Sublexical Processing in Reading Chinese

Xiaolin Zhou and **William Marslen-Wilson**
Department of Psychology
Birkbeck College, University of London
Malet Street, London WC1E 7HX
x.zhou@psychology.bbk.ac.uk

Abstract

The nature of sublexical processing in reading logographic Chinese was investigated in three primed naming experiments. Experiment 1 showed that phonetic radicals in low frequency complex characters are automatically decomposed and used to activate their own phonological representations. Experiments 2 and 3 demonstrated that the semantic properties of these phonetic radicals are also activated. It is argued that sublexical processing in reading Chinese is both a phonological and a semantic event and there is no fundamental difference between sublexical processing of phonetic radicals and lexical processing of simple and complex characters. The implications of these results for theories of lexical processing are discussed.

Introduction

Theories of visual word recognition differ considerably in their assumptions about the nature of sublexical processing. In traditional dual-route models of reading (e.g., Coltheart, Curtis, Atkins, & Haller, 1993), sublexical orthographic units like "_OWN" in real words (e.g., TOWN) or nonwords (e.g., POWN) are used mainly as inputs to grapheme-phoneme conversion rules whose end products are phonemes. In most connectionist models of word reading (e.g., Seidenberg & McClelland, 1989), the sublexical orthographic units are primarily the input to distributed networks which compute a phonological output. Although these units are also the input to orthography-to-semantics networks which compute the semantic output for the whole letter strings, the sublexical orthographic units do not produce detectable semantic output of their own even when these units happen to be words themselves (e.g., OWN in TOWN or BOY in BOYCOTT). In models of visual word recognition that stress the predominant role of phonology in initial lexical access and in access to meaning (e.g., Frost, in press; Lukatala & Turvey, 1994; van Orden, Pennington, & Stone, 1990), sublexical processing of orthographic units is strictly a phonological process, which activates the phonological representation of the whole word before the meaning of the word accessed. Therefore, despite their considerable differences concerning whether the sublexical processing in reading alphabetic words is rule-based and how the computation from orthography to phonology is conducted, most localist and connectionist theories of visual word recognition share the common assumption that the sublexical processing in reading is primarily or strictly a phonological event and it differs rather drastically from lexical processing, which activates both phonological and semantic representations in the lexicon.

In this paper, we argue that this assumption about the nature of sublexical processing cannot be applied without qualification to sublexical processes in reading logographic Chinese. Sublexical processing in reading Chinese, which is restricted here to the processing of phonetic radicals of complex or compound characters, is intrinsically related to orthography, phonology, and semantics. The processing of phonetic radicals, which may provide phonological information in terms of the pronunciation of the complete characters, not only contributes to the mapping from orthography to phonology in lexical access, but also generates activation and competition in the semantic system.

Structural Properties of Chinese Characters and Sublexical Phonological Processing

The Chinese writing system is often described as logographic, where the basic orthographic units, the characters, correspond directly to morphemes and syllables. Chinese characters can be broadly differentiated into two categories: simple and complex, both of which are composed of strokes and arranged in squares of similar size. Simple characters (e.g., 义, yi (4)[1], *righteousness*) constitute about 3% of the total characters in Morden Chinese, and they are holistic visual patterns that do not represent phonological information at the sublexical level. Complex or compound characters, however, do encode phonological information at this level. These characters (e.g., 议 yi(4), *discuss*) can be divided meaningfully into components and most of them are composed of a semantic radical on the left and a phonetic radical on the right, although some arrange their radicals in different ways. Semantic radicals have the function of indicating the semantic category of morphemes corresponding to the complex characters, while phonetic radicals have the function of pointing to the pronunciations of whole characters. However, both functions are not reliable, with exceptions and irregularities littered over the writing system.

[1]Through the paper, the pronunciations of Chinese words are given in Pinyin - the Chinese alphabetic system. Numbers in brackets represent the tones of syllables.

The phonological relations between phonetic radicals and the complex characters they are embedded in can roughly be differentiated into four types (Table 1). Less than 30% of complex characters have the same pronunciations as their phonetic radicals, while about one third of complex characters differ from their phonetic radicals, at least at the segmental level. About another one third of complex characters have either the same segmental templates or the same rhyming parts as their phonetic radicals. However, it should be noted that the same phonetic radicals could be involved in some or all of the four categories, and, more importantly, the phonetic radicals themselves are independent characters, having their own meanings.

Type	Phonetic Radical	Complex Character
Regular	羊 yang(2) *(sheep)*	洋 yang(2) *(ocean)*
Semi-regular	青 qing(1) *(blue)*	情 qing(2) *(affection)*
Rhyming	亡 wang(2) *(die)*	忙 mang(2) *(busy)*
Irregular	白 bai(2) *(white)*	怕 pa(4) *(fear)*

Table 1 The phonological relations between phonetic radicals and complex characters

There is evidence that the Chinese reader does use the phonological information conveyed by phonetic radicals, especially when reading low frequency complex characters. Seidenberg (1985) found that regular complex characters, or characters having the same pronunciations as their phonetic radicals, are named faster than frequency-matched simple characters, but only for low frequency ones. Fang, Horng, and Tzeng (1986) investigated both regularity and consistency effects in naming characters, where consistency was defined according to whether *all* complex characters with a particular phonetic radical are pronounced in the same way as the radical. The authors observed a consistency effect but not a regularity effect. Regular-consistent characters were named faster than regular-inconsistent characters. Subsequent studies in general have found both consistency and regularity effects (Hue, 1992; Peng, Yang, & Chen, 1994).

The typical account of these regularity and consistency effects in reading complex characters was in terms of Glushko's (1979) *activation-synthesis* model and Seidenberg's (1985) time course model of orthographic and phonological activation. According to these models, in initial lexical access, salient orthographic units are extracted from visual input and used to activate in parallel their corresponding orthographic and phonological representations in the lexicon. The phonological activation is synthesized according to the interactive principles to manifest the regularity or consistency effects in naming. The interaction between these effects and frequency is explained by the relative time course of orthographic and phonological activation, where the orthographic processing of high frequency words is sufficiently fast to allow word recognition before phonological activation, but the orthographic processing of low frequency words is not efficient and phonological activation has time to accrue to influence word recognition. To put in a slightly different way, the phonological output for high frequency words is based mainly on their own phonological representations, which are activated efficiently through direct visual mapping and/or via semantics, while the phonological output for low frequency words is influenced by phonological representations of other orthographically similar words activated in parallel from orthography.

The issues for the present research are whether we can obtain more direct evidence of phonological activation of phonetic radicals and, more importantly, whether the sublexical processing of phonetic radicals in reading complex characters is also a semantic event, involving access to semantic representations corresponding to these radicals.

Phonological and Semantic Processing of Phonetic Radicals

We used a primed naming paradigm to investigate the phonological and semantic processing of phonetic radicals in reading complex characters. In this paradigm, prime characters were presented briefly, overwritten immediately by targets. Subjects were required to read aloud the targets and their naming latencies and response errors were recorded. The phonological or semantic relations between primes and targets were systematically manipulated. Complex characters were used as primes in Experiments 1 and 2 and as targets in Experiment 3. The phonetic radicals of complex characters, rather than complex characters themselves, were phonologically (Experiment 1) or semantically (Experiments 2 and 3) related to the targets or primes. In order to isolate any priming effects to the processing of phonetic radicals, we use only irregular complex characters (with left-right structure), since regular and semi-regular complex characters share phonological forms with their phonetic radicals and any priming effects in the present paradigm could be due, directly or indirectly, to the processing of the whole complex characters.

Experiment 1: Phonological processing of phonetic radicals

Experiment 1 investigated the sublexical phonological processing of phonetic radicals in reading (irregular) complex characters. If reading Chinese involves decomposing phonetic radicals from complex characters and using them to access their own phonological representations in the lexicon, we should observe a phonological priming effect between phonetic radicals embedded in complex characters and characters homophonic to the radicals. Moreover, if the automatic decomposition of complex characters is modulated by the frequency of complex characters, we should find the homophone priming effect for low frequency irregular complex characters, but not for high frequency ones.

	Homophone	Complex	Control	Target
Low	卒	粹	恒	族
	zu(2)	cui(4)	heng(2)	zu(2)
	(soldier)	*(pure)*	*(permanent)*	*(clan)*
High	欠	次	比	倩
	qian(4)	ci(4)	bi(3)	qian(4)
	(owe)	*(sequence)*	*(compare)*	*(pretty)*

Table 2 Experiment 1: Design and sample stimuli

As indicated in Table 2, target characters were preceded by complex characters whose phonetic radicals were homophonic to the targets. They were also preceded by the radicals on their own, which served as homophone primes. Two types of complex characters were differentiated according to their frequencies. There were 51 low frequency and 33 high frequency complex characters. Control primes were chosen to match with complex primes on frequency, visual complexity (in terms of the number of strokes), and structure (all left-right composition of semantic and phonetic radicals). There were no orthographic similarities or semantic relations between targets and primes. The SOA (stimulus onset asynchrony) between primes and targets was set at 100 msec. One hundred pairs of unrelated words were used as fillers.

Mean naming latencies are reported in Figure 1. Response errors were too infrequent to allow statistical analysis. The analyses of reaction times showed a significant main effect of prime type ($p<.01$). Post hoc Newman-Keuls tests showed that the mean reaction times in homophone and complex priming conditions were significantly faster than in the control condition ($p<.01$). Moreover, the mean reaction time in the homophone condition was also faster than the mean reaction time in the complex condition ($p<.01$). Although we did not find an significant interaction between prime type and frequency in the main tests, separate analyses for low and high frequency complex primes showed that while the mean naming latency for low frequency complex primes was significantly faster than the control baseline ($p<.05$), the mean reaction time for high frequency complex primes did not differ from the control primes.

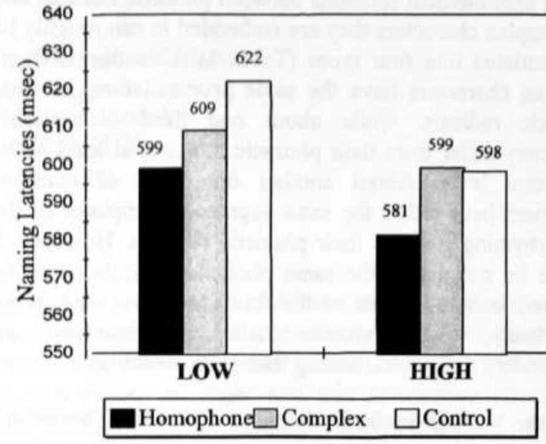

Figure 1 Experiment 1: Mean naming latencies

Three observations can be made from the above analyses. First, the significant homophone priming effect between radical primes and targets suggests that the presentation of homophone primes activates phonological representations they share with the targets. Although there is some doubt about the automaticity of phonological activation in reading Chinese, most evidence suggests this activation is automatic (see Zhou, in press).

Second, since complex characters were neither phonologically nor orthographically nor semantically related to targets, the facilitatory effect between low frequency complex primes and target words could only stem from the fact that phonetic radicals embedded in the complex characters were homophonic to the targets. This effect indicates that in processing complex characters, phonetic radicals are decomposed and used to access the phonological representations of their own. However, the ability to decompose phonetic radicals and gain access to their phonology is modulated by the frequency of complex characters, as indicated by the different patterns of priming effects for high and low frequency complex primes. Either reading high frequency complex characters involves little decomposition, or the transient phonological activation of phonetic radicals embedded in complex characters was suppressed by the phonological activation of complex characters, which are activated faster than low frequency complex characters and hence have more chance to suppress rivalry phonological activation.

Third, the fact that phonological priming by phonetic radicals in complex characters was smaller than priming by phonetic radicals presented alone suggested that in reading complex characters, there was a competition between phonological representations corresponding complex characters and phonological representations corresponding to phonetic radicals. The phonological activation of radicals embedded in complex characters did not reach the same level as when these radicals were standing alone.

These findings replicated the regularity and consistency effects found in single character naming (Hue, 1992; Seidenberg, 1985). They are consistent with the view that sublexical processing of phonetic radicals in reading complex characters is at least in part a phonological event. We believe that access to phonological representations of phonetic radicals is achieved in the same way as access to phonological representations of the whole complex characters, i.e., mapping from orthographic input to phonological representations.

Experiment 2: Semantic processing of phonetic radicals
This experiment examined whether sublexical processing in reading complex characters automatically activates the semantic properties of the phonetic radicals, even though this semantic activation does not contribute to the processing of complex characters themselves. We used the semantic associates of phonetic radicals as targets and the irregular complex characters as primes. If the phonetic radicals were automatically decomposed and used to access their own semantic representations in reading complex characters, the activation of these semantic representations should facilitate the processing of their semantic associates.

Semantic	Complex	Control	Target
青	猜	波	紫
qing(1)	cai(1)	bo(1)	zi(3)
(blue)	*(guess)*	*(wave)*	*(purple)*

Table 3 Experiment 2: Design and sample stimuli

Fifty-one complex characters (e.g., 猜 cai(1), *guess*), whose phonetic radicals (e.g., 青 qing(1), *blue*) have clear semantic associates (e.g., 紫 zi(3), *purple*), were selected. The phonetic radicals and the complex characters had different pronunciations and were not semantically related. Because of restrictions on the selection of stimuli, only low and medium frequency complex characters were used. There were no orthographic and phonological relations between primes and targets. Control primes (e.g., 波 bo(1), *wave*) were chosen to match with complex primes on frequency, visual complexity and structure. Sixty pairs of unrelated characters were used as fillers.

Two SOAs between primes and targets were used, one 57 msec and one 100 msec. Mean naming latencies are reported in Figure 2. Few naming errors were made and are not reported here. Statistical analyses revealed a significant main effect of prime type ($p<.001$) and no interaction between prime type and SOA ($p>.1$). Post hoc tests showed a significant effect of semantic priming, both against control primes and against complex primes ($p<.01$). More importantly, the 10-msec advantage of complex primes, as against control primes, was also significant ($p<.01$).

The strong effects of semantic primes suggested that the semantic activation, due to direct access from orthography,

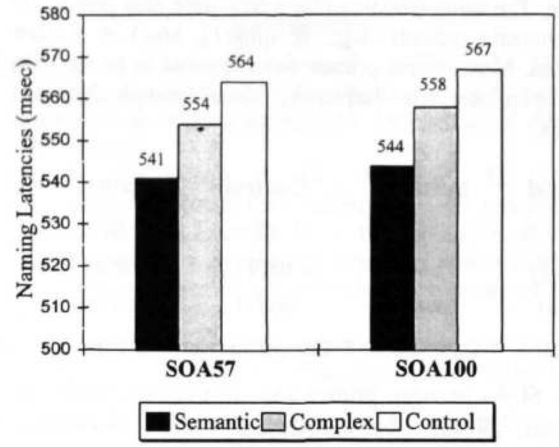

Figure 2 Experiment 2: Mean naming latencies

can efficiently flow to phonology, resulting in the facilitatory effects in naming (Zhou, in press; Zhou & Marslen-Wilson, 1996; Zhou et al., 1996). More importantly, the significant priming effects between complex characters and the semantic associates of their phonetic radicals demonstrated that there is indeed an automatic decomposition process in reading complex characters. The phonetic radicals of complex characters are used to access not only their corresponding form representations in the lexicon, but also their semantic properties, which are irrelevant to the complex characters. The smaller priming effect between complex primes and targets, compared with direct semantic priming, suggests that there is strong competition in the semantic system between representations corresponding to the whole complex characters and representations corresponding to the phonetic radicals. This competition influenced the semantic and hence phonological activation levels of target words.

Experiment 3: Semantic processing of phonetic radicals revisited
The purpose of this experiment was to collect converging evidence for the automatic decomposition of phonetic radicals and the activation of their semantic properties. If reading complex characters involves decomposing phonetic radicals and using them to access their own semantic representations, the presence of semantic primes, related to phonetic radicals but not to complex characters, should facilitate this decompositional process and increase the activation level of semantic representations of phonetic radicals. That will in turn increase the competition between semantic representations corresponding to complex characters and to their phonetic radicals.

Forty-five complex characters (e.g. 猜 cai(1), *guess*) that were used as primes in Experiment 2 were taken as targets in the present experiment, preceded by semantic associates (e.g., 紫 zi(3), *purple*) that were related to their phonetic radicals. The same complex characters were also primed by their phonetic radicals (e.g., 青 qing(1), *blue*) in another condition. New control primes were selected to match with semantic primes on frequency, visual complexity, and structure.

Radical	Semantic	Control	Target
青	紫	裂	猜
qing(1)	zi(3)	lie(4)	cai(1)
(blue)	*(purple)*	*(split)*	*(guess)*

Table 4 Experiment 3: Design and sample stimuli

Two SOAs between primes and targets were used: 100 msec and 200 msec. Mean naming latencies and response error percentages are reported in Table 5. Higher error rates were obtained in the present experiment than in previous experiments, because the complex targets were mainly of low frequency.

	Radical	Semantic	Control
SOA100	689	659	642
	(7.5)	(6.0)	(6.3)
SOA200	694	678	650
	(5.8)	(5.2)	(6.0)

Table 5 Experiment 3: Mean naming latencies and error percentages

Analyses of reaction times showed a significant main effect of prime type ($p<.001$) and no effect of SOA manipulation. Post hoc tests revealed that the mean reaction times in radical priming conditions were significantly slower than the mean reaction times in the other two conditions ($p<.01$). Moreover, the differences between the mean reaction times in semantic and control priming conditions were also significant ($p<.05$).

The significant inhibitory effect of phonetic radicals on the processing of irregular complex characters replicated Zhou (1994) who systematically manipulated the frequency of radicals and complex characters and the phonological relations between them. However, it is difficult from the present data to pinpoint the locus of this effect. The more interesting finding was that the presence of semantic primes related to phonetic radicals of complex characters was able to delay the activation of phonological representations of complex characters. This finding is consistent with the view that there is an automatic decomposition process in reading complex characters and both phonological and semantic properties of phonetic radicals are activated. The presence of semantic primes increases the semantic activation of radicals. This activation either increases its competition with the semantic representations of complex characters or send its activation to the corresponding phonological representations and intensify their competition with the phonological representations of complex characters. Either way, the phonological activation is easily modified by events in the semantic system.

General Discussion

In this study, we observed phonological activation of phonetic radicals in reading complex characters and its interaction with frequencies of complex characters (Experiment 1); and more importantly, we found evidence of automatic decomposition and access to semantic properties of phonetic radicals in reading complex characters (Experiments 2, and 3). These findings demonstrate that there are no fundamental differences between lexical processing and sublexical processing of phonetic radicals in reading Chinese. This view is consistent with the claims of our lexical processing model in reading Chinese (Zhou, in press, Zhou & Marslen-Wilson, 1996; Zhou et al., 1996), which stresses the predominant role of orthography in initial lexical access, the effective mapping from orthography to semantics, the automaticity of phonological activation from orthography, and the efficient flow of semantic activation to phonology.

The findings of automatic activation of phonological representations of phonetic radicals and its interaction with the frequencies of complex characters provide us with an explicit account for the "phonological" processing of pseudo-characters (Fang et al., 1986; Yin & Butterworth, 1992) and the regularity and consistency effects in reading Chinese (Fang et al., 1986; Hue, 1992; Peng et al., 1994; Seidenberg, 1985). The phonological processing of unfamiliar characters or pseudo-characters is nothing more than extracting meaningful orthographic patterns (usually phonetic radicals) from visual input and mapping them onto their corresponding phonological (and semantic) representations. This decomposition strategy is explicitly or implicitly learned and could be under conscious control. It is also useful in reading low frequency complex characters, which tend to be regular. Low frequency complex characters are more likely to be decomposed than high frequency characters, partly because the frequencies of phonetic radicals are usually much higher than the frequencies of complex characters themselves. For regular complex characters, the activation of orthographic representations corresponding to the whole complex characters and their phonetic radicals leads to the activation of the same phonological representations. For irregular complex characters, the activation of orthographic representations corresponding to the whole complex characters and their phonetic radicals leads to the activation of two different phonological representations. The activation of two different semantic representations by complex characters and by their radicals reinforce the competition between phonological representations. The consistency effect is explained by the assumption that the phonetic radicals in complex characters map not only onto their own orthographic and phonological representations, but also the representations of other complex characters having these radicals as components. The competition and/or mutual support among these representations produce the consistency effects observed in the naming studies.

The above account of phonological processing of phonetic radicals in reading complex characters is not exactly the same as the accounts for the regularity and consistency effects in reading alphabetic languages (e.g., Coltheart et al., 1993; Frost, in press; Seidenberg & McClelland, 1989). On the latter accounts, the regularity and consistency effects in naming are due to the "assembly" of phonology using grapheme-phoneme correspondences, either prelexically or lexically. Sublexical processing provides piecemeal information of the whole word phonology. In Chinese, however, the sublexical phonological processing of phonetic radicals provides the whole word phonology, rather than part of it. Moreover, the phonological information of phonetic radicals themselves comes from the lexicon, as "addressed" phonology. In other words, there are such things as "prelexical" phonology in reading Chinese. All phonological information, whether for whole words or for the phonetic radicals, comes from the direct access through orthographic representations in the lexicon (Zhou & Marslen-Wilson, 1996).

The finding of automatic activation of semantic representations of phonetic radicals supports our view that equating sublexical processing to lexical processing in reading Chinese. Although the semantic properties of phonetic radicals have no relations with the semantic properties of the whole characters, they are nevertheless activated to interfere with the semantic processing of the whole characters. This is different from the sublexical processing of orthographic units in reading alphabetic languages. There is no sublexical semantic activation in reading English, even when the orthographic units involved happen to be words. When we ran a primed naming experiment similar to Experiment 2 here, with words (e.g., BOYCOTT) having embedded words (e.g., BOY) as primes and semantic associates (e.g., GIRL) of the embedded words as targets, we found no priming effects at all for the associates.

Thus, although the idea of parallel mapping between visual input and different orthographic, semantic and phonological representations in the lexicon is consistent with interactive and connectionist views of lexical processing (e.g., Seidenberg & McClelland, 1989), it is not clear whether the differences in sublexical processing of alphabetic and logographic writing systems reflect fundamental differences in the way different orthographies representing the sound and the meaning of lexical items. It is a challenge for students of lexical processing to propose a model that would account for both the English and Chinese data.

References

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589-608.

Fang, S. -P., Horng, R. -Y., & Tzeng, O. (1986). Consistency effects in the Chinese character and pseudo-character naming tasks. In H. S. R. Kao and R. Hoosain (Eds.), *Linguistics, psychology, and the Chinese language* (pp. 11-21). Center of Asian Studies, University of Hong Kong.

Frost, R. (in press). Towards a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*.

Hue, C.-W. (1992). Recognition processes in character naming. In H-C. Chen and O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 93-107). Amsterdam: North-Holland.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674-691.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: 1, Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107-128.

Peng, D., Yang, H., & Chen, Y. (1994). Consistency and phonetic-independency effects in naming tasks of Chinese phonograms. In Q. Jing, H. Zhang, & D. Peng (Eds). *Information processing of the Chinese language* (pp. 26-41). Beijing: Beijing Normal University Press.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*, 1-30.

Seidenberg, M. S. & McClelland, J. L (1989). A distributed, developmental model of visual word recognition and naming. *Psychological Review*, *96*, 523-568.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488-522.

Yin, W. G. & Butterworth, B. (1992). Deep and surface dyslexia in Chinese. In H-C. Chen and O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 349-366). Amsterdam: North-Holland.

Zhou, X. (in press). Is there a phonologically mediated access to lexical semantics in reading Chinese? In A. Inhoff, J. Wang, & H.-C. Chen (Eds.) *Reading Chinese script: A cognitive analysis*. NJ: Lawrence Erlbaum.

Zhou, X., & Marslen-Wilson, W. (1996). Direct visual access is the only way to access the Chinese mental lexicon. In G. W. Cottrel (Ed.) *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp.714-719). Mahwah, NJ: Lawrence Erlbaum.

Zhou, X., Marslen-Wilson, W., Shu, H., Bi, Y., & Tang, Y. (1996). *Phonology in reading Chinese*. Manuscript.