Brain & Language 157–158 (2016) 51–62

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Effects of orthographic consistency and homophone density on Chinese spoken word recognition

Wei-Fan Chen [a], Pei-Chun Chao [b], Ya-Ning Chang [a,c], Chun-Hsien Hsu [a], Chia-Ying Lee [a,b,d,e,*]

[a] Institute of Linguistics, Academia Sinica, 128 Academia Road, Section 2, 11529 Taipei, Taiwan
[b] Institute of Neuroscience, National Yang-Ming University, 155 Linong Street, Section 2, 11221 Taipei, Taiwan
[c] Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK
[d] Research Center for Mind, Brain and Learning, National Chengchi University, No. 64, Sec. 2, ZhiNan Rd., Wenshan District, 11605 Taipei, Taiwan
[e] Institute of Cognitive Neuroscience, National Central University, No. 300, Jhongda Rd., Jhongli 32001, Taoyuan, Taiwan

A R T I C L E I N F O

Article history:
Received 7 September 2015
Revised 17 April 2016
Accepted 18 April 2016
Available online 9 May 2016

Keywords:
Homophone density
Orthographic consistency
Spoken word recognition
ERPs
sLORETA

A B S T R A C T

Studies of alphabetic language have shown that orthographic knowledge influences phonological processing during spoken word recognition. This study utilized the Event-Related Potentials (ERPs) to differentiate two types of phonology-to-orthography (P-to-O) mapping consistencies in Chinese, namely homophone density and orthographic consistency. The ERP data revealed an orthographic consistency effect in the frontal-centrally distributed N400, and a homophone density effect in central-posteriorly distributed late positive component (LPC). Further source analyses using the standardized low-resolution electromagnetic tomography (sLORETA) demonstrated that the orthographic effect was not only localized in the frontal and temporal-parietal regions for phonological processing, but also in the posterior visual cortex for orthographic processing, while the homophone density effect was found in middle temporal gyrus for lexical-semantic selection, and in the temporal-occipital junction for orthographic processing. These results suggest that orthographic information not only shapes the nature of phonological representations, but may also be activated during on-line spoken word recognition.

1. Introduction

The abilities to speak and read are two important evolutionary endowments of human beings. Compared to reading acquisition, speech has primacy both in the history of humankind and in an individual's lifetime, and can be acquired without explicit instruction. Acquiring spoken language primarily involves mastering the linkage between phonology and semantics, while learning to read aims to develop efficient mapping of visual symbols (orthography) onto phonological and semantic representations. However, phonology plays a critical role in visual word recognition. For example, heterographic homophones are generally harder to recognize than non-homophonic words (Ferrand & Grainger, 2003; Pexman, Lupker, & Jared, 2001; Xu, Pollatsek, & Potter, 1999) and are more prone to semantic confusion (Tan & Perfetti, 1997; Van Orden, 1987). These homophone disadvantages support the competition driven by mandatory phonological processing during visual word recognition.

Since phonology mediates the establishment of mapping orthography and semantics during literacy acquisition, the automatic phonological activation during visual word recognition is expected. However, it seems absurd to ask whether orthographic word forms would be activated while listening to a spoken word, since there is no obvious benefit in activating orthography at that moment. Nevertheless, an increasing number of studies show that knowledge of orthography influences spoken word recognition. For example, Seidenberg and Tanenhaus (1979) showed that when determining whether two spoken words share the same rhyme, the rhymes with identical spelling (e.g., PIE-TIE) were easier to match than the rhymes with different spelling (e.g., RYE-TIE). Similarly, priming effects across two auditory words were only found to be robust when the phonological overlap also involved an orthographic overlap (Slowiaczek, Soltano, Wieting, & Bishop, 2003). However, the strongest evidence for orthography influencing spoken word recognition comes from studies involving the manipulation of the ***orthographic consistency*** of spoken words, which is defined as the degree of mapping consistency from phonology to orthography (P to O). For example, in English, orthographic consistency can be measured by whether a spoken word contains a rhyme that can be spelled in multiple ways (i.e., /ʌk/ is consistent in that it is always spelled -UCK, while /ip/ is inconsistent since

* Corresponding author at: Institute of Linguistics, Academia Sinica, 128 Academia Road, Section 2, 11529 Taipei, Taiwan.
E-mail address: chiaying@gate.sinica.edu.tw (C.-Y. Lee).

http://dx.doi.org/10.1016/j.bandl.2016.04.005
0093-934X/

it could be spelled either -EEP or -EAP). The ***orthographic consistency effect*** during spoken word recognition has been found in different languages, including Portuguese (Ventura, Kolinsky, Pattamadilok, & Morais, 2008; Ventura, Morais, Pattamadilok, & Kolinsky, 2004), French (Pattamadilok, Morais, Ventura, & Kolinsky, 2007; Ziegler, Petrova, & Ferrand, 2008), Thai (Pattamadilok, Kolinsky, Luksaneeyanawin, & Morais, 2008), and English (Chereau, Gaskell, & Dumay, 2007; Miller & Swick, 2003; Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008). Collectively, these studies report that auditory lexical decisions toward inconsistent orthographic words (whose rhyme could be spelled in multiple ways) were slower than decisions toward orthographically consistent words (whose rhyme could be spelled only one way), and therefore support the concept that orthographic information influences on-line spoken word recognition. However, there are ongoing debates regarding *when* the orthographic consistency occurs within the time course of spoken word recognition, and *how* the orthographic information affects spoken word recognition.

Recent studies have utilized event-related potential (ERP) techniques, which measure the brain's responses on a millisecond scale and provide a set of ERP components that index various stages of the cognitive process, to track the time course of the orthographic consistency effect in a variety of tasks (lexical decision, semantic categorization, rhyming judgment, etc.) (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Pattamadilok, Perre, & Ziegler, 2011; Perre, Bertrand, & Ziegler, 2011; Perre, Pattamadilok, Montant, & Ziegler, 2009; Perre & Ziegler, 2008). The most important issue for these studies was to determine if the orthographic effect would be found before, or at, the N400, an ERP component associated with on-line lexical retrieval, or if the effect would be found at the late positive component (LPC), which is sensitive to deliberate memory retrieval and decision accuracy in the post-lexical stage (Allan & Rugg, 1997; Paller & Kutas, 1992). For example, Perre and Ziegler (2008) manipulated orthographic consistency in early (the first syllable) or late (the second syllable) positions of the spoken words in an auditory lexical decision task. They found that inconsistent words elicited increased negativity than the consistent words did in the 300–350 ms and 400–450 ms bins, at central-posterior sites. Critically, the ERP differences between consistent and inconsistent words were found long before the end of the word and were time-locked to the "arrival" of the orthographic inconsistency in the spoken word. Pattamadilok et al. (2009) also demonstrated the same pattern of orthographic consistency effects that were time locked to the position of orthographic inconsistency (early consistency effect in 300–350 ms, and late consistency effect in 400–425 ms and 450–700 ms) by using the semantic go/no go task. These effects occurred before the onset of frequency and the go/no-go effects, which have been used to index the moment of lexical access and decision-making, respectively. Even so, the orthographic consistency effects were found in frontal to central sites, rather than central to posterior sites. Pattamadilok et al. (2011) examined the orthographic consistency effect and frequency effect with a rhyming judgment task. However, they could not replicate the orthographic consistency effect in the same time window of 300–350 ms, but rather found this effect in the 175–250 ms time window, along with an effect in the 375–750 ms time window in the frontal to central sites. Moreover, the consistent words elicited increased negativities in the late and long lasting later time window than did the inconsistent words, and the pattern was opposite to that demonstrated by previous studies (Pattamadilok et al., 2009; Perre & Ziegler, 2008; Perre et al., 2009). Across these studies, although there are still some controversies regarding the pattern and precise time windows of the orthographic consistency effect, in general these findings support that orthographic information is computed on-line, rather than occurs post-lexically and/or at decisional stage during spoken word recognition.

Another ongoing debate concerns how the orthographic consistency effect emerges during literacy. Two explanations have been proposed to explain the orthographic consistency effect in spoken word recognition; one is the phonological restructuring account, and the other is the on-line activation account (Perre et al., 2009). The phonological restructuring account claims that learning to read alters preexisting phonological representations, and that orthographic consistency plays a major role during the restructuring processes (Muneaux & Ziegler, 2004; Ziegler & Goswami, 2005). To be more specific, the orthographically consistent words develop finer phonological representations than do the orthographically inconsistent words. Therefore, it predicts that the orthographic consistency effect, that reflects the differences in the quality of phonological representation, will be found in brain regions that are responsible for phonological processing, such as the left inferior frontal gyrus (IFG), insula, left superior temporal gyrus (STG), or left supramarginal gyrus (SMG). Alternatively, the on-line activation account assumes that the orthographic information will be activated on-line while processing the spoken words, due to the strong and permanent associations between orthography and phonology that develop through the processes of learning to read. Therefore, the orthographic consistency effect should be found in brain regions responsible for visual-orthographic processing, such as the left ventral occipitotemporal cortex (vOTC) and visual word form area (McCandliss, Cohen, & Dehaene, 2003), in addition to the phonological regions.

Perre et al. (2009) applied the standardized low resolution electromagnetic tomography (sLORETA) to determine the cortical generators underlying the orthographic consistency effect in spoken word recognition. They found that the orthographic consistency effect was localized in a classic phonological area, left BA40, but not in the posterior cortical areas for coding orthographic information. Pattamadilok, Knierim, Kawabata Duncan, and Devlin (2010) applied repetitive TMS over the left SMG and vOTC while participants performed an auditory lexical decision task in which the orthographic consistency of the spoken words was manipulated. The orthographic consistency effect disappeared only when the stimulation was delivered to the SMG, but not to the vOTC. This evidence supports the phonological restructuring hypothesis, rather than the co-activation of orthographic codes. However, other fMRI studies have found activation in the fusiform gyrus during rhyming judgments for auditory words (Booth, Cho, Burman, & Bitan, 2007; Booth et al., 2002; Cao et al., 2010; Cao et al., 2011). Yoncheva, Zevin, Maurer, and McCandliss (2010) found that selectively attending to speech, relative to selectively attending to melody, leads to increased activity in left inferior frontal regions, specifically the left mid-fusiform gyrus near the visual word form area (VWFA) and temporal areas. These findings serve as strong support for the co-activation of orthographic information during auditory lexical processing.

Chinese orthography is often described as a logographic writing system. The basic written unit of Chinese is the character, which consists of strokes or radicals that fit into a square-shaped space. Chinese characters represent monosyllabic (and usually monomorphemic) forms, with the majority consisting of a consonant-vowel (CV) structure. Given the relatively simple syllable structure, most Chinese syllables may represent more than one morpheme, and so are mapped onto more than one orthographic form (characters). The pervasive homophony of Chinese implies that the orthographic form is particularly important for selecting the meaning and escaping homophony in Chinese. Thus, we may expect a greater impact from orthography during spoken word recognition in Chinese than in alphabetic writing systems.