Ergonomics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/terg20

Effects of anomalous characters and small stroke omissions on eye movements during the reading of Chinese sentences

Pingping Liu[a], Weijun Li[b], Buxin Han[a] & Xingshan Li[b]
[a] Centre on Aging Psychology, Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, People's Republic of China
[b] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, People's Republic of China
Published online: 08 Aug 2014.

Ergonomics, 2014
http://dx.doi.org/10.1080/00140139.2014.945492

Effects of anomalous characters and small stroke omissions on eye movements during the reading of Chinese sentences

Pingping Liu[a]*, Weijun Li[b], Buxin Han[a] and Xingshan Li[b]*

[a]*Centre on Aging Psychology, Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, People's Republic of China;* [b]*Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, People's Republic of China*

(*Received 1 December 2013; accepted 7 July 2014*)

We investigated the influence of typographical errors (typos) on eye movements and word recognition in Chinese reading. Participants' eye movements were tracked as they read sentences in which the target words were presented (1) normally, (2) with the initial stroke of the first characters removed (the omitted stroke condition) or (3) the first characters replaced by anomalous characters (the anomalous character condition). The results indicated that anomalous characters caused longer fixation durations and shorter outgoing forward saccade lengths than the correct words. This finding is consistent with the prediction of the theory of the processing-based strategy. Additionally, anomalous characters strongly disrupted lexical processing and whole sentence comprehension, but small stroke omissions did not. Implications of the effect of processing difficulty on forward saccade targeting for models of eye movement control during Chinese reading are discussed.

Practioner Summary: A better understanding of the effects of different types of spelling errors on eye movements and sentence reading could provide valuable evidence to highlight the importance of correct spelling. We find that readers' oculomotor control system may be flexible and able to adapt to unusual presentation effectively.

Keywords: eye movements; typographical errors; saccade length; Chinese reading

1. Introduction

Reading is one of the most complex cognitive activities in which humans routinely engage. Aspects of eye movements, such as when and where to move the eyes next, are widely used to infer cognitive processes in reading (Rayner 1998, 2009). It is well documented that the amount of time that readers spend looking at a word is affected by the ease or difficulty associated with processing the word. In Chinese, many studies have demonstrated that fixation times are shorter on frequent words than on infrequent words (Wei, Li, and Pollatsek 2013; Yan et al. 2006; Yang and McConkie 1999) and that they are shorter on words that are constrained by the preceding context than on unconstrained words (Li et al. 2014; Rayner et al. 2005). However, the nature of saccade target selection and what factors determine where to move the eyes in Chinese remain to be determined (Li et al. 2014; Li, Liu, and Rayner 2011; Shu et al. 2011; Yan et al. 2010; Zang et al. 2013).

Several studies have examined the factors that influence where to move the eyes next in English. Readers initially tend to land on a point between the beginning and the centre of a word, which has been referred to as the preferred viewing location (PVL; Rayner 1979). Prior studies have demonstrated that word boundary information (interword spaces) is the major determinant used to decide where to move, and that saccade length is influenced by both foveal and parafoveal information (Rayner 1998, 2009). Thus, English readers may use a word-centre-based strategy when selecting a saccade target, and the strategy has been successfully incorporated into the E-Z Reader model (Pollatsek, Reichle, and Rayner 2006; Reichle et al. 1998) and other models (Engbert, Longtin, and Kliegl 2002; McDonald, Carpenter, and Shillcock 2005; Reilly and Radach 2006; Richter, Engbert, and Kliegl 2006).

Unlike in English, there are no spaces between printed words in Chinese, and Chinese readers do not always agree on where word boundaries are in sentence processing (Hoosain 1992; Liu et al. 2013). Without the aid of word boundary information, what factors influence where to move the eyes next in Chinese reading?

Prior studies have provided inconsistent results regarding where the eyes move next when reading Chinese. Some studies suggest that similar PVL effects (i.e. near the word centre, as in English) occur for single fixations, whereas the initial positions of multiple fixations are closer to the beginning of a word (Shu et al. 2011; Yan et al. 2010; Zang et al. 2013). In contrast, some studies have reported that Chinese readers do not target any specific position within a word while reading (Li et al. 2014; Li, Liu, and Rayner 2011; Tsai and McConkie 2003; Yang and McConkie 1999). Thus, the question of where to move the eyes while reading Chinese is a difficult one that requires further research.

*Corresponding authors. Email: liupp@psych.ac.cn, lixs@psych.ac.cn

Wei, Li, and Pollatsek (2013) proposed that Chinese readers may use a processing-based strategy to select the saccade target. Readers may attempt to process enough characters as efficiently as possible for each fixation, and they then move their eyes somewhere to the right of those characters. According to this strategy, the processing difficulty of the current fixated words should influence the forward saccade length from that word. Specifically, previous studies have reported that saccade length shortens as the processing difficulty of the words increases (Li et al. 2014; Perea and Acha 2009; Wei, Li, and Pollatsek 2013; White and Liversedge 2006; Yan et al. 2012). For instance, outgoing forward saccade length increases as the frequency and length of the fixated words increase (Li et al. 2014; Li, Liu, and Rayner 2011; Wei, Li, and Pollatsek 2013). White and Liversedge (2006) also found that saccades into word $n + 1$ were significantly shorter if word n was infrequent compared with when it was frequent. Recently, Yan et al. (2012) examined the influence of stroke removal on Chinese character identification and found that 'the more the strokes were removed, the more difficult the text was to process, which caused subjects to make shorter saccades'.

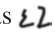
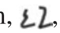
Typographical errors (hereafter referred to as typos), perhaps regarded as infrequent and unconstrained stimuli, are mistakes made in the typing processes of printed materials. It is well known that the occurrence of printing typos in written text is ubiquitous (Chan and Ng 2012; Logan 1999; Zola 1984). The errors existing in a text come in various forms: spelling errors, grammatical errors, semantic violations and others. Crucially, different results may hold for different types of typos (Schotter et al. 2014). While still poorly understood, typos provide a useful window for elucidating how cognitive processes change along with presentation formats when reading for comprehension text that contains typos in terms of stimuli and response measure. In the present study, we explored the effects of two types of typos (specifically anomalous characters and stroke omissions) on word recognition and eye movements in Chinese reading. We will describe the two conditions in detail later.

Furthermore, the findings regarding how different types of typos affect where to move the eyes will help uncover the nature of saccade target selection in Chinese reading. As noted above, most theories of eye movements explore the effects of word frequency and predictability in reading (Kliegl, Nuthmann, and Engbert 2006; Li et al. 2014; Rayner 2009; Rayner et al. 2005; Wei, Li, and Pollatsek 2013). It should be apparent that if some other variables influence when and where to move the eyes, to the extent that such variables can be identified, it should be easier to infer the link between eye movements and cognitive processes in Chinese reading.

Until now, few studies have directly explored how saccade target selection is affected by the processing difficulty caused by typos in Chinese. Previous studies have demonstrated that readers tend to have longer fixation durations on inaccurate words (e.g. Rayner et al. 2004; White and Liversedge 2006; Yang et al. 2012; Zola 1984) than on correct words. Thus, it may be more difficult for participants to process typos than the correct characters in Chinese. According to the processing-based strategy, readers would make shorter outgoing forward saccades from these inaccurate words than correct words. In particular, White and Liversedge (2006) also reported effects of typos on saccades into words (due to differences in orthographic familiarity) in English reading. Such findings raise the important question of whether anomalous characters and stroke omissions might influence saccade targeting to words in Chinese, due to differences in orthographic familiarity of the character components or character sequences. As a result, the first goal of the present study was to explore the effects of typos on saccade target selection by examining forward saccades both in and out of the target words.

Notably, there is no question that some typos can disrupt lexical processing (Kaakinen and Hyönä 2010; Pynte, Kennedy, and Ducrot 2004; Rayner et al. 2004; Schotter et al. 2014; White and Liversedge 2006); the key issue is at what stage in the processing they have an effect in Chinese. Whether typos have an immediate effect on lexical processing and eye movements is elusive. Moreover, we cannot exclude the possibility that typos affect the later stage of lexical processing, such as the post-lexical integration stage. In English, Rayner et al. (2004) found that these anomalous words showed strong effects on gaze duration but not on first-fixation duration. In French, Pynte, Kennedy, and Ducrot (2004) found that the presence of a typo exerted a systematic influence on the later stage of lexical processing: the probability of regressions out of the target word was significantly greater in the typo condition than in the no-typo condition. However, the study did not find that gaze duration on the target word changed as a function of experimental conditions. Meanwhile, it is well established that a number of eye movement measures can be associated with variations in the time course of processing a target word (Rayner 1998, 2009). First fixation duration and gaze duration reflect first-pass reading time and the reader's initial encounter with a target word, whereas total time and regression rates reflect later processing activities and integration processes. Thus, the second goal of the present study was to determine at which point typos have an effect on early and later processing with respect to a specific target word.

Two types of typos were investigated in the present study because they may have different influences on eye movements and word recognition (Schotter et al. 2014). First, the anomalous character condition was designed to explore the effect of severely incorrectly written characters on eye movements during reading. Written Chinese is different from alphabetic languages, such as English, in many important aspects. Most notably, as a logographic script, Chinese is formed by strings of equally spaced boxlike symbols called characters. There are more than 5000 characters according to the Chinese Lexicon

(2003), and individual characters differ greatly in terms of complexity (i.e. the number of strokes). Some Chinese characters have identical radicals or similar appearances; therefore, most typos could be caused by misused characters that are orthographically or phonologically similar to the correct forms (Geng 2010; Hatta, Kawakami, and Tamaoka 2002). For instance, 农 (farm) can be incorrectly written as 衣 (clothes). Thus, the anomalous character condition was designed. Second, the omitted stroke condition was designed to explore how slightly incorrectly written characters influence eye movements and word recognition. In general, characters in Chinese readers' handwriting are likely to have fewer strokes relative to the standard written forms. For instance, 红 (red) could be written as  in handwriting. The new form, , with some strokes omitted is a pseudo-character or typo, but most Chinese readers can identify the character quickly. It was unclear how these omitted stroke characters influence eye movements.

We predicted that the two types of typos may have different influences on eye movements and word recognition. The processes underlying word identification in Chinese are different from those in alphabetic languages (Li et al. 2014; Li, Rayner, and Cave 2009; Rayner, Li, and Pollatsek 2007; Yan et al. 2010; Yang and McConkie 1999). The prevalent assumptions have been that Chinese writing is logographic and that form-to-meaning processing is so important that orthographic encoding in word identification is a core process (e.g. Perfetti, Liu, and Tan 2005). In the omitted stroke condition, only the initial stroke (appropriately 15% of strokes) was removed from the first character of the target word so that participants could guess and recognise the character correctly and quickly. Yan et al. (2012) reported that Chinese characters with 15% of strokes removed were as easy to read as Chinese characters without any strokes removed. Thus, these omitted stroke characters may produce little disruption to eye movements and reading in the present study. However, in the anomalous character condition, the first correct character was replaced by another anomalous character that could not constitute a word with the second character of the target word. The anomalous character may disturb the identification of word boundaries vital for successful word identification (Li, Rayner, and Cave 2009; Liu et al. 2013). Thus, it may be more difficult for participants to process incorrect information in the anomalous condition than in the omitted stroke and normal conditions. The last goal of the present study was to test this prediction.

2. Methods

2.1. Participants

Twenty-four native Chinese speakers (14 females) at universities in Beijing near the Institute of Psychology, Chinese Academy of Sciences, were paid to participate in the experiment. All participants had either normal or corrected-to-normal vision, and they were not aware of the purpose of the experiment.

2.2. Apparatus

Eye movements were recorded with an Eyelink 1000 eye tracker (SR Research Ltd, Osgoode, Canada). Although viewing was binocular, only the right eye was monitored. The sentences were presented on a 21-inch CRT monitor (resolution: 1024 × 768 pixels; refresh rate: 150 Hz) connected to a Dell PC. Participants were seated 58 cm from the monitor, and each character subtended a visual angle of approximately 1.2°.

2.3. Materials

Seventy-two experimental sentence frames were constructed, and Figure 1 shows the example sentences. In the normal condition, 72 two-character target words were selected from the Chinese Lexicon (2003). The average numbers of strokes

Normal	本戏剧的目的是让那些埋头于**农活**的妇女们认识到自己的价值。
Omitted stroke	本戏剧的目的是让那些埋头于**农活**的妇女们认识到自己的价值。
Anomalous character	本戏剧的目的是让那些埋头于**衣活**的妇女们认识到自己的价值。

Notes. The English translation of the sentence is: *The purpose of the drama is to make these women who are busy with farm work to be aware of their value*. The target region has been highlighted by bold font (but not during the actual experiment) for clarification.

Figure 1. Example sentences.

of the first and second characters were 6.9 and 7.0, respectively. The average frequencies of the first and second characters were 915 and 957 occurrences per million, respectively. The pair-wise *t*-tests indicated that the first and second characters were matched exactly on the number of strokes ($p = 0.60$) and frequency ($p = 0.84$). The results indicated the properties of the first and second characters of the target words were similar.[1]

In the omitted stroke condition, the initial stroke was removed from the first character of the target word, and the new form (e.g. 衣) was a pseudo-character. Note that the stroke omission condition necessarily results in different numbers of strokes, resulting in differences in visual complexity across the conditions. In the anomalous character condition, the first character of the target word (e.g. 农 means farm) was replaced by another character (e.g. 衣 means clothes) that could not constitute a word with the second characters (e.g. 活 means work) according to the Chinese Lexicon (2003). All of the anomalous characters in both the anomalous character and omitted stroke conditions were orthographically similar to the first characters of the target words in the normal condition. The pair-wise *t*-tests indicated that the frequency of the first character in the normal ($M = 915$ occurrences per million) and anomalous character conditions ($M = 855$ occurrences per million) did not significantly differ, $t\ (71) = 0.28$, $p = 0.78$.

Experimental sentences ranged from 20 to 29 characters in length ($M = 25.6$). The target words in the normal condition were always low in frequency ($M = 3.3$ occurrence per million) and predictability ($M = 0$) to increase the likelihood that readers would fixate on them. Norming tasks were used to assess the predictability of the target words and the acceptability of the sentences. Predictability ratings for each target word were calculated from cloze task data provided by 12 native Chinese speakers, who were given the beginning portions of the sentence up to the target word and asked to generate the next acceptable word in the sentence. For all of the completions, none were correct. Therefore, none of the target words were predictable from the sentence context. Additionally, acceptability ratings for each experimental sentence were obtained from 12 native Chinese speakers on a scale from 1 (unacceptable) to 7 (perfectly acceptable). The average reported acceptability was 6.4 (ranging from 6.0 to 6.8). None of these participants who had participated in the norming studies took part in the eye-tracking experiment.

In the main experiment, each participant read 16 practice sentences, 72 experimental items and 160 filler sentences. It is possible that the presence of some typos in the experimental sentences alerted participants and lowered the ecological validity of the task (Pynte, Kennedy, and Ducrot 2004). Thus, 160 normal sentences were included as fillers to avoid participants predicting typos within the experimental materials. We constructed three files of materials such that 24 items appeared in each condition in each file and each item appeared in a different condition in each file. Comprehension questions were presented on the screen after 36% of the sentences. All the comprehension questions were designed to avoid the need to comprehend the target words.

2.4. *Procedure*

When participants first arrived for the experiment, they were given instructions and a description of the apparatus. At the start of the experiment, participants performed a calibration procedure by looking at a sequence of three fixation points randomly displayed horizontally across the middle of the computer screen. Following calibration, the gaze position error was smaller than 0.5° of the visual angle. At the beginning of each trial, a white square (approximately 1° × 1°) appeared on the left side of the computer screen, which indicated the position of the first character in the sentence. Once the participant

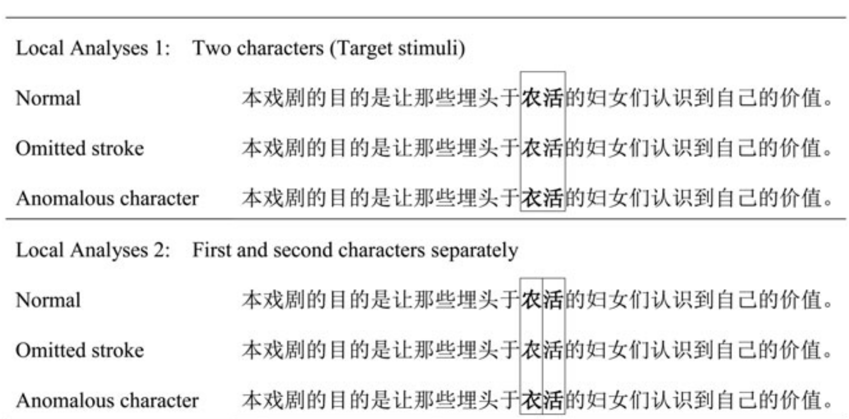


Figure 2. Example of regions of interest for local analyses.

fixated on the white square successfully, a sentence was displayed. Participants read silently for comprehension in a natural way and pressed a button on a keypad when they finished reading the sentence. The eye tracker was checked and recalibrated, if necessary, prior to the presentation of each sentence. Participants were given short breaks between trials to prevent fatigue, and the experiment lasted approximately 50 min.

3. Results and discussion

The average comprehension accuracy was 96%. Approximately 0.5% of total fixations shorter than 80 ms or longer than 1000 ms were excluded from our analyses. Additionally, 4.1% of trials in which there were blinks on the target stimuli or post-target words were discarded prior to analyses.

Three sets of analyses are reported. We first report the reading time for the entire sentence and then report the measures for the target stimulus, which was a two-character region of interest (ROI). We next divided the two-character ROI into two single-character ROIs for analyses (see Figure 2). These analyses provide a more detailed examination of reading behaviour on these different ROIs. Specifically, we reported the *initial landing position* of the eyes in the ROI, the *incoming saccade length* into the ROI (the distance between the initial fixation location in the target word and the previous fixation location) and the *outgoing forward saccade length* (the distance between the last fixation location in the ROI and the next fixation location). Both the *incoming and outgoing forward saccade lengths* refer to forward (rightward) saccades and therefore do not include regressions during first-pass reading. Additionally, a series of standard eye movement measures were reported: (1) *first fixation duration* (the duration of the first fixation on the ROI during first-pass reading), (2) *gaze duration* (the sum of all first-pass fixations on the ROI before moving to another word), (3) *total time* (the sum of all fixations on the ROI, including regressions), (4) *fixation probability* (the probability of fixating on the ROI during the first-pass reading), (5) *refixation probability* (the proportion of cases in which the ROI was refixated before leaving it) and (6) the *probability of regressions out of and into* the ROI. Repeated measures analyses of variance (ANOVAs) were undertaken across the three conditions for both participants' (F_1) and items' (F_2) means.

3.1. *Sentence reading time*

There was a significant difference in sentence reading time among the three conditions, $F_1(2, 46) = 32.53$, MSE = 118,592, $p < 0.001$, $\eta_p^2 = 0.59$, $F_2(2, 142) = 19.21$, MSE = 642,297, $p < 0.001$, $\eta_p^2 = 0.21$. Pair-wise comparisons indicated that sentence reading time was significantly longer (p-values < 0.001) in the anomalous character condition ($M = 5133$ ms, SE = 313 ms) than the other two conditions, but the normal ($M = 4454$ ms, SE = 256 ms) and omitted stroke conditions ($M = 4424$ ms, SE = 227 ms) did not significantly differ (p-values > 0.60). The longer reading times for the anomalous character condition compared to the normal condition clearly showed that the text with an anomalous character was harder to process than the normal text, whereas text with omitted strokes was not.

3.2. *Two-character ROI analyses*

3.2.1. *Saccade length and landing position*

As seen in Table 1, there was a significant difference in outgoing forward saccade length [$F_1(2, 46) = 8.42$, MSE = 0.05, $p = 0.001$, $\eta_p^2 = 0.27$, $F_2(2, 142) = 5.77$, MSE = 0.16, $p = 0.004$, $\eta_p^2 = 0.08$]. Pair-wise comparisons indicated that the outgoing forward saccade length was shorter in the anomalous character condition than in the omitted stroke [F_1(1,

Table 1. Local measures for the two characters.

Measures	Normal	Omitted stroke	Anomalous character
First fixation duration (ms)	308 (8.5)	314 (8.7)	313 (9.2)
Gaze duration (ms)	386 (15.1)	406 (14.3)	482 (25.3)
Total time (ms)	505 (30.3)	520 (26.4)	807 (57.8)
Fixation probability	0.88 (0.03)	0.88 (0.03)	0.89 (0.02)
Refixation probability	0.24 (0.03)	0.26 (0.04)	0.33 (0.04)
Probability of regression out	0.08 (0.02)	0.11 (0.02)	0.15 (0.02)
Probability of regression in	0.18 (0.02)	0.23 (0.02)	0.39 (0.03)
Initial landing position (characters)	0.44 (0.03)	0.45 (0.03)	0.43 (0.02)
Incoming saccade length (characters)	2.36 (0.11)	2.37 (0.13)	2.39 (0.11)
Outgoing forward saccade length (characters)	2.34 (0.12)	2.47 (0.12)	2.22 (0.12)

Note: Standard errors are shown in parentheses.

23) = 15.24, MSE = 0.10, $p = 0.001$, $\eta_p^2 = 0.40$, $F_2(1, 71) = 11.90$, MSE = 0.32, $p = 0.001$, $\eta_p^2 = 0.14$] and normal [$F_1(1, 23) = 4.57$, MSE = 0.09, $p = 0.04$, $\eta_p^2 = 0.14$, $F_2(1, 71) = 2.81$, MSE = 0.36, $p = 0.10$, $\eta_p^2 = 0.04$, though marginally significant in the items' analysis] conditions. The outgoing forward saccade length was marginally significantly longer in the omitted stroke condition than in the normal condition [$F_1(1, 23) = 4.31$, $p = 0.05$, $F_2(1, 71) = 2.88$, $p = 0.09$]. There were no significant differences in initial landing position or incoming saccade length among the three conditions (p-values > 0.75). The findings that anomalous characters did not affect the initial landing position or incoming saccade length but did affect the outgoing forward saccade length indicated that Chinese readers' saccade target selection may be determined by foveal information. Importantly, the result that participants tended to make shorter saccades in the anomalous character condition than in the normal condition is consistent with the prediction of the processing-based strategy.

3.2.2. *Fixation times, fixation probability and refixation probability*

There were significant differences in gaze duration [$F_1(2, 46) = 19.46$, MSE = 3129, $p < 0.001$, $\eta_p^2 = 0.46$, $F_2(2, 142) = 18.99$, MSE = 10,556, $p < 0.001$, $\eta_p^2 = 0.21$] and total time [$F_1(2, 46) = 49.46$, MSE = 14,076, $p < 0.001$, $\eta_p^2 = 0.68$, $F_2(2, 142) = 57.27$, MSE = 36,912, $p < 0.001$, $\eta_p^2 = 0.45$, although there were no differences in first fixation duration or fixation probability among the three conditions (p-values > 0.15). Pair-wise comparisons indicated that gaze duration and total time were significantly longer in the anomalous character condition than the other two conditions (p-values < 0.01). The normal and omitted stroke conditions did not significantly differ (p-values > 0.05). There was also a significant main effect of condition on the probability of refixating the target stimuli [$F_1(2, 46) = 5.99$, MSE = 0.01, $p = 0.005$, $\eta_p^2 = 0.21$, $F_2(2, 142) = 6.98$, MSE = 0.03, $p = 0.001$, $\eta_p^2 = 0.09$]. The target word was more likely to be refixated in the anomalous condition than either the normal or omitted stroke condition (p-values < 0.05), but there was no difference in refixation probability between the normal or omitted stroke condition (p-values > 0.20). The result regarding longer gaze duration in the anomalous character condition showed that anomalous characters caused more refixations and disrupted word recognition when compared with the normal text. However, the omitted strokes did not show the same effects.

3.2.3. *Regression rate*

As shown in Table 1, there were significant differences in the probabilities of regressions out of [$F_1(2, 46) = 6.34$, MSE = 48.62, $p = 0.004$, $\eta_p^2 = 0.22$, $F_2(2, 142) = 8.70$, MSE = 106.47, $p < 0.001$, $\eta_p^2 = 0.11$] and into [$F_1(2, 46) = 38.02$, MSE = 77.32, $p < 0.001$, $\eta_p^2 = 0.62$, $F_2(2, 142) = 29.43$, MSE = 291.79, $p < 0.001$, $\eta_p^2 = 0.29$] the target region among the three conditions. Pair-wise comparisons indicated that there were more regressions out of (p-values < 0.05) and into (p-values < 0.001) the target region in the anomalous character condition than the other two conditions, respectively. The normal and omitted stroke conditions did not differ from each other in the frequency of regressions from the target region (p-values > 0.20), but they marginally differed from each other in the probability of regressions into the target region, $F_1(1, 23) = 3.14$, MSE = 170.53, $p = 0.09$, $\eta_p^2 = 0.12$, $F_2(1, 71) = 3.15$, MSE = 512.55, $p = 0.08$, $\eta_p^2 = 0.04$. The results showed that small stroke omissions slightly disrupted readers' eye

Table 2. Local measures for the first and second characters of target stimuli separately.

	The first character			The second character		
Measures	Normal	Omitted stroke	Anomalous character	Normal	Omitted stroke	Anomalous character
First fixation duration (ms)	308 (9.5)	311 (10.0)	291 (8.7)	288 (7.5)	303 (10.0)	347 (11.0)
Gaze duration (ms)	315 (10.1)	330 (12.0)	308 (11.1)	302 (8.5)	307 (10.5)	384 (13.4)
Total time (ms)	362 (16.9)	386 (18.2)	467 (29.4)	360 (14.7)	363 (17.6)	536 (32.6)
Fixation probability	0.50 (0.03)	0.49 (0.04)	0.51 (0.03)	0.61 (0.03)	0.59 (0.03)	0.65 (0.03)
Refixation probability	0.02 (0.01)	0.03 (0.01)	0.03 (0.01)	0.03 (0.01)	0.02 (0.01)	0.07 (0.02)
Probability of regression out	0.06 (0.01)	0.07 (0.01)	0.08 (0.02)	0.06 (0.02)	0.12 (0.02)	0.17 (0.02)
Probability of regression in	0.13 (0.02)	0.18 (0.03)	0.35 (0.03)	0.13 (0.02)	0.19 (0.02)	0.34 (0.03)
Incoming saccade length (characters)	2.29 (0.15)	2.22 (0.13)	2.30 (0.13)	2.11 (0.10)	2.24 (0.15)	2.09 (0.11)
Outgoing forward saccade length (characters)	2.05 (0.15)	2.04 (0.12)	1.81 (0.12)	2.20 (0.12)	2.30 (0.12)	1.99 (0.11)

Note: Standard errors are shown in parentheses.

movements, but these anomalous characters caused more regressions, which would lead to longer sentence reading time. The results showed that typos affect post-lexical integration processes and sentence comprehension.

3.3. Single-character ROI analyses

The single-character ROI analyses were undertaken to further examine how these typos influenced saccade targeting and word recognition during first-pass reading.

3.3.1. Saccade length

As shown in Table 2, similar to the two-character ROI analyses, there were significant differences among the three conditions in outgoing forward saccade length from the first [$F_1(2, 46) = 4.09$, MSE = 0.10, $p = 0.02$, $\eta_p^2 = 0.15$, $F_2(2, 138) = 3.84$, MSE = 0.34, $p = 0.02$, $\eta_p^2 = 0.05$] and second characters [$F_1(2, 46) = 11.09$, MSE = 0.05, $p < 0.001$, $\eta_p^2 = 0.33$, $F_2(2, 140) = 3.83$, MSE = 0.28, $p = 0.02$, $\eta_p^2 = 0.05$], but there were no differences in incoming saccade length (p-values > 0.15). Pair-wise comparisons indicated that there were similar patterns for the two single-character ROIs. For the first-character ROI, the outgoing forward saccade length was shorter in the anomalous character condition than in the normal [$F_1(1, 23) = 5.67$, MSE = 0.23, $p = 0.03$, $\eta_p^2 = 0.20$, $F_2(1, 69) = 7.69$, MSE = 0.63, $p = 0.007$, $\eta_p^2 = 0.10$] and omitted stroke conditions [$F_1(1, 23) = 7.13$, MSE = 0.17, $p = 0.01$, $\eta_p^2 = 0.24$, $F_2(1, 69) = 4.34$, MSE = 0.65, $p = 0.04$, $\eta_p^2 = 0.06$]. The normal and omitted stroke conditions did not differ from each other (p-values > 0.56). For the second-character ROI, the outgoing forward saccade length was shorter in the anomalous character condition than in the normal [$F_1(1, 23) = 8.00$, MSE = 0.14, $p = 0.01$, $\eta_p^2 = 0.26$, $F_2(1, 70) = 4.04$, MSE = 0.51, $p = 0.048$, $\eta_p^2 = 0.06$] and omitted stroke conditions [$F_1(1, 23) = 22.40$, MSE = 0.10, $p < 0.001$, $\eta_p^2 = 0.49$, $F_2(1, 70) = 6.77$, MSE = 0.61, $p = 0.01$, $\eta_p^2 = 0.09$]. The normal and omitted stroke conditions did not differ from each other (p-values > 0.12). These results are consistent with the prediction of the processing-based strategy: if participants fixated on the characters that were more difficult to identify, they would make a shorter saccade.

3.3.2. Fixation times, fixation probability and refixation probability

Table 2 shows the first fixation duration, gaze duration, and total time data for the first and second characters. For the first-character ROI, there was a significant difference in total time [$F_1(2, 46) = 15.14$, MSE = 4781, $p < .001$, $\eta_p^2 = 0.40$, $F_2(2, 142) = 12.56$, MSE = 17,292, $p < .001$, $\eta_p^2 = 0.15$], although there were no significant differences among the three conditions for first fixation duration or gaze duration (p-values > 0.10). Pair-wise comparisons indicated that total time was significantly longer (p-values < 0.01) in the anomalous character condition than in the normal and omitted stroke conditions, which did not differ significantly (p-values > 0.05). For the second-character ROI, there were significant differences in first fixation duration, gaze duration and total time (p-values < 0.001), although the characters were identical across the three conditions. Likewise, pair-wise comparisons indicated that the three fixation measures were significantly longer (p-values < 0.01) in the anomalous character condition than in the other two conditions, which did not differ from each other (p-values > 0.05).

As shown in Table 2, for the two single-character ROIs, there were no significant differences in fixation probability among the three conditions (p-values > 0.20). There was a significant main effect of condition on the probability of refixating the second-character ROI [$F_1(2, 46) = 8.18$, MSE = 0.002, $p = 0.001$, $\eta_p^2 = 0.26$, $F_2(2, 142) = 9.76$, MSE = 0.005, $p < 0.001$, $\eta_p^2 = 0.12$], but not the first-character ROI [$F_1(2, 46) = 1.30$, $p = 0.29$, $F_2(2, 142) = 1.16$, $p = 0.32$]. The second-character ROI was more likely to be refixated in the anomalous character condition (p-values < 0.05) than in the other two conditions, which did not differ from each other (p-values > 0.05). The results showed that participants fixated longer and were more likely to refixate on the second characters in the anomalous character condition than in the other two conditions, possibly to obtain more exact information for word recognition and sentence comprehension.

3.3.3. Regression rate

As shown in Table 2, there were significant differences among the three conditions in the probabilities of regressions into the first- and second-character ROIs (p-values < 0.001). Pair-wise comparisons indicated that there were more regressions into the two ROIs in the anomalous character condition than in the omitted stroke condition (p-values < 0.001) and more in the omitted stroke condition than in the normal condition (p-values < 0.05). Additionally, there were significant differences in the probabilities of regressions out of the second character (p-values < 0.01) but no differences in the frequency of regressions out of the first character (p-values > 0.20). Pair-wise comparisons indicated that there were more regressions out of the second character in the anomalous character condition than in the omitted stroke condition (p-values < 0.05) and more in the omitted stroke condition than in the normal condition [$F_1(1, 23) = 3.74$, $p = 0.07$, $\eta_p^2 = 0.14$, $F_2(1, 71) = 15.50$, $p < 0.001$, $\eta_p^2 = 0.18$, although marginally significant in the participants' analysis]. The results indicated that more regressions caused longer reading times in the anomalous character condition than in the normal condition.

4. General discussion

The present study aimed to investigate the effect of two types of typos on eye movements and word recognition during Chinese reading. First, we found anomalous characters affected forward saccade targeting. Participants tended to make shorter saccades when the fixated characters were more difficult to process in the anomalous character condition than in the normal condition. Second, with respect to the issue of how early effects appear when processing typos, our results were clear. In terms of the two-character ROI analyses, anomalous characters showed significant effects on gaze duration during first-pass reading. For single-character ROI analyses, anomalous characters showed strong effects on first fixation duration and gaze duration for the second character. Hence, the differences in gaze duration occurred because participants were more likely to refixate on the second characters and fixated longer on the second characters before moving their eyes to another word. Third, the present experiment clearly showed that anomalous characters produced a strong disruption to word recognition and text processing. However, small stroke omissions produced very little disruption to reading. In sum, the detailed measures of eye movements used in this study provide fresh insight into the claim that decisions about when and where to move the eyes are strikingly influenced by how easily words can be identified.

Our study also adds an important finding to the growing literature that readers' oculomotor control system may be flexible and able to adapt to unusual presentation conditions, such as typos, both rapidly and effectively. As described in the methods, we confirmed that participants did not select any special character as the saccade target when they were presented with normal sentences, which is consistent with prior findings (Li et al. 2014; Li, Liu, and Rayner 2011; Tsai and McConkie 2003; Yang and McConkie 1999). Unlike in English writing, the average initial landing position may not be a special position within a word, because there are no spaces between printed words in Chinese reading. However, in the anomalous character condition, participants exhibited more fixations and fixated longer on the second correct character than on the first anomalous character. This may be caused by the fact that readers had to depend on the second character to recognise the target words and comprehend the whole sentence in the anomalous character condition. In a word, these findings indicated that participants' eye movements adapted to these typos quickly.

4.1. Saccade target selection

We found that the incoming forward saccade length into these ROIs did not differ across the three conditions (i.e. null effects on saccade in). There was no evidence that the anomalous characters or stroke omissions were processed in the parafoveal vision, as shown by the lack modulation of saccade targeting to words. However, critically, although the present study showed null effects, character component typos or orthographic familiarity with stronger manipulations may have had additional effects, which should be investigated by further research.

One main finding of the present experiment is that typos in foveal vision strongly affected where to move the eyes. The saccade lengths out of the first and second characters were shorter in the anomalous character condition than the normal and omitted stroke conditions. These results were consistent with the findings reported by previous studies that forward saccade length would shorten with the increase of processing difficulty (Li et al. 2014; Perea and Acha 2009; Wei, Li, and Pollatsek 2013; White and Liversedge 2006; Yan et al. 2012). These results also indicated that foveal processing may play an important role in deciding where to look next in Chinese sentences due to the absence of visual word boundaries (e.g. spaces), and Chinese readers may use a processing-based strategy to select saccade target (Wei, Li, and Pollatsek 2013).

However, note that the shorter outgoing saccade pattern might also be accounted for by effects of foveal load on parafoveal preview (Henderson and Ferreira 1990; Kennison and Clifton 1995). The benefit gained through parafoveal preview decreases as foveal processing becomes more difficult. Reduced preview may result in reduced skipping probability of the following character or short outgoing forward saccade length. Such a mechanism has already been incorporated into models such as E-Z Reader (using two stages of word identification; Reichle, Rayner, and Pollatsek 2003; Reichle, Warren, and McConnell 2009).

Moreover, this main finding could have implications for developing computational models of eye movement control in non-alphabetic languages such as Chinese. Although Rayner, Li, and Pollatsek (2007) first successfully extended the E-Z Reader model to Chinese, some assumptions of these models based on alphabetic languages need to be tested in the context of Chinese reading. The model assumes that where to move the eyes is influenced by low-level information such as word boundaries and word length (Pollatsek, Reichle, and Rayner 2006; Reichle et al. 1998; Reichle, Warren, and McConnell

2009). However, Chinese readers cannot use boundary information to decide where to move the eyes. The present study showed that the ease or difficulty associated with processing a fixated word influences forward saccade length. These findings could provide important evidence for the claims that decisions about where to move the eyes are strikingly influenced by how easily words can be identified in Chinese reading (Wei, Li, and Pollatsek 2013). These results and possible mechanisms could have some implications for developing the basic assumptions of the models of eye movement control in Chinese reading.

4.2. Stages of lexical processing

Similar to the effects of anomalies reported by previous studies (Kaakinen and Hyönä 2010; Rayner et al. 2004), we did find that the typos in foveal vision influenced fixation durations. These anomalous characters strongly affected gaze duration and total time, which indicated that they influenced both the early and later stages of lexical processing. Furthermore, relative to the normal condition, participants were more likely to regress to the first anomalous character and more likely to regress into and out of the second character in the anomalous character condition (see Figure 3). The pattern of regressions showed that anomalous characters guided participants' reanalysis attempts. Moreover, participants may be aware that these first anomalous characters could not provide accurate information and that the second correct characters could be helpful for word recognition; thus, they regressed more and fixated longer on the second characters in the anomalous character condition than in the other two conditions.

Finally, the result that the fixation times on the target word were longer in the anomalous character condition than in the omitted stroke condition was consistent with our prediction. This indicated that anomalous characters rather than small stroke omissions hinder sentence comprehension. Similar to the findings reported by Yan et al. (2012), Chinese characters with initial strokes removed were as easy to read as Chinese characters without any strokes removed. Nevertheless, there were more regressions into the first and second characters in the omitted stroke condition than in the normal condition (see Figure 3), though the fixation times did not differ between the two conditions. The results indicated that omitted strokes that arise in the acceptability of word identification during sentence reading have delayed effects of a smaller magnitude on the eye movements recorded relative to the normal condition. These smaller delayed effects presumably influenced only the later stages of lexical processing in sentence comprehension.

The finding that small stroke omissions produced very little disruption to word recognition and text processing may have some implications for models of word recognition in Chinese. Chinese word recognition can proceed with very little disruption when the characters are not entirely complete; in particular, pseudo-characters with initial stroke omissions could allow participants to guess the complete characters easily. Meanwhile, Li, Rayner, and Cave (2009) argued that Chinese word segmentation and word recognition processes must be dependent on top-down processing. Meanwhile, readers' pre-existing lexical presentations could facilitate identification of the individual characters and allow the reader to segment and recognise the target words from text successfully. Thus, it was easy for readers to recognise characters with small stroke omissions when presented in multicharacter strings. In brief, both character information and pre-existing lexical presentation are important for successful word recognition in Chinese.

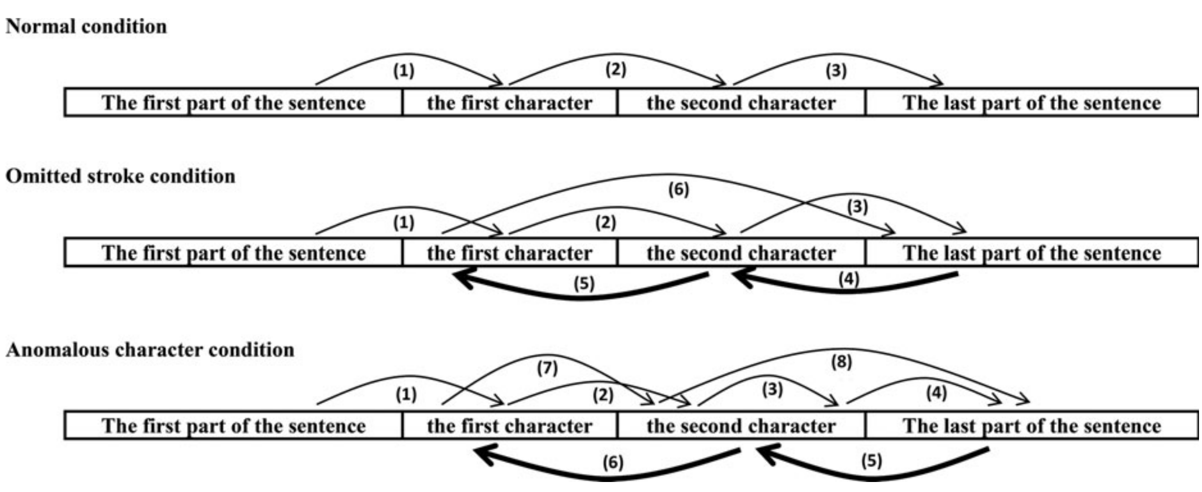


Figure 3. Typical examples of regressions patterns into and out of the two single-character ROIs in the sentence. Notes. The symbol indicates forward saccade, and indicates regressive saccade. The numbers indicate the sequence of saccades.

5. Conclusion

In sum, the present study clearly showed that the processing difficulty of fixated information influenced saccade lengths and fixation times. We found that anomalous characters strongly disrupted saccade target selection, which could have implications for developing models of eye movement control in Chinese. Additionally, the results suggest that an anomalous character disrupts word recognition and entire sentence comprehension, whereas small stroke omissions do not. This provides reliable evidence for the claims that correct spelling is important and that the quality of writing Chinese characters should be enhanced if there are anomalous character typos. However, small stroke omissions, such as those produced during hand writing, result in very little disruption to reading.

Acknowledgements

We thank Jinmian Yang and Meng Zhai for helpful comments on earlier version of this paper.

Funding

This research was supported by the Knowledge Innovation Program of the Chinese Academic Sciences [KSCX2-YW-BR-6]; the National Natural Science Foundation of China [31070904], [31000505]; Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences [Y3CX121005].

Note

1. To ensure that the properties of the first and second characters were comparable in the normal condition, 24 participants' eye movements were recorded when they read 72 normal sentences. None of them took part in the main experiment or norming tasks. The apparatus and procedure were the same as those in the main experiment described. The pair-wise *t*-tests showed that there was no significant difference in fixation probability on the first (0.57) and second (0.60) characters of the target words, $t_1\ (23) = -1.0$, $p = 0.33$, $t_2\ (71) = -1.25$, $p = 0.22$. The results replicated prior findings that Chinese readers do not select any special position within a word as the saccade target (Li et al. 2014; Li, Liu, and Rayner 2011; Tsai and McConkie 2003; Yang and McConkie 1999).

References

Chan, A. H. S., and A. W. Y. Ng. 2012. "Effects of Display Factors on Chinese Proofreading Performance and Preferences." *Ergonomics* 55 (11): 1316–1330.
Chinese Lexicon. 2003 [CD-ROM] Produced by State Key Laboratory of Intelligent Technology and Systems, Tsinghua University and Institute of Automation, Chinese Academy of Sciences. Retrieved from Chinese Linguistic Data Consortium. http://www.chineseldc.org/
Engbert, R., A. Longtin, and R. Kliegl. 2002. "A Dynamical Model of Saccade Generation in Reading Based on Spatially Distributed Lexical Processing." *Vision Research* 42 (5): 621–636.
Geng, L. 2010. "The Research of Network Typos from the Perspective of Pinyin Input Method." Master thesis, Shanghai Normal University.
Hatta, T., A. Kawakami, and K. Tamaoka. 2002. "Errors in Writing Japanese Kanji: A Comparison of Japanese Schoolchildren, Colleage Students and Second-Language Learners of Japanese." *Asia Pacific Journal of Speech, Language and Hearing* 7: 157–166.
Henderson, J. M., and F. Ferreira. 1990. "Effects of Foveal Processing Difficulty on the Perceptual Span in Reading: Implications for Attention and Eye Movement Control." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16 (3): 417–429.
Hoosain, R. 1992. "Psychological Reality of the Word in Chinese." In *Language Processing in Chinese*, edited by H. C. Chen and O. J. L. Tzeng, 111–130. Amsterdam, the Netherlands: Elsevier.
Kaakinen, J. K., and J. Hyönä. 2010. "Task Effects on Eye Movements during Reading." *Journal of Experimental Psychology: Learning Memory and Cognition* 36 (6): 1561–1566.
Kennison, S. M., and C. Clifton. 1995. "Determinants of Parafoveal Preview Benefit in High and Low Working Memory Capacity Readers: Implications for Eye Movement Control." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21 (1): 68–81.
Kliegl, R., A. Nuthmann, and R. Engbert. 2006. "Tracking the Mind during Reading: The Influence of Past, Present, and Future Words on Fixation Durations." *Journal of Experimental Psychology: General* 135 (1): 12–35.
Li, X., K. Bicknell, P. Liu, W. Wei, and K. Rayner. 2014. "Reading Is Fundamentally Similar across Disparate Writing Systems: A Systematic Characterization of How Words and Characters Influence Eye Movements in Chinese Reading." *Journal of Experiment Psychology: General* 143 (2): 895–913.
Li, X., P. Liu, and K. Rayner. 2011. "Eye Movement Guidance in Chinese Reading: Is There a Preferred Viewing Location?" *Vision Research* 51: 1146–1156.
Li, X., K. Rayner, and K. R. Cave. 2009. "On the Segmentation of Chinese Words during Reading." *Cognitive Psychology* 58: 525–552.
Liu, P., W. Li, N. Lin, and X. Li. 2013. "Do Chinese Readers Follow the National Standard Rules for Word Segmentation during Reading?" *PLoS One* 8 (2): e55440.
Logan, F. A. 1999. "Errors in Copy Typewriting." *Journal of Experimental Psychology: Human Perception and Performance* 25 (6): 1760–1773.
McDonald, S. A., R. H. S. Carpenter, and R. C. Shillcock. 2005. "An Anatomically Constrained, Stochastic Model of Eye Movement Control in Reading." *Psychological Review* 112 (4): 814–840.

Perea, M., and J. Acha. 2009. "Space Information Is Important for Reading." *Vision Research* 49 (15): 1994–2000.

Perfetti, C. A., Y. Liu, and L. H. Tan. 2005. "The Lexical Constituency Model: Some Implications of Research on Chinese for General Theories of Reading." *Psychological Review* 112 (1): 43–59.

Pollatsek, A., E. D. Reichle, and K. Rayner. 2006. "Tests of the E-Z Reader Model: Exploring the Interface between Cognition and Eye-Movement Control." *Cognitive Psychology* 52 (1): 1–56.

Pynte, J., A. Kennedy, and S. Ducrot. 2004. "The Influence of Parafoveal Typographical Errors on Eye Movements in Reading." *European Journal of Cognitive Psychology* 16 (1–2): 178–202.

Rayner, K. 1979. "Eye Guidance in Reading: Fixation Locations within Words." *Perception* 8: 21–30.

Rayner, K. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research." *Psychological Bulletin* 124 (3): 372–422.

Rayner, K. 2009. "Eye Movements and Attention in Reading, Scene Perception, and Visual Search." *The Quarterly Journal of Experimental Psychology* 62 (8): 1457–1506.

Rayner, K., X. Li, B. J. Juhasz, and G. Yan. 2005. "The Effect of Word Predictability on the Eye Movements of Chinese Readers." *Psychonomic Bulletin and Review* 12 (6): 1089–1093.

Rayner, K., X. Li, and A. Pollatsek. 2007. "Extending the E-Z Reader Model of Eye Movement Control to Chinese Readers." *Cognitive Science* 31: 1021–1033.

Rayner, K., T. Warren, B. J. Juhasz, and S. P. Liversedge. 2004. "The Effect of Plausibility on Eye Movements in Reading." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30 (6): 1290–1301.

Reichle, E. D., A. Pollatsek, D. L. Fisher, and K. Rayner. 1998. "Toward a Model of Eye Movement Control in Reading." *Psychological Review* 105 (1): 125–157.

Reichle, E. D., K. Rayner, and A. Pollatsek. 2003. "The E-Z Reader Model of Eye-Movement Control in Reading: Comparisons to Other Models." *Behavioral and Brain Sciences* 26: 445–526.

Reichle, E. D., T. Warren, and K. McConnell. 2009. "Using E-Z Reader to Model Effects of Higher-Level Language Processing on Eye Movements during Reading." *Psychonomic Bulletin and Review* 16 (1): 1–21.

Reilly, R., and R. Radach. 2006. "Some Empirical Tests of an Interactive Activation Model of Eye Movement Control in Reading." *Cognitive Systems Research* 7 (1): 34–55.

Richter, E. M., R. Engbert, and R. Kliegl. 2006. "Current Advances in SWIFT." *Cognitive Systems Research* 7 (1): 23–33.

Schotter, E. R., K. Bicknell, I. Howard, R. Levy, and K. Rayner. 2014. "Task Effects Reveal Cognitive Flexibility Responding to Frequency and Predictability: Evidence from Eye Movements in Reading and Proofreading." *Cognition* 131 (1): 1–27.

Shu, H., W. Zhou, M. Yan, and R. Kliegl. 2011. "Font Size Modulates Saccade-Target Selection in Chinese Reading." *Attention, Perception & Psychophysics* 73: 482–490.

Tsai, J. L., and G. W. McConkie. 2003. "Where Do Chinese Readers Send Their Eyes?" In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, edited by J. Hyona, R. Radach, and H. Deubel, 159–176. Oxford: Elsevier Science.

Wei, W., X. Li, and A. Pollatsek. 2013. "Word Properties of a Fixated Region Affect Outgoing Saccade Length in Chinese Reading." *Vision Research* 80: 1–6.

White, S. J., and S. P. Liversedge. 2006. "Foveal Processing Difficulty Does Not Modulate Non-Foveal Orthographic Influences on Fixation Positions." *Vision Research* 46 (3): 426–437.

Yan, G., X. Bai, C. Zang, Q. Bian, L. Cui, W. Qi, K. Rayner, and S. P. Liversedge. 2012. "Using Stroke Removal to Investigate Chinese Character Identification during Reading: Evidence from Eye Movements." *Reading and Writing* 25 (5): 951–979.

Yan, M., R. Kliegl, E. M. Richter, A. Nuthmann, and H. Shu. 2010. "Flexible Saccade-Target Selection in Chinese Reading." *The Quarterly Journal of Experimental Psychology* 63 (4): 705–725.

Yan, G., H. Tian, X. Bai, and K. Rayner. 2006. "The Effect of Word and Character Frequency on the Eye Movements of Chinese Readers." *British Journal of Psychology* 97: 259–268.

Yang, H., and G. W. McConkie. 1999. "Reading Chinese: Some Basic Eye-Movement Characteristics." In *Reading Chinese Script*, edited by J. Wang, A. W. Inhoff, and H. Chen, 207–222. Mahwah, NJ: Lawrence Erlbaum.

Yang, J., A. Staub, N. Li, S. P. Wang, and K. Rayner. 2012. "Plausibility Effects When Reading One- and Two-Character Words in Chinese: Evidence from Eye Movements." *Journal of Experimental Psychology: Learning Memory and Cognition* 38 (6): 1801–1809.

Zang, C., F. Liang, X. Bai, G. Yan, and S. P. Liversedge. 2013. "Interword Spacing and Landing Position Effects during Chinese Reading in Children and Adults." *Journal of Experimental Psychology: Human Perception & Performance* 39 (3): 720–734.

Zola, D. 1984. "Redundancy and Word Perception during Reading." *Perception & Psychophysics* 36 (3): 277–284.