District, Beijing 100101, PR China
[b] Graduate University of Chinese Academy of Sciences, Beijing 100049, PR China

ARTICLE INFO

Article history:
Received 14 June 2011
Received in revised form 10 August 2011
Accepted 16 August 2011

Keywords:
Cross-modal conversion
Event-related potentials
Chinese character
N400
Late positive component

ABSTRACT

In the current study, we explored the effects of ERPs (event-related potentials), related to the cross-modal transfer from visual input to phonological retrieval. Using Chinese single-character words, participants were asked to make orthographic (intra-modal) and phonological (cross-modal) responses to visually presented words. By comparing the cross-modal and intra-modal tasks, we found that both tasks evoke similar activity in the early stage of lexical processing, showing the same pattern of N2 effect (a negative component peaking around 220 ms) and P2 effect (a positive component peaking around 270 ms). However, the effect of the task was significant in the 300–700 ms time window, consisting of a frontal-based N400 effect and a parietal based late positive component (LPC) effect. These findings suggest that the frontal-based N400 is associated with orthography-to-phonology mapping in Chinese, and the LPC reflects greater requirement of maintaining retrieved information in working memory for the cross-modal processing.

The human brain receives plenty of information through visual and auditory modalities. For instance, during language processing, the normal process for a human is reading words using the visual modality and listening to speech via the auditory modality. However, the information received from these two modalities is highly interactive. On a daily basis, humans are required to convert information from one modality to the other, such as reading out a printed word (from visual to auditory) and writing down a heard word (from auditory to visual). These processes are known as cross-modal lexical processing. In contrast, when a reader processes the word-form of a printed word and processes the acoustic information of a speech sound, these processes can be called as intra-modal lexical processing. Compared with intra-modal lexical processing, cross-modal processing may involve more stages or processes over time. Presumably, cross-modal processing starts with input-driven processing (orthographic processing on visual words and phonological processing on auditory words), then moves into a conversion process (orthography-to-phonology conversion or vice versa), and finally ends with access to information related to the target modality. For example, phonological retrieval of visual words involves the following steps: processing the visual words' graphic and orthographic features, mapping orthography onto phonology, and finally, retrieving phonological information.

Neuroimaging and electrophysiological studies have explored the neural basis and temporal characteristics of cross-modal lexical processing. Neuroimaging studies suggest that cross-modal lexical processing evokes activations not only in regions specific to input modality and target modality, but also those regions responsible for cross-modal conversion [4,5,7,32]. For example, Booth et al. [4,5] found that a visual rhyming task evoked activation in the fusiform gyrus for visual word form processing, superior temporal gyrus for rhyming judgment, and posterior hetermodal regions (supramarginal and angular gyri) for cross-modal conversion. To unfold the cross-modal lexical processes over time, the event-related potential technique (ERP) is used [15,18]. In the study of Kiyonaga et al. [18], a cross-modal repetition priming paradigm demonstrated that repetition effect elicited a significantly more negative-going ERP during 300–500 ms (i.e. N400) in a cross-modal condition (visual-to-auditory). The N400 effect may reflect activation of orthographic or phonological representations at whole-word level, and integration with higher-level semantic representation [20].

The aforementioned studies all used alphabetic languages. One of the linguistic characteristics of alphabetic languages is that the word forms are transcripts of speech sounds. The existence of grapheme-to-phoneme correspondence rules may be a major confounding variable in the previous studies [4,5]. In contrast, Chinese characters map onto phonology at the mono-syllable level [31]. For example, a character "献" (contribute) pronounced as/xian4/(the number refers to tone), thus no visual component of this character

* Corresponding author. Tel.: +86 10 6485 2598; fax: +86 10 6487 2070.
E-mail address: dengy@psych.ac.cn (Y. Deng).

0304-3940/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.neulet.2011.08.029

(i.e. left radical "南"/nan2/or right radical "犬"/quan3/) corresponds to a phoneme of the character's pronunciation (/xian4/). Therefore, one key advantage of using Chinese is to exclude the confounding aspect of grapheme-to-phoneme correspondence in the orthography-to-phonology conversion.

However, to our knowledge, no study has directly examined the cross-modal conversion (orthography-to-phonology) in the Chinese language by ERP technique. Some studies have revealed specific effects for orthographic and phonological processing in the Chinese language [3,26,27,32]. For example, by combining the priming paradigm with ERP recordings, Liu et al. [27] reported graphic interferences in the pronunciation task (reduced P200) and in the meaning task (reduced N400), as well as a phonological interference in the meaning task (smaller N400). These findings suggest that visual word processing in Chinese may be processed in the following order: graphic activation, phonological activation and finally semantic activation. Additionally, neuroimaging studies demonstrate that phonological processing of written Chinese characters consistently activates the left middle frontal gyrus (BA9) [25,36,37,39–42]. This region is assumed to be responsible for visuospatial analysis of Chinese characters and orthography-to-phonology conversion in Chinese [25]. These findings indicate that orthography-to-phonology conversion might involve different neural mechanisms across different writing systems. Further investigation into the temporal characteristics of the conversion process from orthography to phonology in Chinese might reveal a possible difference from that in alphabetic languages.

Therefore, in the current study, we aimed to examine temporal characteristics of cross-modal conversion from orthography to phonology in Chinese. Event-related potentials were recorded in both an intra-modal task (making structure judgment on visual characters) and a cross-modal task (making tonal judgment on visual characters). By strictly controlling the perceptual properties of presented stimuli, these two tasks should have great overlap in early processing stages, including perceptual processing and orthographic analysis. However, they might differ in later processes because the cross-modal task requires an additional process of mapping orthographic representation to stored phonological representation; while in the intra-modal task there may be continued analysis of orthographic information during the later stages on information processing. Therefore, by directly comparing ERPs in the cross-modal task to those in the intra-modal task, differences in ERPs can be associated with orthography-to-phonology conversion. Furthermore, we hypothesized that cross-modal processing of Chinese words may present different ERP components and/or different topographic distributions of these ERP components over time, compared to those found in alphabetic languages.

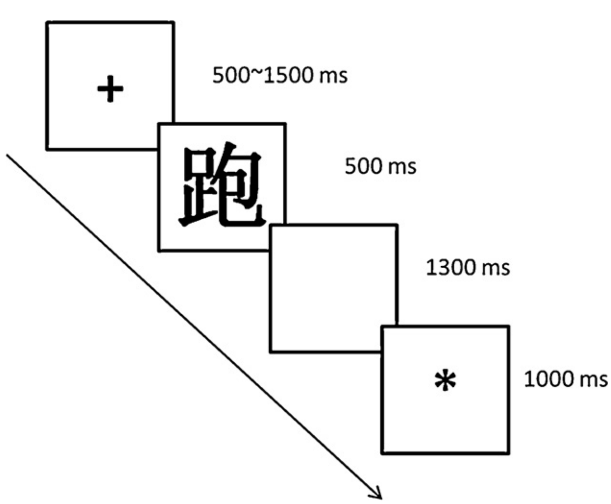

Fig. 1. A timeline representing the experimental procedure for a single trial in the cross-modal and intra-modal tasks.

Twenty native Chinese undergraduates (19–23 years old; 11 females) were paid to participate in this experiment. All participants were all right-handed, free of neurological diseases and psychiatric disorders, and had normal hearing and normal or corrected-to-normal vision. Written informed consent was obtained from all participants prior to the experiment.

Eighty Chinese single-character words were selected from a pool of the 4574 characters most commonly used according to Xiandai Hanyu Pinlu Cidian [Modern Chinese Frequency Dictionary] [44]. These characters were of high frequency (above one hundred times per million words). The average stroke number of these characters was 9.55 (SD = 2.40). The number of strokes reflects the visual complexity of a Chinese character [38]. Because stroke numbers range from 1 to 25 in the Modern Chinese Frequency Dictionary [44] and a stroke number of 8 is the most frequent, so the 80 characters used in the current study had a medium visual complexity. There is only one phonological correspondence for each visual character, i.e. these characters were not polyphones. All characters were used in both cross-modal and intra-modal tasks to rule out the potential confounding of low-level features. Therefore, any difference between these two tasks can only be due to dissimilar cognitive processes elicited by the two tasks. All characters were designed in black against white background in Song font (100 × 100 pixels).

During the collection of ERP data, participants were seated approximately 80 cm away from the computer screen in an electrically and acoustically shielded room. Stimuli were visually presented on a 17-in. monitor.

Participants were instructed to make responses to each visual word based on task instructions. During the intra-modal task, subjects were asked to judge whether presented words had a left-right structure, i.e. whether two major visual components of a character are horizontally configured. For example, a character "理" has a left-right structure because its visual components ("王" and "里") are horizontally configured. In contrast, a character "吞" has an up-right structure rather than left-right one because its visual components ("天" and "口") are configured vertically. In the cross-modal task, the subjects were required to judge whether the pronunciation of presented words has a rising tone (second tone in Chinese).

Before starting the experiment, participants were given 10 practice trials for each task with feedback, which were not included in the analyses. The cross-modal and intra-modal tasks were separated into two independent sessions and stimuli in each session were randomized. The order of these two sessions was counterbalanced across participants and task instructions were presented prior to the start of the testing session. The experimental procedure was identical for the two tasks and a short break was given between two tasks. The experimental procedure was computerized using E-Prime software [35].

To avoid artifacts caused by eliciting decision-related components that may obscure task-related ERP components [23] or those caused by muscle activity, a delayed response paradigm was used (Fig. 1). In each trial, a fixation cross was firstly presented at the center of the screen for a random duration of 500–1500 ms to avoid expectancy. Then, a Chinese character was visually presented for 500 ms followed by a 1300-ms blank. Finally, a red star was shown to prompt participants to respond as quickly and accurately as possible by pressing two keys labeled with "Yes" or "No" with their right or left index fingers, respectively. In the intra-modal task, subjects were told that if the presented character had a left-right structure, they should respond "Yes"; otherwise they should respond "No". During the cross-modal task, subjects were told that