Behavior Research Methods (2022) 54:1989–2000
https://doi.org/10.3758/s13428-021-01730-2

The Beijing Sentence Corpus: A Chinese sentence corpus with eye movement data and predictability norms

Jinger Pan[1] · **Ming Yan**[2,3] · **Eike M. Richter**[4] · **Hua Shu**[5] · **Reinhold Kliegl**[6]

Accepted: 18 October 2021 / Published online: 23 November 2021

Abstract

This report introduces the Beijing Sentence Corpus (BSC). This is a Chinese sentence corpus of eye-tracking data with relatively clear word boundaries. In addition, we report predictability norms for each word in the corpus. Eye movement corpora are available in alphabetic scripts such as English, German, and French. However, there is no publicly available corpus for Chinese. Thus, to study predictive processes during reading in Chinese, it is necessary to establish such a corpus. Also, given the clear word boundaries in the sentences, BSC is especially useful to provide evidence relevant to the theoretical debate of saccade target selection in Chinese. With the large-scale predictability norms, we conducted new analyses based on 60 BSC readers, testing the influences of launch word and target word properties while controlling for visual and oculomotor constraints, as well as sentence and subject-level individual differences. We discuss implications for guidance of eye movements in Chinese reading.

Keywords Corpus analysis · Eye tracking · Chinese reading · Predictability

Eye movement research in reading has developed into a model case for examining the dynamics of cognition and action from a perspective of active vision (e.g., Findlay & Gilchrist, 2003). Core questions about eye movement control during reading are when and where to move the eyes in response to the visual and language-related properties of the reading material and how prediction, constantly under revision in light of new input, moderates these processes. An important limitation of this research, however, is that most of it has been based on reading of alphabetic languages such as English or German, where explicit word boundaries, afforded by inter-word spaces, clearly mark where words begin and end. Against this background, research on reading of unspaced orthographies like Chinese, Japanese, Thai and Tibetan offers a unique window for checking the validity of current assumptions about the interplay of linguistic processing and visuomotor dynamics in reading. Large-scale predictability norms are publicly available in only a few Indo-European languages, but not any Asian languages, preventing researchers from understanding the predictability effect from a cross-language perspective. Additionally, published work on Chinese corpus reading so far has not broadly tested predictability effect on saccade selection. Given these backgrounds, large-scale predictability norms are desirable, and the present work was an endeavor to address such research gaps. We sought to achieve three interrelated goals. First, we provide large-scale predictability norms for a set of Chinese sentences (i.e., the Beijing Sentence Corpus, BSC), the properties of which will be introduced in detail in a later section of the paper. Second, we report additional eye movement data and analyses beyond our first establishment of the BSC (Yan et al., 2010). We discuss the results in light of current models of guidance of eye movement in reading. Finally, we make the complete corpus available to the community for researchers to use the predictability norms

---

✉ Ming Yan
mingyan@um.edu.mo

[1] Department of Psychology, The Education University of Hong Kong, Hong Kong, China

[2] Department of Psychology, University of Macau, Macau, China

[3] Center for Cognitive and Brain Sciences, University of Macau, Macau, China

[4] Department of Psychology and Ergonomics, Technical University of Berlin, Berlin, Germany

[5] Faculty of Psychology, Beijing Normal University, Beijing, China

[6] Department of Psychology, University of Potsdam, Potsdam, Germany

and the eye movement data to explore topics such as lexical processing, memory retrieval in reading, saccade generation, and simulation with and evaluation of computational models of eye movement control during reading.

The corpus-analytic approach

Research on eye movement control in reading uses three complementary approaches: (quasi-)experiments, computational models, and corpus analysis (Kliegl, 2007; Rayner, Pollatsek, et al., 2007b). The (quasi-)experimental approach involves orthogonal manipulations of a small number of factors, allowing researchers to delineate their influences on oculomotor behaviors in the absence of correlations among-predictors. In contrast, corpus analysis delivers information about the significance/reliability of a large number of variables related to the reading process. In this case, predictors are usually correlated (sometimes substantially) and effects must be estimated with statistical control techniques for the influence of the other predictors (i.e., partial effects). The typically much larger number of observations also affords control for and estimations of individual differences among readers, sentences and words (i.e., variance components) as well as of correlation parameters among them (Kliegl et al., 2006).

Both experimental and corpus analytic results deliver benchmark effects that serve as targets for computational models. Eye movement corpora are available for English sentences (Schilling et al., 1998) which were used for the validation of E-Z Reader and its subsequent model variants (Reichle et al., 1998; see also models by Engbert & Kliegl, 2001), for reading of English and French newspapers (Kennedy & Pynte, 2005), and for the validation of the SERIF model (McDonald et al., 2005). Similarly, an eye movement corpus of German sentences (Kliegl et al., 2006) has been used for the validation of the SWIFT model (Engbert et al., 2005; Risse et al., 2014; Schad & Engbert, 2012).

There are two main advantages for the corpus-analytic approach. First, corpus analyses focus on a large number of sentences and words rather than on only a few target words. For instance, the Dundee Corpus (Kennedy et al., 2003; Kennedy et al., 2013) consists of 56,212 tokens from newspapers. As a consequence, independent variables typically cover a much wider range in corpus analysis in contrast to experiments with orthogonal designs, allowing researchers to observe a more complete, ecologically more valid picture of their influences and relative importance. Second, corpus analysis is often based on a much larger number of subjects than experiments and thus provides more reliable estimations of predictors, especially as far as individual differences are concerned.

Eye-tracking corpora have been developed not only for European languages such as English, German, French and Dutch (Cop et al., 2017), Russian (Laurinavichyute et al., 2019), and Turkish (Özkan et al., 2021), but also for Asian languages, including Chinese sentence corpora for adults (Li et al., 2014; Yan et al., 2010) and children (Yan, Pan, & Kliegl, 2019b), Hindi (Husain et al., 2014) and Uighur (Yan et al., 2014). A comparison of critical results in various writing systems has revealed, on one hand, robust similarities in reading across fundamentally different orthographies, indicating language-universal mechanisms in oculomotor control and lexical activation. For instance, properties of the fixated word, including word frequency, predictability and length, are the primary and most frequently documented determinants of fixation duration (Just & Carpenter, 1980; Rayner, 2009). On the other hand, cross-language comparisons may suggest some interesting language-specific influences in reading. For instance, following a debate on whether high-level linguistic information influences saccade target selection in reading and focusing on rich suffixal information in the Uighur language, Yan et al. (2014) reported reliable effects of number of letters and number of suffixes on fixation location, suggesting that saccade generation can be influenced jointly by low-level knowledge of word length and high-level knowledge of morphological complexity. Importantly, the study also served as an example that the corpus-analytic and the experimental control approaches can be in good agreement with one another (see also Hyönä et al., 2021, and Dann et al., 2021, for quasi-experimental evidence on the relevance of morphological processing for Finnish and, with some qualifications, for English, respectively).

The predictability effect

As mentioned above, as one of the most important factors of eye movements in reading, predictability effects are well documented in the existing literature. In general, words that are predicted more from preceding contexts are more likely to be skipped (e.g., Ehrlich & Rayner, 1981; O'Regan, 1979, 1980; Rayner & Well, 1996). When highly predictable words are fixated on, readers typically spend less time on them than on unpredictable ones (e.g., Rayner et al., 2001). Kliegl et al.'s (2006) analysis of the Potsdam Sentence Corpus (PSC) was one of the most comprehensive ones because it included not only variable coding frequency, length, and predictability of the currently fixated word N, but also the corresponding information about its immediate neighbors (i.e., properties of words N − 1 and N + 1). They demonstrated a trade-off between fixation durations on words N and N + 1 as a function of word N + 1 predictability. Counter to a common explanation that high predictability implies easy lexical processing, a highly predictable word N + 1 increased viewing duration on word N. The phenomenon was interpreted as evidence for distributed lexical processing: Memory retrieval of word N + 1 from the prior context shifts part of its lexical processing to the fixation on word N.

In contrast to viewing duration, the predictability effect on fixation location is less clear. While Lavigne et al. (2000) found that readers' eyes traveled further into the words when they were more predictable from prior contexts, other studies reported null effects (Rayner et al., 2006; Vainio et al., 2009).

Following Taylor's (1953) procedure, predictability norm is often evaluated using the cloze test, in which participants are presented with sentence frames prior to target words and are asked to complete the sentences. This procedure is often used to measure the predictability of a single word in a context. The Dundee Corpus (Kennedy et al., 2003; Kennedy et al., 2013) only contains predictability data for a subset of the texts, with approximately 25 guesses per word. Arguably, it takes tremendous effort to obtain and to code predictions for all words in a sentence corpus. The first eye-tracking sentence corpus with complete predictability norms, the PSC, reported cloze predictabilities for all words in 144 German sentences from 272 native German speakers, yielding 83 predications for each word (Kliegl et al., 2004; Kliegl et al., 2006). More recently, like the PSC, the Provo Corpus (Luke & Christianson, 2016, 2018) also contains predictability norms for all words of 134 sentences from five passages based on 470 participants, with each word receiving on average 40 predications (range: 19 to 43).

The Chinese orthography

The Chinese orthography differs fundamentally from English, German and other alphabetic scripts in a number of noticeable ways. The basic writing unit of Chinese is the Chinese character. These characters are square-shaped, monospaced objects that occupy the same horizontal and vertical extents. Each character always corresponds to one syllable, but phonemes are not represented transparently. Chinese characters cover a wide range of visual complexities, which are roughly indexed by their numbers of strokes. For example, the simplest character in Chinese is 一 (pronounced /yi1/, meaning *one*), which has only one single stroke, whereas the most complex character 齉 (pronounced /nang4/, meaning *snuffle*) in simplified Chinese has 36 strokes. Chinese words are typically one to four characters in length, dominated by one- and two-character words. Chinese sentences can be written and read in different directions, horizontally leftwards, horizontally rightwards and vertically downwards. Irrespective of the direction, there are no explicit visual cues, such as spaces, to indicate word boundaries.

When comparing eye movements in reading Chinese and alphabetic scripts, some interesting language-universalities and disparities in oculomotor activities can be noticed. On one hand, reading of different orthographies shares many common mechanisms. Classic Chinese reading experiments (e.g., Chen & Carr, 1926; Peng et al., 1983; Shen, 1927; Sun et al., 1985) have revealed that basic eye movement characteristics are fundamentally similar for Chinese and English (see Tsang & Chen, 2012, for a review). Chinese words that are predictable from prior contexts are also more likely skipped, or fixated on more briefly than those that are not (Rayner, Li, & Pollatsek, 2007a). Additionally, reading Chinese and other scripts involves information processing beyond the currently fixated words and the area of an effective visual field extends to the direction of reading (Inhoff & Liu, 1998; McConkie & Rayner, 1975). On the other hand, oculomotor activities during reading are bound to be shaped by unique features of Chinese. Since the perceptual span in Chinese reading extends rightwards up to four characters (Yan et al., 2015) and Chinese words are typically short, Chinese readers should be able to combine characters within the perceptual span into word units. Indeed, some studies have demonstrated that, despite the lack of explicit word boundaries in Chinese, words are of primary importance in reading. For example, Inhoff and Wu (2005) examined how a sequence of four Chinese characters with word boundary ambiguity was parsed into words. Their results were against a view of strictly serial assignment of characters to words, and they concluded that Chinese readers must segment parafoveal words during online reading (see also Yang et al., 2009, for a similar conclusion).

The Beijing Sentence Corpus

The Beijing Sentence Corpus (BSC) was developed initially to address a theoretical question about saccade generation in Chinese, which we elaborate on below. The sentences were selected from the People's Daily, which is the largest newspaper group and an official newspaper of the People's Republic of China. Some of the selected sentences were modified to remove strong political tones, and possible semantic and word boundary ambiguities. The final BSC sentences contained 7 to 15 words ($M = 11.2$, $SD = 1.6$), corresponding to 15 to 25 characters ($M = 21.0$, $SD = 2.5$). For word category, BSC comprised 702 nouns, 368 verbs, 183 particles, 139 adjectives, 98 adverbs, 60 prepositions, 51 conjunctions, 42 pronouns and 42 others (including numerals, qualifiers, idioms, etc.). For word length distribution, among 1685 word tokens there were 348, 1242, 70 and 25 words of 1 to 4 characters long, respectively. For visual complexity, the numbers of strokes in words varied from 2 to 42 (*M* = 15.6, *SD* = 5.5). The numbers of words with strokes in the ranges of 1–5, 6–10, 11–15, 16–20, and 21 or above were 95, 422, 538, 423 and 207, respectively. For word frequency, the BSC words appeared 5400.8 times per million words (*SD* = 14668.7) based on the Modern Chinese Word Frequency Dictionary (Institute of Linguistic Studies, 1986), or 5262.9 occurrences per million words (*SD* = 13,987.8) based on the SUBTLEX-CH database (Cai & Brysbaert, 2010).

As mentioned above, one dominant feature of the Chinese script is the lack of inter-word spaces and word boundary ambiguity and disagreement (Hoosain, 1991, 1992). More recently, Yen et al. (2009) pointed out that over 80% of Chinese characters can involve word boundary ambiguity. For an examination of the effect of word boundary on saccade planning in Chinese, however, it is critical to provide readers with a set of sentences with relatively clear word boundary information. For this reason, the BSC sentences were selected with as little word boundary ambiguity as possible. In a pretest of word boundary agreement, twenty participants rated whether a word boundary existed after each character. The percentage of word boundary agreement was calculated relative to linguistic segmentation. Across all sentences, the agreement ranged from 91.0% to 100% (*M* = 96.4%, *SD* = 2.6%). The detailed coding method was explained by Yan et al. (2010) (see the Appendix of their work).

The reduction of word boundary ambiguity may incur some cost in terms of representativeness. The effect of boundary ambiguity was explored under a quasi-experimental approach (e.g., Inhoff & Wu, 2005; Yan & Kliegl, 2016; Yen et al., 2012). For instance, Yan and Kliegl (2016) reported longer viewing duration on a target region including four characters with ambiguous than unambiguous word boundaries. They also found an interaction between boundary ambiguity and launch site on fixation location, indicating the influence of word boundary on saccade target selection. Notwithstanding this research, however, these small quantitative effects did not reveal any qualitative shifts in saccade generation. Despite some limitations to the generalizability, the BSC will provide, as a minimum, a useful benchmark against which to evaluate saccade control during reading.

The major purpose of the study was to establish the first reliable predictability norms in Chinese. In the two existing Chinese corpus reports, Yan et al. (2010) did not report large scale predictability norms and Li et al. (2014) collected cloze probability from 61 participants[1]. However, the predictability norms have not been made available publicly. Predictability norms for the PSC were assessed from participants across different ages from 17 and 80 years, including high school students, university students, and older adults. For the Provo Corpus, the participants' ages ranged from 18 and 50 years. In the present study, we focused on a homogeneous sample of participants. A total of 148 students from Beijing Normal University, aged 18–30 years, participated in the predictability test. Each participant completed half (i.e., 75) of the BSC sentences. Therefore, each word received 74 guesses, which is roughly comparable to the PSC. During the collection of the participants' predictability norms, they were seated comfortably in front of a computer and were instructed to guess the first word of a sentence and then to write it down on their answer sheets. After doing so, they pressed the "enter" key on the computer keyboard and the first word of the sentence was displayed on the screen. Based on this first word, participants now guessed the second word of the sentence. They repeated the whole procedure until the end of the sentence. All correct words stayed on the screen. Each participant received a different, randomized order of BSC sentences. The participants were instructed to produce only one word for each guess. A target word was coded as predicted only if an identical word had been guessed by the participant.

Averaged across all participants, word predictabilities ranged from 0 to 0.93 (*M* = 0.17, *SD* = 0.23). Following PSC procedure (Kliegl et al., 2004), the original probability scores were logit transformed using the formula:

$$\text{Logit} = \frac{1}{2} * \ln\left(\frac{p}{1-p}\right)$$

In other words, logits were defined as half of the natural logarithm of the odds p/(1 − p), where p was the original predictability in percentage. Predictabilities of zero were replaced with 1/(2*74), where 74 represents the number of complete predictability protocols (Cohen & Cohen, 1975). For words that were half predictable (p = .50), the odds of guessing were one and the logits were zero. Therefore, words that were predicted successfully by over half of the participants received positive logits, whereas those with less than half correct guesses yielded negative logits. After logit transformation, the mean logit predictability was −1.26 (*SD* = 0.90).

Applications of the Beijing Sentence Corpus

Perhaps one of the most controversial research topics relating to Chinese reading is how a saccadic target is chosen. Current theories on eye movement control in reading generally assume that readers estimate the center of the word which is to be fixated, based on low spatial frequency information (i.e., inter-word spaces) in parafoveal and peripheral

---

[1] This information was acquired via personal communication with the corresponding author.

vision (Engbert et al., 2002; Reichle et al., 1999). The word center serves as the intended location of the next fixation because word processing is assumed to be optimal at this location (McConkie et al., 1989; O'Regan & Lévy-Schoen, 1987). However, given the lack of inter-word spaces, Chinese raises a fundamental challenge to the saccade generation mechanism. Based on empirical findings of flat fixation location distributions in Chinese, McConkie and colleagues concluded that saccade target selection is unlikely to be word-based and left it open whether characters can be used to this end or whether readers simply deploy saccades of a fixed amplitude with some oculomotor error (Tsai & McConkie, 2003; Yang & McConkie, 1999).

So far, existing theories on saccade generation in Chinese can be classified into three main categories. Instead of providing a thorough review on this issue, here we illustrate how analyses based on corpora like the BSC can shed light on this theoretical debate. First, a fixed length saccade strategy in Chinese reading can be derived from work by Yang and McConkie (2004), arguing that neither character nor word is necessary for saccade generation and Chinese readers program saccades to be of a constant amplitude with some distribution of random error, given that most Chinese words are of similar lengths. Following this view, Li et al. (2011) demonstrated that a fixed-length saccade simulation could generate the same shape of the fixation location distribution curve as was found by empirical results in Chinese. However, simulation of this type failed to replicate several other eye movement patterns such as skipping and single-fixation probabilities (Yan et al., 2010). A second type of theory assumes that Chinese readers process as many characters as possible during each fixation and program their saccades to new characters (Liu et al., 2015). However, because such a character-based targeting theory denies any influence from word-boundary knowledge, it may run into difficulty when trying to explain fixation location differences caused by word-boundary ambiguity (Inhoff & Wu, 2005; Yan & Kliegl, 2016) and by word boundary afforded by alternating text colors (Zhou et al., 2018).

The last type of model, proposed by Yan et al. (2010), hypothesized a dynamic switch between a word-based mechanism targeting towards the word center and a character-based mechanism towards the first character of an upcoming word. If readers fail to acquire enough knowledge about the boundary of an upcoming word during previous fixations, the first character of that upcoming word is chosen as the saccade target, leading to fixation location shifting towards word beginning. Alternatively, if a parafoveal word boundary is obtained, a reader generates a saccade aiming at the center of the upcoming word, leading to fixation location shifting away from the word beginning. Extending the E-Z Reader model (Reichle et al., 1998) to simulate Chinese readers' eye movements, Rayner et al. (2007a, b) used predefined word boundary knowledge as a model input and achieved a good simulation. Based on empirical work that Chinese character positional probability (i.e., the likelihood that a character appears in different within-word positions) influenced character-to-word assignment (Yen et al., 2012), a parafoveal word segmentation mechanism was implemented in the Chinese version of SWIFT and word boundary certainty was estimated on the basis of simple statistical information (Richter et al., 2010).

As well as its contribution to the theoretical debates on saccadic generation in Chinese, the BSC has been used for several other research topics. It was used to evaluate the perceptual span in Chinese (Pan et al., 2017; Yan et al., 2015). The clear word boundary, as evident by the high score in boundary agreement rating, makes it an ideal set of materials to explore the effect of word boundary in Chinese. As such, the influence of explicit word boundary afforded by chromatic similarity on the reading of naturally unspaced Chinese sentences was demonstrated by Zhou et al. (2018). In addition, BSC was also used to explore the development of saccade generation in Chinese reading (Yan et al., 2019b). Finally, a slightly modified version of BSC was created for a study of text orientation effect, illustrating that, due to abundant experience in reading vertical texts, Taiwanese readers could sometimes generate saccades more efficiently during their reading of vertically aligned, in comparison to horizontally aligned, sentences (Yan, Pan, Chang, & Kliegl, 2019a). The current BSC predictability norms are based only on young adults. An extension of the sample to include readers from different age groups, such as beginning readers and older adults, is desirable for future work.

Eye-tracking data

Participants A total of 60 students from Beijing Normal University ($M_{age} = 22.0$ year, $SD = 2.6$, 42 female and 18 male) with normal or corrected-to-normal vision participated in the eye-tracking experiment. Data from the first 30 students were reported in work by Yan et al. (2010) on preferred viewing location (PVL) effect (Rayner, 1979) in Chinese reading. The rest of the participants' data have not been reported previously. All participants were native readers of Chinese and had not participated previously in the predictability test. The choice of adding 30 new subjects was supported by a recent simulation work by Kumle et al. (2021), showing that 60 subjects and 120 sentences for corpus analyses can achieve a high statistical power in linear mixed model (LMM) analysis.

Apparatus The participants' eye movements were recorded with an Eyelink II system running at 500 Hz. Each sentence was presented in a single line on a 19-inch ViewSonic G90f

Table 1 Eye movement data variables and descriptions

Column	Variable	Description
Information on the currently fixated word		
1	ID	A unique ID number for each participant
2	SN	A unique ID number for each sentence
3	NW	The number of words in the current sentence
4	WN	The ordinal position of the current word within the current sentence
5	FL	Fixation location, the horizontal position of the fixation relative to word beginning in characters
6	DUR	Fixation duration in milliseconds
7	AO	Outgoing saccade amplitude from the current fixation
8	DIR	Direction of the outgoing saccade. Positive values indicate rightward saccades and negative values indicate regressions
9	O	Incoming saccade amplitude from previous fixation to the current one
10	L	Word length, reciprocal of number of characters
11	F	Word frequency, base 10 logarithm-transformed occurrence per million words
12	P	Logit predictability of the word
13	I	Visual complexity, number of strokes of the word
Information on the word immediately preceding the currently fixated one		
14	WN1	The ordinal position of the last word
15	L1	Word length of the last word
16	F1	Word frequency of the last word
17	P1	Logit predictability of the last word
18	I1	Visual complexity of the last word
Information on the word immediately following the currently fixated one		
19	WN2	The ordinal position of the next word
20	L2	Word length of the next word
21	F2	Word frequency of the next word
22	P2	Logit predictability of the next word
23	I2	Visual complexity of the next word
Information on the last fixated word		
24	WNx1	The ordinal position of the last fixated word
25	FLx1	Fixation location on the last fixated word
26	DURx1	Fixation duration on the last fixated word
27	AOx1	Outgoing saccade amplitude from the last fixation
28	DIRx1	Direction of the last outgoing saccade
29	Ox1	Incoming saccade amplitude to the last fixation
30	Lx1	Word length of the last fixated word
31	Fx1	Word frequency of the last fixated word
32	Px1	Logit predictability of the last fixated word
33	Ix1	Visual complexity of the last fixated word
Information on the next fixated word		
34	WNx2	The ordinal position of the next fixated word
35	FLx2	Fixation location on the next fixated word
36	DURx2	Fixation duration on the next fixated word
37	AOx2	Outgoing saccade amplitude from the next fixation
38	DIRx2	Direction of the next outgoing saccade
39	Ox2	Incoming saccade amplitude to the next fixation
40	Lx2	Word length of the next fixated word
41	Fx2	Word frequency of the next fixated word
42	Px2	Logit predictability of the next fixated word
43	Ix2	Visual complexity of the next fixated word

monitor (resolution: 1024 × 768 pixels; frame rate: 100 Hz) using font Song. The participants were seated 43 cm from the monitor, positioned on a chin rest to minimize their head movements. Each character occupied a 32 × 32 pixel grid and thus subtended approximately 1.5 degrees of visual angle. All recordings and calibrations were done binocularly.

Procedure Gaze positions were calibrated with a nine-point grid (error < 0.5°). Participants' fixation on the initial fixation point initiated the presentation of the next sentence, with its first character occupying the position of the fixation point. The participants were instructed to read the sentences silently for comprehension, then to fixate on a dot in the lower-right corner of the monitor, and finally press a joystick button to signal trial completion. Each participant received the BSC sentences in a different, randomized order. Forty sentences were followed by easy yes/no comprehension questions to encourage the participants' engagement with the reading task. They answered at least 80% of the questions correctly (*M* = 90.4%, *SD* = 5.5%).

Content of eye-tracking data Fixations were determined using an algorithm for saccade detection (Engbert & Kliegl, 2003). Some trials were removed due to participants' blinking, coughing, body movements or tracker errors (n = 927, 10%). We included in the data file a number of commonly used oculomotor indices in eye-tracking research, such as first fixation duration (FFD; duration of the first fixation on a word, irrespective of total number of fixations), gaze duration (GD; the sum of all fixation durations during the first-pass reading of a word), total reading time (TRT, sum of all fixations on a word, including regressive fixations), first fixation location (FL, landing position of the first fixation on a word relative to word beginning) and launch site (LS, the distance between the location of the last fixation and the beginning of the currently fixated word). These indices helped us to illustrate the basic patterns of oculomotor activities in Chinese reading (Yan et al., 2010). The complete content of the data file is described in Table 1. Typically, words with extremely short or long fixations are excluded from data analyses and the criteria may differ among researchers. Therefore, we kept all observations in the data file so that readers can apply their own fixation selection criteria.

Empirical description of eye-tracking data Following classic work on PVL (McConkie et al., 1988; Rayner, 1979), Yan et al. (2010) reported fixation location distributions in Chinese reading. However, they did not simultaneously model oculomotor, visual, lexical, and sentence-level effects on fixation location, mainly due to restrictions in statistical methods when their work was conducted. Hohenstein et al. (2017) pointed out that LMM allows specification of crossed random factors and accepts both categorical and continuous predictors. One major problem in corpus analyses is the correlation among predictors: zero-order regression may be influenced by uncontrolled confounding variables. As such, LMM estimates after statistical control of other variables and removal of between-subject and between-sentence random effects, namely partial effects, can reveal effects of interest without confounding (Yan et al., 2014). These advantages make LMMs well suited for corpus analyses.

As a theoretical consolidation of Yan et al.'s original work, we modeled visual and oculomotor constraints

Table 2 Linear mixed model estimates for word n fixation location.

Fixed effect	Est.	SE	T
(Intercept)	−0.076	0.010	−7.258
k1	0.036	0.009	3.989
poly(lls, 2)1	−35.725	0.826	−43.267
poly(lls, 2)2	2.721	0.465	5.858
l1.c	−0.006	0.004	−1.485
f1.c	0.008	0.002	3.881
lgs1.c	−0.017	0.004	−4.791
p1.c	0.025	0.003	8.580
l.c	−0.026	0.006	−4.302
f.c	0.003	0.002	1.373
lgs.c	−0.035	0.005	−6.827
p.c	0.015	0.003	5.360
k1:poly(lls, 2)1	2.971	1.821	1.631
k1:poly(lls, 2)2	15.914	0.864	18.412
k1:l1.c	−0.064	0.008	−7.749
Variance component		SD	
Word - (Intercept)		0.035	
Sent - (Intercept)		0.013	
Subj - p.c		0.014	
Subj.1 - lgs.c		0.018	
Subj.2 - f.c		0.004	
Subj.3 - l.c		0.032	
Subj.4 - p1.c		0.015	
Subj.5 - lgs1.c		0.014	
Subj.6 - f1.c		0.009	
Subj.7 - (Intercept)		0.070	
Residual		0.183	
Log-likelihood		9023.3	
Deviance		−18046.6	
AIC		−17980.6	
BIC		−17701.6	
N		34689	

k1 = skipping of word n − 1, lls = launch site, l1.c = length of word n − 1, f1.c = frequency of word n − 1, lgs1.c = log stroke count of word n − 1, p1.c = predictability of word n − 1, l.c = length of word n, f.c = frequency of word n, lgs.c = log stroke count of word n, p.c = predictability of word n

Table 3 Linguistic information about all BSC words and descriptions

Column	Variable	Description
1	SN	A unique ID number for each sentence
2	NW	The number of words in the current sentence
3	WORD	The text of the word
4	POS	The part of speech of the word
5	LEN	Word length in number of characters
6	STRK	Number of strokes of the word
7	WF_BLI	Word frequency from the Institute of Linguistic Studies (1986)
8	WF_SUB	Word frequency from Cai and Brysbaert (2010)
9	PRED	Logit predictability of the word
10	CF1_BLI	Character frequency of the 1st character of the word
11	CF2_BLI	Character frequency of the 2nd character of the word
12	CF1_SUB	Character frequency of the 1st character of the word
13	CF2_SUB	Character frequency of the 2nd character of the word
14	C1_STRK	Number of strokes of the 1st character of the word
15	C2_STRK	Number of strokes of the 2nd character of the word
16	C3_STRK	Number of strokes of the 3rd character of the word

(predictors: linear and quadratic trends of launch site, lengths of words N and N – 1 and skipping status of word N – 1) and lexical/contextual effects of launch word and target word difficulty (predictors: frequencies, stroke counts and predictabilities of words N – 1 and N) on fixation location, using the lme4 package (Bates, Maechler, et al., 2015b) in the R programming language environment. We started a model with all main effects introduced above, in addition to all three-covariate interactions, and used likelihood-ratio tests (LRTs) to check changes in goodness of fit after dropping nonsignificant interactions stepwise following the parsimonious LMM principle (Bates, Kliegl, et al., 2015a; Matuschek et al., 2017). Table 2 presents the final model output, and Fig. 1 shows the main effects of launch word frequency (top left), complexity (top middle) and predictability (top right) and target word frequency (bottom left), complexity (bottom middle) and predictability (bottom right).

Launch word effects Considering that a saccade to word N is typically generated during readers' fixations on word N – 1, how launch word properties influence the fixation location on the target word indicates *foveal* influences on saccade target selection. The fixation location on the target word shifted closer to the word center when launch words were more frequent or predictable. In contrast, visually complex launch words prevented the readers from looking further into target words, leading to fixations closer to the word beginning. These results are compatible with a dynamic modulation of perceptual span (Eriksen & St. James, 1986; Henderson & Ferreira, 1990; Inhoff et al., 2000; Inhoff & Rayner, 1986; Risse et al., 2014): When encountering a more difficult foveal word, readers allocate more attentional resources to the current word and reduce parafoveal processing efficiency. Earlier work on the dynamic modulation theory mostly focused on the influence of processing load on fixation duration. Here we propose using fixation location as an alternative measure and the present work indeed extended the foveal load effect to fixation location. This is especially important when considering different saccade models in Chinese. Since the fixed length saccade models hypothesize a constant saccade amplitude with some random noise and do not predict any linguistic and contextual influences (Li et al., 2011; Yang & McConkie, 2004), such foveal load effects would be problematic for these models. The current results, however, agree with the dynamic saccade target model (Yan et al., 2010): when readers are able to obtain more precise parafoveal information, they are more likely to achieve parafoveal word segmentation, leading to more word-based saccades towards the word centers. However, the foveal influences on saccade target selection may not tease apart the dynamic saccade target model (Yan et al., 2010) and the character-based saccade model (Liu et al., 2015), because both models assume the influence of foveal word properties on saccade generation.

Target word effects A target word N is located beyond readers' foveal vision by the time a saccade toward the word is generated (i.e., during a fixation on a launch word). Therefore, estimates of target word properties are influences of upcoming word factors on saccade target selection, or loosely called *parafoveal factors*. Fixation locations on

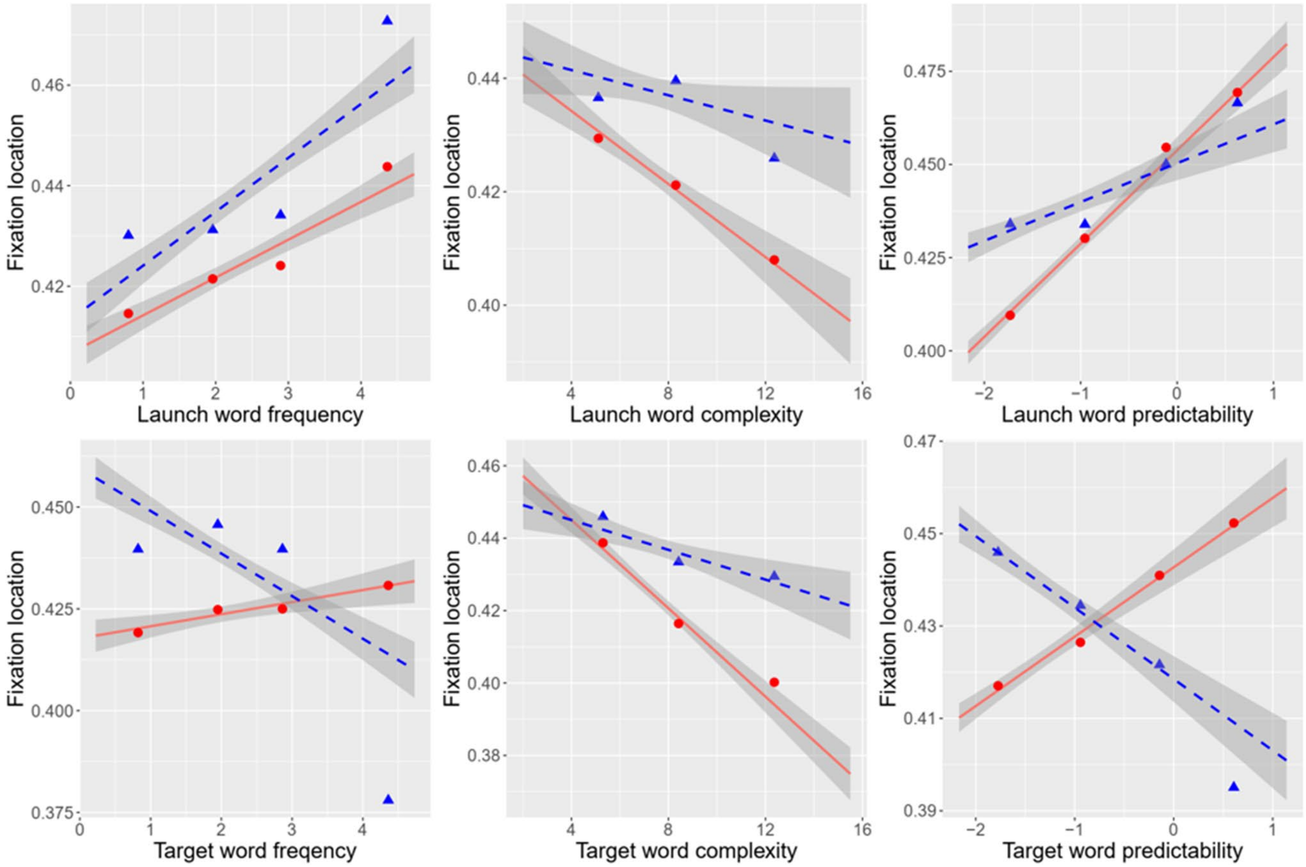

Fig. 1 Partial effects (solid lines) of LMM and zero-order smooths of observed values (dashed lines) of initial fixation location relative to target word length as a function of launch word (top figures) and target word (bottom figures) properties. Shaded bands for partial effects represent 95% CIs based on observation-level LMM residuals.

target words shifted closer to the word centers when the target words were visually simple or predictable. The main effect of target word frequency, however, was not significant. Influences of upcoming words' visual complexity and predictability on saccade target selection indeed echoed the proposal made by Yan et al. (2010, p. 721), that "[l]ow frequency, low contextual predictability, and high visual complexity may contribute to" failure in parafoveal word segmentation, leading to fixation location shifting away from the word center and towards the word beginning. As such, the current study, with large-scale predictability norms, consolidates previous theoretical assumption with novel results.

In contrast to earlier nonsignificant results for word predictability on saccade generation (Li et al., 2014), we found reliable influences of both launch word and target word predictabilities. Possibly, the current larger scale predictability norms make more reliable estimates of these effects. This suggests that nonsignificant findings should be interpreted with caution and that absence of evidence is not evidence of absence. It would be interesting for future studies to explore how Chinese readers' individual word segmentation influences their saccade. For this purpose, a much larger sample of Chinese sentences with word boundary ambiguity and a much larger sample of readers are desirable.

Author Note E21 Humanities and Social Sciences Building, Avenida da Universidade, Taipa, Macau. The authors thank Yingyi Luo for her assistance in checking materials. This research was supported by a Multi-Year Research Grant from the University of Macau (MYRG2020-00120-FSS). Early empirical results were initially presented at the 6th China International Conference on Eye Movements (Kliegl, 2014).

Data availability The BSC is available publicly from the Open Science Framework. Two files can be found and downloaded from this link: https://osf.io/vr3k8/. The file BSC.Word.Info.xlsx is a text file that provides relevant linguistic information as described in Table 3, most critically including the predictability norms. The other file, BSC.EMD.zip, contains the aforementioned eye-tracking data from 60 native readers of Chinese.

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015a). Parsimonious mixed models. arXiv: 1506.04967 [stat.ME].

Bates, D., Maechler, M., Bolker, B., & Walker, S., (2015b). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One, 5*, e10729. https://doi.org/10.1371/journal.pone.0010729.

Chen, L.K., & Carr, H.A. (1926). The ability of Chinese students to read in vertical and horizontal directions. *Journal of Experimental Psychology, 9*, 110–117.

Cohen, J., & Cohen, P. (1975). Applied multiple regression/correlation analysis for the behavioral sciences. Erlbaum.

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eye tracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods, 49*, 602–615. https://doi.org/10.3758/s13428-016-0734-0

Dann, K. M., Veldre, A., & Andrews, S. (2021). Morphological preview effects in English are restricted to suffixed words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001029

Ehrlich, S.F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641–655. https://doi.org/10.1016/S0022-5371(81)90220-6

Engbert, R. & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics, 85*, 77-87.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*, 1035-1045.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*, 621–636.

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

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*, 225–240. https://doi.org/10.3758/BF03211502

Findlay, J.M., & Gilchrist, I.D. (2003). *Active vision: the psychology of looking and seeing*. Oxford University Press.

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

Hohenstein, S., Matuschek, H., & Kliegl, R. (2017). Linked linear mixed models: A joint analysis of fixation locations and fixation durations in natural reading. *Psychonomic Bulletin & Review, 24*, 637–651.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. LEA.

Hoosain, R. (1992). The psychological reality of the word in Chinese. *Advances in Psychology, 90*, 111–130. https://doi.org/10.1016/S0166-4115(08)61889-0.

Husain, S., Vasishth, S., & Srinivasan, N. (2014). Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus. *Journal of Eye Movement Research, 8*. https://doi.org/10.16910/jemr.8.2.3

Hyönä, J., Heikkilä, T. T., Vainio, S., & Kliegl, R. (2021). Parafoveal access to word stem during reading: An eye movement study. *Cognition, 208*, 104547. https://doi.org/10.1016/j.cognition.2020.104547

Inhoff, A.W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 20-34. https://doi.org/10.1037/0096-1523.24.1.20

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

Inhoff, A.W., & Wu, C. (2005). Eye movements and the identification of spatially ambiguous words during Chinese sentence reading. *Memory & Cognition, 33*, 1345-1356.

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 221-246). Elsevier.

Institute of Linguistic Studies. (1986). *Modern Chinese word frequency dictionary*. Beijing Language Institute Publisher. (in Chinese).

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153–168. https://doi.org/10.1016/j.visres.2004.07.037

Kennedy, A., Hill, R., & Pynte, J. (2003). *The Dundee Corpus*. Paper presented at the 12th European Conference on Eye Movement, Dundee, Scotland.

Kennedy, A., Pynte, J., Murray, W.S., & Paul, S.A. (2013). Frequency and predictability effects in the Dundee Corpus: An eye movement analysis. *Quarterly Journal of Experimental Psychology, 66*, 601–618

Kliegl, R. (2007). Towards a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, & Reichle (2007). *Journal of Experimental Psychology: General, 138*, 530-537.

Kliegl, R. (2014, May). *Towards a joint analysis of fixation location and duration during reading of Chinese sentences*. Paper presented at the 6th China International Conference on Eye Movements, Beijing, China.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*, 262-284.

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,*13-35.

Kumle, L., Võ, M.LH., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01546-0

Laurinavichyute, A.K., Sekerina, I.A., Alexeeva, S., Bagdasaryan, K., & Kliegl, R. (2019). Russian Sentence Corpus: Benchmark measures of eye movements in reading in Russian. *Behavior Research Methods, 51*, 1161–1178. https://doi.org/10.3758/s13428-018-1051-6

Lavigne, F., Vitu, F., & d'Ydevalle, G., (2000). The influence of semantic context on initial eye landing sites in words. *Acta Psychologica, 104*, 191-214.

Li, X., Liu, P., & Rayner, K. (2011). Eye movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research, 51*, 1146–1156.

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 Experimental Psychology: General, 143*, 895-913. https://doi.org/10.1037/a0033580

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.

Luke, S.G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology, 88*, 22–60.

Luke, S.G., & Christianson, K. (2018). The Provo Corpus: A large eye-tracking corpus with predictability norms. *Behavior Research Methods,* 50, 826–833. https://doi.org/10.3758/s13428-017-0908-4

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001.

McConkie, G.W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578-586. https://doi.org/10.3758/BF03203972

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28*, 245–253.

McConkie, G.W., Kerr, P.W., Reddix, M.D., Zola, D., & Jacobs, A.M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics, 46*, 245–253.

McDonald, S.A., Carpenter, R.H.S., & Shillcock, R.C. (2005). An anatomically-constrained, stochastic model of eye movement control in reading. *Psychological Review, 112*, 814-840.

O'Regan, J.K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. In: M. Coltheart (Ed.), *Attention and performance. The psychology of reading* (Vol. 12, pp. 363–383). Erlbaum.

O'Regan, K. (1979). Saccade size control during reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics, 25*, 501-509.

O'Regan, J.K. (1980). The control of saccade size and fixation duration during reading: The limits of linguistic control. *Perception & Psychophysics, 28*, 112-117.

Özkan, A., Fikri, F., Kırkıcı, B., Kliegl, R., & Acartürk, C. (2021). Eye movement control in Turkish sentence reading. *Quarterly Journal of Experimental Psychology, 74,* 377-397. https://doi.org/10.1177/1747021820963310

Pan, J., Yan, M., & Laubrock, J. (2017). Perceptual span in oral reading: The case of Chinese. *Scientific Studies of Reading, 21,* 254-263. https://doi.org/10.1080/10888438.2017.1283694

Peng, D.L., Orchard, L.N., & Stern, J.A. (1983) Evaluation of eye movement variables of Chinese and American readers. *The Pavlovian Journal of Biological Science, 18*, 94–102. https://doi.org/10.1007/BF03001861

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception & Psychophysics, 8*, 21–30.

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

Rayner, K., & Well, A.D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*, 504–509. https://doi.org/10.3758/BF03214555

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41,* 943–954.

Rayner, K., Reichle, E.D., Stroud, M.J., Williams, C.C., & 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. https://doi.org/10.1037/0882-7974.21.3.448.

Rayner, K., Li, X., & Pollatsek, A. (2007a). Extending the E-Z Reader model of eye movement control to Chinese readers. *Cognitive Science, 31*, 1021–1033. https://doi.org/10.1080/03640210701703824

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

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., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39,* 4403–4411. https://doi.org/10.1016/S0042-6989(99)00152-2

Richter, E., Yan, M., Engbert, R., & Kliegl, R. (2010, May). *Modeling Chinese reading with SWIFT: How does word segmentation affect targeting?* Paper presented at the 4th China International Conference on Eye Movements, Tianjin, China. Summary available at https://doi.org/10.17605/OSF.IO/VXFYM

Risse, S., Hohenstein, S., Kliegl, R., & Engbert, R. (2014). A theoretical analysis of the perceptual span based on SWIFT simulations of the n+2 boundary paradigm. *Visual Cognition, 22*, 283-308.

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. https://doi.org/10.1080/13506285.2012.670143

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 and Cognition, 26*, 1270-1281.

Shen, E. (1927). An analysis of eye movements in the reading of Chinese. *Journal of Experimental Psychology, 10,* 158–183. https://doi.org/10.1037/h0075609

Sun, F., Morita, M., & Stark, L.W. (1985). Comparative patterns of reading eye movement in Chinese and English. *Perception & Psychophysics 37*, 502–506. https://doi.org/10.3758/BF03204913

Taylor, W.L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly, 30*, 415–433.

Tsai, J.L., & McConkie, G.W. (2003). Where do Chinese readers send their eyes? In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 159-176). Elsevier.

Tsai, J., Lee, C., Tzeng, O.J.L., Hung, D.L., & Yen, N. (2004). Use of phonological codes for Chinese characters: Evidence from processing of parafoveal preview when reading sentences. *Brain and Language, 91*, 235–244.

Tsang, Y.-K., & Chen, H.-C. (2012). Eye movement control in reading: Logographic Chinese versus alphabetic scripts. *PsyCh Journal, 1,* 128-142.

Vainio, S., Hyönä, J., & Pajunen, A. (2009). Lexical predictability exerts robust effects on fixation duration, but not on initial landing position during reading. *Experimental Psychology, 56*, 66–74. https://doi.org/10.1027/1618-3169.56.1.66

Yan, M., & Kliegl, R. (2016). CarPrice versus CarpRice: Word boundary ambiguity influences saccade target selection during the reading of Chinese sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 42*, 1832-1838. https://doi.org/10.1037/xlm0000276

Yan, M., Kliegl, R., Richter, E.M., Nuthmann, A., & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *Quarterly Journal of Experimental Psychology, 63*, 705–725. https://doi.org/10.1080/17470210903114858.

Yan, M., Zhou, W., Shu, H., Yusupu, R., Miao, D., Krugel, A., & Kliegl, R. (2014). Eye movements guided by morphological structure: Evidence from the Uighur language. *Cognition*, 132, 181-215. https://doi.org/10.1016/j.cognition.2014.03.008

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2015). Perceptual span depends on font size during the reading of Chinese sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 209–219. https://doi.org/10.1037/a0038097

Yan, M., Pan, J., Chang, W., & Kliegl, R. (2019a). Read sideways or not: Vertical saccade advantage in sentence reading. *Reading and Writing, 32,* 1911-1926. https://doi.org/10.1007/s11145-018-9930-x

Yan, M., Pan, J., & Kliegl, R. (2019b). Eye movements control in Chinese reading: A cross-sectional study. *Developmental Psychology, 55*, 2275–2285. https://doi.org/10.1037/dev0000819yan.

Yang, H.-M., & McConkie, G.W. (1999). Reading Chinese: Some basic eye-movement characteristics. In J. Wang, A.W. Inhoff, and H-C. Chen (Eds), *Reading Chinese script: A cognitive analysis* (pp. 207-222). Lawrence Erlbaum.

Yang, S.-N., & McConkie, G.W. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology, 16*, 226-261.

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 1192–1204.

Yen, M.-H., Radach, R., Tzeng, O., Hung, D., & Tsai, J.-L. (2009). Early parafoveal processing in reading Chinese sentences. *Acta Psychologica, 131*, 24–33. https://doi.org/10.1016/j.actpsy.2009.02.005

Yen, M.-H., Radach, R., Tzeng, O., & Tsai, J.-L. (2012). Usage of statistical cues for word boundary in reading Chinese sentences. *Reading & Writing, 25*, 1007-1025.

Zhou, W., Wang, A., Shu, H., Kliegl, R., & Yan, M. (2018). Word segmentation by alternating colors facilitates eye guidance in Chinese reading. *Memory & Cognition, 46*, 729–740. https://doi.org/10.3758/s13421-018-0797-5.