

Running head: SEMANTIC PREVIEW BENEFITS IN READING CHINESE

This is a preprint of an article whose final and definitive form was published in the Journal of Research in Reading

Article first published online: 20 MAR 2013

DOI: 10.1111/j.1467-9817.2013.01556.x

Copyright © 2013 UKLA John Wiley & Sons, Inc.

This article may not exactly replicate the final version published in the journal. It is not the copy of record.

**A Validation of Parafoveal Semantic Information Extraction
in Reading Chinese**

Wei Zhou, Reinhold Kliegl, & Ming Yan


ABSTRACT

Parafoveal semantic processing has recently been well-documented in reading Chinese sentences, presumably because of language-specific features. However, due to a large variation of fixation landing positions on pre-target words, some preview words actually were located in foveal vision when readers' eyes landed close to the end of the pre-target words. None of the previous studies have completely ruled out a possibility that the semantic preview effects might mainly arise from these foveally processed preview words. This case, whether previously observed positive evidence for parafoveal semantic processing can still hold, has been called into question. Using linear mixed models, we demonstrate in this study that semantic preview benefit from word N+1 decreased if fixation on pretarget word N was close to the preview. We argue that parafoveal semantic processing is not a consequence of foveally processed preview words.

Keywords: semantic, preview benefit, reading, Chinese

Reading both alphabetic and logographic scripts requires effective extraction of lexical information from the fixated foveal words as well as from the upcoming parafoveal words. In the boundary paradigm (Rayner, 1975), where the type of information from the to-be-fixated target word N+1 is under experimental control during fixations on and prior to the pretarget word N, preview words are replaced by the target word once the eye crosses an invisible boundary located between words N and N+1. Preview benefit (PB) is indicated by a positive difference between fixation durations on target word N+1 when preview is masked over when it is visible. In the past decades, there has been little evidence for the integration of semantic or morpho-semantic information across saccades in alphabetic scripts, especially in English (Altarriba, Kambe, Pollatsek, & Rayner, 2001; Bertram & Hyönä, 2007; Inhoff, 1989; Kambe, 2004; Juhasz, White, Liversedge, & Rayner, 2008; Rayner, Balota, & Pollatsek, 1986; Rayner, White, Kambe, Miller & Liversedge, 2003; but see Hohenstein, Laubrock, & Kliegl, 2010, for positive evidence in German). On the other hand, presumably due to language-specific features which we will elaborate on below, the semantic and morpho-semantic preview effect is well-established in Chinese (Tsai, Kliegl, & Yan, 2012; Yan, Richter, Shu, & Kliegl, 2009; Yan, Risse, Zhou, & Kliegl, 2012a; Yan, Zhou, Shu, & Kliegl, 2012b; Yang, Wang, Tong, & Rayner, 2012; Yen, Tsai, Tzeng, & Hung, 2008). Here we present a study in which we explored the relationship between parafoveal semantic preview benefit and fixation landing position on pretarget words (i.e., preview space) and provide evidence that semantic preprocessing is unlikely to be a consequence of foveally processed preview words.

Relevant Features of the Chinese Writing System and Their Implications for Parafoveal Processing

Chinese is an important script, not only because it is used by the largest number of people, but also because of its fundamental differences from alphabetic scripts. Chinese has been viewed as a logographic writing system with its basic writing units, characters, written in a series of square-shaped objects of identical horizontal size, irrespective of their visual complexity. Characters are formed according to a variety of principles (Feng, Miller, Shu, & Zhang, 2001). The most ancient characters are pictographs and ideograms, which originated from drawings by ancient cave dwellers, or formed through analogy or association. These characters are usually simple in visual make-up and highly frequent because they represent very common concepts. They are henceforth referred to as *simple characters*. As one can imagine, meanings of these two types of characters are easy to obtain, even to people who have little knowledge of Chinese. For example, a modern character “羊”, *sheep*, is developed from an ancient character (i.e., oracle-bone script)  , which resembles the head of a sheep.

On the other hand, most modern Chinese characters are phonograms (compound characters) consisting of two components/radicals, one of which represents the meaning of the whole character, while the other provides a rough clue as to its pronunciation. Nevertheless, as a result of the development of the Chinese writing system over centuries, there are characters whose meanings are unrelated, or only indirectly related to the meanings of their radicals (Shu & Anderson, 1997). Therefore,

Yan et al. (2012b) defined the term *radical semantic transparency* to refer to the degree to which the constituent radicals are semantically related to the meaning of the whole character. Characters such as “骗”, *to cheat*, with a semantic radical “马”, *horse*, are semantically *opaque* from their radicals. Whereas another character with the same semantic radical, “骡”, is *transparent* from its semantic radical because the whole character means *mule*, which is biologically related to horse. Indeed, there is strong evidence indicating that semantics are directly accessible from orthography in Chinese, bypassing phonological mediation under some circumstances (Chen & Shu, 2001; Zhou & Marslen-Wilson, 2000). Against this background of research, we suspect that the well-known sequential activation “orthography-to-phonology-to-semantics” route to word recognition (Van Orden, 1987) with access to word meaning in a relatively late stage might be more dominant in English, but may not be the same in Chinese because its script is known to be optimized for fast and direct semantic extraction (Frost, 2012; Laubrock & Hohenstein, 2012).

Chinese is particularly well suited for demonstrating semantic parafoveal processing not only because of its characteristically strong connection between orthography and semantics, as mentioned above (see also Hoosain, 1991, and Tsang & Chen, 2012, for reviews), but also due to other important factors. For example, most Chinese words are only one or two characters in length, and a Chinese character typically occupies the space of three letters in alphabetic languages (Tsai & McConkie, 2003). As a consequence, word N+1 is on average closer to the fixation

point on word N in Chinese than in alphabetic languages. In addition, the absence of inter-word spaces brings the upcoming word into a much less eccentric position, compared to alphabetic scripts, a fact which may enable Chinese readers to benefit from the higher visual acuity in the near parafovea and further facilitates parafoveal processing. Finally, readers of Chinese are assumed to parafoveally segment characters into word units and target towards word centers (Yan, Kliegl, Richter, Nuthmann, & Shu, 2010; see also Yang, Wang, Xu, & Rayner, 2009, for alternative evidence for parafoveal word segmentation obtained from the boundary paradigm). As a consequence, readers of Chinese may have developed a higher parafoveal processing efficiency.

Recently, reliable semantic preview benefits have been consistently reported in reading Chinese sentences. Yan et al. (2009) used neutral context and simple characters as previews to demonstrate a proof of principle. The semantic preview effect has been extended to traditional Chinese script, which is primarily used in Taiwan and Hong Kong (Tsai et al., 2012), and compound characters (Yan et al., 2012b). Yan et al. (2012b) demonstrated semantic preview benefits from bottom-up processing without top-down contextual support. In addition, using preview characters that were opaque from their semantic radicals as the reference condition, Yan et al. (2012b) found that fixations on target words were significantly shorter in the transparent preview condition. The result suggested semantic preview benefit may go beyond the lexical level and semantic information may also be obtained from semantic radicals of the parafoveal characters.

Landing Position Issues

The reliability of the parafoveal-on-foveal effect (i.e., whether properties of the word beyond the one currently fixated can influence fixation durations on the current word) has been called into question because they were sometimes restricted to fixations close to word N+1 and may therefore be a consequence of saccadic undershoots with intended landing position on word N+1 (Drieghe, Rayner, & Pollatsek, 2008; but see Kennedy & Pynte, 2005). An alternative explanation of the parafoveal-on-foveal effect being restricted to fixations close to the end of word N might be that at least parts of words N+1 are processed in foveal vision. Due to a large variation of fixation landing positions on words (i.e., the preferred viewing location curve), a certain proportion of preview words may actually be located in the foveal vision of the readers when their eyes landed close to the end of the pre-target words (Drieghe et al., 2008). It is reasonable to question the assumption that semantic preview effects might benefit more from these possibly foveally processed preview words and whether previously observed positive evidence for parafoveal semantic processing can still hold. This potentially confounding factor has not been controlled in any previous study on preview effects in reading Chinese sentences.

One way to avoid the confounding factor was reported by Inhoff, Radach, Starr and Greenberg (2000). Inhoff et al. (2000) found a parafoveal-on-foveal effect due to semantic association of words N and N+1: fixation duration on word N is shorter if word N and N+1 are semantically associated. Furthermore, in order to test whether their parafoveal-on-foveal effect came from fixations close to word N+1, they

portioned all trials into two approximately equally-sized groups, a ‘potential error’ group comprised of fixations on the four ending letter spaces of the target, and a likely ‘non-error’ group with fixations on all other letters. If the parafoveal-on-foveal effect was due to fixations close to word N+1, then it should be evident only in the first group, but not in the second. Results from their additional analysis with a factor of fixation location (potential error vs. non-error) did not confirm this possibility.

The Present Study

In the present study, we adopt a similar logic from Inhoff et al. (2000) and reanalyze data reported in Yan et al. (2009) and Yan et al. (2012b) using a linear mixed model (LMM): we specify fixation landing position on pretarget word N as a covariate instead of using it as a factor. The advantage of using a continuous covariate over a dichotomized factor is that the former is known to have stronger statistical power (e.g., Baayen, 2008). Given that readers have faster processing speed in foveal vision as compared to parafoveal vision, the consequence of possibly foveal-processed trials may be twofold: on the one hand, it may facilitate semantic extraction from the preview; on the other hand, activation of the preview word can lead to a competition between the preview word and target word and therefore result in a “preview cost” effect (Yan et al., 2012a). Although the direction of the influence due to these possibly foveal-processed trials is difficult to predict, it translates into a significant interaction between semantic preview benefit and landing position on pretarget words in either case: preview benefits should increase or decrease with preview space. In addition, if parafoveal semantic information can be obtained

independent of landing position on pretarget words, we should observe reliable semantic preview benefits even if those trials in which preview words were foveally processed are excluded from analyses. However, the consequence of removing trials is twofold: it also greatly reduces the number of valid observations, which may lead to decreased statistical power.

METHOD

Subjects

All participants of the two data sets were graduate and undergraduate students from the Beijing Normal University with normal or corrected-to-normal vision, and native speakers of Chinese. Forty-eight participants were tested for Data Set 1, in which we used simple Chinese characters as previews and reported in Yan et al. (2009). An independent sample of a total of 62 participants contributed to Data Set 2 with a manipulation of semantic radical transparency (Yan et al., 2012b).

Material

Forty-eight simple, non-compound characters were selected as targets, which served as the first characters of target words at word N+1 position (i.e., the first word immediately after the boundary for display change) in Data Set 1. Each target character was embedded into a two-character target word. For each target character, five types of preview characters were selected for identical, orthographically related, phonologically related, semantically related, and unrelated preview conditions. In the present paper, we focus on the semantic preview condition.

For Data Set 2, 70 two-character target words with their first characters as target

characters were selected for the preview-type manipulation at word N+1 position. For each target character, four types of preview characters were chosen, including identical and unrelated previews serving as two baseline conditions. Critically, we also had two types of semantically related preview conditions: characters that were either semantically transparent, or opaque from their semantic radicals. To avoid possible orthographic and phonological effects, the three non-identical previews did not share any radicals or syllables with the target.

For both data sets, word-level preview was valid only for the identical condition, whereas the preview characters of the non-identical conditions did not form words with the characters that followed them. As shown in Table 1, the previews were closely matched with respect to visual complexity as indexed by number of strokes, character frequency, as well as predictability from prior contexts. The target characters never appeared among the first three or the last three words.

-- TABLE 1--

The invisible boundary that triggered the display change was located just to the left of the space before word N+1. In order to insure most previews could be processed in parafoveal vision, and given that Chinese readers tend to fixate at a position slightly left to the word centers (Yan et al., 2010), all pretarget words N in both data sets were always two-character words. A sample of sentences is shown in Figure 1. Full details concerning the material are available in Yan et al. (2009) and Yan et al. (2012).

-- FIGURE 1--

Apparatus

Eye movements were recorded with an EyeLink II system (500 Hz). Single sentences were presented on the vertical position one third from the top of the screen of a 19-inch ViewSonic G90f monitor (resolution, 1024 by 768 pixels; frame rate, 100 Hz). Therefore, it took at most 16ms to complete the display change. All recordings and calibrations were binocular. Only 3 subjects reported flashes for only a few trials in Data Set 1 and 18 subjects reported flashes ($M=5$, $SD=4$) in Data Set 2, but none could report exactly what they had seen.

Procedure

Subjects were calibrated with a nine-point grid for both eyes. They were instructed to read the sentences for comprehension, then fixate a dot in the lower right corner of the monitor, and finally press a button to signal completion of the trial. As shown in Figure 1, before readers' eyes cross the invisible boundary located between word N and word $N+1$, they get any one of the four previews at the position of word $N+1$. The preview word is replaced by the target word immediately after the eyes cross this boundary. Fixation on the fixation point initiated presentation of the next sentence. An extra calibration occurred if the tracker did not detect both eyes within a pre-defined window around the initial fixation point.

Data Analysis

Fixations were determined with an algorithm for binocular saccade detection (Engbert & Kliegl, 2003). Analyses were based on right-eye fixations. We distinguish between first-fixation durations (FFDs; the first fixation on a word, irrespective of the

number of fixations), single-fixation durations (SFDs; cases in which a word was inspected with exactly one fixation), and gaze durations (GDs; the sum of fixations during the first reading of the word).

There were four levels of data screening: First, sentences containing a blink or loss of measurement were deleted (18% in Data Set 1 and 7% in Data Set 2). Second, FFDs and SFDs as well as GDs with FFDs shorter than 60 ms or longer than 600 ms were excluded (1% in Data Set 1 and 1% in Data Set 2). Third, trials with regressions from word N or N+1 were excluded (9% in Data Set 1 and 13% in Data Set 2) because they may reflect incomplete parafoveal processing of preview words during fixations on pre-target words N or incomplete foveal processing of target words N+1. Finally, we excluded the trials in which the saccade crossed the boundary during the final 20% of the saccade duration because readers should be more likely to perceive a display change or a flash at this time (15% in Data Set 2)¹.

Inferential statistics are based on planned comparisons for the related and the identical previews with the unrelated preview as a reference condition. Estimates are based on LMMs for durations with crossed random effects for subjects and items using the *lmer* program of the *lme4* package (Bates & Maechler, 2010) in the R environment for statistical computing and graphics (R-Core Development Team, 2010). For each data set, the influence of landing position on pretarget word N was tested in two steps: First, we specify LMMs using fixation landing position and preview duration on pretarget words as covariates and report interactions between semantic preview benefit and these two covariates. Second, we report semantic

preview benefits based on trials in which pretarget words were single-fixated and previews were located beyond 1 degree of visual angle from fixations on pretarget words. Thus we restricted our analyses to a “cleaner” set of parafoveal processing trials. We report log-transformed dependent variables of fixation times in the models because analyses of residuals and inspection of duration distributions strongly suggested that log-transformation is required to meet LMM assumptions (Kliegl, Masson, & Richter, 2010). Analyses for untransformed and log-transformed durations yielded the same pattern of significance.

RESULTS

Data Set 1

A total of 1285, 944, and 1285 observations on the target words contributed to the following FFD, SFD and GD analyses, respectively². In LMMs with covariates, none of the two-way or three-way interactions were reliable in any of the three duration measurements (all $\text{abs}[t\text{-values}] < 1.65$). The failure to find any reliable interaction could be due to the small number of observations in Data Set 1.

In the second step of analysis, we removed trials in which preview words were possibly foveally processed (i.e., a trial is removed as long as the distance between the fixation landing position and the left border of the preview is less than one degree) and a total of 868, 606, and 868 observations remained for FFD, SFD and GD analyses. Relative to unrelated previews, there were significant semantic preview benefits of 21 ms for FFDs and 21 ms for SFDs, with a similar numeric trend of 27 ms for GDs ($b = -.074$, $SE = .029$, $t = -2.56$; $b = -.079$, $SE = .037$, $t = -2.14$ and $b =$

-.066, $SE = .040$, $t = -1.64$, for FFD, SFD and GD analyses, respectively).

-- TABLE 1--

Data Set 2

There were 4719, 3564, and 4719 observations on the target words for FFD, SFD and GD analyses, respectively. In LMMs with the covariates, there was an interesting three-way interaction among semantic preview benefit from transparent preview characters, preview duration and preview space as illustrated in Figure 2 ($b = -.207$, $SE = .080$, $t = -2.60$). This three-way interaction indicates that semantic preview benefit is most pronounced when readers' fixations on pretarget words were far from the previews (i.e., on the first characters of pretarget words) and when they parafoveally processed the previews briefly.

-- FIGURE 2--

Removing trials in which preview words were possibly foveally processed led to 3100, 2222, and 3100 observations from Data Set 2. As expected, semantic preview effects from both transparent and opaque previews were still highly significant (in order of FFD, SFD and GD analyses, for transparent previews: $b = -.056$, $SE = .013$, $t = -4.27$; $b = -.065$, $SE = .016$, $t = -4.13$ and $b = -.076$, $SE = .018$, $t = -4.23$ and for opaque previews: $b = -.045$, $SE = .013$, $t = -3.47$; $b = -.057$, $SE = .016$, $t = -3.60$ and $b = -.042$, $SE = .018$, $t = -2.34$).

In a planned comparison with the opaque condition as the reference category, fixations were estimated to be 13 ms shorter in the transparent condition for GD analysis. This effect was reported as significant in Yan et al. (2012b) and was taken as evidence indicating readers of Chinese could also acquire semantic information from

radicals of parafoveal characters. This sublexical semantic effect was still marginally reliable when we removed trials in which fixations on words N were close to words N+1 from the data set ($b = -.034$, $SE = .018$, $t = -1.90$). We suspect this drop of significance was simply due to greatly reduced statistical power.

DISCUSSION

Although evidence is mixed for parafoveal semantic processing during sentence reading in alphabetic scripts, presumably due to substantial cross-language differences, such effects have been consistently reported in Chinese reading. The main goal of the present study was to rule out a possibility that the semantic preview effects might arise from foveally processed preview words. This potentially confounding factor has not been tested in any of the previous studies on preview benefit during the reading of Chinese. Results from the present study clearly suggested that parafoveal semantic preview benefit is not a consequence of foveally processed preview words. Suffering from loss of statistical power due to removal of foveally processed preview words, semantic preview benefits from both simple and compound preview characters still remained significant. These results certainly validate previously reported parafoveal semantic preview benefits in Chinese and are obviously compatible with the assumption of parallel distributed processing.

Using data from the standard foveal priming paradigm, Zhou, Marslen-Wilson, Taft, and Shu (1999) argued that when the prime is presented briefly, semantic representation of the target word which shares many properties with the prime word is pre-activated and there is little competition between semantic activation of the prime

and target, leading to facilitatory effects in lexical decision. If the prime is presented for a relatively long time, however, non-overlapping semantic representations start to develop and result in competition between prime and target, leading to inhibitory effects. Similarly, experiments using the boundary paradigm, in which the parafoveal preview is manipulated, the semantic and morpho-semantic preview effects have been demonstrated to interact with preview duration. Our explanation (Yan et al., 2012a) that semantic preview benefit from word N+1 is mainly due to short previews, whereas long preview duration leads to a semantic preview cost effect follows the one of Zhou et al. (1999): accumulation of information over a long preview duration may have activated those aspects of non-identical semantically related previews that cancel the preview benefit, presumably even turning a preview benefit into a preview cost. Yen et al. (2008) reported a finding similar to Yan et al. (2012a) in their second experiment: when the prior fixation duration was short (i.e., less than 220 ms), the morpho-semantic preview benefit amounted to 19 ms. This effect was reduced to 4 ms when the prior fixation duration was long.

Cui, Yan, Bai, Hyönä, Wang, and Liversedge (2012) manipulated the preview types of the second constituent characters of two-character Chinese compound words during fixations on the first constituent characters and reported semantic preview benefit in gaze duration when the first characters were of low-frequency category. Their results are basically in agreement with an interaction between semantic preview benefit and preview time (i.e., the fixation duration on the pretarget word), as reported by Yan et al. (2012a): long preview time, strong context and highly frequent pretarget

characters may all lead to more information extraction from the parafovea and accumulated parafoveal semantic information eventually disrupted the processing of the target word.

Kliegl, Hohenstein, Yan, and McDonald (2013) demonstrated a similar effect of preview space as preview time: identical preview benefit increased with both preview time and preview space. This is because, due to a larger preview space associated with fixations close to the target word, more parts of the preview words fall into the perceptual span and readers may acquire more parafoveal information. Therefore, it is reasonable to expect an interaction between preview space and semantic preview benefit. However, such a hypothesis is not supported by Data Set 1 and it turns into an opposite direction in Data Set 2. Fortunately, this seeming contradiction is not difficult to reconcile with our current understanding: Kliegl et al. (2013) tested identical preview from which only correct information is accumulated, whereas in the boundary paradigm using non-identical previews, during fixations on pretarget words, readers accumulate both the converging information between previews and targets that facilitates target word identification, as well as the diverging information that may interfere with target word identification. Here we would like to propose that the modulation effect of preview space on semantic preview benefit works similar as preview duration: when readers' fixations on pretarget words are very close to the previews, they may also easily activate diverging aspects from semantically related previews and cancel the preview benefit. In the end, the issue whether preview duration and preview space yield benefit or cost for processing a target word may

largely be a matter of when critical beneficial and disruptive aspects of preview words become available.

REFERENCES

- 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
- Baayen, R. H. (2008). *Practical data analysis for the language sciences with R*. Cambridge, MA: Cambridge University Press.
- Bates, D. & Maechler, (2010). *lme4: Linear mixed-effects models using S4 classes (R package version 0.999375-32) [Computer software]*. Vienna, Austria: R Foundation for Statistical Computing.
- Beijing Language Institute Publisher (1986). *Modern Chinese Word Frequency Dictionary* (in Chinese). Beijing: Beijing Language Institute Publisher.
- Bertram, R., & Hyönä, J. (2007). The interplay between parafoveal preview and morphological processing in reading. In R. vanGompel, M. Fischer, W. Murray, & R. Hill (Eds), *Eye Movements: A Window on Mind and Brain* (pp. 391-407). New York, NY: Elsevier.
- Chen, H.-C., & Shu, H. (2001). Lexical activation during the recognition of Chinese characters: evidence against early phonological activation. *Psychonomic Bulletin & Review*, 8 (3), 511-518. doi: 10.3758/BF03196186
- Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2012). Processing of compound-word characters in reading Chinese: An eye-movement-contingent display change study. *Quarterly Journal of Experimental Psychology*. Advance

online publication. doi: 10.1080/17470218.2012.667423

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology*, 61, 1239-1249. doi: 10.1080/17470210701467953

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43(9), 1035-1045. doi:10.1016/S0042-6989(03)00084-1

Feng, G., Miller, K., Shu, H., & Zhang, H. (2001). Rowed to recovery: The use of phonological and orthographic information in reading Chinese and English. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27, 1079-1100. doi: 10.1037/0278-7393.27.4.1079

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*. doi:10.1017/S0140525X11001841

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, NJ: Lawrence Erlbaum Associates.

Inhoff, A. W. (1989). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of*

- Memory & Language*, 28, 444-461. doi: 10.1016/0749-596X(89)90021-1
- Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming in reading. In A. Kennedy, R. Radach, D. Heller and J. Pynte (Eds.), *Reading as a visual process*. Oxford: Elsevier.
- Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1560-1579. doi: 10.1037/a0012319
- Kambe, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics*, 66, 279-292. doi: 10.3758/BF03194879
- Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45, 153-168. doi: 10.1016/j.visres.2004.07.037
- Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. (2013). How preview space/time translate into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology*. Advance online publication. doi: 10.1080/17470218.2012.658073
- Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, 18, 655-681. doi: 10.1080/13506280902986058
- Laubrock, J., & Hohenstein, S. (2012). Orthographic consistency and parafoveal

- preview benefit: A resource-sharing account of language differences in processing of phonological and semantic codes. *Behavioral and Brain Sciences*. doi:10.1017/S0140525X12000209
- R Development Core Team (2010). *R: A language and environment for statistical computing*. Wien: R Foundation for Statistical Computing. [Computer software]
- Rayner, K. (1975). The perceptual span and peripheral cues during reading. *Cognitive Psychology*, 7, 65-81. doi: 10.1016/0010-0285(75)90005-5
- 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., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixation in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements* (pp.213-234). Amsterdam, the Netherlands: Elsevier Science.
- Shu, H. & Anderson, R. C. (1997). Role of radical awareness in the character and word acquisition of Chinese children. *Reading Research Quarterly*, 32(1), 78-89. doi: 10.1598/RRQ.32.1.5
- Tsai, J.-L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in Traditional Chinese reading. *Acta Psychologica*, 141, 17-23. doi: 10.1016/j.actpsy.2012.06.004
- Tsai, J.-L., & McConkie, G. W. (2003). Where do Chinese readers send their eyes? In

- J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements* (pp. 159-176). Amsterdam, the Netherlands: Elsevier Science.
- Tsang, Y.-K., & Chen, H.-C. (2012). Eye movement control in reading: Logographic Chinese versus alphabetic scripts. *The PsyCh Journal*. doi: 10.1002/pchj.10
- Van Orden, G. C. (1987). A rows is a rose: spelling, sound and reading. *Memory and Cognition*, 15, 181-198. doi: 10.3758/BF03197716
- Yan, M., Kliegl, R., Richter, E. M., Nuthmann, A., & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *The Quarterly Journal of Experimental Psychology*, 63, 705-725. doi:10.1080/17470210903114858
- Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Chinese readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review*, 16, 561-566. doi: 10.3758/PBR.16.3.561
- Yan, M., Risse, S., Zhou, X., & Kliegl, R. (2012a). Preboundary duration modulates semantic preview benefit for word n+1 and n+2 in Chinese reading. *Reading and Writing*, 25(5), 1093-1111. doi:10.1007/s11145-010-9274-7
- Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012b). Lexical and sub-lexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 38(4), 1069-1075. doi: 10.1037/a0026935
- Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing*, 25, 1079-1091. doi:10.1007/s11145-010-9281-8

- Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word $n + 2$? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 1192-1204. doi: 10.1037/a0013554
- Yen, M.-H., Tsai, J.-L., Tzeng, O.J.-L., & Hung, D.L. (2008). Eye movements and parafoveal word processing in reading Chinese sentences. *Memory and Cognition*, 36, 1033-1045. doi: 10.3758/MC.36.5.1033
- Zhou, X., & Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26, 1245-1265. doi:10.1037/0278-7393.26.5.1245
- Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese. *Language and Cognitive Processes*, 14, 525-565. doi: 10.1080/016909699386185

AUTHOR NOTE

This research was funded by Deutsche Forschungsgemeinschaft Grant KL 955/18. We thank Simon Liversedge and Jukka Hyönä for their comments and discussions with the authors during the 16th European Conference on Eye Movements and the 5th China International Conference on Eye Movements. Data and R scripts used in this paper are available at the Potsdam Mind Research Repository (<http://read.psych.uni-potsdam.de/PMR2/>). Correspondence should be addressed to Ming Yan, Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam-Golm, Germany (E-mail: mingyan@uni-potsdam.de).

FOOTNOTE

1, The display change information was not recorded for data reported in Yan et al. (2009), so we did not apply this filter to Data Set 1.

2, The numbers of observations reported here are different from those from Yan et al. (2009), because we did not remove trials with regressions from words N or N+1.

Table 1

Character Measures

	Data Set 1			Data Set 2			
	Target	Sem.	Unrelated	Target	Trans.	Opaque	Unrelated
Character	永	久	向	考	评	检	养
Translation	forever	long	towards	examine	evaluate	inspect	feed
Sem. radical					讠	木	
					speech	wood	
Frequency	1150 (1728)	1164 (1721)	1163 (1573)	276 (307)	261 (441)	282 (415)	280 (326)
N. Stroke	5.0 (2.1)	5.5 (2.6)	4.9 (1.9)	9.3 (3.0)	9.2 (2.4)	9.2 (2.4)	9.6 (2.3)
Sem. Rating		4.1 (0.6)	1.2 (0.2)		4.1 (0.5)	3.9 (0.5)	1.5 (0.3)
R-C Trans.					3.8 (0.5)	2.1 (0.4)	
R-W Trans.					3.2 (0.9)	2.3 (0.5)	

Note. Means (and standard deviations, in parentheses) of frequency (per million, Beijing Language Institute Publisher, 1986), number of strokes, and rating of semantic relatedness to target of the previews are provided for both data sets. Transparencies of the semantic radicals with respect to the meaning of the character (radical–character semantic transparency rating [R-C trans.]) and entire two-character word (radical-word semantic transparency rating [R-W trans.]) are provided for Data Set 2.

Table 2

Fixation measures

	Data Set 1			Data Set 2			
	Identical	Sem.	Unrelated	Identical	Trans.	Opaque	Unrelated
(A)							
FFD	219 (6)	237 (6)	256 (6)	252 (3)	279 (3)	281 (3)	293 (6)
SFD	219 (7)	240 (8)	257 (7)	252 (4)	280 (4)	284 (4)	297 (6)
GD	255 (12)	310 (13)	333 (13)	289 (6)	341 (6)	354 (6)	366 (9)
(B)							
FFD	216 (8)	245 (8)	266 (7)	256 (3)	282 (4)	287 (3)	301 (6)
SFD	215 (9)	251 (10)	272 (8)	258 (4)	286 (4)	291 (4)	307 (6)
GD	254 (14)	331 (15)	358 (15)	308 (6)	360 (6)	375 (6)	386 (9)

Note. Means (and standard errors in parentheses) of first-fixation duration (FFD), single-fixation duration (SFD), gaze duration (GD) on target words, before (A) and after removing trials in which previews were possibly foveally processed (B).

Figure Captions

Figure 1. A set of sample sentences using the boundary paradigm used in Data Set 1 (A) and Data Set 2 (B). The preview characters occupy the position of the first character of the target word N+1 and are replaced by the target characters as soon as the reader's eyes cross the invisible boundary located between words N and N+1.

Figure 2. Linear regression of FFD on word N+1 as a function of FFD on word N and fixation landing position on word N for Data Set 2. Solid lines stand for unrelated preview, and dashes stand for semantically transparent condition. X-axis is landing position on pretarget word N (in characters). Between-subject and between-item differences for dependent variable and covariance in the LMM were removed prior to regressions. Figure was generated with ggplot2 (Wickham, 2009). Error bands show 95% confidence intervals.

Figure 1

(A)

Semantic preview:

许多人都认为爱情是人类永恒而经典的话题之一。

*

Unrelated preview:

许多人都认为爱情是人类向恒而经典的话题之一。

*

Target sentence:

许多人都认为爱情是人类永恒而经典的话题之一。

↑ * ↑
N N+1

The target sentence translates as *many people think that love is one of the most everlasting and classic topics*.

(B)

Transparent preview:

新颁布的安全生产评核制度从根本上杜绝了各类违章现象。

*

Opaque preview:

新颁布的安全生产检核制度从根本上杜绝了各类违章现象。

*

Unrelated preview:

新颁布的安全生产养核制度从根本上杜绝了各类违章现象。

*

Target sentence:

新颁布的安全生产考核制度从根本上杜绝了各类违章现象。

↑ * ↑
 N N+1

Target sentence translates as *the newly announced evaluation system for production safety has fundamentally prevented violation of regulations.*

Figure 2

