Eye movements in reading and information processing: Keith Rayner’s 40 year legacy

Charles Clifton Jr. a,*, Fernanda Ferreira b,*, John M. Henderson b, Albrecht W. Inhoff c, Simon P. Liversedge d, Erik D. Reichle d, Elizabeth R. Schotter e

a Department of Psychology, University of Massachusetts Amherst, United States
b Department of Psychology, University of California, Davis, United States
c Department of Psychology, Binghamton University, United States
d Department of Psychology, University of Palace, United Kingdom
e Department of Psychology, University of California San Diego, United States

A R T I C L E   I N F O

Article history:
Available online 24 August 2015

Keywords:
Keith Rayner
Eye movement control
Word recognition
Sentences and discourses
Visual scenes
E-Z Reader

A B S T R A C T

Keith Rayner’s extraordinary scientific career revolutionized the field of reading research and had a major impact on almost all areas of cognitive psychology. In this article, we review some of his most significant contributions. We begin with Rayner’s research on eye movement control, including the development of paradigms for answering questions about the perceptual span and its relationship to attention, reading experience, and linguistic variables. From there we proceed to lexical processing, where we summarize Rayner’s work on effects of word frequency, length, predictability, and the resolution of lexical ambiguity. Next, we turn to syntactic and discourse processing, covering the well-known garden-path model of parsing and briefly reviewing studies of pronoun resolution and inferencing. The next section shifts from language to visual cognition and reviews research which makes use of eye movement techniques to investigate object and scene processing. Next, we summarize Rayner and colleagues’ approach to computational modeling, with a description of the E-Z Reader model linking attention and lexical processing to eye movement control. The final section discusses the issues Rayner and his colleagues were focused on most recently and considers how Rayner’s legacy will continue into the future.

Introduction

After 40-plus years of work, many scientists will have published a large number of articles and amassed a great number of citations. This is certainly true of Keith Rayner (1943–2015), but Rayner also did something even more significant and far more rare: He almost single-handedly created a field of investigation and fundamentally changed the way research in a number of fields was conducted. His impact is apparent from the community of scientists he trained and inspired, and in the way even our basic, common-sense understanding of how reading is accomplished has changed since the 1970s, when Rayner first
began to publish. Several of us who were particularly influenced by him and his work prepared this overview to try to convey the significance of this legacy. Perhaps more importantly, we also hope to communicate how this legacy emerged from Rayner’s distinctive approach to science, which emphasized the importance of building theories bottom-up from a large body of solid, replicated findings, and put a premium on trying to understand the kinds of activities and tasks people perform in real-world situations, including during the reading of connected text and the visual processing of other complex stimuli. Rayner’s respect for the facts accounts at least in part for his extraordinary open-mindedness when it came to theoretical assumptions and commitments: As we all know from first-hand experience, when the facts changed, so did his theories. At the same time, Rayner had a healthy skepticism toward scientific trends and fads, and would only change his views in the face of compelling data.

Rayner’s initial research focused on how eye movements were controlled during reading. His early studies put many of the claims found in Huey’s early (1908/1968) analysis of reading on a solid empirical footing. Rayner’s work did in fact answer many of the questions that Huey had raised, providing important evidence, for instance, about the nature of the perceptual span, the region from which useful information is acquired during a single eye fixation. But as Rayner’s career progressed, he extended his work into questions about reading that Huey had never contemplated, and into topics that go well beyond the domain of reading. The following pages review some of these contributions. We begin with a review of his work on eye movement control and its relation to attention and the perceptual span, as well as describing his innovations in eye-tracking methodology. Next we turn to lexical processing, including the implications of this work for reading instruction, and continue on to a survey of Rayner’s research on the processing of sentences and discourse. The following section describes the work Rayner and his students conducted on object processing, scene processing, and visual search. An important aspect of Rayner’s research program is the computational model that integrates these findings and explicitly links attention and eye movement control. The final section reflects on Keith’s style of scholarship and mentorship and how it contributed to his legacy, and highlights the problems and issues his group was focused on in his final years.

Eye movement control and the perceptual span

Keith Rayner’s studies of the perceptual span and eye movement control were predicated on the conviction that the experimental investigation of visual information extraction from text should provide the empirical foundation for building a theory of reading. This ‘bottom-up’ approach fit well with his belief that reading was fundamentally a matter of extracting information from the printed page. Such an approach conflicted sharply with the view that was widely accepted when he and his mentor, George McConkie, started to investigate the perceptual span and eye movement control (e.g., McConkie & Rayner, 1975). According to this view, skilled reading was a linguistic guessing game, primarily a function of effective top-down processing, and perceptual processing was cursory, contributing relatively little to reading (Goodman, 1970; Hochberg, 1970; Levin & Kaplan, 1970). Rayner’s work with McConkie and their students and colleagues largely overturned this view with their demonstrations that experienced readers look at essentially every word in text, and that the ability to extract visual information quickly and efficiently is the foundation of skilled reading.

Eye-movement-contingent display changes

McConkie and Rayner (1975; Rayner 1975) pioneered new experimental techniques that permitted researchers to determine precisely when information is extracted from text during reading (see Rayner, 1978, for a discussion of the relation of these techniques to previously existing techniques). These techniques were based on the long-established fact (Javal, 1878) that, during reading, the eyes briefly fixate on a word so that it falls on the retinal fovea and parafovea, where visual acuity is high, and then jump (“saccade”) to begin a new fixation. The new moving-window technique enabled researchers to monitor eye position in real time and to control the spatial and temporal visibility of text relative to the location of a fixation. In the original study (McConkie & Rayner, 1975), a “window” of intact text was visible around the fixated letter upon fixation onset, and it moved in synchrony with the eyes during reading. Window size was experimentally manipulated, so that it ranged from relatively small, encompassing a single letter, to very large, encompassing a full line of text. Effects of experimentally controlled window sizes were observed through naturally occurring oculomotor responses, and different response components including fixation duration, saccade size, and saccade direction were analyzed to identify how information is extracted and used for the timing and targeting of eye movements.

Rayner (1975) also developed a boundary technique, which also allows researchers to evaluate the type of information extracted from the parafovea during a single fixation. In the boundary technique, text is presented contingent on the execution of a saccade that moves the eyes past an experimentally defined (invisible) spatial boundary. Properties of text that existed at that location prior to and after boundary crossing can be fully controlled and manipulated. In the original study (Rayner, 1975), at the onset of sentence reading and during fixations that preceded boundary crossing, the target word’s location in the parafovea was occupied either by the target itself, e.g., the word chest, a graphemically similar word or non-word, e.g., chart and chovt, respectively, or by a dissimilar nonword, e.g., ekovf. A saccade across the boundary initiated a display change that replaced the original item with the target word. Examination of subsequent target viewing revealed shorter viewing durations when a graphemically similar letter sequence had been visible in
the parafovea than when the preview had been dissimilar. In the 1975 data, the amount of facilitation (often termed preview benefit) depended on how close to the boundary the previous fixation was. When it was within 6 characters of the boundary, first fixation durations were some 30 ms shorter when comparing nonwords with similar letters to nonwords with dissimilar letters, with an additional advantage of approximately 50 ms when the preview was identical to the target word.

Over the past 40 years, numerous studies have used the window and boundary techniques to elucidate the extraction of information during reading fixations and its coordination with eye movement programming. Keith Rayner provided a number of comprehensive reviews (see Rayner, 1978, 1998, 2009b; Rayner & Liversedge, 2011; Rayner & Pollatsek, 1989; Rayner, Pollatsek, Ashby, & Clifton, 2012; Schotter, Angele, & Rayner, 2012). These reviews discussed the spatial area of text from which useful information is extracted, the nature of extracted information, the time course of information extraction, and eye movement programming. The results of the studies covered in these reviews led Rayner and his colleagues to develop a theoretical conception of information extraction and eye movement programming, generally referred to as the sequential attention shift model, a conception that was progressively refined throughout Rayner’s career, culminating in the E-Z Reader model (discussed in detail in a later section).

**Findings from initial work**

Early studies, primarily conducted with the moving window technique, revealed that reading rate was reduced, and that oculomotor activity was perturbed, when readers viewed the centrally fixated character and fewer than sixteen characters to its right and left (McConkie & Rayner, 1975; Rayner & Bertera, 1979; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). Even when viewing was unconstrained, mean saccade size was approximately eight character spaces, implying that the eyes were moved to a parafoveal location of text from which some useful information had been extracted. With small windows, the type of coarse-grain information outside the window mattered (McConkie & Rayner, 1975), suggesting that information uptake was graded, involving the extraction of graphemic detail at and near fixation, and of coarser-grained information from farther into the parafovea. Rayner, Well, Pollatsek, and Bertera (1982) showed that useful information about graphemic details was taken in from the first three letters of the word to the right of fixation, but that providing a preview of the full word (or the two words to the right of fixation) did not further speed up reading. They concluded that the parafoveally obtained graphemic information did not initiate word-level processing.

This work also showed that it was the successful extraction of information during a fixation that determined its duration and the size of an outgoing saccade. Using what is now called the “stimulus onset delay paradigm”, delays of up to 150 ms in the onset of text within the window increased fixation duration by an almost corresponding amount (Morrison, 1984; Rayner & Pollatsek, 1981; Rayner et al., 1981), and the filling of parafoveal interword spaces diminished saccade size, even when the spaces were filled only during the initial 50 ms of a fixation (Morris, Pollatsek, & Rayner, 1990; Pollatsek & Rayner, 1982). Extraction of linguistic information from the fixed segment of text, presumably for visual word recognition, and the extraction of word length information from farther into the parafovea, presumably for the targeting of saccades, thus commenced with fixation onset. Phenomena such as these (and more recent demonstrations, e.g., Dambacher, Slattery, Yang, Kliegl, & Rayner, 2013) provide evidence that eye movements in reading are largely, if not entirely, under the direct control of lexical and other linguistic processing.

Further work showed that decreases in the size of the window hampered reading more when they were imposed to the right rather than to the left of fixation (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980), indicating that the extraction of information was asymmetrically extended toward the right with left-to-right ordered English text. The asymmetry was reversed when reading Hebrew sentences (Pollatsek, Bolozky, Well, & Rayner, 1981) and when reading right-to-left ordered English sentences, and this occurred irrespective of the letter order within words (Inhoff, Pollatsek, Posner, & Rayner, 1989). The range of information extraction was thus determined by reading direction, which extends to encompass upcoming words. The finding that the perceptual span is asymmetric in the direction in which the eyes are about to move is an important piece of evidence linking eye movements and attention and plays an important role in computational models of reading, as we will discuss shortly.

Morrison (1984) provided an integrated and influential summary of this research that served as the framework for the E-Z Reader model reviewed in a later section of this paper. Morrison’s sequential attention shift model claimed that readers extracted linguistic information from a fixed word (N) upon fixation onset until a threshold of processing was reached. After this threshold was crossed, attention shifted to the next word, N+1, a corresponding saccade was programmed, and graphemic information was extracted from the N+1 location until the saccade was executed. Subsequent work using the boundary technique showed that readers extracted substantially more linguistic detail from a parafoveally visible and subsequently fixated word N+1 than was originally assumed. Specifically, converging evidence indicated that linguistic properties of the intact word N+1 preview influenced its usefulness (Briihl & Inhoff, 1995; Inhoff, 1989, 1990; Inhoff & Rayner, 1986; Lima & Inhoff, 1985), and that information extracted from the word at the N+1 location could be used for the activation of its phonological representation (Ashby & Martin, 2008; Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995; Lesch & Pollatsek, 1992, 1998; Miellet & Sparrow, 2004; Pollatsek, Lesch, Morris, & Rayner, 1992). Furthermore, the extraction of useful information from word N+1 was not only a function of its own lexical properties but also of the ease with which word N could be processed (Henderson & Ferreira, 1990), as increases in the linguistic difficulty of foveal word recognition diminished the use of parafoveally visible information.
Current issues

Several aspects of the relationship between information extraction and eye movement control are still under debate, and two sets of issues have shaped much of the recent work in this domain:

(1) **Parafoveal-on-foveal (POF) effects.** If a saccade to a parafoveally visible word is programmed before linguistic information is extracted from it, as maintained by sequential attention shift models, then linguistic properties of upcoming words should not influence the viewing of fixed word N. Nevertheless, such POF effects have been reported in several experimental studies (Inhoff, Starr, & Shindler, 2000; Kennedy, 2000; Starr & Inhoff, 2004; Vitu, Brysbaert, & Lancelin, 2004) and in corpus analyses (in which reading of natural texts is recorded) (Kennedy & Pynte, 2005; Kliegl, 2007; Kliegl, Nuthman, & Engbert, 2006). In defense of the original sequential attention shift assumptions, Rayner, Pollatsek, Drieghe, Slattery, and Reiche (2007) noted that item selection influenced the occurrence of POF effects in corpus analyses. Across studies, experimental POF effects are also elusive, perhaps because they accrue relatively late during the fixation of word N or because they spill over into the next word. Consistent with this finding, Schotter (2013) noted that the eyes may sometimes fail to land on the targeted parafoveal word. POF effects will then occur when the eyes are directed at word N+1 but land on word N instead. Occasional fixation of a parafoveally recognized word and the skipping of short sequences of random letters could occur when saccades under- and overshoot a targeted location, respectively. To date, N+2 preview effects have been obtained with short sentences whose targeting could be liable to error. In addition, measurement error may confound an observed linkage of information extraction and saccade targeting. A saccade that moves the eyes to word location N+1 could be recorded as belonging to location N (Kliegl et al., 2007).

In response to these findings, Rayner, White, Kambe, Miller, and Liversedge (2003; see also Drieghe et al., 2008; Rayner et al., 2007) noted that the eyes may sometimes fail to land on the targeted parafoveal word. POF effects will then occur when the eyes are directed at word N+1 but land on word N instead. Occasional fixation of a parafoveally recognized word and the skipping of short sequences of random letters could occur when saccades under- and over-shoot a targeted location, respectively. To date, N+2 preview effects have been obtained with short sentences whose targeting could be liable to error. In addition, measurement error may confound an observed linkage of information extraction and saccade targeting. A saccade that moves the eyes to word location N+1 could be recorded as belonging to location N (Reiche & Drieghe, 2015).

In summary, the use of two paradigms that allow researchers to manipulate the properties of text contingent on where a reader is looking – the moving window and the boundary technique – has provided important evidence concerning the kinds of information extracted during a fixation from not just the currently fixated word, but those that fall inside the parafoveal region. The results suggest that readers obtain information primarily from the currently fixated word and the word following, and that lexical processing controls attention and the movement of the eyes. The conclusions drawn from these studies exemplify Rayner’s theoretical conviction that attention drives the eye movement system, and that attentional shifts are driven by the extent of lexical processing. These findings provide important data concerning the nature of lexical processing, which is the topic we turn to next.

Lexical processing

Rayner’s conclusion that lexical processing controls attention and eye movements naturally led him to study...
the processes by which written words are recognized. His study of how the eyes move during reading enhanced our understanding of what is involved, in terms of psychological processing, from the earliest stages of a word’s visual encoding right through to the point that its corresponding mental representation is uniquely identified in the lexicon. A key aspect of this work was not only to understand the nature of these processes per se, but also to establish how they mapped onto the fixations that were made on the words of the sentences during normal reading. One of the most significant achievements in this area was the degree of specificity Rayner and colleagues achieved in their work in formulating an explicit mapping of the relationships between the sub-processes of lexical identification and the mechanistic processes associated with eye movement control in reading.

As discussed in the previous section, Rayner was one of the first researchers to recognize, and more importantly, to demonstrate, that visual encoding is initiated during fixations prior to those made directly on a word, but that full lexical identification of that word would likely occur during fixations on the word itself, or even during fixations made on words downstream in the sentence. Such a situation may arise when a word is skipped, or when so-called “spillover effects” occur. Rayner’s work on how lexical identification is distributed across fixations in reading led to the development of a formalized computational account eye movements in reading (Reichle, Pollatsek, Fisher, & Rayner, 1998), discussed in detail in a later section of this chapter.

The Big Three

In Rayner’s view, there are three properties of a word that most strongly influence how easily it can be processed: its frequency (e.g., Rayner & Duffy, 1986), length (e.g., Juhasz, White, Liversedge, & Rayner, 2008) and predictability in context (e.g., Ehrlich & Rayner, 1981). The importance of these properties led Rayner to refer to these as “The Big Three” of lexical processing. Beginning with the first of the Big Three, the basic word frequency effect is the finding that readers spend less time processing words that occur more frequently in the language than words that occur less frequently in the language. Early Rayner papers demonstrating these effects have been frequently and consistently cited in the literature (e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986). Lexical frequency effects are pervasive, occurring not only in word spaced alphabetic languages, but also in unspaced non-alphabetic languages such as Chinese in which words are far less visually obvious units of language (Yan, Tian, Bai, & Rayner, 2006).

Rayner was also instrumental in demonstrating that a word’s length, the second of the Big Three, is also a fundamental characteristic that affects how long a reader spends processing that word. Longer words have more constituent letters, and thus the word’s orthography provides more visual and linguistic information to process. The longer a word is, the longer it takes to process, although this effect is often reflected in refixations rather than extended single fixations. Word length affects not only when to move the eyes; it affects where to move them. In an interesting series of studies (e.g., Juhasz, Inhoff, & Rayner, 2005; Juhasz et al., 2008; White, Rayner, & Liversedge, 2005; see also Juhasz, 2008), the boundary paradigm (see previous section) was used to present either correct or incorrect information about the length of upcoming parafoveal words (e.g., backhund appeared as a preview for back and, and vice versa). The results of these experiments showed that saccades are targeted on the basis of the length of the word in the parafovea. Consequently, when a saccade is made to the parafoveal word, if word length information about that word turns out to be incorrect, reading is disrupted.

The third variable of the “Big Three” is the predictability of a word in context (S. Ehrlich & Rayner, 1981). Words that are more predictable are skipped more, fixated for less time, and refixated after a regression less often than less predictable words (Altarriba, Kroll, Sholl, & Rayner, 1996; Balota, Pollatsek, & Rayner, 1985; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987). More recently, as with the research investigating frequency effects, these effects have been shown to hold for non-alphabetic unspaced languages (Rayner, Li, Juhasz, & Yan, 2005), and it has also been demonstrated that there are differential effects of predictability for readers of different levels of reading skill (Ashby, Rayner, & Clifton, 2005): less-skilled readers show particularly strong effects of discourse context when they fixate on a relatively rare word.

It should be clear from the discussion above that from the earliest stages of Rayner’s career, he recognized the significance of these basic variables that affect the ease with which a word is identified during reading, and therefore, how central they should be to any mechanistic account of reading that might be developed. Given this, it is unsurprising that they are the mainstay of formal computational accounts of eye movement control during reading such as the E-Z Reader model (see below).

Lexical ambiguity effects

Another important aspect of lexical processing that Rayner invested significant effort into understanding was the influence of lexical ambiguity on eye movements in reading. Work by Rayner with Susan Duffy and Robin Morris (Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986), laid the foundations, and in more recent years the work has been extended (e.g., Binder, 2003; Binder & Morris, 1995; Binder & Rayner, 1998, 1999; Dopkins, Morris, & Rayner, 1992; Duffy et al., 1988; Folk & Morris, 2003; Kambe, Rayner, & Duffy, 2001; Pacht & Rayner, 1993; Rayner, Binder, & Duffy, 1999; Rayner & Frazier, 1989; Rayner, Pacht, & Duffy, 1994; Sereno, 1995; Sereno, Pacht, & Rayner, 1992; Wiley & Rayner, 2000). In most of these studies the experimental sentences contained a semantically ambiguous target word. The two meanings of the word could be balanced, that is, each meaning appeared approximately equally often in the language (e.g., case, pitcher), or alternatively, the meanings could be imbalanced, such that one of the meanings occurred less often in the language than the other (e.g., bank, table). One critical finding from this work was that when a word was imbalanced and the prior context provided by the sentence in which it appeared favored the less
dominant meaning, then reading was slowed compared to when the prior context favored the dominant meaning of the target. For example, the ambiguous word *bank* is read more slowly in a context like *The alligator left the ____* than in the context *The robber left the ____*. Rayner and colleagues termed this the “subordinate bias effect”. This finding helped motivate their “reordered access model” of lexical ambiguity processing, which assumes that a context that biases the meaning of an ambiguous word toward the less dominant meaning causes that meaning to compete with the one that is more dominant (e.g., Sereno, Brewer, & O’Donnell, 2003; Sereno, O’Donnell, & Rayner, 2006). Moreover, the logic of this approach, which emphasizes the interaction between frequency and semantic context, helped to inspire some interactive, activation-based models of parsing and sentence comprehension that will be considered in more detail in the next section (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994).

Disappearing text

Here we consider a set of studies that Rayner undertook using a variant of basic contingent change methodology, namely, the disappearing text paradigm. In these experiments the word that the reader fixates disappears (or is masked) a short time after the beginning of the fixation (usually after about 60 ms). In the original study, Rayner et al. (1981) showed that reading was undisturbed when the word disappeared shortly after fixation. This in itself is a surprising finding – intuitively, one might imagine that making the words of a sentence disappear as you are reading them would be disruptive. Even more intriguing, though, were the findings from some later studies that Rayner and his colleagues carried out (Liversedge et al., 2004; Rayner, Liversedge, White, & Vergilino-Perez, 2003). In these experiments, participants read sentences under disappearing text conditions in which there was a target word of either high or low frequency. The critical finding from these experiments was that even under disappearing text conditions, readers continued to fixate a blank space from where the word had disappeared for a period of time that was directly related to the frequency of that word. When the word was high frequency, readers fixated the blank space for less time than when the word was low frequency. This result provided very strong evidence to demonstrate that linguistic, and not visual processing, was the primary influence over decisions of when to move the eyes during reading. This finding also underlined the centrality of word identification as a driver of eye movements in reading, as mentioned earlier, an assumption that is central to the E-Z Reader model.

The disappearing text paradigm proved popular and motivated quite a number of subsequent studies. The paradigm was used to examine parafoveal as well as foveal processing by making the upcoming word, rather than the fixated word, disappear. This manipulation produced significant disruption to reading (Rayner, Liversedge, & White, 2006). More recently, the paradigm has also been used to investigate children’s reading (Blythe, Häikö, Bertram, Liversedge, & Hyöna, 2011; Blythe, Liversedge, Joseph, White, & Rayner, 2009; Rayner, Yang, Castelhano, & Liversedge, 2011). This work built on one of the earliest eye movement studies to investigate children’s reading (Rayner, 1986), which showed that when children read, they make more and longer fixations than adults do. Blythe et al. (2009) used the disappearing text paradigm to follow-up Rayner’s early work into children’s reading and investigated whether children’s increased fixation durations might have been caused by their being less efficient in their visual encoding of words during the early part of a fixation. Counter to this suggestion, Blythe et al.’s results showed that children were just as efficient as adults in visual encoding. This and later research has now led to the generally accepted view that increased fixation durations in children’s reading are primarily caused by less efficient linguistic, rather than visual, processing.

Implications for reading instruction

Not only was Rayner engrossed with issues of reading and eye movement control per se, but he also took time to consider how experimental results from reading research, and in particular, lexical processing in reading, can be used to inform best practice in the politically charged issue of how best to teach children to read. Rayner and colleagues’ authoritative statements regarding best practice in reading pedagogy (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; see also Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2002) provided a concise, but comprehensive and research based, overview of how children’s reading develops. In addition to laboratory-based considerations from experimental psychology, the article also discussed research investigating and evaluating the most effective techniques for teaching reading. One of the central conclusions that emerged from the evaluation of the literature was that, in order for children to become proficient readers, it is necessary for them to master the alphabetic principle, the idea that in the written form of the language, different letters correspond to different sounds. A second important conclusion that the authors formed in the paper was that direct instruction in phonics is an effective technique to allow children to understand the alphabetic principle, while other techniques such as whole word, or whole language approaches that do not adopt this direct approach are less effective. Although they provoked reactions from many individuals with strong views on how reading should be taught, the Rayner et al. (2001, 2002) articles provide excellent examples of scientifically based, translational writing, and they provide a model of how researchers can use findings from basic science to inform discussion and motivate evidence-based practice.

In Rayner’s work, the findings and conclusions concerning lexical processing dovetail closely with the work on eye movement control and the perceptual span. Both emphasize the importance of the word as a basic unit of reading and eye movement control, demonstrating that essentially every word in a text must be properly identified, and that reading processes are facilitated when lexical processing is made easier. In addition, the reciprocal nature of the research strategy is also apparent, in that as more was learned about the way lexical processing influences eye
movement behavior, it became increasingly possible to use eye movement measures to investigate basic language processing. This is an important theme of the next section on the processing of linguistic chunks larger than a single word.

**Sentence and discourse processing**

Rayner recognized that language processing goes beyond the identification of individual words. In includes, minimally, the processing of syntactic structure and the integration of information across sentences. We will begin this section with the topic of predictability, not only because it has links to issues already discussed, but also because prediction has recently become a “hot topic” in psycholinguistics.

**Predictability**

As discussed earlier, Rayner’s early work with Susan Ehrlich (Ehrlich & Rayner, 1981) showed that an expected word in a highly constraining context was read faster than an unexpected (but plausible) one, as well as being read faster and being skipped more often that a word in an unconstraining context. It also showed that deliberate misspellings sometimes went undetected, even when the eyes fixated on the misspelled word; most of these report failures occurred in highly-predictable words. The Ehrlich and Rayner paper did not enter the theoretical debates about interactive vs. bottom-up models of word recognition that were beginning to heat up at the time (e.g., Gough, 1983), or about whether there was really enough predictability in actual-occurring sentences to make predictions helpful (Gough, Alford, & Holley-Wilcox, 1981). Rather, the paper was a report of some empirical observations concerning how context affects perception in a variety of domains, and was aimed at establishing the facts about how context affects word recognition in normal reading. In later work, Rayner and Well (1996) systematically varied the degree of contextual constraint by making use of a modified Cloze procedure. They found that low-predictable words (4%) were read more slowly than medium (41%) and high (86%) predictable words, with no significant difference between the latter two.

Rayner’s empirical approach is nicely illustrated by his response to McDonald and Shillcock’s (2003a, 2003b) demonstration that transitional probabilities in a very low range (from a tenth of one percent to one percent) affected reading speed. Frisson, Rayner, and Pickering (2005) replicated McDonald and Shillcock’s empirical observations, but showed that they could be attributed to differences in Cloze-based predictability (all in the low-predictable range studied by Rayner & Well, 1996), with no residual effect of transitional probability once Cloze predictability is taken into account. For Rayner, this finding did not mean that experience with sequences of words (reflected in transitional probability) was unimportant; it simply meant that transitional probability was likely only one of several sources of information that affected Cloze completions, and for that matter, sentence comprehension processes. It is, for Rayner, an empirical question just what these sources of information are. Nonetheless, Rayner’s work on predictability has turned out to be highly relevant to current theoretical analyses of the role of expectation, construed in various ways, in sentence comprehension (e.g., Hale, 2001; Levy, 2008). Rayner himself did not become embroiled in theoretical debates about, for example, whether pervasive effects of predictability mean that language comprehension is best conceived in terms of memory retrieval rather than rule application, but clearly his findings have had a major impact on the development of theories concerning the nature of anticipatory effects in processing.

**Plausibility**

A wide variety of information sources clearly affect the plausibility of a sentence, and plausibility, in addition to predictability, seem likely to affect sentence comprehension. Rayner, in collaboration with Tessa Warren and others, investigated how the plausibility of basically-unpredictable words affects reading time. Their initial study (Rayner, Warren, Juhasz, & Liversedge, 2004) showed that reading time on a word like carrots was slower when it was anomalous in the context sentence (John used a pump to inflate the large carrots for dinner) than when it was plausible (...used a knife to chop ...) or even when it was implausible (...used an axe to chop ...). The anomaly effect appeared in gaze duration on the target word, with a hint that it appeared even in first fixation duration, meaning that interpretations are identified as anomalous very quickly. However, differences between the plausible and implausible conditions appeared only in late measures of reading, in the total time spent fixating the target word, including regressions back into it, suggesting that plausibility per se has its effect in a later, perhaps integrative, stage. In related work, Warren, McConnell, and Rayner (2008) showed that gaze duration (and marginally, first fixation duration) was lengthened on an “impossible” word (...used a book to teach the tough bread) even in a fantasy context (where e.g., Harry Potter was using the book), but that the fantasy context reduced the time spent going back and re-reading earlier material (and the critical word itself). Rayner and his colleagues interpreted the results in the context of a modular model in which context affects a stage of processing that occurs after lexical and syntactic information have had their effect, but also acknowledged that the findings could be accounted for by a properly developed non-modular model.

**Syntactic processing and the garden-path model of parsing**

One of Rayner’s most influential theoretical ideas concerned the processing of syntactic structure. He and Lyn Frazier (Frazier & Rayner, 1982) used eye movement measures to study the comprehension of sentences that contained a temporary syntactic ambiguity, which led to the development of what they called the “garden-path” theory of sentence processing. One kind of ambiguity involved what they called “late closure” and the other, “minimal attachment.” Late closure is illustrated in Since
Jay always jogs a mile (this) seems like a very short distance to him. If readers’ processing strategy is to keep adding words to a phrase as long as possible, they will take the phrase a mile to be the direct object of the verb jogs. The phrase beginning with this is consistent with this structural analysis. However, when this is absent, the verb phrase must be closed early, violating late closure, and a mile has to be taken as the subject of the main clause of the sentence. The other strategy, minimal attachment, is illustrated in an example such as The second wife will claim the inheritance when it is followed optionally by belongs to her. The minimal analysis of this sentence would take the inheritance to be the direct object of claim. This analysis is minimal because it involves less structure-building than the non-minimal analysis, in which the inheritance is the subject of a complement sentence. If readers prefer late closure and minimal attachment analyses, they should be ‘led down the garden path,’ and their reading should be disrupted when these analyses are disconfirmed (by the absence of this in the late closure example, and by the sentence continuation belongs to her in the minimal attachment example).

Frazier and Rayner recorded eye movements when their subjects read sentences like these. Eye movements were disrupted upon reading the disambiguating regions of both kinds of sentences. Frazier and Rayner interpreted their results as indicating that readers initially assigned a single analysis to a sentence, following what they referred to as late closure and minimal attachment strategies, and then were disrupted when they had to revise their analysis. The term “strategy” has misled some readers to view late closure and minimal attachment as rules for choosing among various available analyses. Rather, they are best viewed as arising out of the pressure to interpret a sentence as quickly as possible: building simpler structures and relating new words to phrases currently in active memory is faster than building new or more complex structures (Frazier & Fodor, 1978).

The Frazier and Rayner (1982) paper led to an explosion of work on sentence processing, having been cited (according to Google Scholar) over 1200 times in the literature. It demonstrated beyond any doubt that language comprehension is essentially incremental, and it enabled researchers to explore how a variety of principles are used in comprehending sentences. It also provoked a debate about modularity (specifically, information encapsulation) that dominated the field of sentence processing for the following couple of decades. To understand its impact, it is important to appreciate how ground-breaking this research was when it was published. Toward the latter half of the 1970s and the early 1980s, the study of syntactic processing was essentially dead. Researchers interested in “the psychology of language” were focused on issues such as word recognition (e.g., Fischer & Bloom, 1979; Marslen-Wilson & Welsh, 1978), inferencing (e.g., McKoon & Ratcliff, 1981; Singer & Ferreira, 1983), and schema-based effects on comprehension (e.g., Bower, Black, & Turner, 1979; Kintsch & Van Dijk, 1978). Any effects of syntax were attributed to a collection of ad hoc decision strategies (Clark & Clark, 1977) or were viewed as so minor as to be undetectable given the measures available at the time for assessing language processing. Even those small effects were predicted to be neutralized in appropriately biasing discourse contexts. Frazier and Rayner’s work, then, made two key contributions. First, the methodological insight was that specific theories of comprehension can only be evaluated using sensitive, online measures that allow the system to operate ‘naturally,’ in the sense that they do not require comprehenders to engage in tasks that are not part of normal comprehension (e.g., monitoring for the occurrence of a particular phoneme; Connine & Titone, 1996). Second, the study provided evidence for two fundamental principles of parsing, minimal attachment and late closure, both of which were consistent with the idea that the language processing system initially adopts the analysis that finishes first. This architecture allows the system to keep up with the input, maintain it in working memory, and integrate it with other sources of information, including context, pragmatic knowledge, and so on.

Beyond syntax

From this perspective, the obvious next step following the discovery of these two parsing principles was to see whether they would hold up to strong semantic and contextual manipulations. Rayner, Carlson, and Frazier (1983) examined sentences with reduced and unreduced relative clauses and with ambiguously attached prepositional phrases. The reduced relative clause sentences had subjects that were more plausible or less plausible on the relative clause reading (The performer sent the flowers was very pleased vs. The florist sent the flowers was very pleased; performers are likely to receive flowers, and florists to send them). Rayner et al. compared these to an unreduced version of the plausible performer sentence (adding who was) and an active version of the same sentence (The performer sent the flowers and was very pleased). Rayner et al. predicted that, if the language processor initially adopts the pragmatically most plausible analysis, readers should be garden-pathed by the reduced relative clause florist sentence and by the active performer sentence but not by the plausible reduced relative clause performer sentence. However, while they observed that their subjects were more likely to correctly paraphrase plausible reduced relative clause sentences than implausible ones, they also observed that reading of the disambiguating region was very was disrupted equally in both the plausible and implausible reduced relative clause versions. That is, plausibility (in the range studied here) could affect one’s eventual interpretation of a sentence, but did not seem to affect the initial analysis of a sentence. Similar results were obtained with ambiguously attached prepositional phrase sentences such as The spy saw the cop with binoculars but the spy didn’t see him versus saw the cop with the revolver. Readers seemed to initially adopt the simpler analysis (in which the prepositional phrase with the revolver modifies the verb, not the preceding noun), even when the resulting interpretation was implausible: their reading of the later portion of the revolver sentence was disrupted. Rayner et al.’s interpretation was that the parser builds structure by following the minimal attachment principle,
but in a second stage of processing revises that structure if it is inconsistent with semantic or pragmatic constraints. Ferreira and Clifton (1986) followed up this study by examining the role of both discourse context and animacy and drew similar conclusions: The parser gets first crack at the input and structures it according to the principle of minimal attachment, with discourse and animacy information coming in at later stages to aid in the reanalysis of those syntactic structures.

These conclusions did not go unchallenged. Much of the following decades of research on sentence comprehension was directed at testing this model and developing alternatives. One line of criticism was to suggest that the semantic and discourse manipulations in these previous studies were weak, and that more strongly biased materials did indeed eliminate the effects of syntactic difficulty (e.g., Altman & Steedman, 1988; Taraban & McClelland, 1988; Trueswell, Tanenhaus, & Garnsey, 1994). Another was to argue that any effects of syntactic difficulty were attributable to the frequency of the less preferred structures rather than to any inherent structural bias (e.g., MacDonald et al., 1994). In many cases, the findings from these studies were used to motivate models of comprehension with radically different architectures from the one assumed by Rayner and Frazier. According to these 'interactive constraint-satisfaction' models, analyses were assumed to be proposed in parallel rather than serially, and high-level constraints were assumed to be used interactively to boost or squelch the activation level of those analyses.

In response, Rayner and his colleagues have provided some evidence that they argue fits better with the approach that he and Frazier proposed than with a constraint-satisfaction approach. For instance, Binder, Duffy, and Rayner (2001) tested what they argued was a strong prediction of the constraint-satisfaction view: that it should be possible to show that syntactically simpler structures are actually more difficult than syntactically complex structures given the right set of semantic constraints. For example, a sentence such as the patient cured the inexperienced doctor should induce a reverse garden-path because the patient cured strongly invites a reduced relative interpretation (based on pragmatic constraints as well as frequency biases), particularly in a strongly biasing discourse context (one that presents two patients who need to be distinguished with a modifier such as a relative clause). In three experiments, however, Binder et al. failed to find any evidence for reversed garden-path effects. Clifton et al. (2003) addressed the possibility that earlier demonstrations that the semantic manipulations that failed to overcome syntactic biases were too weak. They used reduced relative clause materials developed by Trueswell et al., 1994, in which plausibility was manipulated by using animate vs inanimate potential subject noun phrases (the man paid/the ransom paid). According to an interactive model, the inanimate version should allow the parser to predict a relative clause during the earliest stages of parsing. While Trueswell et al. (1994) had shown that their plausibility manipulation marginally reduced the garden-path disruption in some measures (especially first pass reading time) of eye movements, Clifton et al. showed that this was not the case for other measures, especially the tendency for the eyes to regress and re-read earlier material when encountering a syntactic disambiguation. Sentences with animate subjects were generally more difficult than those with inanimate subjects, but this held equally true for whether or not there was any temporary syntactic ambiguity. It appears, then, that plausibility did not affect the initial analysis of a sentence, but instead, affected its eventual comprehensibility.

At this point in the history of the field, it appears that psycholinguistics has accepted the demonstrations by Rayner and many other researchers that a wide variety of factors affect how sentences are read. At the same time, the field has moved on to questions other than whether the language processing system has a modular or fully interactive architecture without fully incorporating some of Rayner's later findings. Nonetheless, no one working in the field today would dispute the idea that Rayner's early work on parsing made it possible for the field to make genuine progress on the question of how people use structural information to quickly and efficiently obtain an interpretation for a sentence.

Higher-level discourse

Rayner and his colleagues contributed to a set of issues in the psychology of language that has been of interest from the 1970s right through to today. These issues have to do with the processes that allow readers to build anaphoric and inferential connections among sentences to create coherent representations of text meaning. Debate has centered on the question of how active comprehenders are: Do they anticipate relationships and connections, elaborating on the text they are reading, or are they more passive? In one of the earliest eye tracking investigations of anaphoric processing, readers were presented with texts in which a pronoun in a sentence referred to an antecedent either right before or several sentences back in the text (Ehrlich & Rayner, 1983). Fixation times on the pronoun revealed that the identification of an antecedent was not always an immediate process; especially for antecedents further back in the text, pronoun resolution processes appeared to take place over several fixations. This finding is consistent with the claim that coherence is established early during text processing, but it also appears that readers sometimes move on to take in more input even as they are still engaged in fleshing out the interpretation of what came before.

One might argue from these results that the text comprehension system is somewhat lazy, allowing some kinds of relations to be only partially specified during online processing. This view is similar to a model of inference which assumed that readers engage only in “minimal inferencing” (McKoon & Ratcliff, 1992), drawing links among concepts in text only when required (and sometimes not even then; see Greene, McKoon, & Ratcliff, 1992). This proposal led to a series of studies in the text processing literature designed to assess whether readers would in fact draw predictive inferences online rather than drawing them only when a piece of text compelled the creation of the inferential link. O'Brien, Myers, Shank, and Rayner (1988) made a key contribution to resolving this debate.
when they demonstrated that elaborative inferences are indeed sometimes drawn online, particularly when they involve anaphoric inferences (e.g., inferring that the term “knife” refers to something referred to as a “weapon” earlier in the text). In a follow-up study, Garrod, O’Brien, Morris, and Rayner (1990) suggested that two types of inferencing need to be distinguished: a passive form in which elaborative inferences arise from passive resonance in memory, and a more active version in which a reader actually predicts an upcoming expression. Again, we see how Rayner’s work anticipated some of the debates that currently dominate the literature on language processing — in this case, the question of how active comprehenders are and the extent to which active prediction normally occurs.

In summary, perhaps the most compelling conclusion that should be drawn from this work on sentence and discourse processing is that eye movements are a critical tool for developing and testing alternative models of how connected text is processed. The studies that have been reviewed demonstrate that not only is it possible to measure how various linguistic variables influence online comprehension, but also that different measures provide insight to different stages of processing: early measures such as gaze duration reflect processes that occur when a word is initially encountered, and later measures such as re-reading time provide information about revision and reanalysis. Although this field is certainly characterized by robust theoretical debate, there is likely almost no controversy about the idea that Rayner’s influence has been evident since the beginning, and that his contributions to our methodological toolkit have been invaluable.

Object and scene processing

After it became clear that reading, a basic visual information processing task, could be understood through the investigation of eye movements, Rayner and his colleagues shifted some of their efforts to the study of other domains of visual information processing, including visual search, object perception, and scene processing.

Visual search

In his early work on visual search, Rayner showed that search normally requires and is facilitated by eye movements, a finding that ran contrary to a common assumption in cognitive psychology that visual search was accomplished entirely by covert attention and that attention could be distributed broadly over wide visual areas. This research therefore made use of tachistoscopic methods in which an entire visual array or scene was shown very briefly, with the viewer required to maintain fixation in the center of the display. Here we can make an important link to Rayner’s work discussed earlier on word processing, which showed that our understanding of lexical processing changed dramatically once researchers moved away from presenting words in isolation and ignoring the role of eye movements. Similarly, in visual search, Rayner and colleagues showed that eye movement measures like fixation duration and saccade amplitude are systematically related to variables generally known to affect the difficulty of search, including item density and target–distractor similarity (Bertera & Rayner, 2000; Greene & Rayner, 2001a, 2001b). Gaze contingent moving window and moving mask studies further indicated that, during search, the amount of information available at the fovea and in the periphery in each fixation had a marked effect on both eye movements and search efficiency (Bertera & Rayner, 2000; Greene, 2006; Greene & Rayner, 2001b; Pomplun, Reingold, & Shen, 2001). For example, using the moving window method, Rayner and Fisher showed that the perceptual span during visual search is influenced by factors like target–distractor similarity (Rayner & Fisher, 1987a, 1987b). The use of eye movements as a method is now commonplace in the study of visual search, and this work is the basis for a variety of theoretically motivated computational models of search in item arrays (Pomplun, Reingold, & Shen, 2003; Zelinsky, 2008) and in natural scenes (Torralba, Oliva, Castelhano, & Henderson, 2006).

Transsaccadic object perception

Rayner’s early work on object perception focused on issues of preview benefits and transsaccadic integration. This work built on the theoretical and methodological approaches that Rayner pioneered in his study of reading. Initial research, developed and conducted in collaboration with Alexander Pollatsek, was directed toward understanding the nature of the representations that are preserved and integrated across saccades during object perception (Pollatsek, Rayner, & Collins, 1984; Pollatsek, Rayner, & Henderson, 1990). The paradigm was based on the boundary technique developed in reading: A preview object was presented outside of the fovea and the participant was asked to execute a saccade to that object. During the saccade the display changed to a target object that the participant was asked to identify (e.g., name) as quickly as possible. By manipulating the relationship between the preview and target object, inferences could be drawn concerning the transsaccadic integration process. The results suggested that relatively abstract representations rather than highly precise visual images are carried across saccades, with the amount of preview benefit much larger in object perception (typically greater than 100 ms) than in reading (Henderson, 1992; Henderson, Pollatsek, & Rayner, 1987, 1989; Henderson & Siefert, 1999, 2001; Pollatsek et al., 1984, 1990).

Scene perception

Rayner and his colleagues extended many of the insights they gained from reading, visual search, and object perception to natural scene perception. In this work, participants view images of scenes on a computer monitor while their eye movements are recorded. At the most basic level, issues under investigation concern the nature of the information that determines where and for how long viewers look at objects and other scene elements (Henderson, 2007, 2011). Theoretical questions have centered on how image variables and cognitive variables combine to control
attention and eye movements in scenes (Foulsham & Underwood, 2008; Henderson, Brockmole, Castelhano, & Mack, 2007; Itti & Koch, 2001; Neider & Zelinsky, 2006; Nuthmann & Henderson, 2010; Parkhurst, Law, & Niebur, 2002; Tatler, Baddeley, & Vincent, 2006; Torralba et al., 2006). Much of this work involves directly manipulating image and cognitive variables and measuring their effects on eye movements. In this literature, eye movement measures like fixation duration and saccade amplitude therefore provide critical evidence for testing theoretical approaches to attentional deployment in scenes. In addition, inspired by Rayner’s work in reading, gaze-contingent techniques in which the display is changed based on the viewer’s eye movements have been extended to the study of scenes (Castelhano & Henderson, 2007; Henderson, McClure, Pierce, & Schrock, 1997; Henderson & Smith, 2009; Luke, Nuthmann, & Henderson, 2013; Rayner, Smith, Malcolm, & Henderson, 2009; Saida & Ikeda, 1979).

Another critical theoretical issue in object and scene perception, and in cognitive psychology more generally, concerns the degree to which perception is influenced by prior knowledge. Rayner and his colleagues initially approached this issue from the perspective of transsaccadic object perception. They showed that the information acquired from an object in the visual periphery prior to fixation could be influenced by the presence of other semantically related nearby objects (Henderson et al., 1987, 1989). This type of work was later extended to scene perception. Rayner and colleagues provided early evidence that scene context is able to exert an influence on the processing of objects in a scene (Boyce, Pollatsek, & Rayner, 1989). In recent years, these same techniques have been extended to more natural scene images including scene photographs and computer-rendered scenes, allowing investigators to ask more subtle questions about the nature of contextual influences on object processing and the degree to which observed effects are due to cognition–perception interaction. Current evidence suggests that local peripheral objects are typically not processed to a semantic level before they are attended (Rayner, Castelhano, & Yang, 2009; Võ & Henderson, 2011). This finding is reminiscent of some reports of a lack of semantic-level parafocal-on-foveal and semantic-level preview benefits in reading, as discussed in the section on eye movement control and the perceptual span (but see the final section of this paper for late developments). This topic remains controversial and continues to generate new research.

Continuing influences of Rayner’s work in object and scene perception

As the study of visual cognition has matured, the eye movement methods that Rayner developed in reading and extended to visual cognition are now commonly used to study theoretical issues in dynamic natural scene perception in videos (Võ, Smith, Mital, & Henderson, 2012) and in the world itself (Land & Hayhoe, 2001). In addition, Rayner and colleagues sought to combine eyetracking and neural data (ERPs) as converging evidence to investigate theoretical issues in word recognition during reading (Sereno & Rayner, 2003; Sereno, Rayner, & Posner, 1998). Initially these measures were collected separately over the same materials, but several laboratories are now working on simultaneous recording and co-registration of eyetracking and neuroimaging (ERP and fMRI) to investigate the neural correlates of visuo-cognitive computations (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011; Henderson & Choi, 2015). Co-registration offers the benefits of the on-line behavioral measures provided by eyetracking along with those of measuring functional neural processes associated with underlying cognitive processes. Co-registration methods are likely to become more common for the study of visual cognition in the coming years as their technical challenges are addressed.

Another extension of this work is the development of the “visual world paradigm” for the study of spoken language processing (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). The literature including studies making use of the visual world paradigm is too extensive to be reviewed here; instead, we simply point out that the technique, which allows people to listen to sentences while simultaneously viewing objects and other visual stimuli, has become increasingly sophisticated as we have learned more about questions such as what kinds of objects tend to attract fixations, how long those fixations tend to be, and how saccades are programmed in the context of images. In addition to telling us a great deal about the interactions between the language and visual information processing systems, the visual world paradigm has led to an explosion of research on a broad range of topics concerning spoken language processing, ranging from speech perception (Tanenhaus, Magnuson, Dahan, & Chambers, 2000) to syntactic priming (Thothathiri & Snedeker, 2008) to conversation (Brown-Schmidt, Gunlogson, & Tanenhaus, 2008). The impact of this research, then, has been significant.

As we hope is evident from this section, Rayner’s influence led directly to the development of an innovative and remarkably successful research program investigating how naturalistic objects and scenes are perceived, identified, and understood. Even though the problems addressed in the field of visual cognition differ from those in the field of reading, many of the insights and methodologies were directly imported from reading to object and scene perception. In the next section, we return to the topic of reading and the creation of a computational model to explain eye movement control during reading and language processing.

A computational model of eye movement control

Rayner’s contributions to our understanding of reading are hardly limited to the empirical studies reviewed thus far. As mentioned earlier, he and his colleagues Erik Reichle, Alexander Pollatsek, and Donald Fisher have made a fundamental theoretical contribution to the field by developing a computational sequential attention shift model of eye-movement control during reading known as E-Z Reader (Reichle et al., 1998). The decision to develop this model was motivated by the absence of a formal account of eye movements in reading, and by the
accumulation of a large number of empirical findings by Rayner and his colleagues demonstrating that the moment-to-moment “decisions” about when to move the eyes during reading were tightly linked to on-going cognitive processing (e.g., word identification).

As discussed in earlier sections of this paper, Rayner’s work showed that lexical identification during the reading of text is a process that is distributed across fixations in reading (Rayner & Liversedge, 2011). This finding is not compatible with earlier computational models of word identification, which had focused on the identification of isolated words. Because of this focus on single word processing, such models implicitly assume that a word is fully available in central foveal vision throughout the entirety of the period during which it is identified. One goal of E-Z Reader was to explore the implications of specific assumptions (building on Morrison’s, 1984, earlier theoretical work) about how attention is distributed between foveal and parafoveal words in reading.

Another goal of E-Z Reader was to address an apparent paradox. Both lexical processing and eye-movement programming have been shown to be relatively time consuming, but individual fixations are often very short in duration. One goal in developing the E-Z Reader model was thus to understand how fixations might be short in duration but still adequate to identify words, by exploring how lexical processing initiates the programming of a saccade that moves the eyes off a word.

A family of models

With that background, E-Z Reader can actually be described as being a “family” of models that has continued to be refined over the years, with each version being developed to explain additional findings relevant to the mental processes that guide the eyes during reading. (Fig. 1 is a schematic diagram of the current version of the model, E-Z Reader 10.) However, all versions of the model have shared two basic theoretical assumptions. The first is that the attention that is necessary to support lexical processing is allocated in a strictly serial manner, to exactly one word at a time. The second is that the completion of a preliminary stage of word identification, called the familiarity check, on a given word is the ‘trigger’ that initiates the programming of a saccade to move the eyes to the next word. Because attention shifts from one word to the next with the subsequent completion of lexical access, the movement of attention is only loosely coupled with the movement of the eyes, allowing the model to explain, among other findings, the fact that fixations tend to be inflated on words following difficult-to-process words (i.e., ‘spillover’ effects; Rayner & Duffy, 1986) and the finding that difficult-to-process words afford less preview of words in the parafovea (i.e., the interaction between foveal load and preview; Henderson & Ferreira, 1990). The remaining assumptions of the model are related to saccadic programming and execution, and were borrowed from well-documented findings. One example is the finding that saccades are programmed in two successive stages—a labile stage that can be canceled if another saccade is initiated, followed by a non-labile stage that cannot be canceled (Becker & Jürgens, 1979). Another is the finding that saccades appear to be directed toward the centers of words (Rayner, 1979), but that the saccades are subject to motor error, so that the distribution of fixation locations is approximately normal in shape (McConkie, Kerr, Reddix, & Zola, 1988).

The original version of the E-Z Reader model (Reichle et al., 1998) was limited to explaining various first-pass eye-movement measures at a relatively coarse grain size—measures of fixation duration and probability of fixation for individual words. Subsequent versions have been developed to explain various aspects of eye-movement behavior at finer levels of details. First, the model was augmented to explain fixation location distributions and the way in which fixations are more likely to occur following an initial fixation near either end of a word (Reichle, Rayner, & Pollatsek, 1999). Subsequent versions of the model have also been modified to include more accurate assumptions about how the frequency and predictability of words influence the times required for their identification (Rayner, Ashby, Pollatsek, & Reichle, 2004), and a pre-attentive stage of visual processing in which update of visual information is delayed by the eye-to-brain lag (Pollatsek, Reichle, & Rayner, 2006a). Most recently, the model has been expanded to include a post-lexical stage of integration that, when it fails, can interrupt the progression of the eyes through the text, resulting in a regression back to the point of integration failure (Reichle, Warren, & McConnell, 2009). The model has also been used to examine a number of important theoretical issues, with some of the most notable including: (1) the extension of the model to several non-reading tasks (Reichle, Pollatsek, & Rayner, 2012); (2) the developmental changes that occur as beginning (e.g., child) readers become skilled (e.g., adult) readers (Mancheva et al., 2015; Reichle et al., 2013) and as adult readers become elderly readers (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006); (3) the extent and time course of parafoveal processing (Schotter, Reichle, & Rayner, 2014; Sheridan & Reichle, in press); and (4) the possibility that fixation-location measurement error can account for apparent parafoveal-on-foveal effects (Reichle & Drieghe, 2015). (For a comprehensive discussion of the model and its theoretical scope, see Reichle, 2011, or Reichle & Sheridan, 2015.)

Impact of the model

Since its initial inception, the E-Z Reader model has motivated a considerable amount of new empirical research (e.g., Inhoff, Eiter, & Radach, 2005; Inhoff, Radach, & Eiter, 2006; cf., Pollatsek, Reichle, & Rayner, 2006b, 2006c). One indication of the model’s success in this capacity is the fact that, as of the writing of this article, the original Reichle et al. (1998) article has been cited almost 800 times, according to Google Scholar. Perhaps even more importantly, the model has inspired the development of new, competitor models such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005), SERIF (McDonald, Carpenter, & Shillcock, 2005), and Glenmore (Reilly & Radach, 2006). Although these models differ from E-Z Reader in a number of ways, one of the most salient of
these differences is related to their assumptions about how attention is allocated in the service of identifying words. This contrast is most apparent if one compares E-Z Reader and SWIFT: Whereas in E-Z Reader attention is allocated in a strictly serial manner, in SWIFT attention is allocated as a gradient that supports the concurrently lexical processing of multiple words. This difference has itself been the subject of heated debate (cf., Radach & Kennedy, 2013; Reichle, Liversedge, Pollatsek, & Rayner, 2009) and has also motivated a large amount of new empirical research on topics related to those mentioned in the first section of this paper on eye movement control and the perceptual span, particularly the processing of words N+1 and N+2 during fixation (e.g., for reviews, see Rayner, 2009a, 2009b). Both this debate and the development of competitor eye-movement models were welcomed by Rayner, who strongly advocated the use of models to motivate new empirical research and the results of that research to adjudicate between models and their assumptions.

Keith Rayner’s legacy

Rayner’s legacy includes many research findings beyond those we have reviewed here. It also, and most importantly, includes the influence he had on those whom he mentored and with whom he collaborated. His approach to mentorship throughout his career reflected a ‘coaching’ style. He was happy for students to branch out into new areas, working with a variety of researchers on a number of different topics. However, Rayner insisted that his collaborators maintain a close connection to past knowledge and findings. He was notorious both as a reviewer and as a member of dissertation committees for insisting on citing the first paper to demonstrate an effect or develop a paradigm. This was not because he had authored that paper, although that was often true. For example, he constantly returned to Dearborn (1906), Dodge (1907), and Huey (1908), pointing out that they had generated a still-relevant hypothesis about reading over a century earlier. He generally started his grant proposals, review papers (e.g., Leinenger & Rayner, in press), and opinion papers (e.g., Rayner & Reingold, 2015) with a quote from Huey. He often marveled at how on target the pioneers were, despite the rudimentary nature of their methods.

Rayner believed that if we lost the knowledge previous researchers had gained, we would reinvent the wheel over and over again – or worse, build wheels that did not turn. It was this insistence on preservation of ideas and on accuracy that led to his final first-authored work. As many readers of this journal will know, in February of 2014, a new app was announced by the company Spritz that claimed to allow people to speed read by presenting words one at a time on their mobile phone or computer (i.e., via Rapid Serial Visual Presentation: RSVP) at rates as high as 1000 words per minute. The announcement generated a great deal of media buzz and Rayner was contacted by many reporters asking for a comment; he replied “Hogwash” to a reporter from the New Yorker (Camp, 2014).

Fig. 1. A schematic diagram of the E-Z Reader model of eye-movement control during reading. The gray arrows (left column of boxes) indicate how information is propagated between the various components of the model, and the black arrows (right column) indicate how control is passed between those components.
The app was developed by a group of engineers and businessmen. None of them had done any research on reading, nor had the team done its due diligence to investigate the existing literature. Within a month Rayner had published a paper in Psychological Science challenging the premise of the app that regressions (i.e., re-reading) were detrimental to the reading process (Schotter, Tran, & Rayner, 2014). In fact, Rayner and his colleagues reported the opposite: that the inability to re-read actually impaired comprehension. Concerned by public excitement around the app despite numerous scientists questioning their claims, Rayner proposed an article to Psychological Science in the Public Interest on the allure of the prospect of speed reading and how consumers should be aware of the psychology of reading in order to evaluate proponents’ wild and inaccurate claims. The article could only appear posthumously (Rayner, Schotter, Masson, Potter, & Treiman, in press), although the first draft was submitted a few days before Rayner passed away.

While Rayner consistently built on past results, he was always willing to question them. Sometimes the conclusions he reached on the basis of existing evidence held up. For instance, a central tenet of his work was that eye movements are under direct cognitive control, as we have discussed throughout this piece (see Rayner & Reingold, 2015): in other words, an eye movement is determined by processes associated with lexical processing of the word being fixated. Not all researchers agreed with this idea, arguing instead that cognitive–linguistic influences on eye movements only appear in very long fixations (Chanceaux, Vitu, Bendahan, Thorpe, & Grainger, 2012; Deubel, O'Regan, & Radach, 2000; McConkie & Yang, 2003; Yang & McConkie, 2001). Despite the difference of opinion he had with other researchers, he maintained an open-lab policy. In fact, many advocates of rival theories would send their students to visit his lab and fruitful collaborations emerged. In the work they did together, the conclusions often came down on his side of the debate. For instance, they found evidence for direct control of eye movements using modifications of the gaze-contingent methods he had developed decades earlier (e.g., Dambacher et al., 2013, using the stimulus onset delay paradigm introduced by Rayner & Pollatsek, 1981; Schad, Risse, Slattery, & Rayner, 2014, using the fast priming paradigm introduced by Sereno & Rayner, 1992).

Sometimes, however, new data convinced him to change his long-held theoretical positions. A good case in point concerns the extraction of parafoveal information during reading. As discussed in the first section on eye movement control, for most of his career, Rayner argued against the idea of semantic preview benefits—the idea that semantic information is acquired from a parafoveal word and is integrated with information acquired from the word once fixated (Rayner et al., 1986). When studies in languages such as German (Hohenstein et al., 2010; Hohenstein & Kliegl, 2014) and Chinese (Yan, Richter, Shu, & Kliegl, 2009; Yan, Zhou, Shu, & Kliegl, 2012; Yang, 2013; Yang, Wang, Tong, & Rayner, 2010) reported evidence consistent with semantic preview benefits, he initially attributed the findings to cross-language differences. However, when one of his students found evidence in English for preview benefits from synonyms (in fact, of a similar magnitude to those obtained from identical previews) Rayner insisted she replicate the effect. And replicate it did—but the follow-up work also demonstrated a lack of preview benefit for semantic associates that were not synonyms (Schotter, 2013; see also Rayner, Schotter, & Drieghe, 2014). In typical fashion, he and his students went on to identify several factors that contributed to semantic parafoveal preview benefits (Rayner & Schotter, 2014; Schotter et al., 2015) and even modeled the effects in E-Z Reader to determine whether they could be accommodated given the assumption of serial attention shift—which, in fact, they can (Schotter et al., 2014).

Rayner was always willing to be wrong. He believed you could not win an argument against convincing data, so if a set of findings did not support what he believed, Rayner would update his beliefs. He insisted that any finding that was inconsistent with a model meant that the model needed to be modified, and that model assumptions could only be tested with good data: “Thus, the models and good data go hand in hand in advancing the field.” (Rayner, 2009a, p. 7). Rayner particularly valued findings obtained from controlled experiments over other kinds of data. For example, he saw both value and danger in corpus analyses of eye movement data, but ultimately favored an experimental approach in which the properties of particular words in sentences are carefully manipulated. By examining the reading of most or all the words in naturally occurring texts, often extended segments of prose such as a short story, corpus analyses are important for generalizing experimental results and for exploring and developing new hypotheses. However, they have all the limitations of correlational analyses in that they show correlations among the variables that are measured and thus are susceptible to the existence of unidentified confounded variables. Ultimately, a field advances best when controlled experimental techniques and naturalistic observational techniques proceed in a mutually informed fashion. We, as Rayner did, look forward to seeing what the data will show us in the future.

In closing, we make the obvious point that Keith Rayner’s lifetime of research could not answer all the interesting questions about how people read and process visual information. Moreover, he left many theoretical controversies to be resolved, including those concerning basic language processing as well as the nature of eye movement control during reading and scene processing. But thanks to his work and his ability to inspire the best scientific ideas and practices in his colleagues, we can point to a large body of knowledge that has enhanced our basic understanding of cognition and which has been applied to problems such as reading pedagogy. The field has lost a great leader and an exceptional human being, but we are all fortunate for what Rayner contributed to the study of cognition during his extraordinary life.

References


