

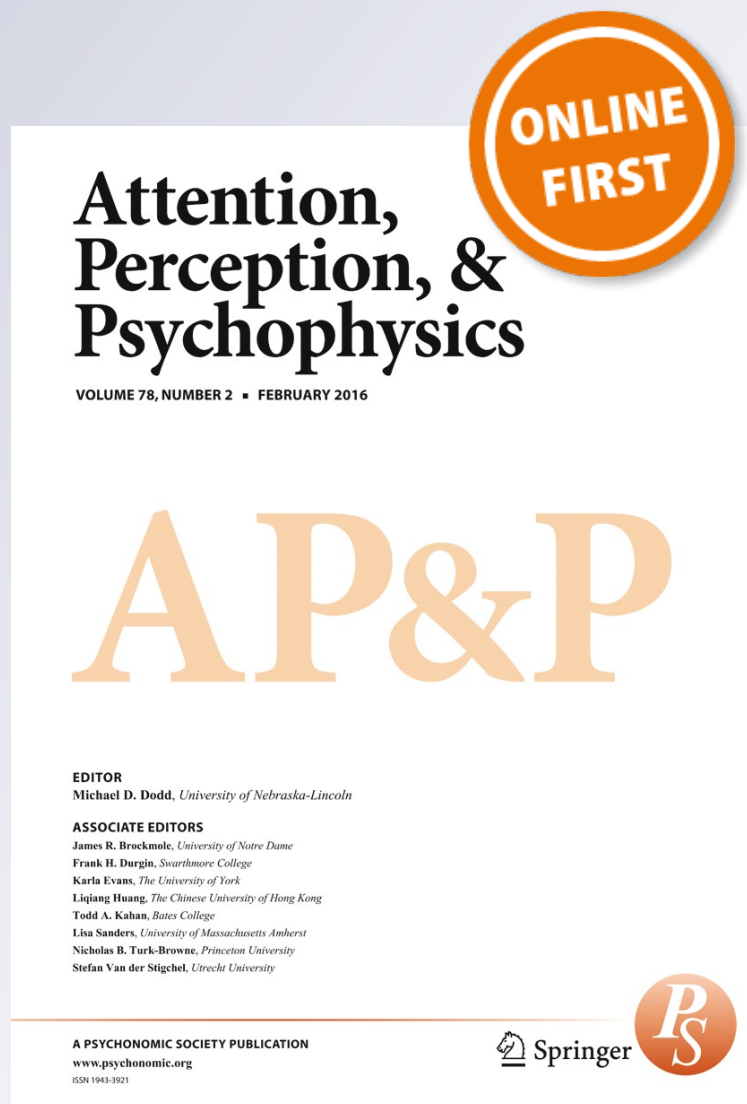
Semantic information mediates visual attention during spoken word recognition in Chinese: Evidence from the printed-word version of the visual-world paradigm

Wei Shen, Qingqing Qu & Xingshan Li

Attention, Perception, & Psychophysics

ISSN 1943-3921

Atten Percept Psychophys
DOI 10.3758/s13414-016-1081-z



Your article is protected by copyright and all rights are held exclusively by The Psychonomic Society, Inc.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Semantic information mediates visual attention during spoken word recognition in Chinese: Evidence from the printed-word version of the visual-world paradigm

Wei Shen^{1,2} · Qingqing Qu¹ · Xingshan Li¹

© The Psychonomic Society, Inc. 2016

Abstract In the present study, we investigated whether the activation of semantic information during spoken word recognition can mediate visual attention's deployment to printed Chinese words. We used a visual-world paradigm with printed words, in which participants listened to a spoken target word embedded in a neutral spoken sentence while looking at a visual display of printed words. We examined whether a semantic competitor effect could be observed in the printed-word version of the visual-world paradigm. In Experiment 1, the relationship between the spoken target words and the printed words was manipulated so that they were semantically related (a semantic competitor), phonologically related (a phonological competitor), or unrelated (distractors). We found that the probability of fixations on semantic competitors was significantly higher than that of fixations on the distractors. In Experiment 2, the orthographic similarity between the spoken target words and their semantic competitors was manipulated to further examine whether the semantic competitor effect was modulated by orthographic similarity. We found significant semantic competitor effects regardless of orthographic similarity. Our study not only reveals that semantic information can affect visual attention, it also provides important new insights into the methodology employed to investigate the semantic processing of spoken words during spoken word

recognition using the printed-word version of the visual-world paradigm.

Keywords Semantic processing · Printed-word paradigm · Spoken word recognition · Chinese · Eye movements

Examining the types of information involved in spoken word recognition and their effects on visual attention can help us better understand how information from the visual and auditory modalities interacts to influence eye movements. In the present study, we investigated whether the activation of semantic information in spoken word recognition can mediate the deployment of visual attention to printed Chinese words.

For our research we utilized the visual-world paradigm, which has been used to study the interplay of cognitive processes such as language processing and attention that have traditionally been investigated separately (Allopenna, Magnuson, & Tanenhaus, 1998; Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; see Henderson & Ferreira, 2004, for a review). In this paradigm, participants listen to an utterance while looking at a visual display, and their eye movements are tracked. It is well-established that the proportions of eye fixation on objects are constrained by semantic, orthographic, or phonological relations between auditory information and object names. For instance, Allopenna, Magnuson, and Tanenhaus (1998) showed participants a screen with four objects: a target referent (“beaker”), a phonological competitor sharing the initial syllables with the target (“beetle”), a phonological competitor sharing the tail syllables with the target (“speaker”), and an unrelated distractor (“carriage”). The participants were asked to follow the instruction of “please pick up the beaker.” Their study showed that both types of phonological competitors attracted more fixations than the unrelated distractors.

✉ Qingqing Qu
quqq@psych.ac.cn

✉ Xingshan Li
lixs@psych.ac.cn

¹ Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Beijing, China

² University of Chinese Academy of Sciences, Beijing, China

Moreover, the onset-match phonological competitor (“bee-
tle”) attracted more and earlier fixations than did the offset-
match phonological competitor (“speaker”). These findings
were in line with the assumptions of the TRACE model of
spoken word recognition (McClelland & Elman, 1986), which
assumes that a set of word candidates sharing the same syl-
lables as target words are activated temporarily upon hearing a
spoken word. More importantly, these findings also showed
that a match established at the phonological level can drive
attention shifts to visual objects, and that the phonological
information of a spoken word is gradually mapped onto fix-
ated objects during spoken word recognition (Dahan,
Magnuson, Tanenhaus, & Hogan, 2001).

Evidence from the visual-world paradigm indicates that
attention shifts to visual objects are driven not only by pho-
nological information, but also by matches built at the seman-
tic level. For example, in Huettig and Altmann (2005), partic-
ipants were asked to listen to a sentence that included a target
word (e.g., “piano”) while looking at a visual display of four
objects without performing any explicit task. Huettig and
Altmann found that the fixation probabilities on semantically
related objects (e.g., “trumpet”) were significantly higher than
those on unrelated objects. The authors argued that the seman-
tic representation of “piano” was activated on hearing the
spoken target word, and that semantic information regarding
the visual object “trumpet” was also activated at the same
time. On the basis of this finding, they proposed that semantic
overlap between the two semantic representations resulted in a
visual attention shift.

Huettig and McQueen (2007) investigated the time course
of information processing during spoken word recognition. In
that study, participants saw four visual object competitors and
heard a spoken target word “beaker” embedded in a neutral
context sentence, such as “eventually she looked at the beaker
that was in front of her.” They found that the proportions of
fixations on the three competitors (a phonological competitor,
a semantic competitor, and a shape competitor) were all sig-
nificantly higher than that on the unrelated distractor. More
importantly, attention shifts to the phonological competitor
occurred earlier than those to the shape competitor, and were
followed by a shift to the semantic competitor. However, the
fixations to all three types of competitors overlapped in time,
which suggested that spoken word recognition is a cascaded
rather than a serial process. In a recent Chinese study, Tsang
and Chen (2010) employed a classic visual-world paradigm
and provided further evidence that a match established at the
semantic level could drive visual attention to visual objects.

Huettig and McQueen (2007) also developed a variant of
the classic visual-world paradigm in which a visual display
depicting objects was replaced with printed words. Like the
classic version, the modified printed-word version is consid-
ered a sensitive tool to investigate phonological processing
during spoken word recognition (see McQueen & Viebahn,

2007; Salverda & Tanenhaus, 2010; Weber, Melinger, &
Tapia, 2007). In McQueen and Viebahn (2007), Dutch partic-
ipants were presented with a target word, a phonological com-
petitor word, and two unrelated distractor words. They found
that the participants fixated more on the phonological com-
petitor words than on the unrelated words. In other words, a
phonological competitor effect similar to that observed in
the classic visual-world paradigm (presented with pictures)
was found in the novel printed-word variant. The phonologi-
cal competitor effect has been validated in Spanish by Weber,
Melinger, and Tapia (2007). They compared various modes of
stimulus presentation (pictorial, printed-word, and the
combination of pictures and words) and found a phonological
competitor effect in all three presentation modes, with a stron-
ger competitor effect in the printed-word display mode than in
the classic pictorial display mode. These findings suggest that
the phonological information of spoken words can mediate a
shift of visual attention, which validates the reliability of the
printed-word display paradigm for investigating phonological
processing during spoken word recognition.

Huettig and McQueen (2007) found a different data pattern
when they replaced visual objects with printed Dutch words.
Only phonological competitor words attracted more fixations
than distractors, whereas no more fixations were attracted by
the semantic competitor than by the distractors. Given the
absence of a semantic competition effect, the authors argued
that semantic information is irrelevant in search of a visual
printed-word display. In addition, they argued that the
printed-word version is less sensitive for investigating seman-
tic processing than is the classic visual-world version.

All of the studies above using the printed-word version of
the visual-world paradigm were conducted with alphabetic
languages. To our knowledge, only one study has used the
printed-word version of the visual-world paradigm to investi-
gate a nonalphabetic language. Shen, Deutsch, and Rayner
(2013) investigated the influence of phonological pitch infor-
mation on Chinese word processing using the printed-word
paradigm. In that experiment, participants were instructed to
listen to a spoken word and select that word from among four
printed words. The researchers manipulated the tone relation-
ship between the printed and spoken words and found that
tone pitch height played an important role in lexical tone
perception.

So far, no study has examined whether visual attention shift
to printed words can be mediated by semantic information in
nonalphabetic languages. A null effect of semantic informa-
tion and a significant effect of phonological information are
perhaps not surprising in the case of alphabetic languages, as
far as the spelling–sound relation and the spelling–meaning
relation are concerned. In nonalphabetic script systems, how-
ever, it is very likely that semantic information plays an im-
portant role in mediating visual attention shifts, given that the
orthographic form–meaning relation is highly transparent due

to the pictographic nature of its origin. In alphabetic languages (e.g., English, Spanish, and Dutch), there is a close relationship between phonology and orthography (Monaghan & Ellis, 2002; Patterson & Morton, 1985). A large number of psycholinguistic studies have supported such close spelling–sound connections by documenting the activation of orthographic information in spoken word recognition and the activation of phonological information in visual-world recognition (e.g., Slowiaczek, Soltano, Wieting, & Bishop, 2003; Ziegler & Ferrand, 1998). Moreover, the spelling–meaning connection is not transparent in alphabetic languages, which may explain why a semantic effect was not observed in previous studies on alphabetic languages.

In contrast to alphabetic languages, Chinese has a strong orthographic form–meaning connection. Studies have shown that the meaning of Chinese characters can be accessed more efficiently through a direct mapping between orthography (i.e., the semantic radical) and semantics (Chen, Flores d'Arcais, & Cheung, 1995; Chen & Shu, 2001; Hoosain, 1991; Hoosain & Osgood, 1983). The majority (85 %) of Chinese characters are composed of a semantic radical that provides information about their semantic category (Zhou, Ye, Cheung, & Chen, 2009). It should be noted that although a semantic radical can carry some semantic cue regarding the whole character (e.g., the semantic category that the whole character belongs to), it does not usually carry the same amount of semantic information as the whole character. For example, the character 仆 (*pu2*, “servant”) includes a semantic radical 亻 (*ren2*, “people”) on its left. By means of this radical, readers can deduce that the character 仆 has a semantic connection with people. However, a Chinese reader cannot deduce the meaning of the character 仆 solely on the basis of the semantic radical 亻, since it is also used in 382 other characters. Moreover, there are many characters with similar meanings in Chinese, but they do not necessarily share the same semantic radical. For example, the characters 走 and 行 have similar meanings (“walking”), but they have different semantic radicals. Therefore, in Chinese, two semantically related characters are not necessarily orthographically similar (i.e., they may not use the same semantic radical).

Besides semantic radicals, most Chinese characters also contain a phonological radical that provides information about the character's pronunciation (Zhu, 1988). However, the orthographic form–sound connection is rather weak. Only 26 % of the characters have the same sound as their phonetic radicals (Fan, Gao, & Ao, 1984; Zhou et al., 2009). Moreover, homophones are common in Chinese. For example, the pronunciation *guan1* applies to many characters, such as 关, 官, 观, 棺, and 冠. Written Chinese has unique characteristics as far as the orthographic form–meaning and orthographic form–sound relations are concerned. Given the opaque orthographic form–sound relation of Chinese characters, retrieving the phonological information of a character may be difficult, whereas

the close orthographic form–meaning relation may ease the activation of semantic information in Chinese.

In the present study, we investigated whether the semantic information of spoken target words can drive visual attention using a printed-word version of the visual-world paradigm. Chinese participants were asked to listen to a spoken target word embedded in a neutral context sentence. A visual display with printed visual words was presented simultaneously. These words were related to the spoken target words in either the semantic (semantic competitor) or the phonological (phonological competitor) dimension, or were unrelated distractors. If the semantic competitor words attracted more fixations than the distractor words, this would imply that the semantic representations of the spoken and visual words overlapped and were activated together to mediate visual attention to printed Chinese words.

In this study, we also aimed to revisit the phonological competitor effect in Chinese. In previous studies conducted on alphabetic languages, a significant phonological competitor effect was observed (Huettig & McQueen, 2007). However, it is highly likely that this effect has been confounded by orthographic processing, given that the two properties are closely linked in alphabetic languages. Indeed, Salverda and Tanenhaus (2010) manipulated the phonological overlap between the target word and a competitor in the display to be either high or low after the orthographic overlap was controlled. They found that the proportions of fixations on the competitor in this case did not differ as a function of phonological overlap. In a different experiment, they directly manipulated the orthographic overlap between the target words and competitor to be either high or low. They found a higher fixation probability on the high-orthographic-overlap competitor than on the low-orthographic-overlap competitor. On the basis of these findings, Salverda and Tanenhaus concluded that what has been conceived of as a phonological competitor effect in the printed-word version of the visual-world paradigm is in fact an orthographic competitor effect. Since orthography and phonology are largely dissociated in Chinese, their corresponding effects can be optimally isolated from each other in this language. In addition, the nonalphabetic nature of the language allows us to use a phonological competitor that is only phonologically related, and not orthographically related, to a spoken target word. In the present study, we thus aimed to investigate the presence of a “pure” phonological competitor effect in Chinese that would not be confounded by a potential orthographic effect.

Another variable commonly manipulated in the visual-world paradigm tasks is the preview time between presentation of the visual display and the auditory target word. Manipulating the onset of the visual display relative to that of the auditory word allows researchers to examine whether the effects are influenced by the preview time. Huettig and McQueen (2007) found that the preview time of a visual

display indeed affected information retrieval. With a long preview time, all three kinds of information (phonology, semantic, and shape) mediated attention shifts to visual objects. However, the phonological competitor effect was not observed with the short preview time. The authors argued that a short preview time might not be enough to retrieve phonological information. We sought to investigate this further by manipulating preview time in Experiment 1.

Experiment 1

Method

Participants Twenty native Chinese speakers (13 females, 7 males) were recruited to participate in the experiment. They were undergraduate students from universities near the Institute of Psychology, Chinese Academy of Sciences, and were paid RMB40 (approximately US\$6) after the experiment. All participants had normal or corrected-to-normal vision, and all were unaware of the purpose of the experiment. The ages of the participants ranged from 19 to 27 years old ($M = 22.50$).

Materials and design Eighty words were selected as the spoken target words. Each visual display consisted of four printed disyllabic words—namely, (1) a semantic competitor (which was semantically related to the spoken target word—e.g., 医生, yīshēng1, “doctor” and 护士, hùshi, “nurse”), (2) a phonological competitor (which shared the same syllable with the first character of the spoken target word—e.g., 衣架, yījiā4, “hanger”), and (3) two unrelated distractors (which were neither semantically nor phonologically related to the spoken target word; see Fig. 1 for an example of the experimental stimuli). We selected the semantic competitors carefully to avoid phonological relations between the semantic competitors and spoken target words. Semantic overlap between the phonological competitors and spoken target words was also avoided. The positions of the four words were fully counterbalanced in the display. The spoken target word was embedded in a neutral context sentence to avoid predictability based on the preceding context. In addition, the four presented visual words were matched closely on word frequency and number of strokes for each display ($F_s < 1$; see Table 1 for the properties of the materials). To evaluate the semantic relatedness between the words, ten participants (who did not participate in the formal experiment) were recruited to judge the semantic relatedness of the word pairs on a 5-point scale (1 = *not related at all*, 5 = *strongly related*). Each spoken target word was paired, respectively, with a semantic competitor, a phonological competitor, and two distractors, to construct four word pairs. The semantic relatedness was significantly higher between the spoken target word and the semantic competitor

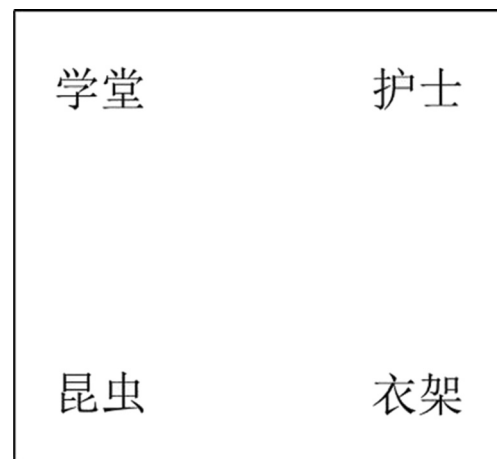


Fig. 1 An example of a printed-word display used in the present study. For a spoken sentence like “在利比里亚人们常把医生叫做天使” (“In Liberia, people often call **doctors** angels”), the printed-word display could consist of a semantic competitor word 护士 (hùshi4, “nurse”), a phonological competitor word 衣架 (yījiā4, “hanger”), and two unrelated distractors 学堂 (xué2tāng2, “school”) and 昆虫 (kūn1chōng2, “insect”) in four different positions on the display

($M = 4.37$) than between the target and either the phonological competitor ($M = 1.22$), $t(79) = 53.14$, $p < .001$, or the distractor words ($M = 1.20$), $t(79) = 65.09$, $p < .001$. See Table 2 for the semantic relatedness scores. An additional ten participants were recruited to perform a predictability rating of the spoken target words. Sentences preceding the spoken target words (target words were not included) were presented to the participants, and they were instructed to write down a possible word that could fit into the sentence. The mean score of the predictability ratings was 0, which suggests that no spoken target word was predicted on the basis of the preceding context.

Furthermore, 80 additional filler trials were presented, to prevent listeners from noticing the manipulations between the visual printed words and the spoken target words. Each filler trial was also composed of four printed words unrelated to any word in the spoken sentence.

To investigate whether the effects were modulated by preview time in Chinese, we included two preview conditions in the work reported below. In the long-preview condition, the visual printed words were presented on average 1,920 ms earlier than the onset of the spoken target words. In the short-preview condition, the visual printed words were presented only 200 ms earlier than the onset of the spoken target words.

Apparatus Eye movements were recorded using an EyeLink 1000 tracker (SR Research, Mississauga, Ontario, Canada). Experimental materials were presented on a 21-in. CRT monitor (Sony Multiscan G520) with a $1,024 \times 768$ pixel resolution and a refresh rate of 150 Hz. The eyetracking system was sampled at 1000 Hz. The participants placed their chins on a chinrest and leaned their foreheads on a forehead rest to

Table 1 Properties of the experimental materials in Experiment 1

		Semantic Competitor	Phonological Competitor	Distractor1	Distractor2
Long-preview condition	Mean word frequency (occurrences per million)	4.76	4.43	4.81	4.82
	Mean stroke number	15.85	16.70	15.45	17.35
Short-preview condition	Mean word frequency (occurrences per million)	3.37	3.40	3.34	3.34
	Mean stroke number	15.35	15.75	16.48	15.63

minimize head movements during the experiment. Although viewing was binocular, eye movement data were collected only from the right eye. Participants were seated 58 cm from the video monitor.

Procedure All spoken sentences were recorded in a neutral tone by a female native Chinese speaker, and the spoken target words in the sentences were not highlighted. The sentences were presented to the participants through a headphone (COMIX, QE6302, China). All of the visual printed words were presented in 30-point Song font in black (RGB: 0, 0, 0) on a gray background (RGB: 204, 204, 204), and each character subtended a visual angle of approximately 1.4° and appeared approximately 7° from the center of the screen at this viewing distance.

After the participants entered the lab, they were given a brief introduction to the eyetracker and the instructions for the experiment. The eyetracker was calibrated and validated at the beginning of the experiment. A calibration was conducted whenever the error was greater than 1° during the experiment. The participants performed this procedure by means of a nine-point calibration, and the validation error was smaller than 1° of visual angle. A drift check was performed at the beginning of each trial, and then a blank screen was presented for 600 ms. In the long-preview condition, the spoken sentence was presented at the same time as the visual display, and the average preview time for the display before the onset of the spoken target word was 1,920 ms (ranging from 1,556 to 2,445 ms). In the short-preview condition, the spoken sentences were presented first and the printed-word display was previewed for 200 ms before the onset of the spoken target word. For one half of the filler trials, the spoken sentences were presented at the same time as the visual display. For the other half of the fillers, the spoken sentences were presented for about 1,800 ms before the onset of the printed-word display. The two preview conditions were

presented randomly to the participants. In addition, the positions of the four printed words were randomized for each trial, and participants performed all trials in random order. Participants were asked to listen to the sentence and to view the printed-word displays without performing any explicit task. They were asked to press the toggle button when the spoken sentence was completed. Six practice trials were presented before the formal experiment to familiarize the participants with the experimental procedure. Each participant performed 40 experimental trials in the long-preview condition, 40 experimental trials in the short-preview condition, and 80 filler trials. The trials were presented in random order, and the entire experiment lasted approximately 45 min.

Data analysis Fixations on a 5° × 5° square centered at the printed word were defined as fixations on the corresponding visual printed word.

Results

We calculated the mean fixation proportions of the semantic competitor, the phonological competitor, and the distractors in each preview condition from 200 ms before to 2,000 ms after the onset of the spoken target word. The entire time period was divided into 22 time windows, with the fixations falling on each visual printed word being measured, and the fixation proportions were calculated every 100 ms. Figure 2 presents the distributions of the mean fixation proportions to the semantic competitor, phonological competitor, and distractors in the long- and short-preview conditions over time.

Figure 2a shows that the fixation proportions in the long-preview condition were almost at similar levels across the four printed words before the emergence of the spoken target words. Thus, no preference was shown to any of the printed

Table 2 Semantic relatedness scores between the spoken target words and the visual printed words in Experiment 1

	Semantic Competitor	Phonological Competitor	Distractors
Long-preview condition	4.46 (0.31)	1.26 (0.36)	1.16 (0.21)
Short-preview condition	4.28 (0.35)	1.18 (0.34)	1.24 (0.31)

The numbers in parentheses are standard deviations

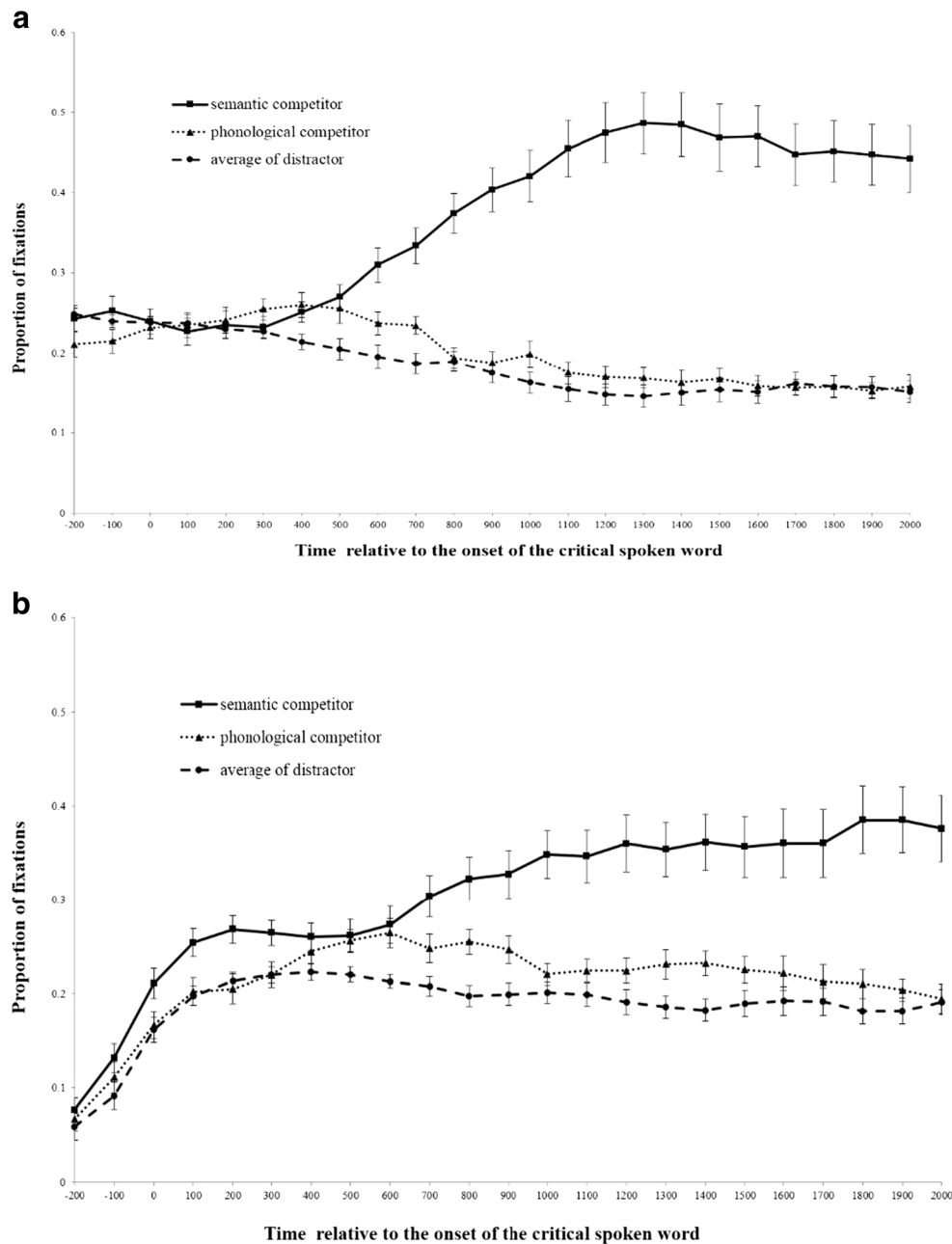


Fig. 2 Graph showing the fixation proportions to semantic competitors, phonological competitors, and the average of the distractors in the long-preview condition (a) and the short-preview condition (b) in Experiment

1, from 200 ms before onset of the spoken target word (0 on the x-axis indicates the onset of the spoken target words)

words before hearing the spoken target words. The fixation probability curve of the semantic competitor started to diverge from those of the phonological competitor and distractors after the onset of the spoken target words. The fixation proportion of the semantic competitor increased continuously, peaking at approximately 1,000 ms after the onset of the critical words. By contrast, the fixation proportion of the phonological competitor increased only modestly after the onset of the spoken target words, and this trend rapidly decreased

to almost the same level as the average for the distractors. A similar pattern was observed in the short-preview condition (see Fig. 2b).

To investigate whether the semantic and phonological competitors indeed attracted more fixations than the distractors, we defined as a dependent variable whether a fixation was made to the printed words within a specific time window of 600 ms, which was time-locked to the onset of the spoken target words (Altmann & Kamide, 2004). Previous eyetracking studies have shown that the time required to plan

and execute an eye movement saccade is about 150–200 ms (Rayner, 1998). Thus, fixations to the printed words were time-locked to the time interval from 200 ms after the onset of the target word to 200 ms after its offset (i.e., from 200 to 800 ms).

The fixation proportion data were analyzed using logit mixed models (Jaeger, 2008; see also Ferreira, Foucart, & Engelhardt, 2013) for the long-preview and short-preview conditions. We included random intercepts for participants and items, and by-participants random slopes for the fixed factor Competitor Type as random effects¹ (Barr, Levy, Scheepers, & Tily, 2013). The competitor type (semantic competitor, phonological competitor, and distractor) was entered as fixed effect. Then, the following comparisons were made: (1) fixations on the semantic competitor versus a distractor and (2) fixations on the phonological competitor versus a distractor. One of the two distractors was randomly assigned as the baseline condition with which the semantic competitor and phonological competitor were compared.² The model was fitted using the `glmer` function from `lme4` package (Version 1.1-7; Bates, Maechler, Bolker, & Walker, 2014) in the R environment (Version 3.1.1; R Development Core Team, 2014). The regression coefficients (b), standard errors (SE), Wald- Z values, and p values are reported.

Model fitting was conducted by first creating a base model including intercepts for participants and items as random effects. The model was then enhanced by adding the fixed factor Competitor Type, and finally by-participants random slopes for Competitor Type. We assessed the model's improvement by conducting log-likelihood ratio tests.

Table 3 shows the proportions of trials on which a fixation was made to the semantic competitor, the phonological competitor, and the distractor in the long-preview and short-preview conditions, respectively. The results showed that in the long-preview condition, the model was significantly improved with the inclusion of competitor type [$\chi^2(3) = 22.70, p < .001$], but adding by-participants slopes for competitor type did not significantly improve the model fit [$\chi^2(9) = 6.72, p > .60$]. Comparison between the semantic competitor and distractor showed that the semantic competitor received more fixations than did the distractor ($b = 0.44, SE = 0.10, Wald-Z = 4.26, p = .001$). Furthermore, the phonological competitor also received more fixations than the distractor ($b = 0.23, SE = 0.10, Wald-Z = 2.25, p = .02$).

In the short-preview condition, the inclusion of competitor type again significantly improved the model fit [$\chi^2(3) = 6.83,$

$p = .07$], and adding by-participants slopes for competitor type did not improve the model fit [$\chi^2(9) = 1.67, p > .90$]. Comparison between the semantic competitor and distractor showed that the semantic competitor received significantly more fixations than the distractor ($b = 0.23, SE = 0.10, Wald-Z = 2.24, p = .02$). However, no difference was found between the fixations on the phonological competitor and the distractor ($p > .80$).

Discussion

Using the printed-word version of the visual-world paradigm, we found a significant semantic competition effect—semantic competitor words were fixated on more often than distractors both when there was only 200 ms of preview time prior to the onset of the spoken target word (i.e., the short-preview condition) and when the displays appeared at sentence onset (i.e., the long-preview condition). These findings demonstrate that semantic information was activated during spoken word recognition and listeners mapped this information onto the printed Chinese words, thereby resulting in a shift of visual attention. Moreover, although the phonological competitor also received more fixations than the distractor, this effect only emerged in the long-preview condition, and was absent in the short-preview condition.

In Experiment 1, orthographic similarity (defined by use of the same radical in both the spoken target word and the printed word) between the target words and the semantic and phonographic competitors were not intentionally controlled. As a result, 17.5 % of the semantic competitors and 20 % of the phonological competitors were orthographically similar to the spoken target words. Could the semantic effect observed in Experiment 1 have resulted not from semantic overlap, but rather from orthographic similarity between the spoken target words and the competitor words? Previous studies using the visual-world paradigm showed that fixations on printed words are sensitive to the degree of orthographic overlap between spoken and printed words (Salverda & Tanenhaus, 2010). In Experiment 1, the orthographic similarity between the semantic competitor and spoken target words in some stimuli made the explanation of “orthographic similarity” plausible. To test this possibility, we conducted Experiment 2, which was identical in most aspects to Experiment 1, but different in that orthographic overlap was manipulated to examine whether the semantic competitor effect was modulated by orthographic similarity.

Experiment 2

In Experiment 2, we varied the degree of orthographic overlap between the spoken target words and semantic competitors. Two conditions were used: The spoken target words were

¹ Given the between-items design used in our study, only by-participants random slopes for the fixed factor were tested.

² In the first round of the data analysis, we randomly assigned one distractor as the baseline condition. Assigning the other distractor as the baseline produced similar results. Therefore, we report only the results from the first round of analysis.

Table 3 Proportions of trials with a fixation to the visual printed words in Experiments 1 and 2

	Semantic Competitor	Phonological Competitor	Distractors
Experiment 1			
Long-preview condition	.48 (.02)	.43 (.02)	.38 (.02)
Short-preview condition	.49 (.02)	.44 (.03)	.44 (.02)
Experiment 2			
S+O+ condition	.54 (.02)	.54 (.02)	.49 (.01)
S+O– condition	.57 (.03)	.53 (.02)	.49 (.01)

The numbers in parentheses are standard deviations

either orthographically related or unrelated to the semantic competitors. However, in both conditions, the spoken target words and semantic competitors were always semantically related. If the semantic competitor effect observed in Experiment 1 resulted from orthographic similarity, we would observe that the semantic effect varied as a function of orthographic overlap—that is, we would observe a semantic effect in the orthographically related condition, but not in the orthographically unrelated condition. An alternative outcome would be the observation of comparable semantic competitor effects in both conditions, regardless of orthographic similarity. Furthermore, the phonological competitors were all carefully chosen so that they were phonologically related with the spoken target words yet shared no orthographic similarity. Our basic premise was that if the phonological competitor effect observed in Experiment 1 was due to orthographic similarity, no phonological competitor effect would be observed in Experiment 2.

Method

Participants Experiment 2 had 20 participants (13 females, 7 males) 20 to 27 years of age ($M = 23$), recruited from the same pool as those of Experiment 1. None of them had participated in Experiment 1. All participants had normal or corrected-to-normal vision and were paid RMB25 (approximately US\$3.90) after the experiment.

Material and design The experimental design was identical to that of Experiment 1, with the following exception: The orthographic similarity between the spoken target word and the semantic competitor was manipulated to be either related

(semantically and orthographically related condition) or unrelated (semantically related but orthographically unrelated condition). For the semantically and orthographically related condition (hereafter the S+O+ condition), the semantic competitor shared a semantic radical with the spoken target word (e.g., 胳膊, ge1bo, “arms” and 大腿, da4tui3, “legs”). For the semantically related but orthographically unrelated condition (hereafter the S+O– condition), the semantic competitor shared no radicals with the spoken target word (e.g., 铅笔, qian1bi3, “pencil” and 橡皮, xiang4pi2, “eraser”). In addition, all phonological competitors were manipulated to be phonologically similar (sharing a same initial syllable and tone) but orthographically dissimilar (sharing no radicals or any other components) with the spoken target words. The other two distractors were unrelated to the spoken target words in the semantic, phonological, and orthographical dimensions. The four printed words were matched across conditions in terms of word frequency and number of strokes (See Table 4 for the properties of the materials).

As in Experiment 1, we recruited ten participants to rate the semantic relatedness of the spoken target words with the other three types of words on a 5-point scale (1 = *not related at all*, 5 = *strongly related*). See Table 5 for the semantic relatedness scores. In the S+O+ condition, the semantic relatedness between the spoken target word and the semantic competitor ($M = 4.49$) was significantly higher than its semantic relatedness to either the phonological competitor ($M = 1.25$), $t(39) = 52.29$, $p < .001$, or the distractor words ($M = 1.20$), $t(39) = 61.01$, $p < .001$. Similarly, in the S+O– condition, the semantic relatedness between the spoken target word and the semantic competitor ($M = 4.36$) was again significantly higher than its relatedness to either the phonological competitor ($M = 1.20$), $t(39) = 33.14$, $p < .001$, or the distractor words ($M =$

Table 4 Properties of the experimental materials in Experiment 2

		Semantic Competitor	Phonological Competitor	Distractor1	Distractor2
S+O+ condition	Mean word frequency (occurrences per million)	4.73	4.53	4.58	4.62
	Mean stroke number	17.80	16.28	16.43	16.40
S+O– condition	Mean word frequency (occurrences per million)	3.64	3.42	3.09	3.11
	Mean stroke number	16.23	16.60	16.55	16.65

1.18), $t(39) = 34.61$, $p < .001$. More critically, the semantic relatedness between the spoken target words and semantic competitors was comparable across the two conditions, $t(39) = 1.19$, $p > .10$. All target words were embedded in neutral context sentences and were unpredictable on the basis of the preceding context.

Furthermore, 12 participants were recruited to assess the plausibility of the preceding context (before the spoken target word) and the printed words on a 5-point scale (1 = *very implausible*, 5 = *very plausible*). The mean scores for the semantic competitor, phonological competitor, and distractor, respectively, were 3.70, 3.74, and 3.94 in the S+O+ condition and 3.71, 3.63, and 3.62 in the S+O− condition. No difference was found across the three competitor types in the S+O+ or the S+O− condition ($ps > .10$ in all cases).

Quintuples of words were selected for 80 experimental trials. Each set consisted of a semantic competitor (half orthographically related and the other half orthographically unrelated), a phonological competitor, and two unrelated distractors. A further 80 quadruples of words were selected for the filler trials, to prevent participants from being aware of the manipulation between the spoken target words and the printed words.

Apparatus The apparatus was identical to that used in Experiment 1.

Procedure The procedure was identical to that of Experiment 1, except that the printed-word display was only previewed for 200 ms before the onset of the spoken target words. We did not include a long-preview condition because the results of Experiment 1 showed that the semantic competitor effect was not affected by the preview time of the visual display.³ The entire experiment lasted approximately 35 min.

Data analysis The data analysis was conducted as in Experiment 1.

Results

As in Experiment 1, we calculated the fixation probability of each corresponding printed word during a time window of 200–2,000 ms after the onset of the spoken target word. Figure 3 shows the fixation proportions of each printed word in the S+O+ condition and the S+O− condition, respectively. As can be seen in panel A, the fixation proportion of the semantic competitor begins to increase

gradually after the onset of the spoken target words. This pattern is similar to that observed in Experiment 1. The fixation proportion of the phonological competitor also increased modestly after the onset of the spoken target words, but then it began to decrease to the same level as that of the distractor. In panel B, we can see a similar data pattern in the S+O− condition.

As in Experiment 1, we analyzed the proportions of fixation to each competitor during a 200- to 800-ms time window using a logit mixed model. Table 3 shows the proportions of trials on which fixations were made on the semantic competitor, phonological competitor, and distractor in the S+O+ and S+O− conditions. To test whether inclusion of any one factor could account for significant variance in the data, we performed model fitting by first creating a base model including only intercepts for participants and items as random effects. Then, the model was enhanced by sequentially adding the fixed factors Competitor Type and Orthographic Similarity, the interaction between the two fixed factors, and finally by-participants random slopes for the fixed factors. The results of the mixed model analysis showed that the model was improved by the inclusion of competitor type [$\chi^2(3) = 21.37$, $p < .001$]. Adding other factors did not improve the fitness of the model [$\chi^2s < 7.3$, $ps > .50$].

Comparison of the semantic competitor with the distractor showed that the semantic competitor received more fixations than the distractor ($b = 0.26$, $SE = 0.07$, Wald- $Z = 3.62$, $p < .001$). In addition, fixations were also significantly higher on the phonological competitor than on the distractor ($b = 0.17$, $SE = 0.08$, Wald- $Z = 2.38$, $p = .02$).

Discussion

In Experiment 2, we manipulated the orthographic similarity between spoken target words and semantic competitors. The results revealed reliable semantic competitor effects, and more critically, the absence of an interaction between semantics and orthography. Significant semantic competitor effects were observed in both the S+O+ and S+O− conditions. This finding clearly demonstrated that orthographic similarity was not the source of the semantic competitor effect observed in Experiment 1. Instead, the results indicate that the activation of semantic information by the spoken target words guided the shift in visual attention. In addition, we did not find evidence that orthographic information mediates visual attention in the Chinese language, since semantic competitors attracted comparable levels of eye movements, irrespective of the orthographic similarity between the competitor and the target.

Additionally, we found a significant phonological competitor effect in Experiment 2. Given that orthographic overlap was avoided in the phonologically

³ It should be noted that the preview time in the long-preview condition was variable, and this variance may also have influenced the results. Therefore, only the short-preview condition was included in Experiment 2.

Table 5 Semantic relatedness scores between the spoken target words and the visual printed words in Experiment 2

	Semantic Competitor	Phonological Competitor	Distractors
S+O+ condition	4.49 (0.30)	1.25 (0.27)	1.20 (0.20)
S+O- condition	4.36 (0.58)	1.20 (0.19)	1.18 (0.19)

The numbers in parentheses are standard deviations

related condition, our observation of a phonological competitor effect indicates that access to phonological

information, rather than orthographic information, mediates visual attention shifts.

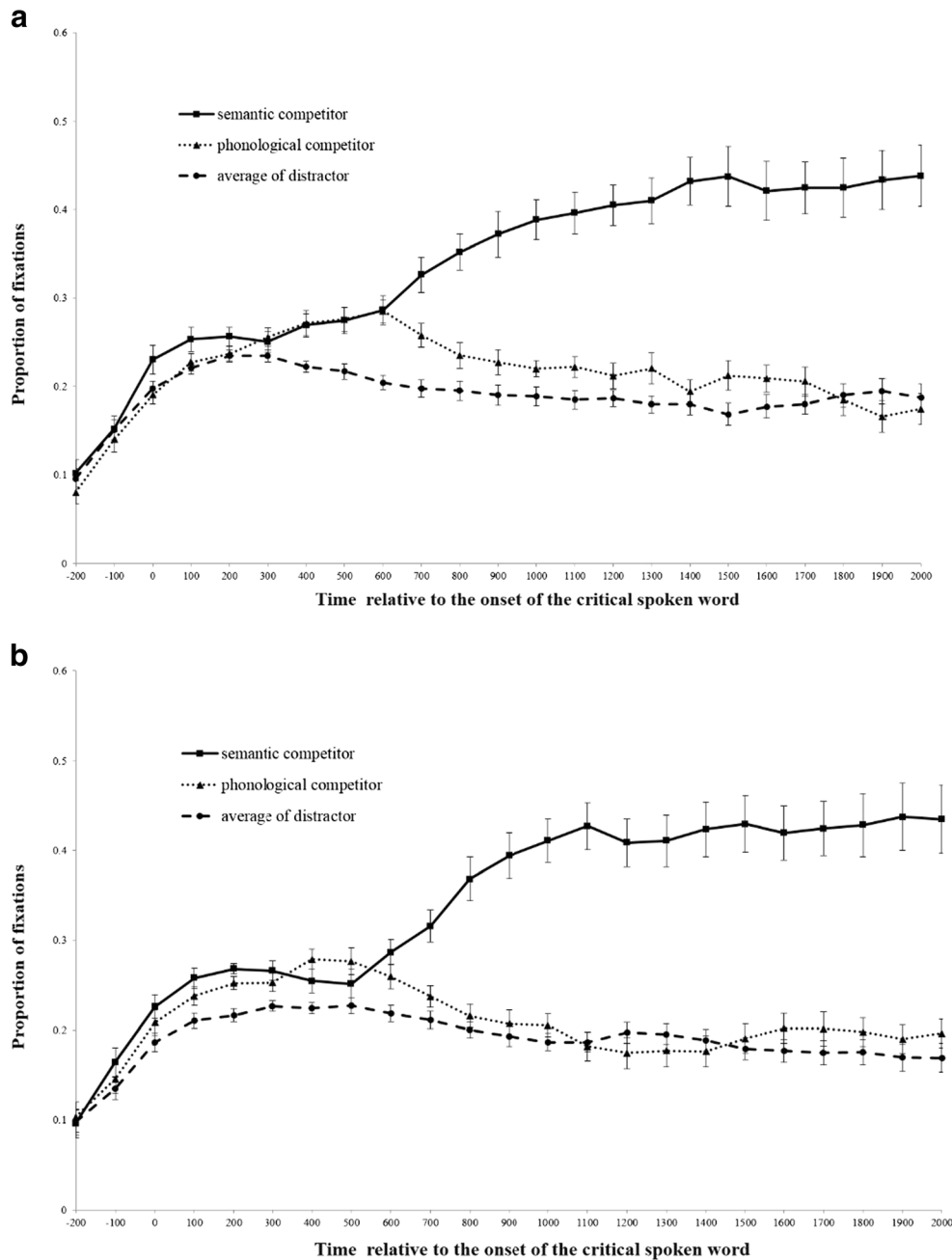


Fig. 3 Graph showing the fixation proportions to semantic competitors, phonological competitors, and the average of the distractors in the S+O+ condition (a) and the S+O- condition (b) in Experiment 2, from 200 ms

before onset of the spoken target word (0 on the x-axis indicates the onset of the spoken target words)

General discussion

In the present study, we investigated whether the semantic information activated in spoken word recognition can influence visual attention shifts to printed words. In the printed-word version of the visual-world paradigm, participants listened to a neutral context sentence including a spoken target word while looking at a printed-word display that included a semantic competitor, a phonological competitor, and two unrelated distractors. The results showed that the semantic competitor attracted more fixations than did the unrelated distractors. This semantic competitor effect was significant, irrespective of the preview time of the printed word display (Exp. 1) and of the orthographic similarity between the semantic competitors and spoken target words (Exp. 2). These findings suggest that semantic information activated in spoken word recognition can mediate visual attention shifts to related printed Chinese words. Moreover, our observation of significant phonological competitor effects in the long-preview condition in Experiments 1 and 2 suggests that phonological competitors attracted more fixations than did unrelated distractors. However, the effect may be less robust, since the phonological competitor effect was absent in the short-preview condition in Experiment 1.

The semantic competitor effect observed in the present study is inconsistent with a previous report by Huettig and McQueen (2007). They found that the semantic competitor did not attract more fixations than did unrelated distractor words. However, our study showed that semantic information was activated and directed eye fixations to a semantic competitor word when participants viewed a display of printed words. One possible reason for the discrepant results is the difference between Chinese and alphabetic languages.

As we previously noted, there is a weak spelling–meaning connection in alphabetic languages but a strong orthographic form–meaning in a nonalphabetical language. For alphabetic languages, it is likely that the semantic representation of the printed words is not activated, and hence no semantic representational match between a visual word and a spoken word is established as processing of the target word unfolds over time. Unlike in alphabetic languages, a close relationship between orthographic form and meaning exists in Chinese. Indeed, previous studies have shown that the semantic information of printed Chinese words can be retrieved more efficiently on the basis of a strong connection between orthographic forms (i.e., a semantic radical) and meaning (Leck, Weekes, & Chen, 1995; Ziegler, Benraïss, & Besson, 1999; Ziegler & Ferrand, 1998). Notably, semantic overlap between spoken word recognition and visual word recognition can develop immediately, which may guide visual attention shifts to the semantic competitor.

We previously noted the strong spelling–sound connections in alphabetic languages, which contrasted with the strong orthographic form–meaning connections in Chinese. However, we should note that the relationship between semantic processing in Chinese and in alphabetic languages may be more complex. Understanding how semantic processing differs between writing systems is beyond the scope of the present study. However, reading studies might provide some insight into this question.

Prior studies have shown that readers can retrieve and access semantic information parafoveally while reading alphabetic languages. This semantic preview effect was found in German (Hohenstein & Kliegl, 2013; Hohenstein, Laubrock, & Kliegl, 2010) and under certain conditions in English (Rayner & Schotter, 2014; Schotter, 2013; Schotter, Lee, Reiderman, & Rayner, 2015). However, most reading studies on English have not revealed evidence of semantic information being processed parafoveally (Altarriba, Kambe, Pollatsek, & Rayner, 2001; Rayner, Balota, & Pollatsek, 1986; Rayner, Schotter, & Drieghe, 2014). In contrast, the semantic preview effect is quite robust and has been found in a number of studies on Chinese (Yan, Richter, Shu, & Kliegl, 2009; Yan, Zhou, Shu, & Kliegl, 2012). These results suggest that semantic processing may occur differently during Chinese reading and the reading of alphabetic languages, and it is very likely that Chinese readers can perceive semantic information more easily because they receive more information through parafoveal vision during sentence reading. These findings are generally consistent with the notion that the Chinese language has stronger orthographic form–meaning connections.

As we discussed in the introduction, eye movements on objects in the visual-world paradigm can be driven by semantic similarity. The *semantic hypothesis* was proposed to explain this eye movement behavior by Huettig and Altmann (2005; see also Huettig & McQueen, 2007). According to this hypothesis, the fixation probability on a particular object reflects the semantic similarity between the conceptual representations accessed by the spoken word and those accessed from the visual objects. On the basis of our findings, we propose that the match established at the semantic level can also mediate visual attention to printed Chinese words.

The results of Experiment 2 clearly excluded the possibility that the observed semantic competitor effect was caused by orthographic similarity between the spoken target words and the semantic/phonology competitors. Previous studies using the printed version of visual-world paradigm showed that fixations on a competitor varied as a function of orthographic overlap (Salverda & Tanenhaus, 2010): Competitors with high orthographic overlap attracted more fixations than did those with low orthographic overlap. Since some of the target words and semantic/phonology competitors in Experiment 1 shared radicals in some stimuli, it was possible that the semantic/

phonology competitor effects observed in that experiment were caused by orthographic similarities. However, our findings in Experiment 2 do not support this claim. In Experiment 2, the semantic competitors and spoken target words were chosen such that they were either orthographically related or unrelated. Semantic competitor effects were found in both conditions, irrespective of orthographic similarity. This finding strongly suggests that the semantic competitor effects observed here are unlikely to have resulted from orthographic similarity.

Our findings contrast with previous reports of significant orthographic effects in alphabetic languages (Salverda & Tanenhaus, 2010). One possible reason for this is the difference in basic linguistic properties between alphabetic and non-alphabetic languages (e.g., Chinese): There is a strong spelling–sound connection in alphabetic languages, but not in non-alphabetic languages. Therefore, the relatively large dissociation between orthography and sound in nonalphabetic languages is likely to be the cause of the limited role of orthography in our semantic competitor effects.

Furthermore, we found that the fixation probability was higher on the phonological competitors than on distractors, consistent with previous findings in alphabetic languages (Huettig & McQueen, 2007; Weber et al., 2007). However, an ongoing debate concerns whether the so-called “phonological competitor effect” can be ascribed to phonological or even orthographic similarity, given the strong spelling–sound connection. McQueen and Viebahn (2007) found that Dutch participants exhibited higher fixation probabilities on phonological competitors than on distractors (see also Huettig & McQueen, 2007; Weber et al., 2007), and thus they proposed a *phonological hypothesis*, which posited that visual attention shifts were driven by phonological matches between the spoken and visual words. However, Salverda and Tanenhaus (2010) claimed that fixations on the word labeled as the “phonological” competitor were due purely to orthographic effects.

Separating phonological from orthographic effects is difficult (if not impossible) in alphabetic languages, because of the close connection between phonology and orthography. Using a nonalphabetic script to examine phonological effect on spoken word processing can provide a clearer answer, because a nonalphabetic script allows for a clean dissociation between orthographic and phonological codes. In Experiment 2, care was taken to avoid any orthographic overlap between the phonological competitors and targets. Thus, the observed phonological competitor effect must have been caused by phonological similarity. Our finding stands in contrast with the claim of Salverda and Tanenhaus (2010) that the “phonological” competitor effect is purely orthographic.

In our present study, the phonological competitor effect was found in the long-preview condition of Experiment 1 and in Experiment 2. However, the effect was absent in the

short-preview condition in Experiment 1, indicating that the phonological competitor effect is not as stable as the semantic competitor effect, which was robust in both experiments. Furthermore, as compared to the large phonological competitor effects found in alphabetic languages (the mean ratios⁴ of phonological competitors were .71 in Exp. 1 and .68 in Exp. 2 of Salverda & Tanenhaus, 2010), the phonological competitor effect observed in our present study was smaller (the mean ratios of the phonological competitors were .52 in both experiments). This indicates that the phonological effect observed in our present study is smaller and less robust. Thus, the results concerning the phonological competitor effect should be interpreted with caution, and studies should be conducted to further investigate this question.

In the present study, we employed a no-target version of the visual-world paradigm (Huettig & McQueen, 2007), in which the spoken target words were not displayed on the visual display. Interestingly, similar competitor effects have been found in the classic, with-target version of the visual-world paradigm (Dahan & Tanenhaus, 2005; Yee & Sedivy, 2006). Huettig and McQueen (2007, p. 477) suggested that the no-target version can “maximize the opportunity to observe competitor effects.” In addition, participants were asked only to listen to the spoken sentence and to view the visual display at the same time, and no other explicit task was required. As compared with the requirements of an explicit task in the classic visual-world paradigm, in which participants are asked to click on a specific referent, the no-task version prevents participants from using any explicit task-specific strategies. However, given that we did not directly compare the results from with-target and no-target versions of the visual-world paradigm in the present study, our findings cannot necessarily be generalized to studies that would include a target stimulus or require a target-selection task.

Furthermore, we found significant semantic competitor effects in both the long- and short-preview conditions in Experiment 1, which suggests that the preview time of the visual display did not influence the semantic competitor effect in the printed-word version of the visual-world paradigm. Given the close relationship between orthographic form and meaning in Chinese printed words, it is relatively easy to retrieve the semantic information of a character on the basis of its orthographic information (Zhou & Marslen-Wilson, 1996). Therefore, even with a limited preview time, skilled Chinese readers can be expected to be able to retrieve semantic information regarding printed words.

As far as methodology is concerned, although the classic visual-world paradigm has been widely used to investigate

⁴ That is, the ratio of the proportion of fixations to the phonological competitor and the sum of the fixation proportions to the phonological competitor and the distractor (for details, see Salverda & Tanenhaus, 2010).

semantic activation in spoken word recognition and its time course, presenting pictures as a visual reference limits the scope of the study, because a large proportion of stimuli cannot be depicted visually. Our study shows that the printed-word version of the visual-world paradigm is sensitive enough for investigating semantic processing, at least in Chinese spoken word recognition.

In conclusion, our findings indicate that both semantic information and phonological information can mediate eye movements to printed words during spoken word recognition in Chinese. In addition, our study shows that the printed-word version of the visual-world paradigm is suitable for

investigating semantic information processing during Chinese spoken word recognition

Author note This research was supported by grants from the Natural Science Foundation of China (Nos. 31571125 and 31400967) and the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (No. Y3CX132005). We thank Liz Schotter and two anonymous reviewers for their helpful comments on an earlier version of the manuscript.

Appendix 1

Table 6 Experimental stimuli used in Experiment 1

Preview Condition	Spoken Target Word	Semantic Competitor	Phonological Competitor	Distractor	Distractor
Long preview	菠萝	香蕉	伯父	石膏	渔民
Long preview	领带	西装	零件	女婿	诗文
Long preview	锤子	斧头	垂柳	毛驴	馅饼
Long preview	火车	轮船	伙计	作文	黄土
Long preview	椅子	桌子	仪器	中部	山地
Long preview	刀子	叉子	导游	虾米	皮筋
Long preview	嫔妃	侍女	贫农	公鸡	林场
Long preview	鼻子	耳朵	笔记	奖金	南非
Long preview	太阳	月亮	泰山	化肥	网球
Long preview	蜜蜂	花丛	秘方	烛台	蚕茧
Long preview	铅笔	橡皮	钱包	衬衣	灾民
Long preview	蓝天	白云	栏杆	车队	坐标
Long preview	钢琴	吉他	杠杆	利器	野猪
Long preview	鼠标	键盘	暑假	毫克	华山
Long preview	香烟	火柴	箱包	碎片	雪花
Long preview	生日	蛋糕	声波	军医	墓地
Long preview	拖把	笤帚	驼铃	桥洞	门齿
Long preview	警察	小偷	颈椎	食盐	臭氧
Long preview	签	酒精	绵羊	丘陵	精英
Long preview	老虎	狮子	劳力	房门	真空
Long preview	国王	王后	锅炉	面具	器械
Long preview	手枪	子弹	首相	茶叶	港口
Long preview	医生	护士	衣架	昆虫	学堂
Long preview	石窟	壁画	食堂	样品	租金
Long preview	信封	邮票	新人	院子	青海
Long preview	雪碧	可乐	雨季	电梯	作文
Long preview	黄豆	豆浆	蝗灾	挂历	酒馆
Long preview	胳膊	大腿	歌谣	汇票	油漆
Long preview	帽子	围巾	毛线	客栈	海带
Long preview	蝗虫	蚂蚱	皇室	分号	果脯
Long preview	墨水	钢笔	茉莉	浪花	镇长
Long preview	烧杯	试管	芍药	鸡汤	气孔
Long preview	士兵	军官	市民	背景	空调
Long preview	手表	闹钟	首富	密使	航标

Table 6 (continued)

Preview Condition	Spoken Target Word	Semantic Competitor	Phonological Competitor	Distractor	Distractor
Long preview	剪刀	针线	茧子	药水	酒会
Long preview	枕头	棉被	诊室	碎石	猛兽
Long preview	电话	铃声	淀粉	囚犯	柠檬
Long preview	苍蝇	蚊子	沧海	财主	鼓楼
Long preview	鲨鱼	鲸鱼	沙砾	胡琴	塔顶
Long preview	马戏	小丑	吗啡	晚宴	修女
Short preview	土豆	白菜	兔子	花朵	护照
Short preview	沙漠	骆驼	傻子	评委	口腔
Short preview	衣服	裤子	医师	步枪	手腕
Short preview	兔子	乌龟	土星	药房	会馆
Short preview	丝绸	麻布	私房	柠檬	饰物
Short preview	杨树	柳树	洋葱	道具	钢板
Short preview	小麦	水稻	肖像	电梯	地道
Short preview	腰果	杏仁	妖怪	火海	车牌
Short preview	手指	指甲	首长	师	简讯
Short preview	书包	课本	叔父	知了	村长
Short preview	桃子	李子	陶瓷	兄长	烟火
Short preview	绳索	铁链	声母	水箱	石狮
Short preview	音响	耳机	阴雨	梧桐	光圈
Short preview	监狱	牢房	尖刀	讲师	经络
Short preview	和尚	尼姑	荷花	银币	鼓声
Short preview	玉石	翡翠	浴室	高僧	单词
Short preview	戒指	项链	界线	瓶颈	老妇
Short preview	蝴蝶	毛虫	胡椒	烧饼	岳母
Short preview	黄金	白银	蝗虫	壮族	烟囱
Short preview	咖啡	牛奶	卡车	日出	福州
Short preview	饺子	汤圆	脚背	账目	门牌
Short preview	石油	煤炭	食物	屏幕	雷达
Short preview	城镇	村落	丞相	帆船	古董
Short preview	宇宙	黑洞	雨衣	郎中	血浆
Short preview	土壤	水分	图形	笑容	射线
Short preview	玫瑰	百合	煤油	帆船	匪徒
Short preview	心脏	大脑	新娘	农药	婚礼
Short preview	豆浆	油条	斗车	凉鞋	内海
Short preview	麻雀	乌鸦	马夫	套装	兄妹
Short preview	大海	小溪	答卷	药方	手心
Short preview	棺材	坟墓	官府	三轮	窗子
Short preview	帐篷	睡袋	账单	果肉	青砖
Short preview	教堂	牧师	轿子	豆腐	助手
Short preview	氧气	氢气	养女	床垫	刀枪
Short preview	面条	米饭	棉布	歌谣	展品
Short preview	花生	瓜子	画卷	匣子	铁丝
Short preview	火炉	木炭	活塞	笛子	吗啡
Short preview	老鹰	大雁	姥姥	鞋匠	手球
Short preview	漫画	卡通	满族	山林	鱼雷
Short preview	荔枝	龙眼	砾石	木桶	帖子

Appendix 2

Table 7 Experimental stimuli used in Experiment 2

Orthographic Similarity	Spoken Target Word	Semantic Competitor	Phonological Competitor	Distractor	Distractor
S+O−	牙齿	舌头	衙门	战友	花卉
S+O−	医生	护士	衣料	牧民	前辈
S+O−	铅笔	橡皮	签证	衬衣	灾民
S+O−	鼠标	键盘	暑假	中秋	华山
S+O−	生日	蛋糕	声波	总监	中尉
S+O−	拖把	笤帚	脱臼	毛裤	分号
S+O−	酒精	棉签	九族	铁矿	童话
S+O−	狮子	老虎	诗词	仓库	椭圆
S+O−	手枪	子弹	首相	茶叶	餐厅
S+O−	石窟	壁画	食堂	报表	老将
S+O−	邮票	信封	游船	寺院	药材
S+O−	可乐	雪碧	渴慕	作文	天线
S+O−	钢笔	墨水	缸瓦	镇长	红利
S+O−	烧杯	试管	稍许	土炕	初赛
S+O−	士兵	军官	事故	专题	导演
S+O−	手表	闹钟	首席	茶碗	技工
S+O−	剪刀	针线	茧子	海棠	鱼干
S+O−	小丑	马戏	晓市	山茶	翠菊
S+O−	沙漠	骆驼	刹车	牌子	插图
S+O−	乌龟	兔子	巫术	贫民	胃炎
S+O−	番茄	鸡蛋	帆船	插图	参数
S+O−	腰果	杏仁	邀功	车牌	二胡
S+O−	铁链	绳索	帖子	蜻蜓	茄子
S+O−	和尚	寺庙	核弹	果汁	鞭子
S+O−	玉石	翡翠	浴场	血脂	花椒
S+O−	戒指	项链	界石	插座	禅师
S+O−	饺子	汤圆	角质	宋词	茅屋
S+O−	宇宙	黑洞	羽毛	假日	职位
S+O−	阳光	土壤	羊毛	文献	礼仪
S+O−	玫瑰	百合	煤炭	匪徒	帆船
S+O−	羚羊	斑马	灵堂	请帖	黑道
S+O−	乌鸦	麻雀	污垢	风琴	字迹
S+O−	棺材	坟墓	关口	椰子	盘子
S+O−	睡袋	帐篷	税种	头饰	苏打
S+O−	公主	王子	功率	歌谣	招标
S+O−	荔枝	龙眼	砾石	猎犬	听筒
S+O−	书包	课本	叔父	知了	甘蔗
S+O−	钢琴	吉他	缸瓦	总兵	唐诗
S+O−	石油	煤炭	食物	屏幕	条规
S+O−	地震	海啸	弟妹	晨练	粮库
S+O+	菠萝	香蕉	播音	石膏	渔民
S+O+	珍珠	玛瑙	真菌	裁缝	画册
S+O+	椅子	桌子	以往	笔记	天线
S+O+	蜂蜜	蜜蜂	风筝	赤道	天皇
S+O+	豆浆	油条	斗牛	金表	春雨
S+O+	胳膊	大腿	歌曲	诗篇	油漆

Table 7 (continued)

Orthographic Similarity	Spoken Target Word	Semantic Competitor	Phonological Competitor	Distractor	Distractor
S+O+	帽子	围巾	茂密	碎石	海带
S+O+	蚊子	苍蝇	闻人	财主	鼓楼
S+O+	鲸鱼	鲨鱼	精兵	胡琴	塔顶
S+O+	蝗虫	蚂蚱	黄疸	兄妹	果脯
S+O+	手掌	拳头	首领	行星	高山
S+O+	夜晚	白天	叶子	套房	武侠
S+O+	沙滩	海岸	杀手	学报	总部
S+O+	古筝	笛子	骨骼	票房	大雪
S+O+	衣服	裤子	医师	步枪	手腕
S+O+	心脏	大脑	新娘	家财	原告
S+O+	杨树	柳树	阳伞	道具	药方
S+O+	瘫痪	瘸子	贪官	星座	黑豆
S+O+	桃子	李子	陶器	兄长	烟火
S+O+	香肠	火腿	箱包	插图	古琴
S+O+	蝴蝶	毛虫	弧度	古庙	方框
S+O+	香蕉	芒果	乡亲	药方	谎话
S+O+	木炭	烧烤	牧人	笛子	吗啡
S+O+	织布	纺纱	脂粉	囚犯	错案
S+O+	江河	湖泊	浆果	梅花	肖像
S+O+	杏仁	核桃	姓氏	战车	角膜
S+O+	蝌蚪	青蛙	磕头	山洞	算盘
S+O+	喜鹊	乌鸦	洗印	财神	衣料
S+O+	黄金	白银	蝗虫	壮族	烟卤
S+O+	氧气	氢气	养女	床垫	刀枪
S+O+	丝绸	纱布	私房	柠檬	鬼怪
S+O+	馒头	馅饼	蛮荒	财团	太爷
S+O+	沼泽	湿地	爪牙	舞会	军衔
S+O+	肝脏	肠胃	甘油	气压	西湖
S+O+	眼睛	眉毛	演技	口供	边陲
S+O+	茉莉	芍药	墨斗	短跑	油井
S+O+	蚯蚓	蚂蚁	秋天	针灸	知府
S+O+	岩石	山脉	颜料	胡椒	尾气
S+O+	绿叶	红花	律条	大肠	阿姨
S+O+	冰雹	雷雨	兵车	鞋匠	土路

References

- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439. doi:10.1006/jmla.1997.2558
- Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish–English bilinguals. *Perception & Psychophysics*, 63, 875–890. doi:10.3758/BF03194444
- Altmann, G. T., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). New York: Psychology Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. doi:10.1016/j.jml.2012.11.001
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7) [Computer software]*. Retrieved from <http://cran.r-project.org/package=lme4>
- Chen, H.-C., Flores d'Arcais, G. B., & Cheung, S. L. (1995). Orthographic and phonological activation in recognizing Chinese characters. *Psychological Research*, 58, 144–153. doi:10.1007/BF00571102

- Chen, H.-C., & Shu, H. (2001). Lexical activation during the recognition of Chinese characters: Evidence against early phonological activation. *Psychonomic Bulletin & Review*, 8, 511–518. doi:10.3758/BF03196186
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6, 84–107. doi:10.1016/0010-0285(74)90005-X
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, 16, 507–534. doi:10.1080/01690960143000074
- Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, 12, 453–459. doi:10.3758/BF03193787
- Development Core Team, R. (2014). *R: A language and environment for statistical computing (Version 3.1.1)*. Vienna: R Foundation for Statistical Computing. Retrieved from www.R-project.org
- Fan, K. Y., Gao, J. Y., & Ao, X. P. (1984). Pronunciation principles of the Chinese character and alphabetic writing scripts. *Chinese Character Reform*, 3, 23–27.
- Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, 69, 165–182. doi:10.1016/j.jml.2013.06.001
- Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.
- Hohenstein, S., & Kliegl, R. (2013). Eye movements reveal interplay between noun class and word capitalization in reading. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2554–2559). Austin: Cognitive Science Society.
- Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1150–1170. doi:10.1037/a0020233
- Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Hillsdale: Erlbaum.
- Hoosain, R., & Osgood, C. E. (1983). Processing times for English and Chinese words. *Perception & Psychophysics*, 34, 573–577. doi:10.3758/BF03205913
- Huetting, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, 96, B23–B32. doi:10.1016/j.cognition.2004.10.003
- Huetting, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57, 460–482. doi:10.1016/j.jml.2007.02.001
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformed or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446. doi:10.1016/j.jml.2007.11.007
- Leck, K. J., Weekes, B. S., & Chen, M. J. (1995). Visual and phonological pathways to the lexicon: Evidence from Chinese readers. *Memory & Cognition*, 23, 468–476. doi:10.3758/BF03197248
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1–86. doi:10.1016/0010-0285(86)90015-0
- McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, 60, 661–671. doi:10.1080/17470210601183890
- Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling–sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 183–206. doi:10.1037/0278-7393.28.1.183
- Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 335–359). Hillsdale: Erlbaum.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. doi:10.1037/0033-2909.124.3.372
- Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology*, 40, 473–483. doi:10.1037/h0080111
- Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 1617–1628. doi:10.1037/a0036763
- Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, 21, 1067–1072. doi:10.3758/s13423-014-0582-9
- Salverda, A. P., & Tanenhaus, M. K. (2010). Tracking the time course of orthographic information in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1108–1117. doi:10.1037/a0019901
- Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, 69, 619–633. doi:10.1016/j.jml.2013.09.002
- Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language*, 83, 118–139. doi:10.1016/j.jml.2015.04.005
- Shen, J., Deutsch, D., & Rayner, K. (2013). On-line perception of Mandarin Tones 2 and 3: Evidence from eye movements. *Journal of the Acoustical Society of America*, 133, 3016–3029. doi:10.1121/1.4795775
- Slowiczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *The Quarterly Journal of Experimental Psychology: Section A*, 56, 233–262. doi:10.1080/02724980244000323
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634. doi:10.1126/science.7777863
- Tsang, Y. K., & Chen, H. C. (2010). Morphemic ambiguity resolution in Chinese: Activation of the subordinate meaning with a prior dominant-biased context. *Psychonomic Bulletin & Review*, 17, 875–881. doi:10.3758/PBR.17.6.875
- Weber, A., Melinger, A., & Tapia, L. (2007). The mapping of phonetic information to lexical representation in Spanish: Evidence from eye movements. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)* (pp. 1941–1944). Dudweiler: Pirrot.
- Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review*, 16, 561–566. doi:10.3758/PBR.16.3.561
- Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition*, 38, 1069–1075. doi:[10.1037/a0026935](https://doi.org/10.1037/a0026935)
- Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1–14. doi:[10.1037/0278-7393.32.1.1](https://doi.org/10.1037/0278-7393.32.1.1)
- Zhou, X., & Marslen-Wilson, W. (1996). Direct visual access is the only way to access the Chinese mental lexicon. In G. Cottrell (Ed.), *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 714–719). Hillsdale: Erlbaum.
- Zhou, X., Ye, Z., Cheung, H., & Chen, H. C. (2009). Processing the Chinese language: An introduction. *Language and Cognitive Processes*, 24, 929–946. doi:[10.1080/01690960903201281](https://doi.org/10.1080/01690960903201281)
- Zhu, X. (1988). Analysis of cueing function of phonetic components in modern Chinese. In *Proceedings of the Symposium on the Chinese Language and Characters* (pp. 85–99). Beijing: Guang Ming Daily Press.
- Ziegler, J. C., Benraïss, A., & Besson, M. (1999). From print to meaning: An electrophysiological investigation of the role of phonology in accessing word meaning. *Psychophysiology*, 36, 775–785. doi:[10.1111/1469-8986.3660775](https://doi.org/10.1111/1469-8986.3660775)
- Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, 5, 683–689. doi:[10.3758/BF03208845](https://doi.org/10.3758/BF03208845)