Research Article

Psychological Science
2018, Vol. 29(2) 278–287

DOI: 10.1177/0956797617734827
www.psychologicalscience.org/PS
SAGE

Eye-Movement Evidence for Object-Based Attention in Chinese Reading

Yanping Liu[1] and Erik D. Reichle[2]
[1]Department of Psychology, Sun Yat-sen University, and [2]Department of Psychology, Macquarie University

Abstract
Is attention allocated to only one word or to multiple words at any given time during reading? The experiments reported here addressed this question using a novel paradigm inspired by classic findings on object-based attention. In Experiment 1, participants ($N = 18$) made lexical decisions about one of two spatially colocated Chinese words or nonwords. Our main finding was that only the attended word's frequency influenced response times and accuracy. In Experiment 2, participants ($N = 30$) read target words embedded in two spatially colocated Chinese sentences. Our key finding here was that only target-word frequencies influenced looking times and fixation positions. These results support the hypothesis that words are attended in a strictly serial (and perhaps object-based) manner during reading. The theoretical implications of this conclusion are discussed in relation to models of eye-movement control during reading and the conceptualization of words as visual objects.

Keywords
attention, Chinese reading, eye movements, lexical decision, open data, open materials

Received 1/10/17; Revision accepted 8/30/17

One of the central questions in reading research concerns how attention is allocated during reading. More specifically, is attention allocated in a serial manner, to only one word at any given time? Or is it instead allocated simultaneously to multiple words? Although these questions might appear straightforward, they have important implications for our understanding of how attention and lexical processing guide the eyes during reading and for the methodology of using eye movements to make inferences about attention and language processing during reading (Rayner, 2009). Both the serial- and parallel-attention theoretical perspectives have proponents who have advanced arguments supporting their positions (see Radach & Kennedy, 2013; Reichle, Liversedge, Pollatsek, & Rayner, 2009). The purpose of this article is to further inform this debate by reporting the results of two experiments designed to provide a novel test of the serial-allocation hypothesis—a test inspired by Duncan's (1984) original demonstration that when viewing two spatially overlapping objects, people can more easily attend to two features of one object than one feature from each of two objects (Ciaramitaro, Mitchell, Stoner, Reynolds, & Boynton, 2011; see Chen, 2012, for a review). Before disclosing how this was done, however, it is first necessary to provide some background.

The debate about attention allocation during reading came to the fore with the development of two classes of computational models of eye-movement control in reading—the serial-attention (e.g., E-Z Reader; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2012) and attention-gradient (e.g., saccade generation with inhibition by foveal target, or SWIFT; Engbert, Nuthmann, Richter, & Kliegl, 2005) models, which instantiate the serial- and parallel-allocation hypotheses, respectively. Both classes of models provide quantitative accounts of eye movements during reading, despite doing so using very different assumptions about how attention and lexical processing are related to ongoing decisions about when and where to

Corresponding Author:
Yanping Liu, Department of Psychology, Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Sun Yat-sen University, 135 Xingang Xi Rd., Guangzhou, China, 510275
E-mail: liuyp33@mail.sysu.edu.cn

move the eyes. For example, according to E-Z Reader, words are attended to and processed in a strictly serial manner; the lexical identification of one word causes attention to shift to the next word, so that its processing can begin. In contrast, the core assumption of SWIFT is that attention is a gradient or flexible "zoom lens" (Schad & Engbert, 2012) that supports the simultaneous processing of a few words, causing the gradient to progress down a line of text as individual words are identified.

Most empirical attempts to discriminate between these two accounts of attention allocation have examined the spatial extent of lexical processing. For example, there have been efforts to show that lexical processing is distributed across multiple words by attempting to demonstrate that variables that influence the lexical processing rate of word $N + 1$ (e.g., its frequency of occurrence in printed text) can influence the time spent fixating on word N. Such parafoveal-on-foveal effects would be illustrated, for example, if fixation durations on word N tended to decrease as the frequency of word $N + 1$ increased, presumably because an increased rate of word $N + 1$ lexical processing reduces the time required for its identification from word N (for a review, see Schotter, Angele, & Rayner, 2012). If demonstrated, such effects would provide evidence against serial-attention models such as E-Z Reader because they posit that only the completion of some stage of lexical processing of word N initiates saccadic programming to move the eyes to word $N + 1$.

Although previous studies have yielded consistent parafoveal-on-foveal effects when, for example, perceptual variables such as the orthographic familiarity of word $N + 1$ affect the fixation duration on word N (Kennedy, 1998, 2000; White, 2008), evidence for lexical parafoveal-on-foveal effects is mixed. Statistical analyses of eye-movement corpora during reading experiments (e.g., Kliegl, Nuthmann, & Engbert, 2006) and nonreading experiments (e.g., Kennedy, 1998, 2000) typically show such effects, but experiments in which lexical properties of target words are controlled during reading typically yield null results (e.g., Drieghe, Rayner, & Pollatsek, 2008; Henderson & Ferreira, 1993; for a recent meta-analysis, see Brothers, Hoversten, & Traxler, 2017) or results that are difficult to interpret (e.g., shorter fixations on word N if word $N + 1$ is low frequency; Kennedy, 1998, 2000). This interpretation is also made difficult because, in natural text, linguistic variables associated with word N and word $N + 1$ covary in complex ways, making inferences based on statistical analyses tentative (see Kliegl, 2007; Rayner, Pollatsek, Drieghe, Slattery, & Reichle, 2007; see also Angele et al., 2015), and because the properties of word $N + 1$ might seemingly influence fixations on word N because of saccadic error, imperfect eye convergence, or eye-tracker measurement error (see Reichle & Drieghe, 2015).

The goal of this article, however, is not to examine the serial versus parallel debate by testing the spatial extent of lexical processing, but rather to examine the issue using a completely novel approach—one that provides a test of the serial hypothesis by asking whether or not people can attend to one of two words that are located in the same region of space, using paradigms inspired by Duncan (1984; see also MacLeod, 1991). These paradigms involved two Chinese word/text stimuli being displayed simultaneously in different colors in the same spatial location (see Fig. 1). In Experiment 1, these stimuli consisted of individual words and nonwords, and participants were instructed to make lexical decisions about one of the two spatially colocated characters. In Experiment 2, the stimuli consisted of target words embedded in whole sentences, which participants were instructed to read for comprehension while their eye movements were recorded. In both experiments, the frequencies of the attended and (according to the serial-attention hypothesis) unattended words were orthogonally manipulated. Because word frequency is known to robustly influence the rate of lexical processing (as measured using lexical decision response times and eye-movement measures; e.g., Schilling, Rayner, & Chumbley, 1998), our factorial manipulation of word frequency provides a novel test of the serial-attention hypothesis. Specifically, if the serial hypothesis is correct, then lexical decision responses and eye-movement measures should be influenced only by the frequency of the attended words and not the frequency of the spatially colocated but unattended (i.e., distractor) words. Such results would suggest that attention is allocated to individual words in an object-based manner. The theoretical implications of this conclusion will be examined in the General Discussion.

Experiment 1

Method

Participants. A total of 18 native-Chinese-speaking undergraduate students (8 males, 10 females) from Sun Yat-sen University were paid 20 yuan for their participation. All had normal or corrected-to-normal vision (including no color blindness), were naive to the purpose of the experiment, and provided informed consent prior to the start of the experiment.

Materials and design. On each trial, participants were instructed to attend to and make lexical decisions about a two-character target word or nonword that spatially overlapped another (unattended) two-character distractor

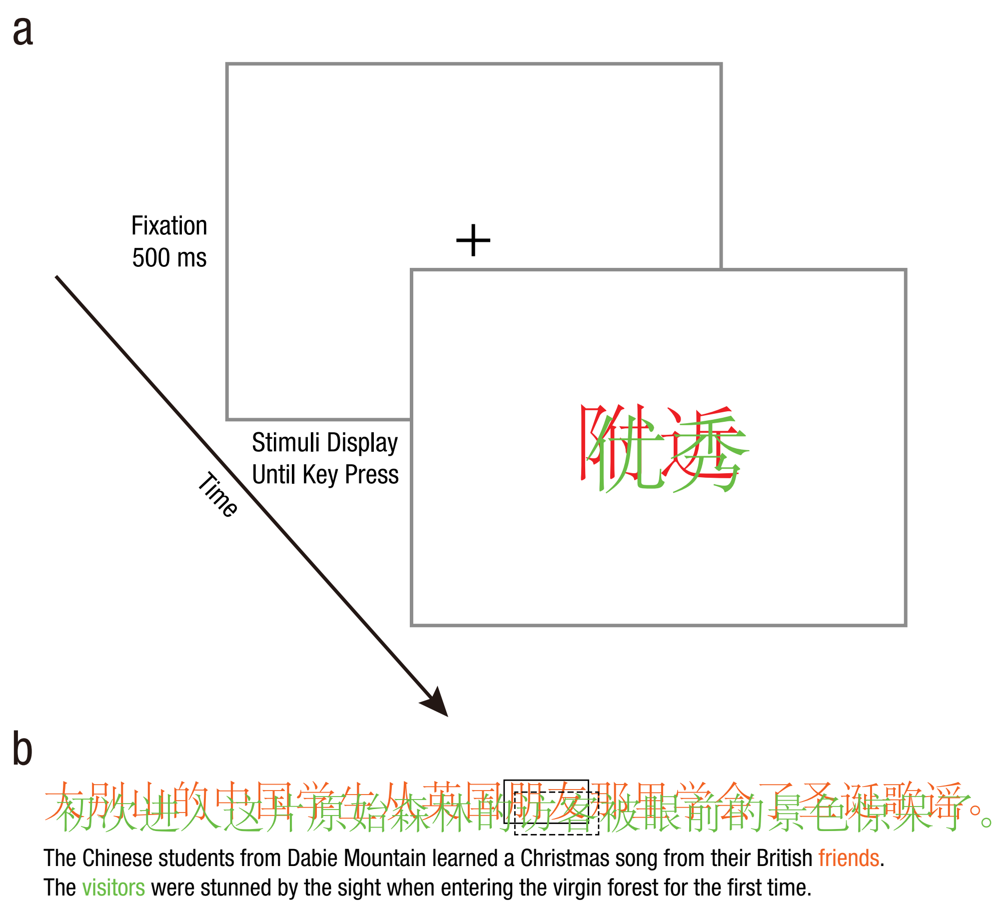

Fig. 1. Example stimuli and procedure used in Experiment 1 (a) and an example sentence (with an English translation) used in Experiment 2 (b). For purposes of illustration, the target and distractor words in (b) are indicated, respectively, by solid and dashed rectangles, and the stimuli are rendered on a white—rather than black—background.

word (see Fig. 1a). Characters were rendered in Song 100 font. Each character subtended about 3° of visual angle, and each pair of characters was diagonally offset by about 0.5° of visual angle. Both the color of the attended character pairs (i.e., red vs. green) and the color of the unattended characters (i.e., red on green or vice versa) were counterbalanced across participants. The experiment consisted of 256 trials, half requiring "yes" responses (if the two characters were a word) and half requiring "no" responses (if the two characters were a nonword). The 128 word trials were constructed using a 2 (target-word frequency: high vs. low) × 2 (distractor-word frequency: high vs. low) within-subjects design, crossing 64 high- and low-frequency target words with 64 high- and low-frequency distractor words. Table 1 shows the mean frequency and complexity (i.e., number of strokes per character) of the words and the complexity of the nonwords. The latter were generated by combining two Chinese characters that do not constitute a two-character word. The experiment was preceded by 20 practice trials, and trial order was completely random for each participant.

Apparatus and procedure. Participants were tested individually in a normally lit room. Stimuli were displayed against a black background on a 27-in. LED video monitor (Model PG27AQ, ASUS, Beitou District, Taipei, Taiwan; resolution = 2,560 × 1,440 pixels, refresh rate = 144 Hz) driven by a Dell computer. Each trial started with a white fixation cross displayed in the center of the monitor for 500 ms, which was then replaced by a stimulus. Participants were asked to make a lexical decision about the upper-left character pair by pressing either the "f" (for words) or "j" (for nonwords) on the keyboard, which also terminated the trial. The next trial started 1,000 ms later. Participants were instructed to respond as quickly and accurately as possible.

Results

To determine whether target- or distractor-word frequency influenced the rate or accuracy of target-word processing, we analyzed response time and accuracy using generalized linear mixed-effects models. Target-word frequency,

Table 1. Word Frequency (per Million) and Complexity (Mean Number of Strokes) of the Characters Used in Experiment 1

	Targets		Distractors	
Trial and stimulus type	Frequency	Complexity	Frequency	Complexity
Word				
HF target, HF distractor	132.68	7.56	131.83	7.52
LF target, HF distractor	2.25	7.55	132.34	7.52
HF target, LF distractor	130.13	7.55	2.34	7.55
LF target, LF distractor	2.18	7.52	2.15	7.53
Nonword				
HF distractor	—	7.56	131.08	7.53
LF distractor	—	7.53	2.09	7.52

Note: HF = high frequency; LF = low frequency.

distractor-word frequency, and their interaction were entered as fixed-effects factors for word trials, but only distractor-word frequency was entered for nonword trials. Response times were log-transformed prior to analysis. To maximize the generalizability of our analyses, we created models that used a maximal random-effects structure; however, the models fail to converge because of overparameterization (Barr, Levy, Scheepers, & Tily, 2013). Therefore, we generated a parsimonious random-effects structure using iterative reduction of insignificant random-effects components from the maximal random-effects models, resulting in a parsimonious random-effects structure including significant variance and covariance components for participants, items, the different slopes of fixed effects for each participant and item, and their correlations (see Bates, Kliegl, Vasishth, & Baayen, 2015, for details). Our results were robust regardless of which random-effects structure was used. The models were fitted using the lme4 package (Version 1.1-14; Pinheiro & Bates, 2000) in R (Version 3.4.2; R Core Team, 2017). The degrees of freedom and p values were estimated using the lmerTest package (Version 2.0-33; Kuznetsova, Brockhoff, & Christensen, 2013).

As can been seen by inspecting mean response time (Table 2) and the linear mixed-effects models (Table 3) for word trials, only target-word frequency influenced response times: Responses were faster for high-frequency targets (M = 713 and 719 ms for the high- and low-frequency-distractor conditions, respectively) than for low-frequency targets (M = 825 and 797 ms for the high- and low-frequency-distractor conditions, respectively), $b = -0.12$, 95% confidence interval (CI) = [−0.17, −0.08], $SE = 0.02$, $t(30.56) = -5.80$, $p < .001$. Neither the effect of distractor-word frequency (p = .383) nor its interaction with target-word frequency (p = .327) was significant. And for nonword trials, response times did not differ significantly between high- and low-frequency-distractor conditions (M = 858 and 864 ms, respectively; p = .863).

As can also be seen in Tables 2 and 3, for word trials, only target-word frequency influenced response accuracy: Accuracy was higher for high-frequency targets (M = .97 and .96 for the high- and low-frequency-distractor conditions, respectively) than for low-frequency targets (M = .86 and .90 for the high- and low-frequency distractor conditions, respectively), $b = 1.49$, 95% CI = [0.82,

Table 2. Mean Response Time (ms) and Accuracy (Proportion Correct) for Each Combination of Stimuli in Experiment 1

Trial type and dependent measure	HF distractor		LF distractor		Frequency effect (LF − HF)	
	HF target	LF target	HF target	LF target	Target	Distractor
Word						
Response time	713 (22)	825 (40)	719 (28)	797 (39)	95	−9
Accuracy	.97 (.01)	.86 (.03)	.96 (.01)	.90 (.02)	−.08	.01
Nonword						
Response time	858 (38)		864 (37)		—	5
Accuracy	.95 (.02)		.94 (.02)		—	−.01

Note: Standard errors are given in parentheses. HF = high frequency; LF = low frequency.

Table 3. Results From the Linear Mixed-Effects Model: Inferential Statistics for Response Time and Accuracy in Experiment 1

Dependent measure and predictor	*b*	*SE*	*t*	*z*	*p*
			Word trials		
Response time					
Intercept	6.59	0.03	$t(19.63) = 195.16$	—	< .001
Target frequency (HF)	−0.12	0.02	$t(30.56) = -5.80$	—	< .001
Distractor frequency (HF)	0.02	0.02	$t(31.93) = 0.89$	—	.383
Target × Distractor Frequency	−0.04	0.04	$t(31.62) = -1.00$	—	.327
Accuracy					
Intercept	3.30	0.27	—	12.38	< .001
Target frequency (HF)	1.49	0.34	—	4.37	< .001
Distractor frequency (HF)	−0.21	0.35	—	−0.60	.550
Target × Distractor Frequency	0.74	0.64	—	1.17	.244
			Nonword trials		
Response time					
Intercept	6.69	0.04	$t(19.34) = 152.63$	—	< .001
Distractor frequency (HF)	0.01	0.03	$t(32.72) = 0.18$	—	.863
Accuracy					
Intercept	3.32	0.28	—	11.92	< .001
Distractor frequency (HF)	0.33	0.27	—	1.21	.228

Note: Accuracy data were analyzed using generalized (i.e., binomial) linear mixed-effects models. HF = high frequency.

2.16], $SE = 0.34$, $z = 4.37$, $p < .001$. Neither the effect of distractor frequency ($p = .550$) nor its interaction with target-word frequency ($p = .244$) was significant. And for nonword trials, accuracy did not differ significantly between high- and low-frequency-distractor conditions ($M = .95$ and .94 ms, respectively; $p = .228$).

Experiment 2

Experiment 1 showed that the time and accuracy when processing an attended stimulus were influenced only by the frequency of that word and not the frequency of a spatially colocated distractor word. This finding suggests that participants can fully comply with instructions and attend to a single word—even if that word happens to overlap almost completely with another word. To examine whether this finding generalizes to natural sentence reading, we conducted an eye-movement experiment in which participants were instructed to read one of two spatially overlapping sentences containing target and distractor words that—as in Experiment 1—could be either high or low frequency.

Method

Participants. A total of 30 native-Chinese-speaking undergraduate students (12 males, 18 females) from Sun Yat-sen University were paid 20 yuan for their participation. All participants had normal or corrected-to-normal vision (including no color blindness), were naive to the purpose of this experiment (e.g., none had participated in Experiment 1), and gave informed consent prior to the start of the experiment.

Materials and design. This experiment used the same 2 (target-word frequency: high vs. low) × 2 (distractor-word frequency: high vs. low) within-subjects design as Experiment 1. Target words consisted of 160 pairs of high-frequency ($M = 120.71$ per million, $SD = 95.02$) and low-frequency ($M = 2.10$ per million, $SD = 1.50$) two-character words, with one word of either type embedded (near the center) within 1 of 160 sentences. The distractor words also consisted of 160 pairs of high-frequency ($M = 133.73$ per million, $SD = 107.62$) and low-frequency ($M = 2.19$ per million, $SD = 1.35$) two-character words, with one word of either type embedded (near the center) within 1 of another set of 160 sentences. The two sentences composing each target-distractor pair were the same length and were matched for naturalness of meaning across all four conditions by 16 additional participants (all ps > .15). Cloze norms from an additional 19 participants also indicated that all target and distractor words were completely unpredictable from their preceding sentences (i.e., no word was predicted by any participant). As Figure 1b shows, stimuli were rendered in Song 30 font (each character subtending approximately 1° of visual angle). On each trial, a distractor-word sentence was diagonally offset (to the lower right) from a target-word sentence by approximately 0.25° of visual angle.

Sentences were rendered in orange and green to control luminance. The location of the attended sentence (upper vs. lower), its color (orange vs. green), and the sentence "layering" (i.e., orange on green or vice versa) was counterbalanced across participants.

Apparatus and procedure. Stimuli were displayed against a black background on a 27-in. LED monitor identical to that used in Experiment 1. Stimulus presentation was controlled by an OpenGL-based Psychophysics Toolbox-3, which incorporates the EyeLink Toolbox extensions (Cornelissen, Peters, & Palmer, 2002) in MATLAB (The MathWorks, Natick, MA). Eye movements were recorded using an EyeLink 1000 Plus eye tracker (SR Research, Kanata, Ontario, Canada) sampling at a 1,000-Hz rate. Viewing was binocular, but eye-movement data were collected only from the right eye.

Participants were given task instructions on arriving at the lab and then seated approximately 58 cm from the video monitor. A chin and forehead rest was used to minimize head movements. An initial three-point calibration and validation procedure was performed until the maximal error was less than 0.4° of visual angle, and recalibration and revalidation was conducted as necessary. During the experiment, participants first read 16 practice sentences (excluded from analysis) and then read the 160 experimental sentences in random order. Each trial consisted of a drift check in the middle of the screen followed by a fixation box (1° × 1°, the size of a single character) at the location of the first character of the sentence. If the initial fixation did not register in the box or the drift check indicated more than a 0.4° error, then the participant was recalibrated; otherwise, a sentence appeared, which participants read silently for comprehension, terminating the trial using a button box. Participants also used the button box to answer a comprehension question that occurred after half of the sentences and to start the next trial.

Results

Accuracy. Mean sentence-comprehension accuracy was 96%, and there were no differences across conditions (all ps > .14).

Eye-movement measures. Approximately 3.4% of trials were removed because an eyeblink occurred during a fixation on, immediately before, or immediately after the target-word region. This region consisted of the union of the target- and distractor word "envelope" because the goal was to determine whether the frequency of the distractor words influenced fixation durations on the target words. Our analyses of these regions were based on four standard eye-movement measures: (a) first-fixation duration, or the duration of the initial fixation on the target word during first-pass reading; (b) gaze duration, or the sum of all first-pass fixations on the target word; (c) total viewing time, or the sum of all fixations on the target word; and (d) fixation position, or the location of the initial fixations on the target words (measured from the left edge of the first character) during first-pass reading. Finally, to test for possible parafoveal-on-foveal effects, we analyzed the duration of last first-pass fixation on pretarget words as a function of both target and distractor word frequency. All fixation-duration measures were log-transformed and, as in Experiment 1, analyzed using linear mixed-effects models with target-word frequency, distractor-word frequency, and their interaction as predictor variables. These models were also analyzed using the same random-effects structure as in Experiment 1.

As can be seen by inspecting the mean fixation-duration measures (see Table 4) and the linear mixed-effects models (see Table 5), all three fixation-duration measures on targets exhibited a similar pattern. First-fixation durations were shorter on high- than low-frequency target words, $b = -0.05$, 95% CI = [−0.07, −0.02], $SE = 0.01$, $t(87.09) = -3.68$, $p < .001$, as were gaze durations, $b = -0.08$, 95% CI = [−0.12, −0.05], $SE = 0.02$, $t(148.52) = -5.17$, $p < .001$, and total viewing times, $b = -0.13$, 95% CI = [−0.17, −0.09], $SE = 0.02$, $t(72.78) = -6.11$, $p < .001$. In addition, neither the effect of distractor-word frequency (all ps > .493) nor its interaction with target-word frequency (all ps > .584) was significant. Moreover, fixation position exhibited a pattern reported previously (e.g., Liu, Reichle, & Li, 2016), being further to the right on high- than low-frequency target words, $b = 0.06$, 95% CI = [0.01, 0.10], $SE = 0.02$, $t(27.94) = 2.59$, $p = .015$, but showing no effect of distractor-word frequency ($p = .366$) or Target-Word Frequency × Distractor-Word Frequency interaction ($p = .316$). Finally, consistent with the findings of Ma, Li, and Rayner (2015), there were no parafoveal-on-foveal effects: The duration of the last fixation on the pretarget word was not influenced by the frequencies of target words and distractor words or their interaction (all ps > .226). These results collectively show that only the frequency of the attended target words influenced when and where the eyes moved during sentence reading.

General Discussion

In this article, we examined attention allocation during reading using two novel experiments inspired by Duncan (1984)—experiments wherein participants attended to and lexically processed one of two spatially colocated words. Experiment 1 showed conclusively that the speed and accuracy of lexical decisions were influenced only by the frequency of the attended target word and

Table 4. Results for the Dependent Measures in Experiment 2

	HF distractor		LF distractor		Frequency effect (LF – HF)	
Dependent measure	HF target	LF target	HF target	LF target	Target	Distractor
First-fixation duration	292 (6)	307 (6)	295 (7)	314 (7)	16	4
Gaze duration	312 (8)	351 (9)	320 (9)	356 (10)	37	6
Total viewing time	350 (9)	416 (12)	355 (11)	421 (16)	65	4
Fixation position	0.95 (0.02)	0.88 (0.02)	0.96 (0.02)	0.92 (0.02)	–0.06	0.02
Pretarget fixation duration	281 (6)	282 (7)	288 (7)	282 (8)	–3	5

Note: The table presents means (with standard errors in parentheses). Fixation position refers to the location of the initial fixations on the target words (measured by the number of characters in the sentence starting from the left edge of the first character). All other values are given in milliseconds. HF = high frequency; LF = low frequency.

not the unattended distractor word. Experiment 2 likewise showed that fixation times and positions on target words during sentence reading were modulated only by target-word frequencies. Together, these results suggest that attention can be allocated to one of two spatially colocated words. We will now speculate about why this might be true.

In both experiments, our measures of target-word processing time were influenced only by the frequencies of those words, consistent with the serial-attention hypothesis. The failure to observe distractor-word frequency effects or Target-Word Frequency × Distractor-Word Frequency interactions is also inconsistent with the attention-gradient hypothesis (e.g., SWIFT; Engbert

Table 5. Results From the Linear Mixed-Effects Model: Inferential Statistics for the Dependent Measures in Experiment 2

Dependent measure and predictor	*b*	*SE*	*t*	*p*
First-fixation duration				
Intercept	5.64	0.02	$t(35.17) = 328.60$	< .001
Target frequency (HF)	–0.05	0.01	$t(87.09) = -3.68$	< .001
Distractor frequency (HF)	–0.004	0.01	$t(122.60) = -0.34$	.733
Target × Distractor Frequency	–0.01	0.03	$t(100.06) = -0.40$	.694
Gaze duration				
Intercept	5.72	0.02	$t(36.73) = 282.10$	< .001
Target frequency (HF)	–0.08	0.02	$t(148.52) = -5.17$	< .001
Distractor frequency (HF)	–0.01	0.02	$t(166.79) = -0.69$	.489
Target × Distractor Frequency	–0.02	0.03	$t(216.05) = -0.56$	.579
Total viewing time				
Intercept	5.82	0.03	$t(37.28) = 232.58$	< .001
Target frequency (HF)	–0.13	0.02	$t(72.78) = -6.11$	< .001
Distractor frequency (HF)	–0.001	0.02	$t(123.40) = -0.03$	.980
Target × Distractor Frequency	0.001	0.04	$t(80.46) = -0.02$	.985
Fixation position				
Intercept	0.92	0.01	$t(28.30) = 71.48$	< .001
Target frequency (HF)	0.06	0.02	$t(27.94) = 2.59$	.015
Distractor frequency (HF)	–0.02	0.02	$t(30.21) = -0.92$	.366
Target × Distractor Frequency	0.04	0.04	$t(157.61) = 1.01$	.316
Pretarget fixation duration				
Intercept	5.58	0.02	$t(33.70) = 276.05$	< .001
Target frequency (HF)	0.01	0.01	$t(78.17) = 0.60$	.549
Distractor frequency (HF)	–0.01	0.01	$t(58.62) = -1.23$	.226
Target × Distractor Frequency	–0.01	0.03	$t(52.27) = -0.41$	.683

Note: HF = high frequency.

et al., 2005); according to such accounts, distractor-word processing should draw attention resources away from target-word processing, resulting in an interaction between the two variables.[1] This predicted interaction, which is a direct consequence of the assumption that attention is a limited resource that can be allocated only to a small number of words or letters, does not reflect the precise shape of the attention gradient (e.g., a linear vs. nonlinear distribution of attention; see Liu, Reichle, & Gao, 2013). Our results therefore support a strong version of the serial-attention hypothesis—one in which, at any time, attention is allocated to only one word, even if that word is spatially colocated with another word.

Of course, one might ask whether other aspects of eye-movement behavior, such as where saccades are directed during reading, might shed light on the serial-versus-parallel debate. Although much is known about saccadic programming and its relation to attention (e.g., covert attention probably shifts to locations of impending saccade targets; Deubel & Schneider, 1996), saccade targeting will not be likely to inform the debate because both serial-attention (e.g., E-Z Reader; Reichle et al., 2012) and parallel-attention models make similar assumptions about saccade targeting—that saccades are directed toward the centers of word "targets" using peripheral visual information (i.e., blank spaces demarcating word boundaries).[2] Thus, although our observed target-word frequency effects on fixation positions might seemingly favor serial-attention models, they are probably also consistent with parallel-attention models.

Although neither E-Z Reader nor SWIFT were explicitly designed to explain the admittedly artificial situation of reading one of two spatially colocated sentences, the demonstration that attention can be allocated completely to one of two colocated words is naturally concordant with the core assumption of serial-attention models such as E-Z Reader (Reichle, 2011). In contrast, although the "zoom-lens" assumption of SWIFT (Schad & Engbert, 2012) might conceivably allow it to explain our results by adopting the assumption that readers focus their attention on individual words in a serial manner, this account amounts to making the parallel-attention model a serial one. The main findings of our experiments, in combination with the lack of evidence for lexical parafoveal-on-foveal effects in Experiment 2, are thus more parsimoniously explained by the assumption that readers attend to and process words in a serial manner.

Finally, it is worth speculating about how an object-based view of attention might be reconciled with serial-attention theories of attention allocation in reading, especially given that the reported experiments were conducted in Chinese—a writing system that does not demarcate word boundaries. According to Duncan's (1984) original conception, visual objects are identified in two stages: a preattentive stage, in which the features in the visual field are segmented into objects, followed by an attentive stage, in which individual objects are identified (Chen, 2012; Ciaramitaro et al., 2011). This view is consistent with the interpretation of word identification offered by the E-Z Reader model (see Reichle, 2011) if words are conceptualized as visual objects. In the context of reading alphabetic languages such as English, this conceptualization is intuitive because individual words are demarcated by blank spaces, effectively rendering each word as a distinct visual object. This may be more complex in languages such as Chinese, however, because individual words are not separated by boundaries but must instead be segmented from continuous lines of characters (Zang, Liversedge, Bai, & Yan, 2011). This segmentation may be linked to word identification through a highly interactive process whereby the activation of candidate lexical representations constrains earlier stages of visual processing (e.g., segmentation), delimiting the number of characters that probably constitute any given attended word (Li, Rayner, & Cave, 2009), but possibly resulting in some information leaking in from spatially adjacent characters.

Although this leakage hypothesis was not directly examined in the present research, it suggests that attention allocation is more complex than traditionally acknowledged. With writing systems having word boundaries and many short words (e.g., English), attention may be very well approximated by serial-attention models. However, with writing systems lacking clear word boundaries (e.g., Chinese) or containing many long, polymorphemic words (e.g., Finnish), we suspect that this might be less true, with some amount of leakage from spatially adjacent characters or morphemes resulting from the requirement to segment and identify spatially contiguous strings of characters or morphemes. However, even with the possibility that attention control is imperfect, readers probably attempt to allocate attention in a serial manner to the maximal extent possible in the service of identifying meaningful units.

Action Editor

Rebecca Treiman served as action editor for this article.

Author Contributions

Both authors conceived and designed the study. Testing, data collection, and data analysis and interpretation were performed by Y. Liu. Both authors drafted the manuscript and approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was supported by the National Natural Science Foundation of China (31500890), the Fundamental Research Funds for the Central Universities (17wkpy64), and the U.S. National Institutes of Health (RO1HD075800).

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/tgmf6. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617734827. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. This is how SWIFT accounts for the reduced word $N + 1$ preview typically observed when the difficulty of processing word N increases (Henderson & Ferreira, 1990).
2. In nonspaced writing systems such as Chinese, saccade length is also modulated by lexical processing difficulty (Liu, Huang, Gao, & Reichle, 2017; Liu, Reichle, & Li, 2015, 2016).

References

Angele, B., Schotter, E. R., Slattery, T. J., Tenenbaum, T. L., Bicknell, K., & Rayner, K. (2015). Do successor effects in reading reflect lexical parafoveal processing? Evidence from corpus-based and experimental eye movement data. *Journal of Memory and Language*, *79–80*, 76–96.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv:1506.04967*. Retrieved from https://arxiv.org/abs/1506.04967

Brothers, T., Hoversten, L. J., & Traxler, M. J. (2017). Looking back on reading ahead: No evidence for lexical parafoveal-on-foveal effects. *Journal of Memory and Language*, *96*, 9–22.

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, *74*, 784–802.

Ciaramitaro, V. M., Mitchell, J. F., Stoner, G. R., Reynolds, J. H., & Boynton, G. M. (2011). Object-based attention to one of two superimposed surfaces alters responses in human early visual cortex. *Journal of Neurophysiology*, *105*, 1258–1265.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*, 613–617.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837.

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

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501–517.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813.

Henderson, J. M., & Ferreira, F. (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*, 417–429.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology*, *47*, 201–221.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–179). Oxford, England: Elsevier.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *53*, 429–455.

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007). *Journal of Experimental Psychology: General*, *136*, 530–537.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*, 12–35.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests in linear mixed effect models. Retrieved from https://cran.r-project.org/web/packages/lmerTest/index.html

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*, 525–552.

Liu, Y., Huang, R., Gao, D.-G., & Reichle, E. D. (2017). Further tests of a dynamic-adjustment account of saccade targeting during the reading of Chinese. *Cognitive Science*, *41*, 1264–1287.

Liu, Y., Reichle, E. D., & Gao, D.-G. (2013). Using reinforcement learning to examine dynamic attention allocation during reading. *Cognitive Science*, *37*, 1507–1540.

Liu, Y., Reichle, E. D., & Li, X. (2015). Parafoveal processing affects outgoing saccade length during the reading of Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1229–1236.

Liu, Y., Reichle, E. D., & Li, X. (2016). The effect of word frequency and parafoveal preview on saccade length during the reading of Chinese. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1008–1025.

Ma, G., Li, X., & Rayner, K. (2015). Readers extract character frequency information from nonfixated-target word at long pre-target fixations during Chinese reading.

Journal of Experimental Psychology: Human Perception and Performance, *41*, 1409–1419.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology*, *66*, 429–452.

Rayner, K. (2009). Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General*, *136*, 520–529.

R Core Team. (2017). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org

Reichle, E. D. (2011). Serial attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 767–786). New York, NY: Oxford University Press.

Reichle, E. D., & Drieghe, D. (2015). Using E-Z Reader to examine the consequences of fixation-location measurement error. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 262–270.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences*, *13*, 115–119.

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.

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

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition*, *20*, 391–421.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*, 1270–1281.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*, 5–35.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 205–223.

Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 961–978). New York, NY: Oxford University Press.