Read Writ
DOI 10.1007/s11145-014-9521-4

Additive effects of stimulus quality and word frequency on eye movements during Chinese reading

Pingping Liu · Xingshan Li · Buxin Han

Abstract Eye movements of Chinese readers were recorded for sentences in which high- and low-frequency target words were presented normally or with reduced stimulus quality in two experiments. We found stimulus quality and word frequency produced strong additive effects on fixation durations for target words. The results demonstrate that stimulus quality and word frequency affect different stages of processing (e.g., visual processing and lexical processing). These results are consistent with the findings of previous single-word lexical decision studies, which showed that stimulus quality manipulation primarily affects the early pre-attentive stage of visual processing, whereas word frequency affects lexical processes. We discuss these findings in terms of the role of stimulus quality in word recognition and in relation to the E-Z Reader model of eye movement control.

Keywords Eye movements · Stimulus quality · Word frequency · Lexical processing · Chinese reading

Introduction

In order to infer the structure of information processing stages during reading, many studies have examined the joint effects of stimulus quality (whether a printed word is degraded) and word frequency on word recognition in single-word response tasks such as lexical decision and naming (e.g., Balota & Abrams, 1995; Bangert,

P. Liu · B. Han
Center on Aging Psychology, Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

X. Li (✉)
Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
e-mail: lixs@psych.ac.cn

Published online: 10 June 2014

Abrams, & Balota, 2012; Becker & Killion, 1977; O'Malley & Besner, 2008; O'Malley, Reynolds, & Besner, 2007; Plourde & Besner, 1997; Scaltritti, Balota, & Peressotti, 2013; Stanners, Jastrzembski, & Westbrook, 1975; Yap & Balota, 2007; Yap, Balota, Tse, & Besner, 2008). However, few studies have investigated the joint effects of these two variables in the context of natural reading. An important question is whether stimulus quality affects an early stage of lexical processing or the pre-attentive stage of visual processing that forms part of the influential E-Z Reader model (Reichle, Pollatsek, & Rayner, 2012; Reichle, Warren, & McConnell, 2009; Sheridan & Reingold, 2013; White & Staub, 2012). To test this question directly, the present study explored the joint effects of stimulus quality and word frequency on eye movements for target words during natural reading.

According to Sternberg's additive factors logic (Sternberg, 1969), two variables that have additive effects affect different stages of processing, whereas two variables that have interactive effects may affect at least one common stage of processing (but see McClelland, 1979, for an alternative explanation of additive effects). Many single-word recognition studies have investigated the influence of stimulus quality and word frequency on RT and the distribution of RT responses, and interpreted these results based on the additive factors logic.

However, this research has not produced consistent results with regard to the effects of stimulus quality and word frequency on word recognition. In studies using semantic categorization and pronunciation tasks, several researchers have reported an interactive effect of these two variables on RT (e.g., O'Malley et al., 2007; Yap & Balota, 2007, Experiment 2 & 3), which suggests that the two variables influence a common stage of lexical processing. In contrast, in the lexical decision task, numerous experiments have shown additive effects of these two variables on RT or the distribution of RT responses (Balota & Abrams, 1995; Becker & Killion, 1977, Experiment 3 & 4; O'Malley et al., 2007; Plourde & Besner, 1997; Stanners et al., 1975; Yap & Balota, 2007, Experiment 1). The indication, therefore, is that stimulus quality and word frequency affected different stages of processing (e.g., feature level and letter level) in this task. In order to reconcile these findings, Scaltritti et al. (2013) proposed that the joint effects of these two variables vary as a function of the experimental task or the type of stimuli, and that the patterns of additive or interactive effects that are typically observed support the notion of an adaptive and flexible lexical processor.

Most previous studies used single words as stimuli in lexical decision, semantic categorization or pronunciation tasks. These tasks are very different from normal reading. Hence, how the two variables influence lexical processing during normal sentence reading remains unclear. The present study therefore investigates the joint effects of these two variables in natural reading using measures of eye movements, as these provide an excellent online indication of the cognitive processes that underlie reading (Rayner, 1998). The findings from this research will reveal the influence of these variables on different stages of processing during reading, and are likely to have important implications for models of eye movement control in reading, such as the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle et al., 2009, 2012).

A central assumption of the E-Z Reader is that lexical processing determines when the eyes move forward from a word. The model postulates two separate stages of lexical processing, namely, L1 (the familiarity check) and L2 (the completion of lexical access). In addition, the model includes an early pre-attentive stage of visual processing during which visual features on the printed page are propagated from the retina to the brain. The E-Z Reader model has successfully accounted for a broad range of empirical findings in English, and has been successfully extended to account for eye movement control while reading in Chinese (Rayner, Li, & Pollatsek, 2007). Thus, the present study tests the influence of stimulus quality and word frequency on different stages of the processing of words when reading in Chinese reading within the framework of the E-Z Reader model.

Several recent studies have investigated the effect of stimulus quality on eye movement behavior by manipulating the contrast of a single target word within a sentence (Drieghe, 2008; Jordan, McGowan, & Paterson, 2012; Reingold & Rayner, 2006; Wang & Inhoff, 2010; White & Staub, 2012). Unsurprisingly, fixation durations are increased when a word is visually degraded. Reingold and Rayner (2006) tested a core assumption of the E-Z Reader model by manipulating the visual contrast of target words. According to the model, the factors that influence L1 for a given word affect the viewing times on that word, but not on the next word. Reingold and Rayner found that reduced contrast increased viewing times on the target words, but not on the next words. These results therefore support the assumption that stimulus quality may influence the initial stage of lexical processing, which is known as L1 within the E-Z Reader model. The results have since been replicated by several studies (Drieghe, 2008; White & Staub, 2012), and so it seems clear from this research that stimulus quality may affect at least the initial stage of lexical processing.

More recently, Sheridan and Reingold (2013) explored the conjoint effects of stimulus quality and word frequency on eye movements during reading. They found that stimulus quality and word frequency produced interactive effects on fixation time measures during first-pass reading, whereas these variables produced additive effects on later fixation time measures. In addition, Paterson, McGowan, and Jordan (2012) reported the interactive effects of stimulus quality and lexical frequency on eye movements in English reading by using a moving, gaze-contingent foveal-filtering technique. According to Sternberg's additive factors logic, these findings are consistent with stimulus quality and word frequency affecting a common stage (L1) of lexical processing.

However, Sheridan and Reingold (2013) reported that they could not rule out the possibility that stimulus quality may also affect the pre-attentive visual processing stage in the E-Z Reader model (see Fig. 1A). Thus, by using different contrast manipulations, the present study aimed to determine whether stimulus quality affects an early lexical-processing stage or an even earlier pre-attentive visual processing stage when reading Chinese. One key question is which measures could be used as indexes of different processing stages. 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, single fixation duration and gaze duration reflect first-pass reading time and the reader's initial encounter with a target word, whereas go-past time, total time and regression rates reflect later processing activities and integration processes. Thus, in the present study, these eye movement measures could provide good indications of effects that are due to an earlier pre-attentive visual processing stage.

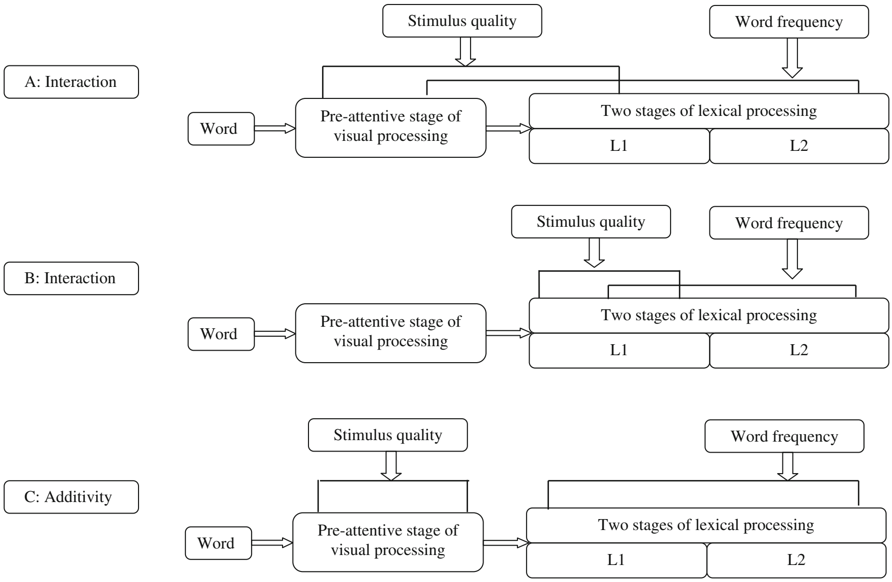

Fig. 1 Illustrations of the three assumptions of how stimulus quality and word frequency affect stages of processing with the framework of the E-Z Reader model

In the context of Chinese reading, how stimulus quality and word frequency affect word recognition is unknown, partly because there are lots of differences between Chinese and English. First, written Chinese is formed of strings of equally spaced box-like symbols called characters. Each character is formed from a series of individual strokes, and characters may differ in visual complexity due of variations in the number of strokes. Several studies have shown that the properties of characters (e.g., character frequency, character complexity) strongly affect the eye movements of Chinese readers (Li, Bicknell, Liu, Wei, & Rayner, 2014; Yan, Tian, Bai, & Rayner, 2006; Yang & McConkie, 1999). Second, unlike in English, there are no spaces between words in written Chinese, and Chinese readers segment and identify words based on lexical knowledge and contextual information (Hoosain, 1992; Li, Rayner, & Cave, 2009; Liu, Li, Lin, & Li, 2013). Given these differences between English and Chinese, it is quite possible that stimulus quality and word frequency might exert a qualitatively different influence on the processing of words in these two languages, and so the present research set out to examine the conjoint influence of these variables in Chinese.

While there are lots of differences between Chinese and English, several studies have shown the similar effects of word properties on eye movements between two orthographies (Inhoff & Rayner, 1986; Li et al., 2014; Rayner, Li, Juhasz, & Yan, 2005; Yan et al., 2006). For example, low frequency words are fixated longer than high frequency words (Inhoff & Rayner, 1986; Li et al., 2014; Paterson et al., 2012; Sheridan & Reingold, 2013; Yan et al., 2006), and words that are less predictable in context are fixated longer than more predictable words (e.g., Rayner et al., 2005; Rayner, Reichle, Stroud, & Pollatsek, 2006). Given these similarities in the findings on eye movements, as in English, it is possible that stimulus quality and word frequency might have an interactive effect in Chinese.

As noted above, word frequency produces a robust effect on lexical processing and eye movements in both English and Chinese (e.g., Inhoff & Rayner, 1986; Li et al., 2014; Reingold, Yang, & Rayner, 2010; Yan et al., 2006). Some studies have shown that word frequency may be associated with the processing of both L1 and L2 stages of word recognition in the E-Z Reader model (Reichle et al., 2009; Reingold et al., 2010; Staub, 2011). However, how stimulus quality affects the different stages of processing in Chinese reading remains unclear. According to Sternberg's additive factors logic, an interactive effect of stimulus quality and word frequency on fixation time measures would replicate the findings of Sheridan and Reingold (2013), which appear to show that stimulus quality can influence lexical processing (see Fig. 1A, B). In contrast, if these two variables were to produce an additive effect (see Fig. 1C), then such a pattern would support the assumption that the influence of stimulus quality is limited to the pre-attentive visual processing stage (Reichle et al., 2009, 2012; White & Staub, 2012).

To recapitulate, the present study used the same participant pool and stimulus set to directly explore the joint effects of word frequency and stimulus quality in Chinese reading with two different manipulations of stimulus quality. In Experiment 1, following Sheridan and Reingold's (2013) manipulation of visual contrast, participants' eye movements were recorded while they read sentences in which high- and low-frequency target words were presented normally (i.e., the normal condition) or with reduced stimulus quality (i.e., the faint-target condition). All of the other characters in the sentence frames were always presented normally in Experiment 1. By comparison, in Experiment 2, both the target words and sentence frames were degraded in the faint condition (see Table 1).

Experiment 1

Methods

Participants

Thirty-two native Chinese speakers (16 females; average age = 22.8 years; range 19.9–26.3 years) from several universities near the Institute of Psychology in Beijing were paid to participate in the experiment. All participants were proficient in Chinese and had either normal or corrected-to-normal vision.

Apparatus

Eye movements were recorded with an SR Research Ltd. Eyelink 1000 eye tracker. Viewing was binocular, but only the movements of the right eye were monitored. The sentences were presented in the center of the screen on a 21-inch CRT monitor (resolution: 1,024 × 768 pixels; refresh rate: 150 Hz) connected to a Dell PC. Participants were seated at a viewing distance of 58 cm from the computer monitor, and each Chinese character subtended a visual angle of approximately 1.2°.

Table 1 Example sentences

Condition	Sentence
Normal_HF	目前有两家民营企业正在*准备*开办小型航空公司。
	Currently, two private companies *prepare* to set up small airlines.
Normal_LF	目前有两家民营企业正在*策划*开办小型航空公司。
	Currently, two private companies *plan* to set up small airlines.
Experiment 1	
Fain-target_HF	目前有两家民营企业正在*准备*开办小型航空公司。
	Currently, two private companies *prepare* to set up small airlines.
Faint-target_LF	目前有两家民营企业正在*策划*开办小型航空公司。
	Currently, two private companies *plan* to set up small airlines.
Experiment 2	
Faint_HF	目前有两家民营企业正在*准备*开办小型航空公司。
	Currently, two private companies *prepare* to set up small airlines.
Faint_LF	目前有两家民营企业正在*策划*开办小型航空公司。
	Currently, two private companies *plan* to set up small airlines.

Note. Target words appear in italics (but not during the experiment). HF = high-frequency targets; LF = low-frequency targets. In Experiment 1, only the target words were presented at low contrast in the faint-target condition. In Experiment 2, the entire sentences were presented at low contrast in the faint condition.

Materials

Seventy-two pairs of high-frequency and low-frequency target words were embedded in sentence frames. Experimental sentences were 20–28 characters in

Table 2 Properties of the target words used in the study

	HF	LF	*t*	*p*
Word frequency	168.13	7.73	10.28	<.001
Word complexity	15.71	15.94	−.94	.35
Word predictability	.17	.17	.50	.62
First character complexity	7.74	7.75	−.04	.97
Second character complexity	7.97	8.19	−.75	.46
First character frequency	1,539	1,346	.93	.35
Second character frequency	1,429	1,222	1.00	.32
Acceptability rating	5.52	5.58	−1.00	.34

Word and character frequency are in occurrences per million. The number of individual strokes in a character is treated as the index of its visual complexity. Word complexity is the sum of the number of strokes of the first and second characters. Acceptability ratings were collected on a scale of 1 (unacceptable) to 7 (perfectly acceptable). The *t* values were generated from a paired-samples *t* test between items that belong to the HF group and items that belong to the LF group

HF = high-frequency targets, *LF* = low-frequency targets

length, and obtained from an online corpus.[1] Some of the sentences were revised slightly in order to remove semantic ambiguities. All of the target words were 2-character words, which were listed as words in the *Chinese Lexicon* (2003). The frequency of the high-frequency words was above 50 occurrences per million ($M = 168.13$), whereas that of the low-frequency words was below 16 occurrences per million ($M = 7.73$).

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 37 native Chinese speakers, who were given the sentence frame up to, but not including, the target word and asked to generate the next acceptable word in the sentence. Acceptability ratings for each sentence were obtained from 20 native speakers of Chinese on a scale of 1 (unacceptable) to 7 (perfectly acceptable). None of the participants who participated in the norming studies took part in the eye-tracking experiment. Finally, the high- and low-frequency words were paired and matched exactly (see Table 2) based on word complexity, word predictability, character complexity and character frequency (all ts < 1.1).

The sentences were displayed on a single line in 24-point Song font. For all conditions, the luminance of the background was 51.67 cm/d^2. The normal text, presented at a high contrast, had a luminance of 1.94 cm/d^2. The faint target words, presented at low contrast, had a luminance of 44.71 cm/d^2.

Procedure

Participants were tested individually. At the start of the experiment, participants performed a calibration procedure by looking at a sequence of three fixation points

[1] http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp?dir=xiandai. Center for Chinese Linguistics PKU.

randomly displayed horizontally across the middle of the computer screen. Following calibration, the gaze position error was smaller than .5° of visual angle. At the beginning of each trial, a white square (about 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 fixated on the white square successfully, a sentence was presented. Participants were instructed to read silently for comprehension and to press a button on a keypad when they finished reading the sentence. Comprehension questions were presented on the screen after 42 % of the sentences.

Design and data analysis

A 2 (stimulus quality: normal vs. faint) × 2 (frequency: high vs. low) counterbalanced design was used to enable each participant to view an equal number of sentences in each condition but read only one sentence for each paired item. Each participant read eight practice sentences and seventy-two experimental sentences. The experimental sentences were presented in a pseudorandom order for each participant, with a maximum of two items in the same condition appearing consecutively. Data were analyzed through repeated measures analyses of variance (ANOVAs) on both participants' means (F_1) and items' means (F_2).

Results and discussion

The average comprehension accuracy was 93.5 %, indicating that participants read and understood the sentences well. Approximately 3.7 % of fixations were excluded due to blinks and track losses, and 3.7 % were removed because they were shorter than 80 ms or longer than 1,000 ms. 11.7 % of trials in which there were blinks on the target words were discarded prior to analyses.

The following measures were used to examine processing times for the high- and low-frequency target words in the normal and faint conditions (see Table 3): (a) *first fixation duration* (the duration of the first first-pass fixation on the target word), (b) *single fixation duration* (cases in which the reader made only one first-pass fixation on the target word), (c) *gaze duration* (the sum of all first-pass fixations on the target word before moving to another word), (d) *go-past time* (the amount of time that the reader looks at the target word as well as any time spent rereading earlier parts of the sentence before moving ahead to inspect new portions of the sentence), (e) *total time* (the sum of all fixations on the target word, including regressions), and (f) *incoming regression rates* (the probability of regressions into the target word). The data of fixation durations were subjected to Bayesian analysis of variance (ANOVA; Rouder, Morey, Speckman, & Province, 2012; Wetzels, Grasman, & Wagenmakers, 2012), which allowed us to demonstrate the strength for the interpretation of null effects in the ANOVA model. To test for each of the main effects and for the interaction effects, we calculated the Bayes factor by comparing the larger model with the null model (e.g., BF[MFSW: MN]). Following Rouder et al. (2009), we considered odds greater than 10:1 to provide strong evidence for the larger model and for odds less than 3:1 to provide little evidence for this model.

Table 3 Eye movement measures for the target word for each condition

Measures	Normal		Faint	
	HF	LF	HF	LF
Experiment 1				
First fixation duration (ms)	272 (6.7)	281 (7.1)	368 (12.2)	392 (12.5)
Single fixation duration (ms)	269 (6.7)	284 (8.5)	392 (14.0)	428 (14.8)
Gaze duration (ms)	317 (11.8)	342 (14.1)	480 (16.6)	548 (20.5)
Go-past time (ms)	426 (22.6)	440 (22.0)	571 (21.4)	651 (23.4)
Total time (ms)	576 (30.7)	630 (33.9)	859 (45.1)	898 (46.8)
Incoming regression rates	.18 (.03)	.22 (.02)	.32 (.03)	.32 (.03)
Experiment 2				
First fixation duration (ms)	268 (6.6)	283 (7.8)	308 (7.4)	323 (7.5)
Single fixation duration (ms)	266 (6.5)	285 (8.4)	306 (6.8)	321 (8.2)
Gaze duration (ms)	308 (8.8)	344 (11.7)	352 (13.2)	371 (9.5)
Go-past time (ms)	394 (16.9)	426 (19.0)	447 (21.4)	475 (20.7)
Total time (ms)	494 (33.3)	550 (33.9)	572 (33.6)	604 (39.6)
Incoming regression rates	.17 (.02)	.18 (.02)	.17 (.02)	.16 (.02)

Standard errors are shown in parentheses

HF = high-frequency targets, *LF* = low-frequency targets

Fixation times and regression rates

There were significant main effects of stimulus quality and word frequency for all of the reading time measures (see Table 3). Reading times were significantly longer in the faint-target condition than in the normal condition for all five measures [first fixation duration: $F_1(1, 31) = 103.68$, $MSE = 3{,}345.78$, $p < .001$, $\eta_p^2 = .77$, $F_2(1, 71) = 293.62$, $MSE = 2{,}737.79$, $p < .001$, $\eta_p^2 = .81$; single fixation duration: $F_1(1, 31) = 106.87$, $MSE = 5{,}342.58$, $p < .001$, $\eta_p^2 = .78$, $F_2(1, 71) = 321.40$, $MSE = 4{,}253.09$, $p < .001$, $\eta_p^2 = .82$; gaze duration: $F_1(1, 31) = 109.24$, $MSE = 9{,}960.75$, $p < .001$, $\eta_p^2 = .78$, $F_2(1, 71) = 478.14$, $MSE = 5{,}071.01$, $p < .001$, $\eta_p^2 = .87$; go-past time: $F_1(1, 31) = 87.46$, $MSE = 11{,}593.40$, $p < .001$, $\eta_p^2 = .74$, $F_2(1, 71) = 122.78$, $MSE = 17{,}806.16$, $p < .001$, $\eta_p^2 = .63$; total time: $F_1(1, 31) = 86.98$, $MSE = 27{,}912.93$, $p < .001$, $\eta_p^2 = .74$, $F_2(1, 71) = 199.02$, $MSE = 26{,}995.42$, $p < .001$, $\eta_p^2 = .74$]. Reading times were significantly longer for low-frequency words than for high-frequency words for all five measures [first fixation duration: $F_1(1, 31) = 7.79$, $MSE = 1{,}201.94$, $p = .009$, $\eta_p^2 = .20$, $F_2(1, 71) = 9.87$, $MSE = 1{,}789.74$, $p = .002$, $\eta_p^2 = .12$; single fixation duration: $F_1(1, 31) = 21.38$, $MSE = 988.08$, $p < .001$, $\eta_p^2 = .41$, $F_2(1, 71) = 8.69$, $MSE = 3{,}271.20$, $p = .004$, $\eta_p^2 = .11$; gaze duration: $F_1(1, 31) = 32.29$, $MSE = 2{,}096.74$, $p < .001$, $\eta_p^2 = .51$, $F_2(1, 71) = 24.51$, $MSE = 6{,}561.77$, $p < .001$, $\eta_p^2 = .26$; go-past time: $F_1(1, 31) = 13.39$, $MSE = 5{,}287.19$, $p = .001$, $\eta_p^2 = .30$, $F_2(1, 71) = 11.71$, $MSE = 13{,}825.88$, $p = .001$, $\eta_p^2 = .14$; total time: $F_1(1, 31) = 6.71$, $MSE = 10{,}240.40$, $p = .01$, $\eta_p^2 = .18$, $F_2(1, 71) = 4.32$, $MSE = 31{,}383.18$, $p = .04$, $\eta_p^2 = .06$]. Although the stimulus quality $\times$ word

Table 4 Ratio of Bayes factors for the two-way ANOVA

Measures	BF (M_{FSW}: M_N)	BF (M_{SW}: M_N)	BF (M_{S+W}: M_N)	BF (M_S: M_N)	BF (M_W: M_N)
Experiment 1					
First fixation duration	2.43×10^{76}	.22	1.18×10^{77}	2.30×10^{76}	1.52
Single fixation duration	1.88×10^{94}	.11	4.61×10^{94}	1.56×10^{93}	2.01
Gaze duration	1.76×10^{94}	1.44	1.18×10^{94}	3.48×10^{88}	6.81×10^{3}
Go-past time	1.74×10^{35}	1.94	1.43×10^{35}	4.70×10^{33}	11.9
Total time	1.12×10^{45}	.07	1.56×10^{46}	1.14×10^{46}	.66
Experiment 2					
First fixation duration	2.59×10^{14}	.071	3.58×10^{15}	4.52×10^{14}	4.91
Single fixation duration	7.03×10^{11}	.078	1.29×10^{13}	4.49×10^{11}	14.45
Gaze duration	4.20×10^{4}	.131	2.99×10^{5}	2,758	84.56
Go-past time	5.22	.069	79.59	79.33	.828
Total time	426	.085	4,452	969	3.66

M_{FSW} = the full model that contains the main and interaction effects; M_{SW} = the model that contains only the interaction effects; M_{S+W} = the model that contains only the main effects; M_S = the model that contains only the effects of stimulus quality; M_W = the model that contains only the effects of word frequency; M_N = the null model

frequency interaction was not significant for first fixation duration [$F_1(1, 31) = 1.61$, $p = .21$, $F_2(1, 71) = 1.18$, $p = .28$], single fixation duration [$F_1(1, 31) = 2.82$, $p = .10$, $F_2(1, 71) = 2.00$, $p = .16$] or total time [$F_1(1, 31) = .21$, $p = .65$, $F_2(1, 71) = .36$, $p = .55$], it was significant for gaze duration [$F_1(1, 31) = 7.42$, $MSE = 1{,}994.86$, $p = .01$, $\eta_p^2 = .19$, $F_2(1, 71) = 4.65$, $MSE = 6{,}583.75$, $p = .04$, $\eta_p^2 = .06$] and go-past time [significant by participants, $F_1(1, 31) = 5.25$, $MSE = 6{,}682.72$, $p = .03$, $\eta_p^2 = .15$, but not by items, $F_2(1, 71) = 2.68$, $MSE = 21{,}748.18$, $p = .11$, $\eta_p^2 = .04$]. Additionally, the effect of stimulus quality on incoming regression rates was significant, $F_1(1, 31) = 21.87$, $MSE = .02$, $p < .001$, $\eta_p^2 = .41$, $F_2(1, 71) = 36.21$, $MSE = .03$, $p < .001$, $\eta_p^2 = .34$. There were more regressions into the target words in the faint-target condition than in the normal condition. None of the other effects were reliable.

Bayesian analyses

We further evaluated significant interactions by calculating the Bayes factors (Rouder et al., 2012; Wetzels et al., 2012) to enable comparisons of models within an ANOVA design (see Table 4). We calculated the Bayes factor between the model which contains only the main effects and the full model. The Bayes factors for single fixation duration, gaze duration and go-past time were just 2.45:1, .67:1 and .82:1, respectively, which cannot support what model was preferred (Rouder, Speckman, Sun, Morey, & Iverson, 2009). Nevertheless, models excluding the interaction were preferred for first fixation duration (4.86:1) and total time (13.93:1).

Thus, according to the first fixation duration which could be used as an index of the earlier pre-attentive visual processing stage, it was apparent from the findings of Experiment 1 that stimulus quality and word frequency produced additive effects in Chinese reading. However, the interactive effects of the two variables in single fixation duration and gaze duration were very weak according to these Bayesian analyses. The experiment did not replicate the findings of Sheridan and Reingold (2013), and Paterson et al. (2012), which state that stimulus quality affects the stages of lexical processing (see Fig. 1A, B). We found that stimulus quality may affect an even earlier pre-attentive visual processing stage (see Fig. 1C). In order to confirm this finding, another contrast manipulation was used in Experiment 2. In this experiment, in the faint condition, the sentence frames and the target words had the same luminance.

Experiment 2

Methods

Participants

Thirty-two native Chinese speakers (17 females; average age = 22.0 years; range 18.9–25.5 years) were recruited from the same participant pool as that in Experiment 1 and were paid to take part in the experiment. None of them had participated in Experiment 1.

Apparatus, material, procedure and design

The apparatus, materials, procedure and design were identical to those used in Experiment 1, except that the sentence frames and the target words had the same luminance for each condition.

Results and discussion

The average comprehension accuracy was 94.6 %. Approximately 1.3 % of fixations were excluded due to blinks and track losses, and 2.2 % were removed because they were shorter than 80 ms or longer than 1,000 ms. 6.5 % of trials in which there were blinks on the target words were discarded prior to analyses.

Fixation times and regression rates

First fixation duration, single fixation duration, gaze duration, go-past time and total time produced very similar data patterns, and so are discussed together. Main effects of stimulus quality and word frequency were obtained for all three measures (see Table 3), but the stimulus quality × word frequency interaction was not significant for any of these measures (all Fs < 1.8). Reading times were significantly longer in the faint condition than in the normal condition for all five measures [first fixation duration: $F_1(1, 31) = 51.18, MSE = 991.10, p < .001, \eta_p^2 = .62, F_2(1, 71) = 75.44$, $MSE = 1{,}511.81$, $p < .001$, $\eta_p^2 = .52$; single fixation duration: $F_1(1, 31) = 40.18$, $MSE = 1{,}126.88$, $p < .001$, $\eta_p^2 = .56$, $F_2(1, 71) = 57.25$, $MSE = 1{,}805.40$, $p < .001$, $\eta_p^2 = .45$; gaze duration: $F_1(1, 31) = 22.50$, $MSE = 1{,}828.70$, $p < .001$, $\eta_p^2 = .42$, $F_2(1, 71) = 26.79$, $MSE = 3{,}232.58$, $p < .001$, $\eta_p^2 = .27$; go-past time: $F_1(1, 31) = 26.20$, $MSE = 3{,}213.66$, $p < .001$, $\eta_p^2 = .46$, $F_2(1, 71) = 11.97$, $MSE = 17{,}721.93$, $p = .001$, $\eta_p^2 = .14$; total time: $F_1(1, 31) = 19.04$, $MSE = 7{,}227.41$, $p < .001$, $\eta_p^2 = .38$, $F_2(1, 71) = 13.23$, $MSE = 25{,}055.94$, $p = .001$, $\eta_p^2 = .16$]. Reading times were also significantly longer for low-frequency words than for high-frequency words [first fixation duration: $F_1(1, 31) = 13.53, MSE = 532.73$, $p = .001, \eta_p^2 = .30, F_2(1, 71) = 9.62, MSE = 1{,}641.44, p = .003, \eta_p^2 = .12$; single fixation duration: $F_1(1, 31) = 14.40$, $MSE = 671.46$, $p = .001$, $\eta_p^2 = .32$, $F_2(1, 71) = 9.90, MSE = 2{,}142.17, p = .002, \eta_p^2 = .12$; gaze duration: $F_1(1, 31) = 27.76$, $MSE = 888.90, p < .001, \eta_p^2 = .47, F_2(1, 71) = 10.23, MSE = 5{,}758.66, p = .002$, $\eta_p^2 = .13$; go-past time: $F_1(1, 31) = 9.21$, $MSE = 3{,}151.68$, $p = .005$, $\eta_p^2 = .23$, $F_2(1, 71) = 4.11, MSE = 19{,}710.80, p = .05, \eta_p^2 = .06$; total time: $F_1(1, 31) = 8.13$, $MSE = 7{,}574.27, p = .008, \eta_p^2 = .21, F_2(1, 71) = 6.05, MSE = 28{,}671.65, p = .02$, $\eta_p^2 = .08$]. Additionally, incoming regression rates did not differ across conditions (all Fs < 1.0).

Bayesian analyses

We further evaluated non-significant interactions by calculating the Bayes factors to enable comparisons of models within an ANOVA design (see Table 4). The model (i.e., M_{S+W}) which includes main effects of stimulus quality and word frequency produced the highest Bayes factor for all three measures (first fixation duration: 3.58×10^{15}:1; single fixation duration: 1.29×10^{13}:1; gaze duration: 2.99×10^{5}:1; go-past time: 79.59; total time: 4,463:1). We also calculated the Bayes factor between the model with only main effects (i.e., M_{S+W}) and the full model (i.e., M_{FSW}), whose interpretation is straightforward. Models excluding the interaction were preferred (first fixation duration: 13.82:1; single fixation duration: 18.35:1; gaze duration: 7.12:1; go-past time: 15.25:1; total time: 10.45:1). Thus, both the ANOVAs and Bayesian analyses showed the reliable additive effects of stimulus quality and word frequency which affect separate stages of processing during Chinese reading.

General discussion

In the present study, we explored the joint effects of stimulus quality and word frequency on word recognition in Chinese with two different contrast manipulations. The main finding of this study is that stimulus quality strongly affected fixation times on a target word independently of the frequency of the target word during reading. This finding is consistent with previous findings from single-word recognition studies that used lexical decision tasks to demonstrate additive effects of stimulus quality and word frequency on RT (e.g., Balota & Abrams, 1995; Bangert et al., 2012; O'Malley et al., 2007; Plourde & Besner, 1997; Stanners et al., 1975; Yap & Balota, 2007). The detailed measures of eye movements used in this study provide fresh insight into the claim that decisions about when to move the eyes are strongly influenced by how easily words can be identified. Fixation times increased when a word was visually degraded or of low lexical frequency. According to the additive factors logic, the additivity of stimulus quality and word frequency implies that the two variables affect the different stages of processing.

The results showed that stimulus quality produced reliable effects on eye movement behavior in sentence reading. An important finding from the present study was that fixation durations were longer when a word was faint than when it was presented normally. The result also showed that high-quality foveal input is essential for a range of complex visual tasks, especially reading (Drieghe, 2008; Jordan et al., 2012; Reingold & Rayner, 2006; Sheridan & Reingold, 2013; Wang & Inhoff, 2010; White & Staub, 2012).

In addition, the present study indicates that the length of time that the eyes remain on a word is strongly influenced by the frequency of the fixated word. This finding replicates the findings from a number of prior studies in Chinese (Li et al., 2014; Yan et al., 2006; Yang & McConkie, 1999), and provides further evidence that cognitive/linguistic processing determines decisions about when to move the eyes when reading in Chinese.

Finally, the present study showed additive effects of stimulus quality and word frequency on word recognition. A comparison of Bayes factors provides no reason to believe that an interaction between the two variables was present in our data. These findings indicate that stimulus quality also influences an early stage of lexical processing in Chinese. In both experiments, our results are inconsistent with the findings of Sheridan and Reingold (2013) and Paterson et al. (2012), who observed interactions between stimulus quality and word frequency on first fixation durations. Unlike in studies conducted in English, we did not observe strong interactive effects on fixation times on target words in both experiments.

One plausible interpretation for this inconsistency may be the different characteristics of reading Chinese compared to English. As noted above, unlike in English, there are no spaces between words in Chinese, and the number of possible word candidates cannot be constrained by visual cues to word boundaries. Chinese readers do not always agree on where word boundaries are, and they tend to use the lexical knowledge and sentence context to segment sentences into individual words (Hoosain, 1992; Li et al., 2009; Liu et al., 2013). Li et al. (2009) proposed a word segmentation and recognition model which assumes that Chinese word recognition involves multiple levels of processing consisting of a visual perception level, a character recognition level, and a word segmentation and recognition level. Word segmentation and recognition are interactive processes. In Experiment 1, only the target words were degraded in the faint-target condition. This contrast manipulation may facilitate the word segmentation of these target words which could improve the processes of word recognition. Hence, there was a tendency of interactive effect of stimulus quality and word frequency, although the effect was weak. However, in Experiment 2, both the target words and sentences were degraded in the faint condition. This manipulation of stimulus quality did not provide cues to word segmentation which must be vital for successful word identification. Thus, only additive effects of the two variables were observed in the second experiment. Furthermore, the additivity may reflect serially organized processes in which some processes start only after the previous process has finished. Indeed, this early stage of visual processing may be relatively similar for words that are presented clearly and words that are visually degraded, and this may explain why similar effects of word frequency are observed in fixation durations for these words in Experiment 2.

Moreover, these two experiments suggest that the influence of stimulus quality on word recognitions is different when the contrast manipulation is applied to all the words in a sentence rather than just a single target word. Indeed, White and Staub (2012) suggested that manipulating contrast of only a single word in a sentence might have an artifactual influence on word processing. According to this account, a target word may be emphasized if the target word has a different visual format from that of the other words in a sentence, and this may influence how the target word is processed. Our results showed that fixation durations were longer when lower contrast was applied to just the target word (Experiment 1) than when applied to the entire sentence (Experiment 2). One possibility is that when entire sentences were degraded, readers could anticipate the contrast of the target words and this may have affected their processing fluency, and also their word recognition strategies.

Based on the additive factors logic, these two variables could have affected two separate, discrete, and serially organized stages. When considered within the framework of the E-Z Reader model, our results imply that the manipulation of stimulus quality primarily affects an early pre-attentive visual processing stage, whereas word frequency primarily affects lexical or higher-level processes. Indeed, the present findings seem to be compatible with the model depicted in Fig. 1C. Because the present study relies on Sternberg's additive factors logic, it is important to remember there are some limitations to this approach. Separate stages imply additive effects, but additive effects may not necessarily imply separate stages. For example, a class of cascaded interactive models can also produce additive effects (McClelland, 1979). Roberts and Sternberg (1993) discussed this issue in detail and suggested additional tests to rule out cascaded processing. More importantly, they proposed that cascaded models had difficulty in accommodating additivity in the higher order moments, such as variance and skewness. However, our claim that stimulus quality influences the pre-attentive visual processing stage could be examined further by performing the appropriate computations simulations in E-Z Reader. Indeed, in general it seems clear that further work is required to fully establish the effects of stimulus quality on eye movements while reading.

The present findings may also have further implications for the future development of the E-Z Reader model. The model relies heavily on word frequency as an input for simulating fixation times in reading, and assumes that the two stages of lexical processing (L1 and L2) in the model are strongly influenced by word frequency (Reichle et al., 1998, 2009, 2012; Staub, 2011). It should be apparent that if other variables, such as stimulus quality, also influence fixation times on words, then such factors should also be included in the model in order to even better predict fixation times during reading. Indeed, establishing how stimulus quality influences fixation durations on words could make an important contribution to developing a more comprehensive model of eye movement control in reading.

In summary, prior studies provided an important step forward in understanding how stimulus quality and word frequency conjointly influence eye movement behavior when reading English. Here we addressed the same issue in the context of the logographic writing system by using two different contrast manipulations. Unlike in Sheridan and Reingold's study, the results from the present study suggest that these two variables can have additive effects on word recognition and eye movement behavior in Chinese reading. We have argued that the effects of these visual and linguistic variables provide insights into the nature of the different stages of processing during Chinese reading, and also may place important constraints on models of eye movement control during reading.

Acknowledgments This research was supported by the Knowledge Innovation Program of the Chinese Academic Sciences (KSCX2-YW-BR-6), by a Grant from the Natural Science Foundation of China (31070904), and by a grant from Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (Y3CX121005). We thank Kevin Paterson for very helpful comments on an earlier version of this article. We thank Keith Rayner and an anonymous reviewer for helpful input on this paper. We also thank Yanping Liu for his helpful comments concerning statistical issues addressed in this article.

References

Balota, D. A., & Abrams, R. A. (1995). Mental chronometry: Beyond onset latencies in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1289–1302. doi:10.1037//0278-7393.21.5.1289.

Bangert, A. S., Abrams, R. A., & Balota, D. A. (2012). Reaching for words and nonwords: Interactive effects of word frequency and stimulus quality on the characteristics of reaching movements. *Psychonomic Bulletin & Review, 19*, 513–520. doi:10.3758/s13423-012-0234-x.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 389–401. doi:10.1037/0096-1523.3.3.389.

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/.

Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin & Review, 15*, 856–860. doi:10.3758/Pbr.15.4.856.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 111–130). North-Holland: Elsevier.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40*, 431–439.

Jordan, T. R., McGowan, V. A., & Paterson, K. B. (2012). Reading with a filtered fovea: The influence of visual quality at the point of fixation during reading. *Psychonomic Bulletin & Review, 19*, 1078–1084. doi:10.3758/s13423-012-0307-x.

Li, X., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (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*, 895–913. doi:10.1037/a0033580.

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58*, 525–552. doi:10.1016/j.cogpsych.2009.02.003.

Liu, P., Li, W., Lin, N., & Li, X. (2013). Do Chinese readers follow the national standard rules for word segmentation during reading? *PLoS One, 8*, e55440. doi:10.1371/journal.pone.0055440.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 186*, 287–330. doi:10.1037/0033-295X.86.4.287.

O'Malley, S., & Besner, D. (2008). Reading aloud: Qualitative differences in the relation between stimulus quality and word frequency as a function of context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1400–1411. doi:10.1037/A0013084.

O'Malley, S., Reynolds, M. G., & Besner, D. (2007). Qualitative differences between the joint effects of stimulus quality and word frequency in reading aloud and lexical decision: Extension to Yap and Balota (2007). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 451–458. doi:10.1037/0278-7393.33.2.451.

Paterson, K. B., McGowan, V. A., & Jordan, T. R. (2012). Eye movements reveal effects of visual content on eye guidance and lexical access during reading. *PLoS One, 7*, e41766. doi:10.1371/journal.pone.0041766.

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology, 51*, 181–194. doi:10.1037/1196-1961.51.3.181.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1–56. doi:10.1016/j.cogpsych.2005.06.001.

Rayner, K. (1998). Eye movements in reading and information processing: 20 Years of research. *Psychological Bulletin, 124*, 372–422. doi:10.1037//0033-2909.124.3.372.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506. doi:10.1080/17470210902816461.

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review, 12*, 1089–1093.

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z reader model of eye movement control to Chinese readers. *Cognitive Science, 31*, 1021–1033.

Rayner, K., Reichle, E. D., Stroud, M. J., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*, 448–465. doi:10.1037/0882-7974.21.3.448.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125–157. doi:10.1037/0033-295X.105.1.125.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye-mind link. *Psychological Review, 119*, 155–185. doi:10.1037/a0026473.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z reader to model effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*, 1–21. doi:10.3758/PBR.16.1.1.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17*, 742–746. doi:10.1111/j.1467-9280.2006.01775.x.

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1677–1683. doi:10.1037/a0019959.

Roberts, S., & Sternberg, S. (1993). The meaning of additive reaction-time effects: Tests of three alternatives. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (pp. 611–653). Cambridge, MA: MIT Press.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356–374. doi:10.1016/j.jmp.2012.08.001.

Rouder, J. N., Speckman, P. L., Sun, D. C., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237. doi:10.3758/Pbr.16.2.225.

Scaltritti, M., Balota, D. A., & Peressotti, F. (2013). Exploring the additive effects of stimulus quality and word frequency: The influence of local and list-wide prime relatedness. *Quarterly Journal of Experimental Psychology, 66*, 91–107. doi:10.1080/17470218.2012.698628.

Sheridan, H., & Reingold, E. M. (2013). A further examination of the lexical-processing stages hypothesized by the E-Z Reader model. *Attention, Perception, & Psychophysics, 75*, 407–414. doi:10.3758/s13414-013-0442-0.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior, 14*, 259–264. doi:10.1016/S0022-5371(75)80069-7.

Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General, 140*, 407–433. doi:10.1037/A0023517.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315. doi:10.1016/0001-6918(69)90055-9.

Wang, C. A., & Inhoff, A. W. (2010). The influence of visual contrast and case changes on parafoveal preview benefits during reading. *Quarterly Journal of Experimental Psychology, 63*, 805–817. doi:10.1080/17470210903147494.

Wetzels, R., Grasman, R. P. P. P., & Wagenmakers, E. J. (2012). A default bayesian hypothesis test for ANOVA designs. *American Statistician, 66*, 104–111. doi:10.1080/00031305.2012.695956.

White, S. J., & Staub, A. (2012). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 603–617. doi:10.1037/A0025338.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97*, 259–268. doi:10.1348/000712605X70066.

Yang, H., & McConkie, G. W. (1999). Reading Chinese: Some basic eye-movement characteristics. In J. Wang, A. W. Inhoff, & H. Chen (Eds.), *Reading Chinese script* (pp. 207–222). Mahwah, NJ: Lawrence Erlbaum Associates.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 274–296. doi:10.1037/0278-7393.33.2.274.

Yap, M. J., Balota, D. A., Tse, C. S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: Evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 495–513. doi:10.1037/0278-7393.34.3.495.