Can Native Sentence Reading Processing Mechanisms Predict Second Language Reading Behavior? Computational Modeling of GECO-CN Eye-Tracking Data in Chinese-English Bilinguals

Fang Yifei
College of Chinese Language and Culture, Nankai University, China
fangyifei@mail.nankai.edu.cn

Abstract

This study investigates whether cognitive parameters extracted from first-language (L1, Chinese) reading can directly predict second-language (L2, English) reading behaviors and whether such parameters reflect interlanguage development. As a methodological innovation, we adopt a learner-corpus-based computational approach that integrates eye-tracking data from the GECO-CN corpus, which includes 30 Chinese–English bilinguals reading a full novel, with individualized E-Z Reader model fitting. By applying L1-fitted parameters to L2 texts for zero-parameter transfer prediction, we found that L1 processing routines do not transfer directly to L2 reading; the transfer model yielded significantly higher prediction errors than a native-English baseline. Crucially, online processing dynamics (specifically changes in word-frequency sensitivity and saccadic preparation) were stronger predictors of L2 reading comprehension accuracy than offline vocabulary size (LexTALE), with the former showing significant positive and negative correlations, respectively. This integration of learner corpora with computational modeling reveals a dissociation between lexical competence and procedural reading skills, suggesting that static vocabulary knowledge alone cannot account for real-time reading efficiency. The findings offer empirical insights for L2 reading instruction, demonstrating that computational modeling of naturalistic eye-tracking data can uncover online processing parameters that are more predictive of reading comprehension than static vocabulary measures alone, thus highlighting the need to attend to procedural aspects of reading beyond lexical accumulation.

Keywords: learner corpus; eye-tracking; computational modeling; second language reading; Chinese-English bilinguals

1. Introduction

Reading is a fundamental skill for extracting content from written text. Sentence prediction refers to the cognitive process in which readers pre-activate potential upcoming lexical, semantic, or syntactic structures based on available contextual information during language processing. Such prediction reduces subsequent lexical processing load and further influences eye movement behaviors in natural reading, including fixation duration, skipping probability, and saccadic rhythm(Wu Xiaogang & Chen Qi, 2026). Investigating the cognitive processes underlying reading can help learners improve reading efficiency and attain desired proficiency levels, which is critical for both second language acquisition and the international promotion of Chinese language education.

Whether discussing sentence reading or discourse reading, one cannot separate the discussion from the typological properties of the language itself, as these properties inevitably exert distinct influences on natural reading. With respect to the Chinese writing system, it differs markedly from the Indo-European language family, of which English is representative. The first issue concerns morphology: the English writing system consists of alphabetic letters, and word length may vary according to the number of letters; however, Chinese characters are square-shaped logographs composed of strokes, with each character occupying the same square space and equal spacing. A character is composed of radicals, and radicals are themselves formed through strokes arranged in specific ways. Altering the number of strokes changes the visual complexity of the character but does not change its length. For example, the character"十"has two strokes, whereas"数"has thirteen strokes, yet both words have a character length of one. Consequently, Chinese characters carry a high density of visual information, and readers rely heavily on the holistic processing of character combinations and parafoveal boundary segmentation during reading(Siegelman et al., 2022). This visual complexity affects word recognition to some degree. Previous research has found that fixation durations are longer for highly complex characters(Liversedge et al., 2014).

Second, as an alphabetic script, English has clear physical word boundaries and systematic orthography-to-phonology correspondence rules, whereas Chinese does not employ interword spaces to mark word boundaries(Li & Pollatsek, 2020). This lack of explicit segmentation can make reading more difficult, resulting in longer fixation durations and more complex eye movement patterns. Furthermore, the character composition of Chinese is highly flexible, allowing a limited set of characters (approximately 5,000 commonly used characters) to form a vast number of words (approximately 56,000 common words)(Li et al., 2022), and to create new words that are broadly acceptable and comprehensible, that is, polyphonic characters, such as "将来"and "将军", which in English are represented by entirely different words: future and captain. Models based on alphabetic languages have also emphasized the importance of spaces influencing eye landing positions. Although considerable debate remains regarding the unit of Chinese characters and the principles governing recognition within characters, some research has suggested that word representations can, to some extent, circumvent the interference of word boundaries, thereby revealing the underlying mechanisms of Chinese reading(Li et al., 2014). Therefore, the influence of the internal structure of Chinese characters as a variable on reading mechanisms is not within the scope of the present study.

For a long time, empirical investigations into how the native language influences second-language reading have largely remained at the macro level, focusing on correlations among behavioral test scores, lacking direct examination of real-time cognitive processing mechanisms. There are a large number of Chinese-English bilinguals worldwide, including many international students learning a second language, yet how they manage these two entirely different languages, and whether this management incurs any adverse effects, remains unclear.

Eye-tracking is a tool well suited for achieving real-time measurement of mechanisms in natural bilingual reading. During natural reading, readers' eye movement trajectories, such as fixation duration, skipping rate, and saccade amplitude. They can directly reflect the underlying lexical processing load and oculomotor mechanisms(Rayner, 1998). Several theories and studies have addressed eye movements in natural reading. Koda's (2005) cross-linguistic transfer hypothesis proposes that metallinguistic awareness and automatized decoding strategies developed during native language acquisition profoundly intervene in and reshape saccadic rhythm and lexical access during second-language reading. Perfetti & Harris's (2013) Lexical Quality Hypothesis and the reading universality theory further argue that, despite the marked differences in surface visual features across writing systems, the underlying lexical access and oculomotor control mechanisms of skilled reading are fundamentally shared. For instance, highly predictable words tend to receive shorter fixations, higher skipping rates, and more stable saccadic trajectories. Thus, the brain's lexical prediction mechanism is directly involved in oculomotor control during natural reading. However, most studies have relied heavily on highly controlled isolated-sentence reading paradigms, which, while rigorously controlling experimental variables, eliminate the influence of contextual information and extended discourse, making it difficult to generalize findings to real-world natural reading scenarios.

Therefore, although reliance on conventional eye movement measures can reveal changes in processing load during reading, these measures are ultimately behavioral outcome data and cannot easily distinguish dynamic mechanisms across different cognitive stages. For example, an equivalent lengthening of fixation duration may originate from lexical recognition difficulty, may reflect integration costs following prediction failure, or may even involve alterations in saccadic programming and attentional allocation strategies(Zang et al., 2013). Consequently, purely statistical comparisons often fail to directly elucidate the sources and hierarchical relationships of underlying cognitive mechanisms during reading. Wu (2026), Guan(2021), and others have found that lexical properties, interword information, and parafoveal processing in Chinese reading all influence eye movement patterns, and these complex effects cannot be captured by a single eye movement measure. In recent years, a growing body of research has emphasized the importance of computational modeling in reading research. The core advantage of this approach lies in its ability to decompose continuous, complex eye movement behaviors into a set of latent parameters with psychological interpretability, and further to simulate interactions among different cognitive mechanisms. Mainstream computational models of eye movement control, such as E-Z Reader and SWIFT, have successfully formalized lexical processing, attentional shifting, and saccadic control processes, and have achieved predictive power for natural reading behavior through parameter fitting(Reichle et al., 2003). Relative to traditional statistical analyses, computational modeling not only enhances comparability across experiments, but also directly tests the applicability of different theoretical mechanisms to cross-linguistic reading, making it particularly suitable for investigating transfer mechanisms in bilingual reading.

Despite efforts over recent decades to explain the mechanisms of Chinese reading, the reproducibility of relevant studies remains limited and requires further validation. Moreover, although the reading characteristics of Chinese and English texts differ substantially, they do share certain signature phenomena that affect reading, such as frequency effects and predictability effects(Whitford & Titone, 2015). If this is the case, then how do people read when their two languages are entirely different? Do their two languages interfere with each other during reading?

The present study focuses on cognitive parameters that reflect core reading mechanisms, including lexical recognition efficiency, the ability to utilize word frequency information, saccadic preparation rhythm, and visual attention span. We then attempt to combine individually fitted parameter models with cross-language zero-parameter transfer prediction models, with the aim of investigating: (1) whether individual cognitive parameters extracted from natural L1 Chinese reading can directly predict eye movement behaviors in L2 English reading; (2) if not, which reading parameters undergo significant changes during the cross-linguistic transfer from L1 Chinese to L2 English. We hypothesize that eye movement control strategies from the native language will significantly persist in second language processing, and that proficient bilinguals will exhibit higher reading comprehension efficiency in L2 discourse.

2. Method

2.1 Materials

The data for the present study were derived from the GECO-CN (Ghent Eye-Tracking Corpus for Chinese, Sui et al., 2022), a bilingual eye-movement corpus established by researchers at Ghent University, with open access available at https://osf.io/pmvhd/?view_only=77def2827a514254957cc846e14826cf. This corpus records continuous eye movement trajectories of Chinese–English bilinguals while reading the full text of the British mystery novelist Agatha Christie's work, The Mysterious Affair at Styles. The corpus includes data from 30 proficient bilinguals who were native speakers of Chinese and late learners of English (M_{age} = 25.33 , *SD* = 2.60), with poor data excluded. Their explicit English vocabulary level, as assessed by the LexTALE test, was at an intermediate-to-high level (*M* = 75.75%, *SD* = 12.07%), and their daily exposure to Chinese and English was relatively balanced. Under no imposed experimental task pressure, participants achieved a reading comprehension accuracy of 65.83% (*SD* = 13.27%) for the full English novel, indicating that they possessed the cognitive ability and engagement to sustain long-form natural L2 reading.

Among them, 15 participants were asked to read Chapters 1–7 of the novel in Chinese and Chapters 8–13 in English, whereas the other 15 read in the reverse order. Each participant read approximately 5,000 sentences in total. The Chinese version contained 59,403 words (tokens) with 5,053 distinct types, and the English version contained 56,841 words (tokens) with 5,363 distinct types. Finally, they were also asked to complete 1–6 paper-based multiple-choice reading comprehension questions about the chapter content, with four options per question, to determine whether they had understood the material. The language of the questions matched the language of the chapters, and the number of questions was proportional to the chapter length.

GECO-CN also preserves authentic features of natural reading, such as long-range discourse prediction and skipping of high-frequency function words. In addition, the segmentation of the Chinese novel text was performed manually by researchers according to the Modern Chinese Dictionary, avoiding the inadequacies of automatic segmentation. These data facilitate our modeling of cross-language reading mechanisms.

	L1(Chinese)	L2(English)	*t* L1–L2
(%)	M [SD]	M [SD]	F
Comprehension score	74.67 [11.74]	65.83 [13.27]	3.79 [29]***
Subjective language exposure	49.23 [23.44]	45.03 [22.71]	0.51 [29]
Vocabulary judgment score		41.73 [16.84]	
LexTALE score		75.75 [12.07]	
Spelling score		60.24 [12]	
HSK	95.47 [3.01]		
Age of acquisition (years)		7.73 [3.23]	

Table 1 Descriptive statistics of the GECO-CN Chinese–English bilingual sentence reading eye-movement corpus. *Note.* *p* < .05, *p* < .01, ***p* < .001. The first two columns present the mean (M) with standard deviation (SD*) in brackets; the third column presents the *F*-statistic with degrees of freedom in brackets. Subjective language exposure refers to the percentage of time participants currently spend in a given language environment, as surveyed in the LEAP-Q questionnaire. Because some participants were also exposed to languages other than Chinese and English, the total subjective language exposure does not necessarily sum to 100%.

2.2 Data Preprocessing

To ensure the reliability of individual modeling, strict cleaning criteria were applied during data preprocessing. Trials containing invalid values or abnormal fixations (e.g., NaN) were excluded, and words with extremely low frequency (Zipf ≤ 0) were filtered out. In addition, each participant was required to have more than 50 valid words with usable fits in each language. After cleaning, all participants met the modeling threshold.

2.3 Parameter Estimation for Modeling

Because the prediction process itself cannot be directly observed, computational modeling is often employed to map latent cognitive processing mechanisms onto quantifiable empirical eye-movement parameters. In this way, processes that are otherwise difficult to measure directly, such as prediction efficiency, lexical access speed, and attentional resource allocation, can be indirectly reflected in fixation durations, skipping behaviors, and saccadic rhythms. The modeling approach of the present study primarily follows the eye-movement lexical modeling study by Dias et al.(2021), published in Schizophrenia Bulletin. That study applied computational modeling to eye-movement and ERP (event-related potential) data from schizophrenia patients, and its methodology has considerable relevance for second-language reading as well.

Definitions of variables and indicators	
Lexical measures	
Word_Length	Number of Chinese characters or number of English letters
Word frequency (Zipf)	Converted from raw logarithmic frequency to Zipf constant
Eye-movement measures	
First Fixation Duration (FFD)	Reflects early lexical identification and saccade execution, in ms
Gaze Duration (GD)	Reflects depth of overall lexical processing, in ms
Skip Rate	Reflects visual attention span, in %
Modeling parameters	
Commonly termed lexical access time (a_1)	Simulates baseline time to identify any word, in ms, i.e., intercept in the formula; a larger a1 indicates more effortful reading
Lexical frequency sensitivity(a_2)	Simulates the reader's ability to use frequency to accelerate reading, i.e., slope in the formula; a larger a2 indicates greater ability to extract high-frequency words.
Saccadic preparation time (m_1)	Simulates the minimum preparation time needed for the eye to move from the current fixation to the next word, in ms; a larger m1 indicates longer preparation and slower oculomotor rhythm.
Visual attention span (θ)	Simulates the effective range (in number of characters or letters) over which the eye can process information.

Table 2 Definitions of the indicators used in the present eye-movement modeling. Lexical and eye-movement measures are empirical values; modeling parameters are simulated values.

The variables and parameters selected for the computational model, along with their meanings, are shown in Table 2. First, we extracted two lexical features from the corpus as independent variables: word length and word frequency (Zipf). The dependent eye-movement measures were: (1) First Fixation Duration (FFD), (2) Gaze Duration (GD), and (3) Skip Rate.

Subsequently, we separated the L1 and L2 data from the GECO-CN database by individual participant (30 participants), allowing us to extract dynamic reading-related cognitive characteristics at the individual level. The modeling combines elements of the Monte Carlo stochastic simulation framework from E-Z Reader with the algebraic heuristic modeling approach proposed by Farrell & Lewandowsky (2017). The modeling logic is as follows:

We assume that for each participant there exists a baseline lexical identification time (a1), a lexical frequency sensitivity (a2), a saccadic preparation time (m1), and a visual attention span (θ). Based on prior theory and research, we set plausible bounds for these four parameters to prevent extreme deviations: a1 ∈ [100, 350] ms, a2 ∈ [1, 100] ms, m1 ∈ [100, 250] ms, and θ ∈ [0.5, 8.0] characters or letters.

Next, we defined a set of equations as the forward simulation model. According to cognitive resource allocation theory, the time needed to recognize a word is not constant but is heavily constrained by word frequency. Thus, the theoretical lexical processing time (T_{lex}, in ms) is determined jointly by baseline identification speed and frequency effects, as shown in Equation (1):

$$T_{lex} = \max(a_1 - a_2 \times \text{Zipf}, 50) \quad (1)$$

In Equation (1), T_{lex} represents the baseline lexical processing time minus the time saved due to high frequency. The constant 50 ms denotes the minimum physical transmission time for neural electrical signals(Rayner, 1998). In other words, even when identifying a high-frequency, basic word such as "的" , the processing time cannot fall below the 50-ms floor.

The simulated values for the eye-movement measures (simulated Gaze Duration, simulated First Fixation Duration, and simulated Skip Rate) were obtained using the following equations (Reichle et al., 2006), shown as (2)–(4):

$$\text{Simulated GD} = T_{lex} \quad (2)$$

Equation (2) computes the simulated Gaze Duration for non-skipped words, which equals the lexical processing timeT_{lex}.

$$\text{Simulated FFD} = \min(T_{lex}, m_1) + 20 \quad (3)$$

Equation (3) computes the simulated First Fixation Duration for non-skipped words. When T_{lex} is large (e.g., the participant takes 300 ms to identify a low-frequency word) but the saccadic preparation time m1 has already elapsed, the FFD is determined by m1; conversely, if T_{lex} is small (e.g., when identifying "的"), the FFD is driven by T_{lex}. The constant 20 ms represents the purely mechanical time required for eye muscle contraction (Hautala et al., 2022; Parker et al., 2019).

$$\text{Simulated Skip Rate} = e^{(-\text{Word\_Length}/\theta)} \quad (4)$$

Equation (4) computes the simulated Skip Rate. In mathematics, any continuous process in which the rate of change is proportional to the current quantity itself—such as visual acuity degradation with distance or radioactive decay—can be modeled by the exponential function based on the natural constant *e* (Veldre & Andrews, 2018). Suppose the visual span θ is 4; for a one-character word (e.g., "I"), the skip probability is exp(−1/4) ≈ 77.8%, meaning there is a 77.8% chance of skipping; for a seven-letter word (e.g., "kindness"), exp(−8/4) ≈ 13.5%, so skipping is unlikely.

Next, the model takes the empirical lexical and eye-movement measures from the eye-tracking database for each participant and performs an inverse optimization (i.e., Differential Evolution, DE) against the simulated values. The algorithm initializes random combinations of the parameters within their respective bounds and computes a total loss function, which is a weighted sum (with weights determined empirically by the authors) of the Root Mean Square Errors (RMSE) for the three simulated eye-movement measures, see Equations (5)–(8):

$$RMSE_{FFD} = \sqrt{\text{mean}((\text{Real\_FFD} - \text{Sim\_FFD})^2)} \quad (5)\text{, for non-skipped words}$$

$$RMSE_{GD} = \sqrt{\text{mean}((\text{Real\_GD} - \text{Sim\_GD})^2)} \quad (6)\text{, for non-skipped words}$$

$$RMSE_{Skip} = \sqrt{\text{mean}((\text{Real\_Skip} - \text{Sim\_Skip})^2)} \quad (7)$$

$$\text{Total Loss} = \frac{RMSE_{FFD}}{200} + \frac{RMSE_{GD}}{250} + \left(RMSE_{Skip} \times 5\right) \quad (8)$$

Through thousands of iterations of mutation and competition, the algorithm converges when it finds the parameter combination that minimizes the total loss. This global minimum-error parameter set is then taken as the individual cognitive profile for that participant.

In brief, based on theoretical and prior research, we established a set of equations. Given the empirical eye-movement data for each participant, we generated numerous random initial values for the internal parameters (a_1, a_2, m_1, θ) and iterated until we obtained a set of parameters for which the simulated outcome values minimized the discrepancy from the empirical values (operationally defined in this study as simulated = empirical). This process yields the individual cognitive constants for that participant within our model framework.

2.4 Data Analysis

The data modeling and analyses for the present study were conducted using Python scripts (primarily employing the pandas, NumPy, and SciPy libraries) in VScode (version 1.120.0), and all visualizations were generated using RStudio software (2026.04.0+526, mainly with the dplyr, tidyr, readr, and ggplot2 packages). All code has been open-sourced on the author's personal GitHub repository.

2.4.1 Parameter optimization for individual L1 readers

For each participant, the Chinese reading eye-movement data were fed into the above equations, and parameter optimization was performed using the Differential Evolution algorithm(Storn, 1997). The objective function was defined as the combined Root Mean Square Error (RMSE), with specific empirical weights assigned to the RMSEs of FFD, GD, and Skip Rate to minimize the overall deviation between simulated and empirical values. This procedure yielded a unique set of cognitive parameters （$L1_{a1}, L1_{a2}, L1_{m1}, L1_{\theta}$） representing each participant's native-language reading habits.

2.4.2 Cross-language zero-parameter transfer prediction

This step directly tested the core hypothesis of the present study. The L1 cognitive parameters, without any re-fitting, were directly applied to the English text properties to predict the individual's L2 eye-movement behavior (the transfer model). As a benchmark, we also established a baseline model based on typical eye-movement norms for native English readers （a_1 = 200, a_2 = 10, m_1 = 150, θ = 4.0） (Hautala et al., 2022). It is important to note that, because of the substantial spatial orthographic differences between Chinese characters and English letters, the L1 visual span was converted strictly according to physical retinal size before transfer (1 Chinese character ≈ 2.5 English letters in width) (Inhoff & Liu, 1998). If the RMSE of the transfer model was significantly lower than that of the baseline model, this would confirm the cross-language retention of underlying processing habits.

2.4.3 Re-fitting of L2 parameters

To examine which specific aspects posed the greatest difficulty for bilinguals when reading in English, we again applied the Differential Evolution algorithm to re-fit the English eye-movement data for each individual, following the same procedure as for L1, to extract L2-specific parameters ($L2_{a1}, L2_{a2}, L2_{m1}, L2_{\theta}$). By computing the differences between paired parameters (e.g., $\Delta a_2 = L2_{a2} - L1_{a2}$), we aimed to identify the core factors contributing to L2 reading difficulty.

2.4.4 Individual differences and validity checks

After obtaining the cross-language discrepancy values for each parameter, we conducted Pearson correlation analyses between these differences and the participants' external language characteristics (i.e., the static LexTALE vocabulary scores and text comprehension scores reported in Table 1). This part of the analysis allowed us to explore how individual vocabulary proficiency influences real-time L2 reading efficiency in bilinguals.

3. Results

3.1 Zero-parameter transfer prediction test

We extracted the individually fitted L1 parameters from the Chinese reading data and directly applied them to the L2 corpus without any re-optimization to perform zero-parameter forward prediction (the transfer model). The fitting error (RMSE) of this transfer model was then compared against that of a baseline model, which was based on parameters derived from typical native English readers. The results of a paired-samples t-test (see Figure 1) showed that the prediction error of the baseline model (*M* = 3.75, *SD* = 0.34) was significantly lower than that of the zero-parameter transfer model (*M* = 3.98, *SD* = 0.24), $t(29) = 4.08, p < .001, Cohen's\ d = 0.74$ (see Table 3A). This result rejected the hypothesis of L1 assimilation transfer. The significant increase in RMSE prediction error indicates that, when Chinese–English bilinguals engage in natural L2 reading, their L1 reading processing mechanisms cannot adequately predict their L2 reading behavior.

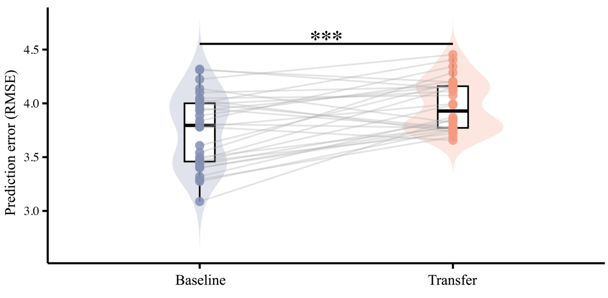

Fig.1 Results of the cross-language zero-parameter transfer prediction from L1 Chinese to L2 English

	L1 Baseline M *(SD)*	L2 MTransfer M *(SD)*	t(29)	*p* (two-tailed)	Cohen's d
A. Transfer prediction error					
Prediction error (RMSE)	3.75 (0.34)	3.98 (0.24)	4.08	< .001***	0.80
B. Paired-samples t-test results for cognitive parameters					
Baseline lexical identification time a_1 (ms)	305.76 (43.85)	350.00 (0.00)	5.53	< .001***	
Lexical frequency sensitivity a_2	13.37 (4.53)	6.27 (6.04)	4.65	< .001***	1.35
Saccadic preparation time m1 (ms)	190.96 (24.41)	217.13 (24.13)	9.93	< .001***	1.08
Visual attention span θ	9.89 (2.16)	3.22 (1.06)	21.51	< .001***	4.00
C. Pearson correlations between cognitive parameter discrepancies and individual language proficiency measures					
Δa_1 vs. text reading comprehension accuracy			r = 0.19	.325	
Δa_2 vs. text reading comprehension accuracy			r = 0.42	.020*	
Δm_1 vs. text reading comprehension accuracy			r = −0.50	.005**	
Δa_1LexTALE_EN			r = 0.27	.153	
Δa_2 vs. LexTALE_EN			r = 0.28	.141	
Δm_1 vs. LexTALE_EN			r = 0.05	.806	

Table 3 A. Transfer prediction error; B. Paired-samples t-test results for L1 and L2 cognitive parameters; C. Pearson correlations between cognitive parameter discrepancies and individual language proficiency measures.

3.2 Cross-language differences in individual cognitive parameters

As the zero-parameter transfer prediction yielded poor results, we proceeded to re-fit the L2 eye-movement data for each participant and compared the resulting L2 parameters against their L1 counterparts through paired comparisons (see Table 3B). Figure 2 illustrates the differences between L1 and L2 for lexical frequency sensitivity and saccadic preparation rhythm, respectively. In the figure, green indicates a decrease and red indicates an increase; each of the 30 dots represents one participant, and the connecting line shows the trend between L1 and L2 for the same individual.

For baseline lexical identification time (a_1), all L2 parameters reached the model's upper bound of 350 ms, which was significantly higher than the L1 value (*M* = 305.76, *SD* = 43.85), t(29) = 5.53, *p* < .001. Notably, 10 participants had already reached this upper bound in L1. For lexical frequency sensitivity (a_2), the L2 value (*M* = 6.27, *SD* = 6.04) was significantly lower than the L1 value (*M* = 13.37, *SD* = 4.53), t(29) = 4.65, *p* < .001 (see Figure 2a). For saccadic preparation time (m_1), the L2 value (*M* = 217.13, *SD* = 24.13) was significantly longer than the L1 value (*M* = 190.96, *SD* = 24.41), t(29) = 9.93, *p* < .001 (see Figure 2b). For visual attention span (θ), after conversion of retinal size between Chinese characters and English letters (1 Chinese character ≈ 2.5 English letters in width), the L2 value (*M* = 3.22, *SD* = 1.06) was significantly smaller than the adjusted L1 value (*M* = 9.89, *SD* = 2.16), t(29) = 21.51, *p* < .001. Although the finding for visual attention span (θ) is in line with expectations, its interpretability requires further examination because English words are on average longer than Chinese words; therefore, this measure is not presented in this section.

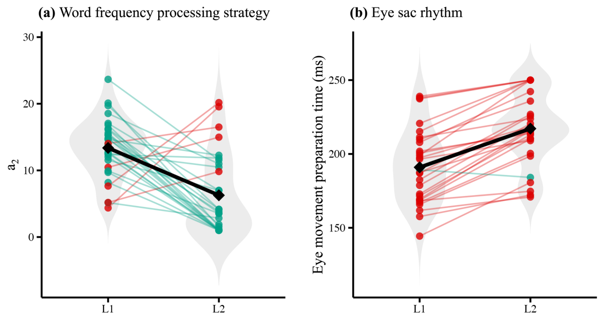

Fig.2 Individual differences in core cross-language cognitive parameters.(a) Paired-samples t-test results for lexical frequency sensitivity (a_2) between L1 and L2; (b): Paired-samples t-test results for saccadic preparation time (m_1) between L1 and L2

3.3 Relationship between cross-language changes in individual cognitive parameters and individual vocabulary proficiency

To examine whether the cross-language changes in parameters were associated with actual reading performance, we computed discrepancy values for each parameter ($\Delta = L2 - L1$) and conducted Pearson correlation analyses between these discrepancies and participants' reading comprehension accuracy on the English passages (COMPR_EN) (see Fig.3). The change in lexical frequency sensitivity (Δa_2) was significantly positively correlated with text reading comprehension accuracy $r = 0.42$, $p = .020$), whereas the change in saccadic preparation time (Δm_1) was significantly negatively correlated with text reading comprehension accuracy ($r = -0.50$, $p = .005$). As a discriminant validity check, none of the cognitive parameter discrepancy measures showed significant correlations with individual vocabulary knowledge (LexTALE_EN) (Δa_1: $r = 0.27$, $p = .153$; Δa_2: $r = 0.28$, $p = .141$; Δm_1: $r = 0.05$, $p = .806$; $\Delta\theta$: $r = -0.24$, $p = .202$). This pattern indicates that changes in natural-reading cognitive processing parameters are associated with actual text comprehension, but not with individuals' vocabulary knowledge (see Table 3C).

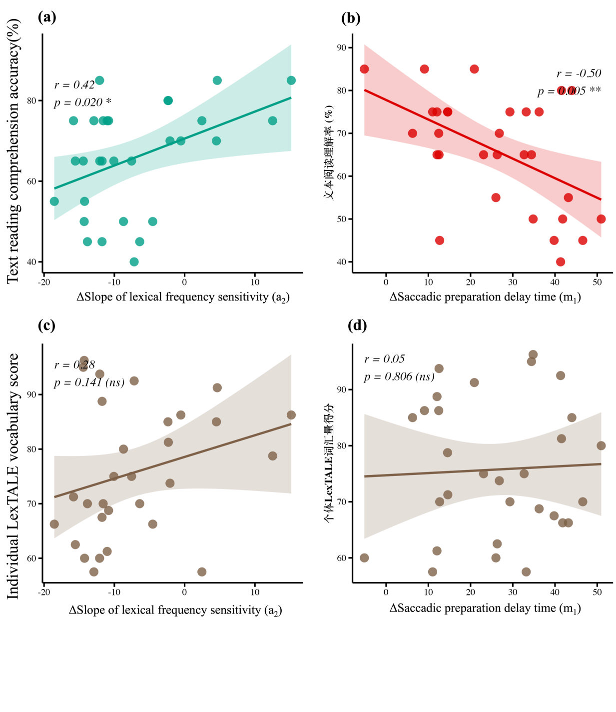

Fig.3 (a)Relationship between frequency processing and text reading comprehension accuracy; (b)Relationship between oculomotor rhythm and text reading comprehension accuracy; (c)Relationship between frequency processing and individual vocabulary proficiency; (d)Relationship between oculomotor rhythm and individual vocabulary proficiency.

4. Discussion

4.1 Native language reading processing mechanisms have limited predictive power for second-language reading behavior

The primary finding of this study is that the zero-parameter transfer model, which was based on native Chinese reading parameters, did not predict L2 English reading behavior as well as the baseline model. This indicates that cognitive parameters developed through native-language reading cannot be directly applied to second-language reading contexts. In bilingual reading, the processing systems of the two languages are not simply additive or directly translatable; rather, they interact in a complex manner that involves both simultaneous influence and mutual constraint. Chinese reading lacks physical word boundaries, and readers rely more heavily on contextual and frequency information to engage in broader pre-processing, with relatively flexible fixation selection(McBride, 2016; Snell & Grainger, 2019). In other words, in the absence of physical boundary constraints, the brain uses high-frequency words to filter fixation targets, resulting in a very large a2 value in Chinese reading and an expanded parafoveal pre-processing span, allowing readers to skip across word units, phrases, and even lines and paragraphs during natural reading. In contrast, the English reading system, which has explicit spaces as boundaries, relies on grapheme-to-phoneme linear conversion at the underlying level(Cui, 2023). In this context, Chinese–English bilinguals tend to read in a word-by-word or syllable-by-syllable manner, thereby diminishing the effectiveness of lexical frequency sensitivity (a_2).

4.2 Reading cognitive processing parameters generally predict individual text comprehension

The present study found that cognitive processing parameters during natural reading significantly predict text comprehension. First, our findings are broadly consistent with the general literature on frequency effects, which have been tested across various alphabetic scripts and Chinese characters. For example, for two-character words, Yan et al. (2006) argued that high-frequency words are stored as whole units in the reader's mental lexicon, whereas low-frequency words are stored as individual characters. Studies on compound words in alphabetic languages such as English, Dutch, and Finnish have also found that the frequency of constituent morphemes significantly influences compound word processing, thus demonstrating a morpheme-level frequency effect (Pollatsek & Hyönä, 2005). In sum, the presence of high-frequency words not only accelerates morpheme-level identification but also enhances the coherence of semantic integration, thereby improving comprehension accuracy at the sentence and text levels. The Lexical Quality Hypothesis also states that the efficiency of word identification is modulated not only by word frequency but also by the degree of consistency among phonological, orthographic, and semantic representations. Therefore, lexical frequency sensitivity is highly correlated with semantic integration and contextual prediction ability, which also complements the findings discussed in (3). Second, we found that saccadic preparation time (m1) was negatively correlated with text comprehension, suggesting that individuals with slower oculomotor rhythms performed worse in discourse comprehension. This may reflect the influence of processing delays or inefficient attentional allocation on comprehension.

4.3 Individual vocabulary knowledge and natural reading procedures may be dissociated

Traditional second-language acquisition theories hold that larger vocabulary size leads to more fluent reading. However, the results of the present study showed that LexTALE, a static measure of explicit English vocabulary breadth, failed to predict any of the cognitive processing parameters (Fioravanti & Siyanova-Chanturia, 2024). This finding diverges from traditional SLA accounts. One possible explanation, as suggested by Kiliańska-Przybyło & Ślęzak-Świat (2021) is that the core of proficient bilingual reading lies in the efficiency of retrieval during processing, rather than in the sheer size of the vocabulary store. A more plausible explanation, in our view, is that discourse reading provides rich contextual information; even when participants encounter unfamiliar words, they may still infer meaning through contextual cues. That is, some words may carry such strong contextual weight that they allow readers to maintain reading efficiency even without fully recognizing other words, thereby significantly reducing the direct impact of individual word prediction ability on reading efficiency. For instance, Liu Zhifang et al. (2020) showed through three experiments that contextual predictability operates at both the whole-word access stage and the within-word final-character processing stage, whereas the interaction between contextual predictability and word frequency or first-character frequency was not significant. Contextual predictability can pre-activate lexical representations independently of lexical knowledge, thus improving reading efficiency. Although the present study did not explore this issue in depth, the possible dissociation between individual vocabulary knowledge and natural reading procedures, as well as the potential competition with contextual inference ability, warrants further investigation. Future research could further examine the competition or facilitation among these abilities.

4.4 Differences in cognitive models between native and second language may help investigate different stages of interlanguage development

The methods and findings of this study also have implications for the study of interlanguage development. Pélissier et al. (2023)showed that during sentence reading, bilinguals exhibit mutual inhibition among form-related lexical neighbors, and that the degree of inhibition is modulated by L2 proficiency, indicating that L2 ability itself influences cross-linguistic competition and processing strategies. As L2 proficiency increases, individuals gradually enhance their ability to manage cross-linguistic competition, leading to changes in reading strategies and oculomotor control patterns. The GECO-CN corpus used in this study consists of university-level participants with intermediate-to-high English proficiency. To further investigate the mutual influence between native Chinese and L2 English reading processing mechanisms, it would be necessary to include data from native English speakers and L2 Chinese learners as control groups. By introducing reading parameters from native English speakers or using integrated Chinese–English bilingual parameters to systematically predict L2 reading behavior, it may be possible to reveal differences in individual interlanguage processing mechanisms under different integration approaches, thereby reflecting different stages of interlanguage development. This suggests that computational modeling, compared with traditional experimental designs, can better reveal the interplay of multiple factors during natural discourse reading, and thus more closely approximates real-world reading situations.

4.5 Limitations and future directions

The present study benefited from rich individual eye-movement data; however, some degree of overfitting occurred after modeling. This study represents an initial attempt by the authors to apply eye-movement modeling to second-language acquisition research, and the parameter settings are not yet fully mature and require further adjustment, optimization, and validation in future work. Finally, the current heuristic algebraic model made certain simplifications in probability distributions to achieve individual fitting. Future studies may consider incorporating deep learning algorithms to enable cross-language parameter tracking within a stochastic simulation framework using more complex computational models.

5. Conclusion

Using the GECO-CN bilingual eye-movement corpus, the present study applied E-Z Reader model fitting and zero-parameter transfer prediction, and found that L1 Chinese reading parameters had limited predictive power for L2 English reading behavior. The results showed that among the parameter discrepancy measures, changes in lexical frequency sensitivity and saccadic preparation time were associated with reading comprehension accuracy, whereas changes were unrelated to static vocabulary size. This suggests that, in natural reading, online processing parameters may play a stronger role in predicting comprehension performance than offline lexical knowledge alone. In addition to vocabulary accumulation, L2 instruction may also need to focus on how learners utilize frequency information and adjust oculomotor rhythm during actual reading. Given the limitations of the sample scope and model simplifications, this study provides only a preliminary investigation of intermediate-to-advanced bilinguals. Future research could include participants at different proficiency levels and employ more refined computational models to further examine the relationship between parameter changes and reading performance.

6. Bibliographical References

Cui, Y. (2023). Eye movements of second language learners when reading spaced and unspaced Chinese texts. Frontiers in Psychology, 14, 783960. https://doi.org/10.3389/fpsyg.2023.783960

Dias, E. C., Sheridan, H., Martínez, A., Sehatpour, P., Silipo, G., Rohrig, S., Hochman, A., Butler, P. D., Hoptman, M. J., Revheim, N., & Javitt, D. C. (2021). Neurophysiological, oculomotor, and computational modeling of impaired reading ability in schizophrenia. *Schizophrenia Bulletin*, *47*(1), 97–107. https://doi.org/10.1093/schbul/sbaa107

Farrell, S., & Lewandowsky, S. (2017). *Computational modeling of cognition and behavior*. Cambridge University Press.

Fioravanti, I., & Siyanova-Chanturia, A. (2024). Eye movements in the investigation of different properties of multi-word expressions: A systematic review. *Research Methods in Applied Linguistics*, *3*(1), 100092. https://doi.org/10.1016/j.rmal.2023.100092

Guan Qun, Li Yifei, & Li Xingguo. (2021). Eye-tracking research on vocabulary and text integration in English second language reading. *Journal of PLA Foreign Languages Institute*, *44*(1), 19–26, 159.

Hautala, J., Hawelka, S., Loberg, O., & Leppänen, P. H. T. (2022). A dynamic adjustment model of saccade lengths in reading for word-spaced orthographies: Evidence from simulations and invisible boundary experiments. *Journal of Cognitive Psychology*, *34*(4), 435–453. https://doi.org/10.1080/20445911.2021.2011895

Inhoff, A. W., & Liu, W. (1998). The PerceptualSpan and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1,), 20.

Kiliańska-Przybyło, G., & Ślęzak-Świat, A. (2021). Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide. Aline godfroid. Digital Scholarship in the Humanities, 36(1), 243–246. https://doi.org/10.1093/llc/fqaa075

Koda, K. (2005). Insights into second language reading: A cross-linguistic approach (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139524841

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(2), 895–913. https://doi.org/10.1037/a0033580

Li, X., Huang, L., Yao, P., & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. Nature Reviews Psychology, 1(3), 133–144. https://doi.org/10.1038/s44159-022-00022-6

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. Psychological Review, 127(6), 1139–1162. https://doi.org/10.1037/rev0000248

Liu Zhifang, Tong Wen, Zhang Zhijun, & Zhao Yajun. (2020). The impact of contextual predictivity on word processing during reading: Eye-tracking evidence. Acta Psychologica Sinica, 52(9), 1031–1047.

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. Visual Cognition, 22(3–4), 441–457. https://doi.org/10.1080/13506285.2014.889260

McBride, C. A. (2016). Is Chinese special? Four aspects of chinese literacy acquisition that might distinguish learning chinese from learning alphabetic orthographies. Educational Psychology Review, 28(3), 523–549. https://doi.org/10.1007/s10648-015-9318-2

Pélissier, M., Haugland, D., Handeland, B., Urland, B. Z., Wetterlin, A., Wheeldon, L., & Frisson, S. (2023). Competition between form-related words in bilingual sentence reading: Effects of language proficiency. Bilingualism: Language and Cognition, 26(2), 384–401. https://doi.org/10.1017/S1366728922000529

Perfetti, C. A., & Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. Language Learning and Development, 9(4), 296–316. https://doi.org/10.1080/15475441.2013.813828

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of finnish compound words. Language and Cognitive Processes, 20(1–2), 261–290. https://doi.org/10.1080/01690960444000098

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372. https://doi.org/10.1037/0033-2909.124.3.372

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. Behavioral and Brain Sciences, 26(4), 445–476. https://doi.org/10.1017/S0140525X03000104

Siegelman, N., Schroeder, S., Acartürk, C., Ahn, H.-D., Alexeeva, S., Amenta, S., Bertram, R., Bonandrini, R., Brysbaert, M., Chernova, D., Da Fonseca, S. M., Dirix, N., Duyck, W., Fella, A., Frost, R., Gattei, C. A., Kalaitzi, A., Kwon, N., Lõo, K., … Kuperman, V. (2022). Expanding horizons of cross-linguistic research on reading: The multilingual eye-movement corpus (MECO). Behavior Research Methods, 54(6), 2843–2863. https://doi.org/10.3758/s13428-021-01772-6

Snell, J., & Grainger, J. (2019). Readers are parallel processors. Trends in Cognitive Sciences, 23(7), 537–546. https://doi.org/10.1016/j.tics.2019.04.006

Storn, R. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization, 11(4), 341–359. https://doi.org/10.1023/A:1008202821328

Sui, L., Dirix, N., Woumans, E., & Duyck, W. (2022). GECO-CN: Ghent eye-tracking COrpus of sentence reading for Chinese-English bilinguals. Behavior Research Methods, 55(6), 2743–2763. https://doi.org/10.3758/s13428-022-01931-3

Veldre, A., & Andrews, S. (2018). How does foveal processing difficulty affect parafoveal processing during reading? Journal of Memory and Language, 103, 74–90. https://doi.org/10.1016/j.jml.2018.08.001

Whitford, V., & Titone, D. (2015). Second-language experience modulates eye movements during first- and second-language sentence reading: Evidence from a gaze-contingent moving window paradigm. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41(4), 1118–1129. https://doi.org/10.1037/xlm0000093

Wu Xiaogang & Chen Qi. (2026). A cross-level eye-tracking study of sentence prediction processing patterns among Chinese English learners. Modern Foreign

Languages, 49(2), 218–228. https://doi.org/10.20071/j.cnki.xdwy.20251219.006
Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. British Journal of Psychology, 97(2), 259–268. https://doi.org/10.1348/000712605X70066
Zang, C.-L., Zhang, M.-M., Guo, X.-F., Liu, J., Yan, G.-L., & Bai, X.-J. (2013). Several effects on Chinese lexical processing: Evidence from eye movements: several effects on chinese lexical processing: evidence from eye movements. Advances in Psychological Science, 20(9), 1382–1392. https://doi.org/10.3724/SP.J.1042.2012.01382