Eye movements in reading: Some theoretical context

Ralph Radach1 and Alan Kennedy2
1General and Biological Psychology, University of Wuppertal, Wuppertal, Germany
2School of Psychology, University of Dundee, Dundee, UK

The study of eye movements has proven to be one of the most successful approaches in research on reading. In this overview, it is argued that a major reason for this success is that eye movement measurement is not just a methodology—the control of eye movements is actually part and parcel of the dynamics of information processing within the task of reading itself. Some major developments over the last decade are discussed with a focus on the issue of spatially distributed word processing and its relation to the development of reading models. The survey ends with a description of two newly emerging trends in the field: the study of continuous reading in non-Roman writing systems and the broadening of the scope of research to encompass individual differences and developmental issues.

Keywords: Eye movements; Models of reading; Parafoveal processing; Serial processing; Parallel processing.

In 2004, Radach and Kennedy wrote a review article describing the state of the art in the area of eye-movement-based experimental reading research (Radach & Kennedy, 2004). This Special Issue provides a welcome opportunity to update that review and discuss new developments in a rapidly moving field, at the same time pointing to unresolved issues for further research. The review is selective, in terms of both the issues addressed and the depth of discussion for each topic. The limited number of references cited should therefore be seen simply as examples in terms of both empirical findings and arguments supporting a certain theoretical position. Interested readers are referred to several detailed state-of-the-art reports, covering in more detail many of the issues that have been central to the field in recent years (Hyönä, Radach, & Deubel, 2003; Kennedy, Radach, Heller, & Pynte, 2000; Liversedge, Gilchrist, & Everling, 2011; Radach, Kennedy, & Rayner, 2004; Underwood, 2005; van Gompel, Fischer, Murray, & Hill, 2007). Rayner (1998, 2009) provides a condensed overview in his seminal review articles, and an excellent introduction can be found in the comprehensive text book authored by Rayner, Pollatsek, Ashby, and Clifton (2012).

In the earlier review, it was argued that research on eye movements in reading had been conducted...
from three different, albeit partially overlapping, perspectives: First, there had been a focus on visual processing and sensorimotor control. This sees reading as one domain of “active vision” (Findlay & Gilchrist, 2003) in which basic issues on perception, attention, and motor control can be examined within an environment that is far more ecologically valid than many laboratory tasks (see Vitu, 2011, for a recent overview). Related to this approach was the key question of how selective attention (Helmholtz, 1896/1989) comes into play in reading. It is possible, for example, that the term may simply serve as a handy umbrella for functionally very different mechanisms of visual processing such as saccade preparation, letter and/or word recognition, navigation on a page of text, and so on (Radach, Reilly, & Inhoff, 2007).

The second perspective is more related to a cognitive science tradition, where reading is seen as a domain of human information processing, similar to understanding a visual scene. At this level of inquiry, the principal question is how different levels and stages of processing unfold over time and interact with each other, from the acquisition of letter information all the way to the formation of a text representation. So far, the main focus of this approach has been on word-level processing, providing the empirical base for most of the existing computational models of oculomotor control in reading. A substantial part of the present paper is devoted to the discussion of research issues within this tradition of research (see, e.g., Pollatsek & Rayner, 1989, for a discussion of reading as one key domain within cognitive science).

Finally, it was clear at the time of our 2004 review (judged by the number of publications) that the most prominent and productive approach was the use of eye movement measurement to develop and test psycholinguistic hypotheses about the processing of written language. Most of this work focused on the level of sentence processing, as is evident in the seminal compilation by Clifton, Staub, and Rayner (2007) of results from 100 studies, documenting how eye movement analyses can be used to examine syntactic, semantic, pragmatic, and world-knowledge factors manifested in reading.

The present article first provides an updated introduction to some basic principles of oculomotor control, combined with a brief sketch of the most important eye movement measures. It then turns to consider some key issues related to the main topic of the present Special Issue: the spatially distributed processing of words in continuous reading and the related question of serial versus parallel word processing. Given the topic of this Special Issue, our attempt to set the context of the debate has sought to give equal weight to both approaches. The reader should be warned, however, that attempting even-handedness is itself perilous in the context of a strongly polarized debate. In particular, our efforts may be construed as unduly reflecting our own (somewhat parallel) theoretical leanings. There is no ready response to this charge, other than to admit it. Readers in search of arguments strongly in favour of attaching priority to serial processing can consult Reichle, Liversedge, Pollatsek, and Rayner (2009) or, more recently, Schotter, Angele, and Rayner (2012).

The discussion continues with consideration of two major new developments in the field that are beginning to enrich all three of the research perspectives sketched above: (a) the use of eye movement methodology to study reading in non-Roman writing systems. It will be argued that these innovative approaches to the dynamics of reading in Chinese, Japanese, Korean, and Thai deliver more than just an extension of the ongoing research agenda. By addressing issues like visual character complexity, parafoveal linguistic processing, word segmentation, and even semantic analysis in this context, we gain valuable new insight into the important question as to what extent information processing in reading is universal as opposed to language- and script-specific. (b) The second is the use of eye movement analyses to study the development of, and individual variation in, reading skill. After several decades of basic research using mostly college-age populations of skilled readers, it seems opportune to broaden the scope of research to encompass the wide and largely unknown landscape of inter- and intraindividual variations of reading and to bridge the gap to educational and clinical application.
A review of oculomotor control and eye movement measures in reading

Rather than repeat what was said in 2004, the reader is referred to that paper for a detailed discussion of basic methods in eye movement recording and data analysis. In the intervening years, the field has established a number of informal standards, which are now almost a necessary precondition for publication. One example is the set of eye movement measures commonly reported (see Radach & Kennedy, 2004, Tables 1 and 2, for details). Most important from a cognitive science perspective on reading are word-based viewing time measures, including initial fixation duration, gaze duration, and total viewing time, which are assumed to reflect different stages in the time line of word processing (see Reingold, Reichle, Glaholt, & Sheridan, 2012, for a recent discussion of processing dynamics in relation to fixation times). The duration of the first fixation on a word is assumed to be related to early orthographic (and other prelexical) processing, while the summed duration of all fixations before leaving the word (gaze duration) more plausibly reflects later stages of word processing, including lexical access. Finally, the mental effort of postlexical integration of meaning on the sentence level is related to the time spent rereading previously attended words, as expressed in a variety of total viewing time measures (see Inhoff & Radach, 1998; Inhoff & Weger, 2003; Rayner, 1998, for detailed discussions). Rayner and Liversedge (2011) have recently provided a very informative discussion of cognitive factors that influence word viewing time measures such as fixation and gaze durations on the lexical, sentence (syntactic and semantic) and discourse level.

A number of measures have been specifically developed as plausible indices of later, high-level, cognitive effects in research on the sentence or even passage level. Such omnibus measures are inevitably complex, or even ambiguous, due to the fact that, when processing difficulties arise, there are several possible responses, including an increase in viewing times on the present word or region and/or the triggering of regressive saccades, leading to the rereading of prior segments of text (Frazier & Rayner, 1982; Liversedge, Paterson, & Pickering, 1998; Murray, 2000). Regression path duration (often called “go-past time”), computed as the summed duration of all fixations from first fixating a critical word or region during first-pass reading until leaving it to the right (i.e., including regressive fixations up to that point), has become a standard high-level viewing time measure (see, e.g., Warren, White, & Reichle, 2009, for a recent discussion of clause-level word meaning integration based on this measure).

Turning to the spatial parameters of saccades, the most important measures are obviously their amplitude and landing position. A large body of research has shown that the fixation positions for both initial fixations and immediate refixations within the same word are largely determined by low-level visual constraints such as the length of the target word, the launch distance of the incoming saccade (McConkie, Kerr, Reddix, & Zola, 1988; O’Regan, 1990; Rayner, 1979), and, interestingly, the position of the word on the line of text (Kuperman, Dambacher, Nuthmann, & Kliegl, 2010; Radach, 1996; Vitu, Kapoula, Lancelin, & Lavigne, 2004). Another factor that codetermines saccade landing position is the pattern of preceding fixations: Landing position shifts to the left when the prior word is not fixated and to the right when the preceding fixation was a refixation on the prior word (Krügel & Engbert, 2010; Radach & Kempe, 1993).

1 In recent years, there has been steady progress in the development of methods and software for eye movement data analysis. For example, Lans, van der Wedel, and Pieters (2012) suggested an interesting algorithm to identify fixations from eye movement records of both eyes using individual eye velocity thresholds. Tang, Reilly, and Vorstius (2012) developed flexible data analysis and data visualization software for reading and made their work available to the research community. Baptista, Bohn, Kliegl, Engbert, and Kurths (2008) developed an algorithm capable of reconstructing eye position during blinks, addressing a major source of data loss, especially with problematic populations such as children or older adults. The question of whether and how blinks are related to or coordinated with the acquisition and processing of visual information in reading is as yet unresolved.
In conclusion, there is general agreement on what the important oculomotor measures are. There is also a general consensus on the set of general principles governing the control of eye movements. In line with a large body of evidence from basic oculomotor research (see Findlay & Walker, 1999, for an overview), the decision of where to move has long been considered to be independent of the decision of when to move the eyes (Morrison, 1984; Rayner & McConkie, 1976). It would be impossible in a brief review to arbitrate on the success of different models in capturing details of spatial control. Given that the topic remains somewhat controversial, our conclusion is more modest—namely, that in the computational models discussed below, different processing modules exist for the timing and triggering of a saccade, as opposed to the selection of its target and the specification of the intended saccade’s amplitude. We should emphasize, however, that independence of when and where decisions does not mean that there are no relations between saccade amplitude and fixation duration. As demonstrated by Pollatsek, Rayner, and Balota (1986) and Radach and Heller (2000), fixation durations are generally increased after longer incoming (progressive) saccades, presumably because of less parafoveal preprocessing when the previous fixation position was more remote (see Vitu, McConkie, & Zola, 1998, for a discussion of different results in the context of regressions).

With regard to the spatial aspect of eye movement control, there is an emerging consensus that the selection of a target for the next saccade is generally word based (although see McConkie & Yang, 2003; Vitu, 2008, 2011; Yang & McConkie, 2004, for a contrary view). That is, in most cases a saccade will be launched with the goal of reaching a particular word. Note that since most saccades do not go very far, in the vast majority of cases the decision as to which word comprises the target for the next saccade is limited to a few likely targets: the current word (a refixation saccade); the next few words in the right parafovea (a progressive interword saccade); or the preceding one or two words (a short-range regression). The fact that these spatial decisions are codetermined by low-level visuomotor constraints (e.g., launch distance and word length) and ongoing linguistic processing can be nicely demonstrated by the frequency of fixating the next word to the right. The phenomenon of not fixating the next word is often referred to as “word skipping”, although this carries the suggestion that, as in spoken language where each word is necessarily pronounced in serial order, in reading all words “should” be fixated in the right order (see Kennedy & Pynte, 2009, for a discussion of reading as a surrogate for spoken language). As demonstrated by Brysbaert, Drieghe, and Vitu (2005) in an elegant meta-analysis of published studies, most of the variance in fixation probability can be accounted for by word length. However, there is also evidence for a smaller, but significant, cognitive component. For example, longer words may be “skipped” more often when they are highly predictable from the prior context (Rayner, Slattery, Drieghe, & Liversedge, 2011).

A second, equally interesting, demonstration of the interplay between visual and cognitive factors concerns the immediate refixation during first-pass reading within the same gaze duration. In a classic paper, McConkie, Kerr, Reddix, Zola, and Jacobs (1989) first described the U-shaped functional relationship between refixation probability and initial landing position (see O’Regan, 1990, for precursors in the single word recognition literature). This curve has its minimum close to the word centre where word viewing is generally optimal, but rises the closer initial fixation position is to the word beginning or ending. Interestingly, for low-frequency words the refixation frequency curve is elevated but its steepness parameter remains constant, suggesting that visual and cognitive influences on refixations may be independent (e.g., Radach & McConkie, 1998; Vitu 1991). Indeed, this hypothesis has recently been confirmed by Hutzler, Braun, and Jacobs (2008) with the help of concurrent electroencephalography (EEG) and eye movement recordings in the variable viewing position paradigm (O’Regan & Jacobs, 1992). Hutzler and colleagues concluded from the time course of measured event-related potential (ERP) lexicality effects that an early
cohort of refixations corrects for suboptimal initial fixation positions without being influenced by lexical characteristics of the fixated word. If this observation can be replicated in continuous reading, it would provide direct evidence for a class of saccades in reading not directly driven by cognitive processing.

Depending on which word is selected as the next target (including the current word or a word to the left), a saccade is generated to bring the eyes to this target (see also our discussion above on amplitude as a saccade parameter). However, there are systematic error tendencies leading to deviations between intended and attained landing positions such that near targets are (occasionally) overshot, and far targets are often undershot (McConkie et al., 1988). Indeed, a fair proportion of saccades, ranging from 10 to as much as 30%, depending on word length, appear not to land on the intended word (Engbert & Nuthmann, 2008; Engbert, Nuthmann, & Kliegl, 2007). These systematic error tendencies do not appear to exist for regressive saccades, which are remarkably accurate (Inhoff, Weger, & Radach, 2005; Vitu, 2005), perhaps because they use spatial memory for target positions in the process of saccade generation (Kennedy, 1992; Kennedy, Brooks, Flynn, & Prophet, 2003; Kennedy & Murray, 1987).

There is also a small but significant influence of orthographic processing on saccade amplitude. Saccades land further to the right when the orthographic regularity of the initial letters of words in the parafovea is higher (Hyönä, 1995; White & Liversedge, 2004, 2006). It is important to note that this effect is graded so that there is a rightward shift in the landing position of the incoming saccade both from low to medium and from medium to high orthographic regularity of the target word beginning, suggesting that the effect is cognitive and not due to visual attraction or “pop out” by irregular letter clusters (Radach, Heller, & Inhoff, 2004). Higher order processing of parafoveal words on the lexical or semantic level is generally assumed to have no influence on saccade landing positions, presumably because processing results cannot be utilized before the saccade amplitude is programmed (minimally about 80 to 100 ms before execution, see Deubel, O’Regan, & Radach, 2000, for a discussion). Testing the limits of this assumption, Lavigne, Vitu, and d’Ydewalle (2000) have suggested that landing sites may shift to the right under the most favourable conditions, when words can be predicted from strong semantic context and when the prior fixation is at a close location so that the target word visibility is optimal.

So far the discussion has centred on facts about eye movement control in reading that are more or less undisputed. However, differences in opinion quickly emerge when the time course of word processing and its relationship with saccade generation are considered in detail. Positions in this debate can be ordered along two axes, describing the amount of cognitive control in saccade generation and the degree of parallelism in word processing. Indeed all published models of eye movement control can be located within the space defined by these dimensions (see Radach et al., 2007, for a more detailed discussion). According to one view on saccade generation, fixation durations are almost exclusively determined by the cognitive workload of linguistic processing (e.g., Reichle et al., 2009; Reichle, Rayner, & Pollatsek, 2003), while other researchers believe that saccades are generated somewhat autonomously, and hence the duration of fixation is the product of combined low-level and cognitive processing (e.g., Engbert, Longtin, & Kliegl, 2002; Reilly & Radach, 2006; Yang & McConkie, 2001).2

---

2 It should be noted that both dimensions—autonomous saccade generation versus cognitive control and sequential versus parallel word processing—are really necessary for a meaningful model classification (Jacobs, 2000; Radach et al., 2007). This can be illustrated (i.e., without further specific commitment to the plausibility of any particular model) using the example of Mr. Chips, an ideal observer model of reading developed by Legge, Klitz, and Tjan (1997; see also Legge, Hooven, Klitz, Mansfield, & Tjan, 2002) where both letter and word processing are parallel across word boundaries but saccade control is exclusively determined by cognitive processing. One the other hand, there are models such as SERIF (McDonald, Carpenter, & Schilckoo, 2005), and the competition-interaction model (Yang, 2006; Yang & McConkie, 2001), where there is only modest, and indirect, cognitive influence on saccade control.
The second, closely related focus of debate is on whether there can be temporal overlap between the processing of words (limited parallel processing—see Kennedy, Pynte, & Ducrot, 2002, for a discussion of the several meanings of the term “parallel processing”) within the perceptual span. This span, the area around the current fixation position within which useful information can be acquired, includes up to about 4 letters to the left and 15 letters to the right (with full letter discrimination up to a distance of about 10 letters) so that it can often encompass, for readers of English, the next two or three words to the right (see Rayner, 1998, 2009, for detailed reviews). Within this region, spatially distributed (foveal and parafoveal) letter and word processing is taking place during each fixation.

In all sequential attention shift (SAS) models, a processing window one word wide, usually referred to as “attention”, is assumed to move in a strictly sequential fashion from word to word (e.g., Pollatsek, Reichele, & Rayner, 2006; Reichele et al., 2003). In contrast, in several alternative processing gradient (PG) conceptions, linguistic processing encompasses several words within a gradient of letter processing or “field of activation” (Engbert, Nuthmann, Richter, & Kliegl, 2005; Reilly & Radach, 2006). This controversy forms the core theme for most contributions to the present Special Issue, with several articles reporting data supporting limited parallel processing, while others maintain that the processing of consecutive words occurs in a strictly sequential fashion. The following section is devoted to a sketch of this complex debate, intended to illustrate some of the major lines of argument.

Spatially distributed word processing and models of reading

When Reichele, Pollatsek, Fisher, and Rayner (1998) presented E-Z Reader as the first realistic algorithmic model of eye movement control in reading, this represented the opening of a new phase in the development of the field. The model was used to fit a corpus of reading data obtained by Schilling, Rayner, and Chumbley (1998) and even in its first instantiation successfully accounted for a wide range of effects, including variation in viewing time as a function of word frequency, the effects of contextual predictability, fixation probability, and spillover effects. The E-Z Reader model, and more generally the way in which empirical data generation and modelling act together in its development, has set a standard against which alternative theoretical frameworks have to compete.

Despite its success, the model immediately drew criticism focusing on some of its core assumptions, most prominently the strictly sequential (“leave-on-completion”) nature of word processing that is a necessary feature of all SAS models. This critique has subsequently led to the development of an alternative PG modelling approach, which allows for a limited amount of parallel word processing within the perceptual span. Before discussing the critical issue of spatially distributed word processing, it is necessary to point out that despite their many differences, both modelling approaches have substantial similarities. For example, E-Z Reader, SWIFT, and Glenmore all implemented the visuomotor mechanisms underlying the computation of saccade metrics suggested by McConkie et al. (1988). Furthermore, in all three there are a whole host of modules operating in parallel, as is the case for lexical processing and saccade preparation in E-Z Reader. Finally, a common assumption across the entire field is that letters within (short) words are processed in parallel (Reicher, 1969; see Adelman, Marquis, & Sabatos-DeVito, 2010, for a recent discussion and experimental demonstration). So the difference in opinion on sequential versus parallel processing centres entirely on the specific question of whether letters within the perceptual span can be processed in parallel across a word boundary.

The first line of critique against SAS models emerged even before E-Z Reader was first published and was primarily directed against the earlier noncomputational model proposed by Morrison (1984). Kennedy (1998, 2000) presented evidence for a processing trade-off between the word currently fixated and the word immediately to its right. In a typical experiment (Kennedy
et al., 2002), participants were asked to read a sequence of five words, looking for words denoting articles of clothing. The length and frequency of “foveal” words (Word 3 in the sequence) and the length, frequency, and initial-letter constraint of “parafoveal” words (Word 4) were manipulated. Results indicated that gaze duration on foveal words was systematically modulated by properties of as-yet unfixated parafoveal words. However, the exact form of such early parafoveal-on-foveal effects differed from experiment to experiment, and these early studies were immediately criticized for not using a more natural reading task (e.g., Rayner, White, Kambe, Miller, & Liversedge, 2003). Crucially, the exact form of these effects differed as a function of the length of the foveal word. When this was short, lexical properties of a parafoveal word modulated foveal viewing time; when it was long, parafoveal initial letter constraint had an influence. A failure to control for foveal word length (and hence parafoveal visibility) may have contributed to discrepant results in some early experimental studies. Lexical constraint (i.e., the number of words in the lexicon sharing a given word’s initial letters) apparently occupies an interesting “sublexical” middle ground in terms of processing level.

This early work has stimulated an enormous amount of further research in terms of both corpus analyses and controlled experimental sentence reading studies (see Drieghe, 2011; Schotter et al., 2012, for comprehensive reviews). Overall, there is now widespread agreement that parafoveal-on-foveal effects are real on the orthographic level, while the evidence on lexical effects is somewhat mixed (and strongly disputed by proponents of SAS models.) Apparently lexical parafoveal-on-foveal effects can be traced reliably in massive corpora of natural reading data (e.g., Kennedy & Pynte, 2005; Klegl, Nuthmann, & Engbert, 2006) while their experimental demonstration appears to require favourable viewing conditions, as is the case when the foveal word is relatively short and of high frequency (Kennedy & Pynte, 2005, term this the “visibility hypothesis”). One line of defence against lexical parafoveal-on-foveal effects has been the suggestion that they may be due to mislocated fixations (Rayner et al., 2003). This account has been strongly advocated by Drieghe, Rayner, and Pollatsek (2008) and equally strongly rejected by Kennedy (2008). In any case, the distinction between orthographic and lexical parafoveal-on-foveal effects is intriguing and is returned to below.

The discussion of sequential versus parallel word processing took a new turn when Rayner, Juhasz, and Brown (2007) first examined whether a parafoveal preview benefit could be obtained from the preview of word $N+2$. To enable the nonspecialist reader to understand the rationale of this work, it is necessary to introduce the methodology used to collect the data and relate this to the logic of linguistic processing within the E-Z Reader framework. The method in question is the boundary technique of eye movement contingent display changes (Rayner, 1975). An invisible boundary is defined to the right of a currently fixated word $N$ and to the left of a parafoveal target word $N+1$. The target word, while in the parafovea, is replaced with a mask to deny or modify its preprocessing, but is restored to its original as soon as the eyes cross the boundary. The change itself is usually unnoticed due to saccadic suppression during the movement out of word $N$. If subsequent viewing durations while fixating $N+1$ are increased relative to a suitable control, this is taken as evidence that the prior masking had hampered the acquisition of parafoveal information (see, e.g., Inhoff, Radach, Eiter, & Juhasz, 2003, for differential effects of linguistic and spatial mask manipulations). In work examining parafoveal processing of word $N+2$, the boundary is placed to the right of a word two positions left of this target so that the distance between fixation position and target word during parafoveal information acquisition is markedly increased.

In the processing framework of the E-Z Reader model, any acquisition of linguistic information from word $N+2$ (while fixating word $N$) would require the following processing operations: Full lexical processing of word $N$, a shift of attention to word $N+1$, full lexical processing of word $N+1$, a shift of attention to word $N+2$, and some initial lexical processing of word $N+2$. THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2013, 66 (3)
Since all these stages of processing must be completed serially, this is a very unlikely scenario within the timing constraints of lexical processing and saccade generation (Sereno, Rayner, & Posner, 1998; Deubel et al., 2000; Inhoff, Eiter, & Radach, 2005; Kliegl, Dambacher, Dimigen, Jacobs, & Sommer, 2012). A similar argument can be made against semantic preprocessing of parafoveal words because this requires more time than lexical processing and hence should only very rarely occur within any SAS framework (see below). In a processing gradient architecture, however, it is assumed that letter information is acquired concurrently from all positions with the perceptual span, so that parallel lexical processing is a possibility.

The results from a number of $N+2$ preview experiments are quite mixed. Rayner, Juhasz, et al. (2007) as well as Angele, Slattery, Yang, Kliegl, and Rayner (2008), reported no benefit of $N+2$ previews on subsequent $N+2$ viewing. Similarly, Angele and Rayner (2011) did not obtain lexical $N+2$ preview effects, even when $N+1$ was the word “the”, a short function word with the highest possible word frequency in the English language. However, there are also some reports of positive outcomes. For example, Kliegl, Risse, and Laubrock (2007), using three-letter $N+1$ words, obtained effects of $N+2$ previews on $N+1$ viewing durations. Radach, Inhoff, Glover, and Vorstius (2013) combined the use of short and high-frequency $N+1$ words with a contextual manipulation such that $N+2$ became predictable, resulting in reliable $N+2$ preview effects. It would appear that $N+2$ preview experiments really test the limits of spatially distributed linguistic processing. This difficulty demonstrating effects driven by word $N+2$ is to be expected, given the limited extent of the perceptual span within which letter information can be discriminated (Rayner, 1998). In any processing gradient architecture the rate of information acquisition must be a function of letter eccentricity (in line with an asymmetric perceptual span) so that distant parallel processing will be quite limited (see Reilly & Radach, 2006, for a detailed discussion).

A similar conclusion can be drawn with respect to recent attempts to demonstrate semantic preview effects. Hohenstein, Laubrock, and Kliegl (2010) examined the issue in German sentence reading, using a combination of the boundary and fast priming paradigms. The fast priming technique allows for the examination of the acquisition of information in foveal vision by manipulating letters during a fixation (Sereno & Rayner, 1992). In the study by Hohenstein et al., a random consonant string was initially presented at the location of a target word. When the eyes entered the target region, a prime was presented for a varying time interval and was then replaced by the target word. As one main result, they observed a semantic preview effect when a semantically related parafoveal word was available during the initial 125 ms of a fixation on the pretarget word. This finding suggests that, at least under specific favourable conditions, some parafoveal extraction of semantic information may be possible. The finding, however, stands against a number of failed attempts to obtain semantic preview effects in English (Schotter et al., 2012). On the other hand, there are several studies reporting robust semantic preview effects when reading Chinese and Korean (see below, Section 3).

Taken together, with respect to all areas of empirical work on sequential versus parallel word processing, the state of affairs appears to be that evidence for parallel word processing is not easy to obtain. The existing data reinforce proponents of PG models in their quest to make the case for limited parallel word processing, while proponents of SAS models maintain their lines of defence. This situation will only be resolved following the accumulation of more empirical evidence; consequently the controversy may continue for some time yet. It may therefore be interesting to step back and consider some theoretical arguments that provide a more general context for the ongoing debate. Recently, Reichle et al. (2009) argued why sequential attention allocation in reading is, in their view, more plausible than any alternative. Their position includes the following three lines of argument:

1. A general assumption that has been repeatedly made is “that attention is necessary to “bind” the features of words—like those of other visual objects—into unitary representations that can...
then be used by higher level cognitive systems” (Reichle et al., 2009, p. 117; see also Pollatsek et al., 2006). Reichle et al. (2009) refer to evidence from visual search studies in support of their opinion. One example is the location of an object defined by two features (e.g., a red X) in an array of objects constructed from the same features (e.g., red Ts and green Xs). In such a task, the number of objects affects task difficulty, with larger arrays slowing search and increasing error rates (e.g., Treisman, 1988; Treisman & Gelade, 1980). The suggestion is that attention must be focused on each object so that its shape and colour can be “bound” into a single representation; since attention is limited and can only be allocated to one location at a time, the task becomes more effortful and time consuming as the number of objects increases.

The problem with this form of argument is that the suggested analogies between reading and visual search can be phrased in different ways, depending on which aspects are considered critical. As an example, Treisman (1998, p. 1299) has also suggested that when “target features are known in advance and when the relevant features are highly discriminable, subjects can use a feature-based grouping strategy to bypass the binding process”. In skilled reading, target features are certainly known in advance and also highly discriminable, so that the above description of search without binding appears as convincing as the alternative. Moreover, Treisman and Gormican (1988) even went a step further, suggesting the possibility of binding features from multiple objects within the same focus of attention.

2. According to SAS models, one major advantage of the sequential progression of attention is that the temporal order of word recognition directly defines grammatical word order as needed for sentence comprehension. Because in processing gradient models words can be recognized “out of order”, parallel processing is assumed to require buffering so that supralexical knowledge can be used to restore canonical word order. This argument needs to be seen in the context of empirical studies of what happens when this canonical inspection order of words is violated during reading. As an example, Kennedy and Pynte (2008) examined atypical noun adjective sequences in the French part of the Dundee Corpus. They found that violations to canonical reading order produced increased word viewing durations only for fixations on the two critical words and the immediately following word, while further downstream there were no longer lasting effects. Taken together with the fact that canonical word order is relatively rarely honoured in normal reading (Hogaboam, 1983, suggests in only around 20% of cases; Kennedy & Pynte, 2008, using a different measure, suggest 15%). This apparent lack of severe processing disruption consequent on “nonserial” inspection is as problematic for SAS as for PG models—both fail to solve what Pollatsek et al. (2006) call the “run home” problem. If the reader is to understand these two words, it matters in what order they are processed. Unfortunately, the standard serial model provides no mechanism for translating between temporal and spatial adjacency (“comes after” does not imply “comes next to”). The problem appears equally intractable for parallel models like SWIFT. Not surprisingly, Rayner, Pollatsek, Liversedge, and Reichle (2009) strongly dispute the interpretation of the data offered by Kennedy and Pynte (see Kennedy & Pynte, 2009, for a reply).

Apart from the interpretation of empirical data in cases where word viewing mismatches physical word order, a more theoretical point can be added to this discussion. One could argue that visually presented words on a line of text offer an immediate and permanent external source of information (as part of “the world as an outside memory”; O’Regan & Noe, 2001). Within this frame of reference, the relative spatial location of a word defines word order unambiguously, even when (and possibly even because) there is temporal overlap in the processing of adjacent words. From this it follows that it would not be necessary to rely on the temporal order of lexical processing completion to define grammatical word order.

3. A further argument made by Reichle et al. (2009) is that the lexical processing of multiple words is “not consistent with any existing model of word identification” (p. 117). They cite the “triangle” model (Seidenberg & McClelland, 1989),
where a word is identified via activating orthographic input units and then propagating this activation further towards meaning output units. Critically, “if one were to simultaneously activate orthographic units for two words, this would produce noisy output corresponding to neither word—not the meanings and/or pronunciations corresponding to two separate words, as required by attention-gradient models” (Reichle et al., 2009, p. 117).

This argument can be disputed at both a general and a specific level. Generally, a design decision to focus on single word output does not preclude the possibility that multiple word candidates are active over time. For example, the dual-route cascaded (DRC) model of Coltheart, Rastle, Perry, Langdon, and Ziegler (2001) routinely activates multiple word candidates during the time course of processing a single printed input word. At a more specific level, it can be noted that models of word identification exist that explicitly deal with the concurrent processing of multiple words. The BLIRNET model (Mozer, 1991; Mozer & Behrmann, 1990) has a hierarchical feedforward architecture and explicitly models the role of attention in processing multiword visual inputs. In fact, any model of single word recognition inspired by McClelland and Rumelhart’s (1981) interactive activation (IA) architecture should be capable of handling the simultaneous activation of multiple words (see, e.g., McClelland & Rumelhart’s own tentative account for the processing of two adjacent words based on explicit letter position coding). The Glenmore model (Reilly & Radach, 2006) uses a variant of this interactive activation architecture, permitting the processing of multiple word candidates at both the letter and the word level. During the acquisition of information from an input vector representing the perceptual span, several word candidates become active and compete for dominance based on their frequency and the strength of their bottom-up letter-level activation. Once one of the candidate words crosses a threshold of activation, it is considered “recognized” and is removed from competition.

Eye movement analyses of reading in non-Roman writing systems

Over the last decade there has been a remarkable expansion of eye movement work on reading in non-western writing systems. By far the most attention has been given to Chinese, where a very active and productive research community is beginning to emerge (see Zang, Liversedge, Bai, & Yan, 2011, for a detailed review). In addition, there is now solid empirical work on several other non-Roman scripts including Arabic (Abubaker, McGowan, White, Jordan, & Paterson, 2011); Hebrew (Deutsch, Frost, Pelleg, Pollatsek, & Rayner, 2003); Japanese (Kajii, Nazir, & Osaka, 2001; Sainio, Hyöna, Bingushi, & Bertram, 2007); Korean (Kim, Radach, & Vorstius, 2012); and Thai (Reilly, Aranyanak, Yu, Yan, & Tang, 2011; Winskel, Radach, & Luksaneeyanawin, 2009). Considering the fact that there are approximately 80 different writing systems (Everson, 2002), including many with very large communities of readers, it is obvious that we are only just beginning to get a feel for the diversity of human reading and writing. Progress in this respect will depend on the ability to systematically cover examples from broad classes of writing systems based on dimensions such as alphabetic versus non-alphabetic, logographic versus syllabic, or spaced versus nonspaced (see below). Such a development would shed light on the fundamental question as to the degree to which the processing of written language is determined by universal principles as opposed to properties of specific languages and writing systems (Reilly & Radach, 2012).3

3 The current stream of studies on Chinese reading rests on pioneering work from two leading researchers in the field, George McConkie and Albrecht Inhoff, who, together with young Chinese researchers, started more than two decades ago to examine problems of oculomotor control and word processing while reading Chinese. Since 2004, Keith Rayner, Deli Shen, Guoli Yan, and other Chinese colleagues have organized the biannual China International Conference on Eye Movements (CICEM), which has developed into a very effective engine of productive international collaboration.
It is not surprising that most of the research issues studied in the current wave of studies on non-Roman scripts parallel prior lines of work on reading in European languages like English, German, or French. Although it may initially appear not very interesting to re-address similar issues in just another script, it turns out that this can offer very useful new perspectives for ongoing theoretical debates. The extra value is to a large extent due to the specific ways in which alternative writing systems code linguistic information. As an example, written Chinese is a largely logographic system, with characters composed of strokes comprising basic visual features such as lines, curves, and dots. Importantly, Chinese characters can usually be divided into subcharacters generally referred to as radicals, which in most cases carry specific phonological and semantic information (Hoosain, 1991; see Zang et al., 2011, for an informative introduction). An important convention in the writing of Chinese (as well as Japanese and Thai) is that there are no spaces between word-level units. Despite this lack of visual word segmentation, reading rates for equivalent text in English and Chinese appear to be nearly identical (Sun & Feng, 1999), suggesting that the basic rate of information extraction is equivalent (see also Inhoff & Liu, 1998). Similarly, Rayner, Li, Juhasz, and Yan (2005) found effects of word predictability on viewing time measures that were very similar to those previously reported for English reading by Rayner and Well (1996; see also Kliegl, Grabner, Rolfs, & Engbert, 2004, for comparable data in German). The perceptual span in Chinese reading extends 1 character to the left and 2–3 characters to the right of fixation during reading Chinese, which is visually more compressed than the perceptual span in English but comparable in terms of the number of words from which parafoveal information can be extracted (Inhoff & Liu, 1998; Sun & Feng, 1999).

An important issue that has been addressed in Chinese (and to some extent in Japanese) concerns the difficulty of Chinese character and word processing. As an example, Yang and McConkie (1999) showed that the complexity of characters within two-character words (expressed as the number of strokes) affected gaze duration, fixation probability, and the number of refixations on a word. Yan et al. (2012) examined in more detail how the visual composition and complexity of characters determines their recognition. At the level of lexical processing, Yan, Tian, Bai, and Rayner (2006) have shown that viewing time measures were sensitive to both word and character frequency and that these frequency effects are similar in size to those reported in many previous studies for reading in English. More specifically, when reading two-character Chinese words, individual character frequency modulated the overall word frequency effect, with the first character having a larger effect than the second; on the other hand, the effect of character frequency was less pronounced with high-frequency target words.

Presenting readers with unspaced text in English incurs substantial costs in terms of extra word-processing time (Rayner, Fischer, & Pollatsek, 1998). In contrast, readers of unspaced non-European writing systems are quite effective at extracting word information, so that the question arises of how the segmentation of words is accomplished and the degree to which it is part of spatially distributed (parafoveal) processing. A useful approach to address this issue can be derived from analyses of long multimorphemic words in European languages that allow for productive compounding, such as German (Inhoff, Radach, & Heller, 2000) and Finnish (Bertram, Pollatsek, & Hyönä, 2004). Statistical cues for word boundaries may be a major factor contributing to word segmentation in the parafovea. When fixating a word, the presence of a character that is often located in the word-final position may therefore indicate that there is indeed a word boundary to its right. In contrast, extra processing costs may be incurred when the word-final character is generally used as a word beginning.

Evidence for such a mechanism was obtained by Yen, Radach, Tzeng, and Tsai (2012), who used the boundary technique (see Section 3) to manipulate the “diagnostic” value of the second character of a two-character word seen in the parafovea. However, in the context of these results it should be noted that only a minority of Chinese characters...
unambiguously signal a word boundary. As noted by Yen, Radach, Tzeng, Hung, and Tsai (2009), about 18% of 5,915 unique Chinese characters appear exclusively in one within-word position (including single-character words), the beginning characters, or the ending characters of multicharacter words, while as much as about 49% of characters can occur in all within-word positions. It follows that parafoveal processing of cues based on the statistics of letter positions, even though effective, can only provide limited and/or preliminary information that may have to be verified during subsequent fixations.

This leads to a very interesting consequence for the targeting of saccades in nonspaced writing systems. In alphabetic scripts employing word spacing, the landing positions of incoming progressive saccades tend to cluster about halfway between the word beginning and the word centre, honouring the so-called preferred viewing position (Rayner, 1979). This phenomenon has been taken as strong evidence for word-based eye movement control, and the underlying component mechanisms are considered to be among the most important visuomotor constraints on eye movements in alphabetic script reading (McConkie et al., 1988).

Surprisingly, when Yang and McConkie (1999; see also Tsai & McConkie, 2003) examined landing positions in a corpus of Chinese text, they found no clear evidence for a preferred viewing position effect: The distributions of saccade landing positions across words were almost flat. More recently, Yan, Kliegl, Richter, Nuthmann, and Shu (2010) presented participants with sentences in which characters at word boundaries were not ambiguous, and in this case clear evidence for a preferred viewing position emerged, with a tendency for saccades to land near the word centre in single-fixation cases, whereas in cases with multiple fixations initial landing position was closer to the word beginning (see Kajii et al., 2001, for evidence of preferred landing at word beginnings in Japanese reading). Seen in conjunction with the evidence on the use of statistical letter information discussed earlier, this pattern can be seen as evidence for the proposition that skilled Chinese readers target the beginning or centre of a word as a consequence of success (or failure) in the attempted segmentation of the word carried out from prior fixation positions (but see Li, Liu, & Rayner, 2011, for an alternative account of similar data). Flexible saccade targeting strategies have also been found in the reading of Japanese, where in mixed hiragana and kanji text the visually more distinctive kanji characters appear to provide useful parafoveal segmentation cues (White, Hirotani, & Liversedge, 2012). A comparative analysis of how visuomotor constraints and orthographic information act together in Thai and Chinese to determine saccade landing positions has recently been reported by Reilly et al. (2011).

In Chinese script, characters are written in a delimited area irrespective of their complexity, so that a significant amount of visual information is encompassed within a small and regular space. Given this high density of information, the question arises as to whether such a writing system presents more favourable conditions for parafoveal linguistic processing beyond word N + 1. Indeed, Yang, Wang, Xu, and Rayner (2009) have reported N + 2 preview effects as well as parafoveal-on-foveal effects in an experiment using two-letter high-frequency target words. More recently, Yang, Rayner, Li, and Wang (2012) reported a new study in which they placed a relatively low-frequency N + 1 word in front of critical N + 2 target words. Under these conditions there was no longer a preview effect from word N + 2. This work leads to the conclusion that words within the perceptual span compete for processing resources and that a sufficient amount of such resources becomes available for more distant words only when the processing demands of more proximal words are relatively low. This reasoning is in harmony with work by Yan, Kliegl, Shu, Pan, and Zhou (2010), who also suggested that the processing load from words close to the fovea modulates the perceptual span in Chinese (see, e.g., Henderson & Ferreira, 1990; Kennison & Clifton, 1995, for evidence on similar foveal-on-parafoveal effects in English).

Just as in European alphabetic scripts, in non-Roman writing systems the issue of whether high-level information such as semantic or...
sentence-level syntactic information can be processed in the parafovea has recently become the focus of intense scientific debate. In contrast to the situation described in Section 3 for research conducted in English or German, there is considerably more evidence for a parafoveal semantic processing effect in Chinese reading. As an example, Yan, Richter, Shu, and Kliegl (2009) reported parafoveal preview effects (and parafoveal-on-foveal effects) on fixation and gaze duration for Chinese preview characters, which were semantically related to targets (see also Yan, Kliegl, Shu, et al., 2010, and Yan, Risse, Zhou, & Kliegl, 2012). Yang, Wang, Tong, and Rayner (2012) combined semantic relatedness between preview and target with a potentially important high-level factor, the contextual plausibility of the preview within the current sentence. They demonstrated that the plausibility of the preview produced a stronger preview effect than semantic relatedness per se, although the latter did have some impact on short single fixations (similar to results reported by Yan et al., 2012). In conclusion, there appears to be good evidence that semantic information is regularly processed from parafoveal words within the relatively compact perceptual span provided by Chinese script.

Kim et al. (2012) have added a new angle to this debate, presenting evidence of high-level parafoveal processing in Korean reading. The Korean writing system, hangul, is similar to Chinese in being spatially compact, but is alphabetic in nature. The letters are arranged in uniform blocks, which, in turn, usually correspond to syllables. Kim et al. manipulated case markers unique to hangul (and, in a similar form, Japanese) where specific character suffixes are used to indicate the case role of a noun (e.g., subject, object, or topic). In one critical condition, a boundary paradigm was used to display contextually incorrect case markers, effectively creating a semantic mismatch between preview and target. Importantly, all previews were legal, so that processing of their appropriateness was exclusively based on prior contextual information. The results indicated that previews of inappropriate case-marked characters resulted in elevated reading times of the target word, specifically for late viewing time measures (gaze duration and total viewing time). As case markers are very frequently used in hangul script, it can be concluded from these findings that Korean readers routinely acquire high-level linguistic information available in the parafovea. In addition to the spatial compactness of the hangul script, this may be based on the fact that Korean is a left-branching language, where information critical for the assignment of meaning is often delayed (e.g., when a direct object is placed before a verb), creating high demand for disambiguating information early in a sentence.4

Individual differences and reading development

One of the “issues for the future” discussed by Radach and Kennedy (2004) was the need for more research on individual variation in reading skill and on the effects of task demands. Fortunately, the last decade has seen a number of attempts to make progress in this direction. One relatively straightforward methodological approach that can yield very informative results is to divide a sample of readers into groups with low versus high reading skill (or simply fast and slow readers) and analyse the influence of this dichotomy on a set of oculomotor measures. Following the main theme of the discussion so far, the focus here is on a few examples of parafoveal word processing. Chace, Rayner, and Well (2005) examined the effect of reading skill (based on a standardized assessment) on the parafoveal acquisition of phonological and orthographic information. The results suggested that less skilled readers do not use phonological codes to integrate information across eye movements. Perhaps more fundamentally, their findings also indicated that the group of less skilled readers overall showed only weak preview benefit effects. In

---

4 As Korean shares this structural feature with many other left-branching languages like Turkish, Japanese, Tamil, and Basque, one may speculate whether there might be more writing systems in which the simultaneous foveal and parafoveal processing of semantic and/or syntactic information is the rule and not the exception in routine reading behaviour.
line with this conclusion, Rayner, Slattery, and Bélanger (2010) reported that slower readers had substantially smaller perceptual spans. This argues strongly for the inclusion of interindividual difference measures like reading skill (Ashby, Rayner, & Clifton, 2005), reading speed (Inhoff, Connine, Eiter, Radach, & Heller, 2004), or working memory capacity (Kennison & Clifton, 1995), either as design factors or as random variables in regression model analyses.

Taking this idea a step (or, indeed many steps) further, Kuperman and van Dyke (2011) collected eye movement data during sentence reading from a large sample (n = 71) of young adult readers with relatively low level of formal education. In addition, participants completed a battery of 18 assessments, intended to measure component skills of reading such as phonological awareness, simple and complex memory span, rapid automatized naming, word and nonword reading (decoding), reading comprehension, and so on. A fundamentally important finding was that the effect of individual differences far exceeded the role of word and text properties in accounting for variability in measured eye movements. Several interesting associations emerged between verbal skills and the time-line of linguistic processing as reflected in specific eye movement measures. For example, vocabulary size appeared to be important in the early stages of lexical processing, whereas comprehension skills were more engaged at the stage when word meaning is integrated with the meaning of the sentence. A very interesting applied aspect of this work is the use of eye movements to validate hypothesized component skills, very few of which were actually found to be reliable predictors of word-based eye movement parameters.5

Another set of recent studies has used eye movement data to examine reading in populations with mental illness or impairments in sensory and language function such as acquired dyslexia in the form of central (deep vs. surface) dyslexia (Schattka, Radach, & Huber, 2010) or alexia with letter-by-letter reading (Ablinger, Huber, Schattka, & Radach, 2012; Johnson & Rayner, 2007; Rayner & Johnson, 2005). Examining paraverbal processing in skilled deaf readers, Bélanger, Slattery, Mayberry, and Rayner (2012) found that these participants had a larger perceptual span than would be predicted on the basis of their reading ability, suggesting that their generally enhanced attention allocation is utilized in reading. The first application of the moving window technique (McConkie & Rayner, 1975) in patients with schizophrenia provided evidence for a reduced perceptual span, which was associated with deficits in phonological processing and reduced saccade amplitudes (Whitford et al., 2012). The authors of this intriguing work suggest that specific deficits in language, oculomotor control, and cognitive control contribute to the observed impairment of reading in schizophrenia.

In addition to considering differences between readers, there is widespread agreement that variation within one and the same reader, related to phenomena like intention, motivation, global strategies, and so on, are important for both the process of reading and its outcome in terms of comprehension (Heller, 1982; Rayner et al., 2012; Tinker, 1958). Yet, to date, there is very little work directly addressing this fundamental issue. One exception is a study by Radach, Huestegge, and Reilly (2008), who presented sentences either at random or as part of integrated paragraphs. They found that this format manipulation alone substantially influenced word-based eye movement parameters (which may explain some differences between sentence- and corpus-based research). More importantly for the present discussion, these authors also attempted to manipulate depth of processing by asking readers either to respond to simple

---

5 The importance of this line of research can hardly be overstated. Anyone who has been at a meeting like the annual conference of the Society for the Study of Reading (SSSR) knows that the community of reading researchers working with psychometric assessments of reading skills is far larger than the community engaged in cognitive-science-based reading research. Unifying these two “styles of science in the study of reading” (Stanovich, 2003) would greatly extend the impact of eye movement work within the education science community and pave the way for more applied work.
word verification items or to deal with comprehension questions reflecting semantic relations within whole sentences. In addition to the expected overall increase in word viewing time measures in the comprehension condition, there was also a substantial increase in the size of the word frequency effect in all measures of word viewing time. Interestingly, this difference turned out to be more pronounced in single-sentence than in corpus reading. Another interesting type of strategic adjustments in response to subtle changes in text properties has been reported by Carminati, Stabler, Roberts, and Fischer (2006), who provide the first analysis of how readers of poetry respond to changes in subgenre and rhyme scheme. Kaakinen and Hyönä (2010) examined changes in reading strategy as a result of a proofreading versus comprehension instruction.

A very important intraindividual variation concerns the question to what extent oral reading differs from reading silently (see Radach, Schmitten, Glover, & Huestegege, 2009, for a discussion from a developmental perspective). Ashby, Yang, Evans, and Rayner (2012) have recently provided the first detailed comparison of the perceptual span in both modes of reading. This issue will certainly get more attention in the future, given the importance of oral reading in reading development and instruction. However, a full understanding of reading aloud will require a detailed understanding of the interplay between its word processing and language production aspects, as it becomes evident in analyses of the eye–voice span (Buswell, 1920; Inhoff, Solomon, Radach, & Seymour, 2011).

Of course, by far the most fundamental intraindividual change in reading is its development over the lifespan. In recent years there has been a surge in work on reading in older adults, demonstrating, to give just a few examples, that age-related decline has negative consequences in syntactic processing (Kemper & Liu, 2007) and that older readers respond differently to distracting information (Rozek, Kemper, & McDowd, 2012). Rayner, Reichle, Stroud, Williams, and Pollatsek (2006) simulated age differences in response to word frequency, predictability, and font difficulty within the framework of the E–Z Reader model, concluding that lexical processing is slowed, leading to adjustments in terms of a “more risky reading strategy”.

Looking at parafoveal word processing, Rayner, Castelhano, and Yang (2009) reported that older readers appear to have slightly smaller and less asymmetric perceptual spans. In a follow-up study using the boundary technique, the same group of authors found that while preview benefit was similar in first-fixation and single-fixation durations, in later processing measures such as gaze duration and go-past time parafoveal processing appeared attenuated (Rayner, Castelhano, & Yang, 2010). Building on this work, Risse and Kliegl (2011) provided a detailed analysis of age effects with regard to parafoveal preview and parafoveal-on-foveal effects, reporting an intricate combination of subtle deficits and compensation mechanisms.

Considering the vast number of publications using eye tracking to study reading in adults, it is quite striking how little attention has been paid to development of the skill at an individual level. This is surprising given the fact that normal reading development is a prerequisite for adequately developed literacy skills, which in turn are a necessary condition for leading a successful life in a modern society. The relatively small number of studies on normal reading in children stands in sharp contrast to the massive body of published research on developmental reading disabilities, notwithstanding the fact that understanding how successful reading evolves is a necessary precondition for a causal understanding of developmental delays and disabilities (McConkie et al., 1991; Radach et al., 2009).

In a recent review on children’s eye movement during reading (Blythe & Joseph, 2011), a total of 10 studies were listed, comparing results from different age groups, 8 of which were carried out in English, 1 in German, and 1 in Finnish. Of these studies, 6 were published within the last decade, which is very good news, and it is to be hoped that the pace of work in this area will accelerate even more. One intriguing finding common to several of these studies is the fact that progress
in reading development co-occurs with a sharp reduction in the time spent rereading prior sections of text, pointing to postlexical integration effort as a major obstacle to fluent reading at a very young age (Huestegge, Radach, Corbic, & Huestegge, 2009). In line with this and similar findings, there is currently a move towards integrating lower level fluency skills with higher level comprehension skills in education-science-related reading research. As one example, a large longitudinal project is currently underway at the Florida Center for Reading Research as part of the US “Reading for Understanding” Research Initiative. This work combines eye movement data collected from children in Grades 1 to 5 with an extensive battery of psychometric reading skill and comprehension measures.

Considering spatially distributed processing, Rayner (1986) reported that while the perceptual span was smaller for developing readers than for adults, the typical asymmetry (more parafoveal processing occurring to the right of fixation) was already present in very young readers. These results were recently confirmed and extended by Häikiö, Bertram, Hyönä, and Niemi (2009) for Finnish readers. They reported that 8-year-old readers can identify about 5 letters to the right of fixation, and 10-year-olds about 7 letters, whereas 12-year-old readers have already reached the level of adults, with a rightward letter identification span of 9 letters. Interestingly, the perceptual span was larger for faster readers of all ages, confirming the hypothesis that slower readers focus more processing resources on the foveal region. At present there are no studies directly addressing the issue of serial versus parallel word processing in younger readers. The existing spatial and temporal processing limitations of young children (see Blythe & Joseph, 2011, for a thoughtful discussion), including the well-known reliance on sublexical units, lead to the hypothesis of a more “serial” mode of processing in early phases of reading development, but this has yet to be tested.

Concluding remarks
This overview attempted to provide an introduction into some important issues in current eye movement research on reading. The endeavour was largely (but not exclusively) focused on the theme of the present Special Issue, the serial versus parallel processing of words within the constraints of the perceptual span. Special attention has been given to phenomena that have recently emerged as new issues of debate: the occurrence of parafoveal-on-foveal effects, parafoveal processing of distant (N + 2) words, and the possibility of parafoveal processing at a semantic level. It appears safe to conclude that, no matter which model architecture is preferred as an explanation, the processing of words beyond the one next to the current fixation is really testing the limits of the system. It will remain difficult to demonstrate such effects, and the fact that they tend to only occur under favourable conditions will be taken by one side of the debate as evidence that they are atypical, while the other side will insist they disprove important claims about sequential processing.

A much more robust situation has emerged with respect to parafoveal-on-foveal effects. While such effects remain controversial at the lexical level and beyond, they are more generally accepted for sublexical and orthographic levels of processing. This leads to a shift in the ongoing debate: If lexical parafoveal-on-foveal effects are considered problematic for sequential attention models, while effects at “lower” levels are not (Schotter et al., 2012) the question arises what exactly happens at the transition. It would appear that the processing of infrequent letter combinations can only take place on the basis of statistical knowledge about the co-occurrence of letters within a given language. Does this not mean it is part of an initial phase of lexical processing as postulated in SAS models? The only alternative would seem to involve assuming a form of “preattentive” implicit memory for very rare letter combinations. This discussion may profit greatly from a clarifying consideration and precise definition of the notion of attention in relation to the ongoing processing of orthographic and lexical information (Kennedy & Pynte, 2009; Radach et al., 2007).

In the earlier discussion of non-Roman writing systems, the absence of visual word segmentation...
was identified as a key feature of Chinese, Japanese, and Thai scripts (see, e.g., Reilly et al., 2011, for an analysis of cognitive influences on saccade targeting in Thai). This raises the question of how the problem of serial versus parallel word processing should be addressed in this context. The existing extension of the E-Z reader model for Chinese reading (Rayner, Li, & Pollatsek, 2007) maintains processing assumptions originally developed for reading English, including the mechanisms of word-based attention shifts and saccade targeting. The extended model was able to approximate Chinese reading behaviour very well, but the success came at the expense of not explicitly addressing the problem of word segmentation. Given the general agreement that letter processing within words is parallel, it is a reasonable starting assumption that Chinese characters within the perceptual span are also processed in parallel (Li, Rayner, & Cave, 2009). If this is indeed the case, word segmentation may become more of an outcome than a starting point of lexical processing, precluding a word-by-word allocation of attention. It represents an interesting case of how the constraints of a writing system may alter the weight (or even order) of processing stages so that the common goal of comprehension is achieved in fundamentally different ways.

An important conclusion of the present paper is that the investment of more energy into the study of individual differences and reading development would be worthwhile. Currently the fact that reading skill varies within an individual tends to be glossed over. Little is known about how such top-down adjustments may translate into changes on the microlevel of linguistic processing and ocularmotor control. This may have important consequences for the debate over serial versus parallel word processing; reading “depth” (i.e., a more or less shallow processing mode) may covary with more or less parallel processing. One step in this direction is illustrated in the groundbreaking work by Wotschak and Kliegl (2013). They demonstrated that a difference in reading strategy induced by manipulating the frequency and difficulty of comprehension questions can strongly modulate parafoveal-on-foveal effects. A broad stream of similar research on the varieties of reading strategy and with it serial versus parallel processing may emerge over the next decade, even if this means saying farewell to the convenient construct of “normal reading”.

Original manuscript received 13 November 2012
Accepted revision received 14 November 2012
First published online 8 January 2013

REFERENCES


Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in


I. The location of initial eye fixations on words. *Vision Research, 28*, 1107–1118.


Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement...


