Comparative Analysis of Commonly Used Eye Movement Control Models in Reading Research

Authors: Chen Songlin, Chen Xinwei, Li Huangxia, Yao Panpan, Yao Panpan

Date: 2023-09-12T00:00:00+00:00

Abstract

Based on serial processing theory, parallel processing theory, and interactive activation theory, some classic eye movement control models have simulated and investigated common eye movement behaviors, experimental effects, and their underlying cognitive mechanisms of information processing during reading. Five classic eye movement control models in reading research—E-Z Reader 10th, SWIFT, Glenmore, OB1 Reader, and CRM—exhibit both similarities and unique characteristics in model architecture, basic model logic, explanation of common eye movement behaviors, and explanation of common experimental effects. Based on the comparative analysis of the aforementioned models, future models need to examine issues of post-lexical integration, word order, and extralinguistic factors, attempt to provide some explanation for the latest empirical research findings on preferred viewing location, establish unified standards for model comparison, and explore the cross-linguistic explanatory power of each model.

Full Text

A Comparative Analysis of Commonly Used Eye Movement Control Models in Reading Research

Chen Songlin, Chen Xinwei, Li Huangxia, Yao Panpan
School of Psychology, Beijing Language and Culture University, Beijing 100083, China

Abstract

Based on sequential processing theory, parallel processing theory, and interactive activation theory, several classic eye movement control models have been constructed to simulate common eye movement behaviors and experimental effects during reading, thereby exploring the underlying cognitive mechanisms of information processing. Among the five classic eye movement control models in reading research—E-Z Reader 10th, SWIFT, Glenmore, OB1 Reader, and CRM—there exist both similarities and unique characteristics in model structure, basic logic, explanations of common eye movement behaviors, and interpretations of typical experimental effects. Building upon this comparative analysis, future models need to examine issues concerning post-lexical integration, word order, and extra-linguistic factors. They should also attempt to explain recent empirical findings regarding preferred viewing location, establish unified standards for model comparison, and explore the cross-linguistic explanatory power of each model.

Keywords: E-Z Reader, SWIFT, Glenmore, OB1 Reader, CRM, eye movement control
Classification Number: B842.5

Reading is a sophisticated cognitive activity that coordinates a series of distinct cognitive processes, including visual information processing, character and word recognition, attention allocation and shifting, and oculomotor control. Investigating the cognitive mechanisms underlying reading processes represents a core topic in psycholinguistics, while uncovering the principles of eye movement control during reading has become a particularly active area of research. In recent years, based on extensive eye-tracking experimental data, researchers have developed various computational models to simulate information processing during reading (Li, Huang et al., 2022; Mézière et al., 2021; Yu et al., 2021; Zhang, Yao et al., 2022). This paper systematically analyzes five classic computational models of eye movement control and compares their structures, logic, and explanations of common reading-related eye movement behaviors and experimental effects. A thorough understanding of these models will enhance our knowledge of actual reading processes.

Based on differences in visual attention allocation within the perceptual span, computational models of eye movement control can be broadly categorized into two types: sequential attention shift models and parallel graded processing models. These two theoretical camps have remained controversial since their inception (Reichle, Liversedge et al., 2009; Snell & Grainger, 2019a; Snell & Grainger, 2019b; for reviews see Murray et al., 2013; Rayner et al., 2003; Ma & Li, 2012; Sui et al., 2013; Wu & Mo, 2008). The fundamental assumption of sequential processing models is that attention can only be allocated to one word at a time and shifts from one word to the next in a strictly serial order. The typical representative of this approach is E-Z Reader (Pollatsek et al., 2006; Rayner et al., 2005; Reichle et al., 1998, 1999, 2003, 2006; Reichle, Warren et al., 2009). In contrast, parallel processing models assume that multiple words can be processed simultaneously, with the processing degree of each word modulated by a gradient of visual spatial attention. Typical representatives include SWIFT (Engbert et al., 2002, 2005; Richter et al., 2006), Glenmore (Reilly & Radach, 2006), and OB1 Reader (Snell et al., 2018). However, all these models were constructed based on empirical findings from alphabetic writing systems (English and German). Due to language-specific characteristics, whether these models possess explanatory power for non-alphabetic scripts such as Chinese remains to be verified. Chinese writing consists of square-shaped characters without explicit word boundary markers, posing challenges for existing models in simulating word segmentation and word recognition. Addressing these specific features of Chinese writing, Li and Pollatsek (2020) proposed the Chinese Reading Model (CRM). This model advocates parallel processing of characters within the visual span and parallel processing of words, though the activation strength of word n+1 is generally lower, resulting in slower processing.

Existing domestic literature has provided detailed introductions to E-Z Reader, SWIFT, and Glenmore (Chen & Deng, 2006; Hu et al., 2007; Liu et al., 2006; Shen et al., 2002; Sui et al., 2018), but lacks coverage of the updated E-Z Reader (Reichle, Warren et al., 2009), OB1 Reader, and CRM. Moreover, previous review articles have focused on introducing individual models without conducting systematic comparative analyses, potentially leading to isolated and one-sided understanding. This paper introduces the latest versions of these models and provides a systematic comparative analysis, offering prospects for future developments in eye movement control modeling. For a more systematic and intuitive comparison, we have summarized the characteristics of each model in Table 1 .

2.1 E-Z Reader

E-Z Reader represents the prototypical sequential processing model, proposed by Reichle, Pollatsek, Rayner, and colleagues based on English reading research (Pollatsek et al., 2006; Rayner et al., 2005; Reichle et al., 1998, 1999, 2003, 2006; Reichle, Warren et al., 2009). Since its initial proposal in 1998, the model has evolved through ten versions to accommodate new experimental findings and has been widely used to explain various eye movement phenomena during reading (Bordag & Opitz, 2022; Tschense & Wallot, 2022; Veldre et al., 2023; Yang et al., 2021). E-Z Reader's development has progressed from simple to complex. Specifically, Model 1 (Reichle et al., 1998) examined the effect of word frequency on lexical recognition. Model 2 (Reichle et al., 1998) added the influence of predictability. Model 3 (Reichle et al., 1998) incorporated the mechanism of multiple fixations on a single word. Model 4 (Reichle et al., 1998) introduced the factor of eccentricity. Model 5 (Reichle et al., 1998) proposed different mechanisms of eccentricity effects, assuming that eccentricity has a greater impact on lexical access. Model 6 (Reichle et al., 1999) clarified the influence of various variables on reading eye movement control at the character level while refining the refixation mechanism. Model 7 (Reichle et al., 2003) added an early visual processing stage and improved formulas and parameters based on new experimental and simulation data. Models 8 (Rayner et al., 2005) and 9 (Pollatsek et al., 2006; Reichle et al., 2006) simulated results from new experimental paradigms (e.g., moving window, disappearing text) and adjusted parameters to better reflect reality. Model 10 (Reichle, Warren et al., 2009) introduced a post-lexical integration stage to explain the influence of higher-level factors on reading, making it the only model among the five discussed here that addresses post-lexical integration after word recognition. In this series, we focus on the latest version, E-Z Reader 10th (hereinafter referred to as E-Z Reader, with its structure shown in Figure 1 [Figure 1: see original paper]). This model comprises two components: lexical processing and saccade control. Lexical processing includes three stages: familiarity check, lexical access completion, and post-lexical integration. Saccade control includes three phases: labile saccade programming, non-labile saccade programming, and saccade execution. These two components interact: completion of the familiarity check triggers saccade programming to the next word, lexical access completion shifts attention to the next word, and post-lexical integration modulates saccade control. The model successfully explains how word frequency, word length, visual acuity, predictability, and post-lexical integration affect readers' eye movements during reading.

2.2 SWIFT

SWIFT represents the prototypical parallel processing model, constructed based on German reading data by Engbert, Nuthmann, Richter, and Kliegl (Engbert et al., 2002, 2005; Richter et al., 2006). It has been updated to its second version (model structure shown in Figure 2 [Figure 2: see original paper]), with the second version providing more detailed elaboration of the first. This model has also been used to simulate and explain numerous experimental effects (Antúnez et al., 2022; Gregg et al., 2023; Snell et al., 2023). Unlike E-Z Reader's assumption that attention can only be allocated to one word at a time with strictly serial lexical processing, SWIFT posits that attention is distributed in parallel across all words within the perceptual span (typically four words). Multiple words are processed simultaneously, though their processing degree and speed differ (influenced by factors such as visual acuity, frequency, and predictability). These words compete, and the winner becomes the saccade target. Similar to E-Z Reader, SWIFT includes lexical processing and saccade control components. Lexical processing comprises a pre-processing stage and a lexical access stage, while saccade control includes labile and non-labile stages. The pre-processing stage handles basic natural properties of words, with lexical activation gradually increasing from 0 to maximum. During lexical access, activation decays from maximum until lexical processing is complete (activation returns to 0), at which point attention resources are reallocated based on activation levels of multiple words in the perceptual span. SWIFT assumes all saccades are generated randomly, meaning there are random intervals between saccades during reading, though saccade planning and execution are influenced by current lexical processing. Factors such as word frequency and predictability of the currently processed word affect saccade planning within a certain range (foveal target inhibition mechanism). The model successfully explains how word frequency, predictability, and visual acuity influence reading eye movements.

2.3 Glenmore

Both E-Z Reader and SWIFT begin modeling from lexical recognition without addressing sub-lexical information (letter) processing. To address this limitation, Glenmore was developed by Reilly and Radach (Reilly & Radach, 2006; model structure shown in Figure 3 [Figure 3: see original paper]). Based on German data, this model employs parallel processing theory and an interactive activation architecture. It has also been used to simulate and explain numerous experimental effects (Brossette et al., 2022; Gordon et al., 2020; Schwalm & Radach, 2023). Specifically, Glenmore includes three processing layers—visual input, letter, and word layers—and several modules including a saliency map (distribution of activation values for different processing units within the perceptual span), a fixation center, and a saccade generator. For word recognition, Glenmore assumes that visual information is transmitted to letter and word layers, completing letter-level and word-level processing of fixated words. The letter and word layers interact: information from the letter layer facilitates word recognition, while the word layer provides feedback to maintain letter activation. Simultaneously, multiple activated words compete, with each word's activation strength and speed influenced by its frequency. For saccade control, the model assumes target selection occurs through parallel processing and competition in the saliency map. Similar to SWIFT, visual and letter information from the currently fixated word is converted into a saliency map that influences saccade target selection. Actual saccade triggering is controlled by a fixation center module that receives input from multiple cognitive processing pathways (letter layer activation, word layer activation that indirectly affects the fixation center via the letter layer, reading task, material difficulty). Once the fixation center's activation exceeds an adjustable threshold, a saccade is triggered. Additionally, to explain spillover effects, the model assumes that after a fixation, if a word is accessed, its activation becomes zero; if not accessed, its activation can be carried over to the next fixation.

2.4 OB1 Reader

Although Glenmore incorporates letter processing, the connections between letters and words are predetermined—letters in the letter layer are pre-assigned to specific words rather than activating all possible words as described in traditional interactive activation models. The same applies to connections from the word layer to the letter layer. This operation does not align with cognitive reality. Moreover, existing parallel processing models (SWIFT, Glenmore) have not adequately resolved the conflict between word activation order and sentence word order. To address these issues, Snell, Leipsig, Grainger, and Meeter constructed OB1 Reader in 2018 based on German data, combining parallel processing theory with interactive activation models (Snell et al., 2018; model structure shown in Figure 4 [Figure 4: see original paper]). This model innovatively incorporates a spatial-topic information representation module, resolving the conflict between activation order and word order and sparking extensive investigation into word order issues in reading mechanisms (Brossette et al., 2022; Dufour et al., 2022; Mirault et al., 2022; Pegado & Grainger, 2020; Primativo et al., 2022; Zhang, Wang et al., 2022). Specifically, the model assumes that visual input activates open bigram nodes and spatial-topic information representations containing word length information, with open bigram activation modulated by visual acuity, crowding, and attentional weight. Open bigram nodes activate multiple lexical nodes that compete and are matched to spatial-topic information representations. Successful matching leads to lexical access, while failed matching results in non-access. For saccade control, the model assumes saccade triggering is random, with target selection depending on word recognition: when an activated word successfully matches spatial-topic information, a saccade is planned to the most salient word in the visual field (its saliency being the sum of constituent letter activation values); if matching fails, a regression is planned to unidentified words on the left. The model also assumes attentional span size depends on whether a word is accessed: span expands upon successful access and contracts upon failure. OB1 Reader explains effects such as word length and orthographic neighborhood size.

2.5 CRM

The four models described above were constructed based on research from alphabetic writing systems (English and German). Given Chinese language specificity, these models cannot directly explain eye movement phenomena in Chinese reading. Unlike alphabetic scripts, Chinese lacks inter-word spaces to mark word boundaries. The absence of word boundaries prevents readers from using spaces for word segmentation and saccade target selection, challenging the four models' ability to explain Chinese reading. How word segmentation occurs in Chinese reading becomes a primary issue that must be addressed in building Chinese reading models. Additionally, Chinese uses characters as writing units. These square-shaped, meaning-bearing characters consist of different strokes, exhibit substantial complexity differences, and far outnumber letters in alphabetic systems (Li et al., 2017). These characteristics make character-level processing a central focus in Chinese reading, presenting another challenge for existing models.

Some researchers attempted to extend E-Z Reader to Chinese reading, simulating effects such as word frequency (Rayner et al., 2007). However, this attempt did not consider Chinese specificity—namely, the absence of clear word boundary markers to assist saccade target selection—making it inconsistent with actual Chinese reading. It also failed to consider the impact of character recognition on Chinese reading.

Li and Pollatsek constructed a Chinese reading model to address these issues based on Chinese eye movement data (Li & Pollatsek, 2020; model structure shown in Figure 5 [Figure 5: see original paper]). CRM provides a theoretical foundation for investigating Chinese reading mechanisms (Li, Huang et al., 2022; Liao et al., 2022; Sui et al., 2022; Yao, Alkhammash et al., 2022; Yao, Slattery et al., 2022; Yao, Staub et al., 2022; Zhang, Bai et al., 2022; Zhang, Yao et al., 2022). The model employs interactive activation logic and comprises two modules: lexical processing and eye movement control. Similar to Glenmore, CRM's lexical processing module includes visual, character, and lexical layers. The character and lexical layers influence each other, while activated nodes within each layer compete and inhibit one another. Specifically, all characters within the perceptual span are activated in parallel, and all possible words composed of these characters are also activated. Activated words compete until a single winner emerges. Once a word wins, it is segmented from the string, completing lexical recognition. Thus, CRM posits that word segmentation and word recognition occur simultaneously and are inseparable in Chinese reading. The eye movement control module includes saccade units, fixation word units, character activation maps, and saccade target selection, with the first two addressing when the eyes move and the latter two addressing where they move. CRM assumes that when to move is influenced by processing time for the currently fixated word: greater activation strength of the fixated word (higher activation value of the fixation word unit) leads to greater activation strength of the saccade unit, and a saccade is executed when a threshold is reached. Where to move is determined by a processing-efficiency-based strategy: the eye movement control module searches the character activation map from left to right, selecting the first character with activation below a certain threshold as the next saccade target. Regarding the fundamental issue of parallel versus sequential processing, CRM advocates parallel processing of characters and words, though word n+1 generally has lower activation strength and thus slower processing. The model explains how visual acuity, word length, word frequency, and predictability affect Chinese word segmentation and lexical recognition.

3.1 Model Structure

Overall, all five models contain lexical recognition and eye movement control components, differing only in structural details due to varying assumptions. For lexical recognition, E-Z Reader and SWIFT both include pre-processing (called familiarity check in E-Z Reader) and lexical access completion stages. Models built on interactive activation (Glenmore, OB1 Reader, CRM) are more complex, implementing lexical recognition through interactions between different processing layers—generally including visual, letter/character, and word layers. For eye movement control, all models include structures for saccade target selection and saccade generation. For target selection, parallel processing models (SWIFT, Glenmore, OB1 Reader, CRM) implement this through a saliency map structure, while the sequential processing model E-Z Reader assumes the default target is the next word after the currently fixated word, thus requiring no specific target selection mechanism. For saccade generation, Glenmore and CRM differ from other models by controlling saccades through independent fixation center or saccade unit modules, with whether these modules reach activation thresholds determining saccade execution. Finally, beyond lexical recognition and eye movement control, E-Z Reader uniquely includes a post-lexical integration stage to explain how semantic integration after word recognition affects eye movements. This structure simulates semantic integration during reading—a relatively advanced and complex cognitive process compared to independent lexical access that other models have not addressed.

The core logical difference among the five models concerns whether visual attention distribution within the perceptual span is serial or parallel. Regarding whether multiple words can be processed simultaneously, E-Z Reader represents sequential processing, while SWIFT, Glenmore, OB1 Reader, and CRM represent parallel processing. Additionally, CRM advocates parallel processing at the character level—all characters within the perceptual span are activated simultaneously, and all possible words they can form compete until one exceeds the threshold, completing lexical recognition. For word processing, CRM posits that non-overlapping words do not compete and can be processed in parallel, though word n+1 generally has lower activation strength and slower processing. Furthermore, although E-Z Reader advocates sequential word processing, it holds that saccade programming and attention shifting are dissociable: after completing lexical access of word n, attention can shift to word n+1 to begin preprocessing (familiarity check). This allows the sequentially-based E-Z Reader to explain preview effects.

3.3 Letter/Character Recognition

E-Z Reader and SWIFT both begin modeling from word recognition without explicit assumptions about letter recognition. Glenmore, OB1 Reader, and CRM all include letter/character recognition and assume visual acuity influences recognition: highest at the fovea and decreasing with eccentricity. These three models also adopt interactive activation logic, assuming lexical layer activation influences letter/character activation: high-activation words facilitate activation of corresponding letters/characters while inhibiting other letters/characters at the same position.

Based on empirical findings (Inhoff et al., 2003), Glenmore assumes word length information affects average letter activation levels: average letter activation decreases as word length increases. OB1 Reader incorporates effects of attentional weight and crowding on letter recognition: letters receiving higher attentional weight are easier to recognize, and isolated letters and letters at word edges are easier to recognize. In CRM, character recognition is achieved through similarity matching between input character images and prior character templates. Characters at the same position compete, and the winner is recognized.

3.4 Lexical Recognition

Regardless of sequential or parallel processing assumptions, all five models assume lexical recognition occurs when a word's activation reaches a certain threshold, and all agree that factors such as word frequency influence the recognition process. However, models differ in how these factors affect lexical recognition.

E-Z Reader, SWIFT, and OB1 Reader assume lexical recognition is facilitated by lowering the threshold, while Glenmore and CRM assume facilitation increases the word's activation value. Models also show subtle differences in factors influencing lexical recognition. All models agree that visual acuity and word frequency affect word recognition. Except for Glenmore, all models agree that predictability influences word recognition. E-Z Reader and CRM assume predictability only affects recognition of the next unit after the current processing unit is recognized. SWIFT separates word frequency and predictability effects, assuming word activation difficulty can be measured by word frequency alone, with predictability used to adjust processing efficiency, potentially making predictability's effect on processing earlier than word frequency's. Additionally, E-Z Reader and OB1 Reader assume word length affects word recognition. Based on empirical results (Inhoff et al., 2003), Glenmore assumes long words gain better activation from constituent letters and compete better with short words, thus word length also influences recognition. Glenmore also allows task difficulty to modulate word recognition top-down via the "fixation center" module. As mentioned, since OB1 Reader and CRM include letter/character recognition, competition among multiple possible words formed by similar letters/characters also affects lexical recognition. Although Glenmore is a parallel processing model, it focuses on explaining saccade target selection mechanisms and does not model activation and competition among multiple candidate words during recognition. Finally, because Chinese lacks inter-word spaces marking boundaries, CRM's lexical recognition process simultaneously includes word segmentation. Models built for alphabetic scripts assume readers can obtain word boundary information from primary visual input (inter-word spaces) and thus do not require a separate word segmentation module.

3.5 Lexical Integration

Among the five models discussed, only E-Z Reader includes a semantic integration process for recognized words. This process is incorporated as the third stage of lexical processing (post-lexical integration) to reflect the semantic integration required for readers to incorporate recognized words into higher-level representations during online processing—for example, integrating words into syntactic structures, generating semantic representations matching context, and incorporating meaning into discourse models. E-Z Reader provides detailed discussion of the immediacy of word integration and phenomena such as regressions and refixations resulting from integration failure (see Sections 4.2 and 4.3).

3.6 Saccade Target Selection (Where to Move the Eyes)

Parallel and sequential processing models differ substantially in saccade target selection. Sequential models (E-Z Reader) assume the next saccade target is the first unrecognized word after the currently fixated word, thus involving no complex target selection process. Notably, E-Z Reader assumes word n+1 can be skipped during reading (see Section 4.4). Parallel models (SWIFT, Glenmore, OB1 Reader) can activate multiple words simultaneously within the perceptual span, thus requiring target selection. The basic logic involves competition among simultaneously activated words: Glenmore and OB1 Reader assume the word with highest activation becomes the next target, while SWIFT assumes the word with highest activation has the highest probability of becoming the target. CRM's target selection is character-based: the model uses a processing-efficiency-based strategy to determine eye movement location, with the eye movement control module searching the character activation map from left to right to identify the first character with activation below a certain threshold as the next saccade landing position.

3.7 Saccade Programming and Execution (When to Move the Eyes)

Models differ considerably in how they program and execute saccades, which can be explained from two perspectives. First, regarding saccade generation mechanisms, Reingold et al. (2012) distinguished between direct control and indirect control, with direct control further divided into triggering and interference mechanisms. Direct control means processing of the currently fixated word influences the initiation of the next saccade locally and immediately. Indirect control means features and attributes of the currently fixated word do not directly influence saccade initiation but involve delayed adjustment of fixation duration (non-real-time), typically determined by average processing difficulty of encountered reading materials. The triggering mechanism in direct control initiates saccade programming when processing reaches a certain degree, while the interference mechanism holds that saccade initiation is unrelated to completion of specific processing stages but results from varying degrees of inhibition of saccade latency by processing difficulty.

From the perspective of lexical factors influencing saccade generation, the five models can be classified as direct control, with the sequential model (E-Z Reader) using a triggering mechanism and parallel models (SWIFT, Glenmore, OB1 Reader, CRM) using interference mechanisms. Specifically, in sequential models, completing the familiarity check triggers saccade programming. In parallel models, saccade occurrence is controlled by word processing difficulty, with different difficulty levels producing different inhibitory effects. Regarding non-lexical factors, E-Z Reader can be classified as direct control, triggering new saccades to correct positions when fixating incorrectly. SWIFT, Glenmore, and CRM can be classified as both direct and indirect control. All three models assume visual encoding difficulties trigger saccade inhibition (direct control). Meanwhile, SWIFT and CRM introduce free parameters for saccade generation to simulate the phenomenon of automatic saccade triggering after a period when no information is perceived or lexical processing is zero (time out), representing indirect control.

Second, regarding model architecture, Glenmore and CRM uniquely control saccades through independent fixation center or saccade unit modules, with whether these modules reach activation thresholds determining saccade execution. However, although both models use saccade control units, their input information differs. CRM's saccade unit connects only to the lexical layer, making fixation word activation strength the decisive factor for eye movement. Glenmore's fixation center connects to both lexical and letter layers and can be influenced by top-down processing (e.g., task difficulty), making its determinants more numerous and complex. E-Z Reader, SWIFT, and OB1 Reader lack independent saccade modules; saccade programming and execution are triggered by a random distribution while saccade latency is influenced by word recognition. Additionally, E-Z Reader and SWIFT both assume saccade programming can be divided into labile and non-labile stages, with labile saccades being cancellable and non-labile saccades non-cancellable. The two models differ in specific handling of these stages. As a sequential model, E-Z Reader assumes the labile stage begins after completing the familiarity check of the currently fixated word n, after which attention shifts to word n+1. The saccade program only enters the non-labile stage and executes after word n completes lexical access. SWIFT does not specify a clear temporal relationship between word recognition and saccade programming.

4.1 Fixations

Fixations involve two measures: fixation location and fixation duration. Fixation location is closely linked to saccade target selection. Fixation duration is influenced by factors such as word frequency and predictability. Additionally, saccades correlate with fixation duration, as saccade execution means fixation point shifting and the previous fixation ending. Models differ slightly in handling this mechanism. Glenmore and CRM assume these factors affect fixation duration by influencing fixation word activation strength, which adjusts saccade unit activation timing, producing longer or shorter fixations. E-Z Reader, SWIFT, and OB1 Reader assume saccade programming originates from a random distribution that automatically executes after a certain time without other influences. Word frequency, predictability, and other factors promote or inhibit saccades by changing random distribution values, producing longer or shorter fixations. Additionally, because E-Z Reader includes a post-lexical integration stage, semantic integration also affects fixation duration. For example, when semantic integration fails rapidly, readers extend current saccade latency or produce refixations, increasing fixation duration on the current word.

4.2 Regressions

Regressions refer to eye movements in the opposite direction of normal reading to reprocess information. As a special form of saccade, different models offer different explanations. Parallel models (SWIFT, Glenmore, OB1 Reader) assume that if unidentified words remain before the currently fixated word and win in saccade target selection competition, a saccade toward that word occurs—a regression. Sequential models (E-Z Reader) assume that during reading, there is a certain probability of semantic integration failure for word n, causing comprehension difficulty. If a saccade to the next word has already been executed or its program has entered the non-labile stage, a new regression saccade is programmed back to where comprehension difficulty first occurred (generally word n). A limitation of the sequential model's explanation is its inadequate handling of long-distance regressions: although E-Z Reader acknowledges some probability of regressing to content before word n, the actual model only simulates regressions to word n-1, not earlier positions. CRM assumes regressions are influenced by higher-level cognitive language processing, but since the model does not include a higher-level cognitive processing module, it does not address regressions.

4.3 Refixations

Refixations are another special form of saccade, referring to second or multiple fixations on a processing unit before the fixation point moves rightward past it. Parallel models (SWIFT, Glenmore, OB1 Reader) assume that if current word recognition is difficult, refixation probability increases. SWIFT also assumes correlation between saccade amplitude and refixation probability: shorter saccades have higher refixation probability. In sequential models (E-Z Reader), refixations occur in two situations: first, the distance between word center and initial landing position affects refixation probability—greater distance means higher probability; second, post-lexical integration can cause refixations when integration fails rapidly (while saccade programming remains labile), helping semantic integration of the currently fixated word. As mentioned, CRM's saccade target selection is based on processing efficiency: the eye movement control module searches the character activation map from left to right to find the first character with activation below a threshold. Following this logic, if the currently fixated character's activation remains below threshold, a refixation on that character occurs.

4.4 Skipping

Skipping refers to not fixating on a processing unit during reading, which can be divided into skipping during first-pass reading and skipping during the entire reading process, with the former being the typical reference. Skipping is somewhat proprietary to sequential models because in parallel models' spatial saliency framework (SWIFT, Glenmore, OB1 Reader), the concept of word skipping is meaningless since there is no default saccade to word n+1. According to E-Z Reader, when word n+1 completes lexical access and integration in the parafovea (e.g., "the," Angele & Rayner, 2013), readers cancel the default saccade to word n+1 and skip directly to word n+2.

5.1 Word Frequency Effect

The word frequency effect refers to high-frequency words being skipped more easily and having shorter reading times than low-frequency words (Kuperman et al., 2023; Li, Li et al., 2022; Liu et al., 2020). As a robust experimental effect, all eye movement control models simulate it. E-Z Reader, SWIFT, and OB1 Reader assume high frequency lowers the activation threshold, while Glenmore and CRM assume high frequency increases activation value. Both approaches allow high-frequency words to reach activation threshold faster, shortening processing time.

5.2 Predictability Effect

The predictability effect refers to high-predictability words being skipped more easily and having shorter reading times than low-predictability words (Chang et al., 2020; Cui et al., 2022; Liu et al., 2020; Yao, Staub et al., 2022). Models explain predictability similarly to word frequency, except Glenmore does not simulate predictability effects. E-Z Reader further assumes that when word predictability exceeds readers' ability to guess word n from prior context, familiarity check processing time becomes zero, making whole-word processing time zero.

5.3 Word Length Effect

The word length effect refers to short words being skipped more easily and having shorter reading times than long words (Kuperman et al., 2023; Li, Li et al., 2022; Zang et al., 2018). Models explain this similarly to word frequency effects. However, because Chinese lacks inter-word spaces, word length and boundary information are unknown before lexical recognition, so word length effects in Chinese reading must occur after word segmentation. According to CRM, word segmentation and lexical recognition are unified processes, so post-segmentation word length effects may result from long words requiring activation of more character information during segmentation and recognition.

5.4 Preview Effect

The preview effect refers to readers processing information in the parafoveal region (Chang et al., 2020; Cui et al., 2022; for review see Zhang, Zang, & Bai, 2020). Because parallel models (SWIFT, Glenmore, OB1 Reader) assume visual attention is distributed across words in the perceptual span, the preview effect is a direct result of this parallel processing. Sequential models (E-Z Reader) assume attention shifts serially, allocating resources to only one word at a time, offering a different explanation. Specifically, E-Z Reader's preview effect results from attention shifting. The model assumes lexical recognition comprises familiarity check and lexical access completion stages. After completing lexical access, attention shifts to the next word. Since no actual saccade occurs, this attention shift allows readers to perform familiarity check processing on the next word through parafoveal vision. A related issue concerns the depth of preview processing. Parallel models (SWIFT, Glenmore, OB1 Reader) assume preview processing of word n+1 can reach semantic levels, while in E-Z Reader, preview processing only involves primary information, not semantics.

5.5 Parafoveal-on-Foveal and Foveal-on-Parafoveal Effects

These effects extend preview effects: readers process parafoveal information, and parafoveal and foveal processing influence each other (Zhang et al., 2019; Zhang, Zang, Xu et al., 2020). When examining parafoveal processing effects on foveal processing, it's the parafoveal-on-foveal effect; the reverse is the foveal-on-parafoveal effect. Regarding parafoveal-on-foveal effects, sequential (E-Z Reader) and parallel (SWIFT, Glenmore, OB1 Reader) models make opposite claims. In sequential models, parafoveal processing of word n+1 occurs after lexical access of word n, with clear temporal order, so sequential models argue no parafoveal-on-foveal effect exists. In parallel models, simultaneous processing of multiple words leads to attentional resource competition, producing parafoveal-on-foveal effects. This mechanism also explains foveal-on-parafoveal effects. For sequential models, this effect can exist: post-lexical integration of word n affects parafoveal processing of word n+1. When word n encounters post-lexical integration difficulty, readers reduce preprocessing of word n+1 and shift attention back to the currently fixated word n.

5.6 Spillover Effect

The spillover effect refers to processing effects on word n being delayed and manifesting when fixation lands on word n+1 (Pollatsek et al., 2008; Rayner & Duffy, 1986). All five models provide some explanation. In E-Z Reader, when fixating word n+1, readers may simultaneously perform post-lexical integration of word n, and because predictability factors in word n+1's familiarity check are closely related to word n's post-lexical integration, processing of word n affects word n+1. SWIFT assumes saccades are random but inhibited by foveal processing (word n). In some cases, because foveal word recognition is slower than saccade generation, word recognition's influence on the saccade system has a time delay—when word n's inhibitory effect occurs, the eyes have already moved to word n+1. This delayed foveal inhibition explains spillover effects. Glenmore explains this effect by assuming activation values for letters or words can be carried from the current fixation to the next, making spillover effects straightforward. OB1 Reader and CRM base their explanations on parallel processing logic: when fixating word n+1, its processing time is affected by parafoveal word n. If word n is easy to process, readers can allocate more cognitive resources to word n+1, shortening its processing time.

5.7 Preferred Viewing Location

Preferred viewing location refers to the position of maximum distribution of first-fixation locations on a word. In English research, this position is generally slightly left of word center (McConkie et al., 1988; Rayner, 1979), though recent research advocates flexibility in fixation location (Cutter et al., 2017, 2018). Based on the former view, E-Z Reader, SWIFT, and OB1 Reader assume saccades target word centers, while Glenmore assumes they target the most salient word. Systematic and random errors then cause the preferred viewing location to be slightly left of center. Because Chinese lacks inter-word spaces and word boundaries are unclear, determining word center positions is challenging, making it controversial whether Chinese reading shows a consistent preferred viewing location effect (for review see Li, Liu, & Ma, 2011; Li et al., 2017). Addressing Chinese specificity, some researchers propose a processing-efficiency hypothesis (Li et al., 2015; Ma et al., 2015; Wei et al., 2013): readers first attempt to process as much information as possible at a given fixation, only moving when processing efficiency decreases sufficiently. Others propose a parafoveal word segmentation hypothesis (Zhou et al., 2018; Bai et al., 2012): if readers complete word segmentation in the parafovea, they fixate word centers. Recent research proposes a dynamic adjustment hypothesis (Li et al., 2015; Liu et al., 2019; Liu et al., 2019a, 2019b; Xia et al., 2023; Wang et al., 2018), assuming fixation location relates to processing load in foveal and parafoveal regions and adjusts dynamically. Specifically, more information obtained from the parafovea leads to longer saccades. Foveal processing load affects parafoveal processing—smaller foveal load enables more parafoveal processing and longer saccades. However, this view remains controversial. Thus, preferred viewing location in Chinese reading remains disputed and requires deeper investigation. CRM's view on this issue is that Chinese readers do not select saccade targets at fixed word positions. Chinese reading saccade target selection is character-based, employing a character processing-efficiency strategy.

5.8 Word Segmentation

Word segmentation refers to readers using cues to segment target words from continuous strings for further processing. Alphabetic-based models (E-Z Reader, SWIFT, Glenmore, OB1 Reader) assume readers obtain word boundary information from primary visual input (inter-word spaces) without requiring special segmentation. Because Chinese lacks inter-word spaces, an additional segmentation process is needed (Huang et al., 2021; Li et al., 2009; Liu et al., 2019). CRM assumes lexical processing and word segmentation are unified: all characters within the perceptual span are activated simultaneously, all possible words composed of these characters are activated and compete, and the unique winner completes lexical recognition while being segmented from the string.

5.9 Individual Differences and Task Difficulty

Different reader populations or task difficulties may lead to different eye movement behaviors (Mak & Willems, 2019; Staub, 2021). Among the five models, SWIFT, Glenmore, and OB1 Reader discuss reader differences or task difficulty factors beyond linguistic factors, though with different emphases. OB1 Reader focuses on how individual differences affect attention distribution, while SWIFT and Glenmore focus on effects on saccades. Specifically, OB1 Reader assumes attentional focus has variable width, with reader proficiency modulating this distribution: higher proficiency increases word recognition success, expanding attentional width, and vice versa. SWIFT assumes saccade programming intervals are random but have a predefined average time related to individual reading rate: faster readers have shorter average saccade intervals. Glenmore assumes task difficulty affects saccade planning and execution by influencing the fixation center module's activation threshold: higher difficulty raises the threshold, making saccades more difficult.

6 Comparative Summary

A visual comparison of the models is summarized in Table 1. Overall, the models share many commonalities while also showing numerous differences. Each model explains eye movement behaviors and experimental effects based on its own logic: for example, all introduce factors such as visual acuity and word frequency in letter/character recognition and lexical recognition; all explain common eye movement behaviors such as fixations, regressions, and saccades; and all simulate common experimental effects such as word frequency and preview effects.

The core logical difference among the five models concerns whether visual attention distribution within the perceptual span is serial or parallel. This attention distribution difference leads to divergent claims about letter/character and word recognition: sequential models cannot process multiple words simultaneously, while parallel models can. This difference also leads to divergent explanations of common eye movement behaviors: sequential models attribute regressions to post-lexical integration, while parallel models attribute them to lexical recognition. Attention distribution differences similarly cause divergent explanations of effects such as preview and parafoveal-on-foveal effects: sequential models assume no semantic preview effect or parafoveal-on-foveal effect, while parallel models assume the opposite. Finally, each model can explain phenomena/effects that others cannot: CRM reasonably explains word segmentation and preferred viewing location in Chinese reading; E-Z Reader explores post-lexical integration after word recognition; and SWIFT, Glenmore, and OB1 discuss non-linguistic factors.

Table 1 Comparison of Five Models

Feature	E-Z Reader	SWIFT	Glenmore	OB1 Reader	CRM
Basic Logic	Lexical processing + saccade control	Lexical processing + saccade control	Lexical processing + saccade control	Lexical processing + saccade control	Lexical processing + saccade control

Feature	E-Z Reader	SWIFT	Glenmore	OB1 Reader	CRM
Structure	Familiarity check + lexical access completion + post-lexical integration	Preprocessing + lexical access	Visual input + letter + word layers	Visual input + letter + word layers + spatial-topic information	Visual input + character + word layers
Saccade Control	Saccade programming + execution	Unstable + stable stages	Saliency map + fixation center + saccade generator	Saccade target selection + execution	Saccade unit + fixation word unit + character activation map + target selection + execution
Letter/Character Recognition	Not explicitly addressed	Not explicitly addressed	Visual acuity, lexical feedback, attentional weight, crowding	Visual acuity, lexical feedback, attentional weight, crowding	Visual acuity, lexical feedback, template matching
Lexical Recognition	Threshold reduction	Threshold reduction	Activation increase	Threshold reduction	Activation increase
Predictability Effect	Threshold reduction	Threshold reduction	Not simulated	Threshold reduction	Activation increase

Feature	E-Z Reader	SWIFT	Glenmore	OB1 Reader	CRM
Word Length Effect	Threshold reduction	Threshold reduction	Long words gain more activation from letters	Threshold reduction	Word segmentation requires processing more characters
Preview Effect	Attention shifting result	Parallel processing result	Parallel processing result	Parallel processing result	Parallel processing result
Parafoveal-on-Foveal Effect	Does not exist	Attentional competition	Attentional competition	Attentional competition	Attentional competition
Foveal-on-Parafoveal Effect	Post-lexical integration causes attention shifting	Attentional competition	Attentional competition	Attentional competition	Attentional competition
Spillover Effect	Post-lexical integration difficulty	Delayed foveal inhibition	Activation carried to next fixation	Resource allocation	Resource allocation
Preferred Viewing Location	Systematic and random errors cause left-of-center bias	Systematic and random errors cause left-of-center bias	Systematic and random errors cause left-of-center bias	Systematic and random errors cause left-of-center bias	No preferred location; character-based processing efficiency

Feature	E-Z Reader	SWIFT	Glenmore	OB1 Reader	CRM
Word Segmentation	Obtained from primary visual input (spaces)	Obtained from primary visual input (spaces)	Obtained from primary visual input (spaces)	Obtained from primary visual input (spaces)	Unified with lexical processing
Non-linguistic Factors	Not addressed	Affects average saccade interval	Affects fixation center threshold	Affects attention distribution width	Not addressed

Future Directions

Based on the theoretical debate between sequential and parallel processing, this paper systematically compared five models (E-Z Reader, SWIFT, Glenmore, OB1 Reader, and CRM) across model structure, basic logic, explanations of common eye movement behaviors, and interpretations of experimental effects. After systematically understanding these influential eye movement control models, we propose the following future directions:

1. **Further examine post-lexical integration issues.** Currently, only E-Z Reader addresses semantic integration after word recognition, and its exploration merely introduces a high-level semantic processing stage to explain issues previous versions could not. How post-lexical integration and overall sentence framework construction proceed has not been deeply investigated. With advances in syntactic and pragmatic research, semantic integration will become a direction for model development.

2. **Address word order issues.** Sequential models provide relatively straightforward answers to word order, while parallel models require detailed explanations. OB1 Reader partially resolved the conflict between word activation order and sentence word order by introducing spatial-topic information representation. However, this word-length-based solution is clearly unsuitable for Chinese reading. Explaining Chinese word order phenomena within a parallel processing framework requires attention.

3. **Incorporate more extra-linguistic factors to improve explanatory power.** SWIFT, Glenmore, and OB1 Reader discuss reader factors or task difficulty beyond linguistic factors. Future models could incorporate age, gender, intelligence, attention, and language proficiency to enhance practical explanatory and applicative capabilities.

4. **Explain recent empirical findings on preferred viewing location.** Alphabetic models assume saccades target left-of-center word positions

and cannot explain flexibility findings. Chinese reading landing positions are more complex. CRM provides a character processing-efficiency solution but cannot explain recent findings about foveal and parafoveal processing load (Xia et al., 2023). Future models should address this to comprehensively understand saccade targeting mechanisms.

5. **Establish systematic evaluation standards for model comparison.** The core purpose of computational models is simulating real reading mechanisms and deepening understanding of this advanced cognitive behavior. Beyond developing existing or building new models, systematic evaluation criteria should be established. Current models are built on specific empirical data with key parameters varying across simulated effects, making quantitative comparison difficult. Future research could build unified large-scale experimental effect databases to analyze falsifiability of key differences and compare explanatory power across models using standardized data.
6. **Explore cross-linguistic explanatory power of each model.** Existing models are built on specific languages (E-Z Reader on English, SWIFT on German, CRM on Chinese). Research on models' explanatory power for other languages is limited and has not fully considered how language specificity might affect model application. Future research could explore whether models built for specific languages apply to other languages after analyzing cross-linguistic specificity and commonality. For example, whether CRM's character processing-efficiency-based saccade target selection mechanism could extend to alphabetic reading.

References

[The references section appears to be complete and properly formatted in the original, so it should be preserved as-is in the translation. However, the user has not explicitly asked for the references to be translated, and they are already in a mixed format. Since the instruction says to preserve citations exactly, I will keep the reference list in its original form but ensure the formatting is clean.]

[The full reference list from the original Chinese text is preserved here with proper formatting…]

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.