ORIGINAL ARTICLE

The role of global top-down factors in local eye-movement control in reading

Ralph Radach · Lynn Huestegge · Ronan Reilly

Abstract Although the development of the field of reading has been impressive, there are a number of issues that still require much more attention. One of these concerns the variability of skilled reading within the individual. This paper explores the topic in three ways: (1) it quantifies the extent to which, two factors, the specific reading task (comprehension vs. word verification) and the format of reading material (sentence vs. passage) influence the temporal aspects of reading as expressed in word-viewing durations; (2) it examines whether they also affect visuomotor aspects of eye-movement control; and (3) determine whether they can modulate local lexical processing. The results reveal reading as a dynamic, interactive process involving semi-autonomous modules, with top-down influences clearly evident in the eye-movement record.

Introduction

In recent years, there has been a steady growth in research using eye movements to study information processing in reading. In much of this literature, eye tracking is used to test specific psycholinguistic hypotheses, predominantly with respect to processing at the word and sentence level (Clifton, Staub, & Rayner 2007, for a review of current issues). Another significant part of this literature is focused on the development of eye-movement control models that attempt to explain how the linguistic and visuomotor processing streams act together to produce the observable oculomotor reading behavior (see Radach, Reilly, & Inhoff, 2007, for a discussion). Radach and Kennedy (2004) have argued that although the development of the field has been impressive, there are a number of issues that will require much more attention in future research. One of these issues concerns the variability of skilled reading both inter-individually and within the same reader. In this paper, we propose to explore the topic of variability in three ways: (1) quantify the extent to which top-down factors influence the temporal aspects of reading as expressed in word-viewing duration measures; (2) examine whether top-down factors influence visuomotor (spatial) aspects of eye-movement control; and (3) determine whether top-down factors can modulate local lexical processing.

Two research cultures

Everything from the nature of the encoding through to the age, experience and domain knowledge of the reader is implicated in the facility with which he or she can extract meaningful information from the text. An important goal for both educators and cognitive scientists is to quantify a reader's understanding of a text and the ease or otherwise with which it is achieved. However, given the size and complexity of the task of understanding this singular skill, there has been an inevitable division of labor between those

R. Radach (✉)
Department of Psychology and Florida Center for Reading Research, Florida State University, Tallahassee, FL, USA
e-mail: radach@psy.fsu.edu

L. Huestegge
Department of Psychology, RWTH Aachen University, Aachen, Germany

R. Reilly
Department of Computer Science,
National University of Ireland, Maynooth,
Kildare, Ireland

concerned with its high-level features (discourse structure, text cohesion, sentence comprehension) and its more low-level aspects (word recognition, orthographic effects, eye-movement control). This division has been further exacerbated by the use of different methodologies and technologies by the respective research groups. The traditional methodologies associated with reading comprehension involve, for example, tests of memory for gist or detail (Butcher & Kintsch, 2003 for a comprehensive review), whereas the spatio-temporal aspects of eye movements are the key data for researchers more focussed on the information processing aspects of reading.

Currently, these broad methodological approaches now represent two distinct research sub-cultures (Stanovich, 2003). Comprehension reading researchers tend to be more concerned with issues of language processing and the educational implications of their research; eye movement reading researchers are more focussed on issues of perception and oculomotor control. Supporting this division is the remarkable fact that much of the variance in eye movement data can be accounted for by low-level visual and lexical features of the text. The relative tractability of modeling oculomotor and lexical level phenomena compared to more high-level comprehension processes has helped foster the development of several successful computational models of eye-movement control in reading (e.g., McDonald, Carpenter, & Shillcock, 2005; Reichle, Rayner, & Pollatsek, 2003; Reilly & Radach, 2006; Engbert, Nuthmann, Richter, & Kliegl, 2005). Nonetheless, the very success of these models may have to some degree served to reinforce the already entrenched cultural divisions that exist.

Correlational studies of top-down effects

There is a substantial body of research that has examined the relationship between text and task complexity and various coarse grain reading rate measures. One example is Haberlandt and Graesser (1985) who used a moving window paradigm to measure word reading times of subjects reading texts of various levels of difficulty (Just & Carpenter, 1980). Results from these and similar studies demonstrate significant interactions between word-level reading times and text complexity measures. Aaronson and Ferres (1984, 1986) observed that the global reading task influences the semantic and syntactic processing of the text. On the one hand, skilled readers use a more structure-oriented "recall strategy" when their reading task is a simple verbatim recall of sentences, and a more meaning-oriented "comprehension strategy" when they have to give true-false responses to statements. These results are in line with the hypothesis of subject-controlled reading strategies at least in skilled readers.

Factors such as age and reading ability have also been shown to interact with comprehension measures and gross reading times. In a study of recall and reading time as a function of discourse structure, Bisanz et al. (1992) found a complex and at times counterintuitive interplay between reading ability and reading times. Poor readers seemed to be able to compensate for poor bottom-up skills by making more effective use of context than good readers. This had the effect of speeding up their reading times for some sentences compared to those of more skilled readers. These findings provide support for the interactive-compensatory model of Stanovich (1980), who argued that poorer readers with under-developed bottom-up skills placed more reliance on contextual cues when these were available to them.

The above-mentioned studies used overall sentence reading time as their major dependent measure. Clearly, there is a need for more fine-grained analysis of reading time such as that afforded by eye movement data. Moreover, the studies so far discussed used correlational analysis as the main tool for inferring causal relationships. More controlled experiments employing moment-to-moment measures of the comprehension process are a logical next step.

Individual variation and top-down effects on eye movements

While the number of studies of reading comprehension using coarse-grained reading times is large, there have been comparatively fewer studies using more fine-grained eye movement data, especially with respect to local word-related processing measures (e.g., Rayner, Chace, Slattery, & Ashby, 2006; Cook & Meyers, 2004; Garrod & Terras, 2000). Ashby, Rayner, and Clifton (2005) in a recent study explored the eye movement data of average and highly skilled readers as determined by the Nelson–Denny test, a timed assessment of comprehension and vocabulary. They focussed in particular on viewing times for words that were varied independently in terms of their predictability and frequency. As expected, less skilled readers were slower, made more regressions and spent more time re-reading regions in the text that had been previously fixated. Interestingly, Ashby et al. found that word frequency effects on viewing duration measures were more pronounced in average readers, due to inflated fixation and gaze durations for low frequency words (Haenggi & Perfetti, 1994; Hyönä & Olson, 1995). It also appeared that average readers often had not achieved lexical access by the time their eyes moved off an unpredictable target word, as indicated by a lack of a word frequency effect on gaze durations. Furthermore, the results indicated that the lexical processes of the highly skilled readers were less influenced by

predictability than those of the average readers, thus lending further support to the interactive-compensatory theory mentioned in the previous section. Taken together, these results suggest substantial variation in the cognitive processes involved in lexical access and post-lexical integration.

Pynte and Kennedy (2006), using a regression study of various viewing time measures (inter-word saccade latency, saccade extent, skipping probability, first fixation, gaze duration, and number of fixation) found significant effects from independent measures not immediately related to words in the current perceptual span (e.g., mean length of adjacent words). They also identified a global word-length effect when comparing English and French reading (French words are, on average, longer than English ones). This appeared to influence the viewing strategies of the different reader populations. For example, while the gaze duration of English readers was sensitive to the lexical frequency of word $n + 1$, this effect was absent in French readers. In fact, the informativeness of the initial letters in word $n + 1$ seemed to perform an analogous role to frequency for readers of French. Pynte and Kennedy (2006) argue that this shows readers' sensitivity to the word length statistics of their language. Because of the greater length variability of French, experienced readers do not rely on being able to identify parafoveally word $n + 1$ and consequently adjust their word-viewing strategy accordingly.

In addition to the relatively static, intrinsic factors of text structure, mean word length, and levels of reading ability, there are extrinsic factors, such as the task demands placed on the reader. So even for the same text and the same level of reading skill and domain knowledge, one can envisage a situation in which the reader has to read the same text for different purposes. However, in comparison to inter-individual differences, this type of variability within readers has attracted far less experimental attention. This is somewhat surprising, as it has long been known that skilled readers are able to adjust their behavior to task demands. For example, Tinker (1958) noted in his seminal review on the state of the art in reading research: "The mature reader… will change his pace (*reflected in eye movements*) to fit the purpose of reading and the nature and the difficulty of the material. He will read rapidly when that is appropriate. In certain other situations he will employ slow analytical reading" (p. 223, our italics). Following this line of reasoning, Heller (1982) also emphasized the potential modulation of reading as a function of the intended level of linguistic processing: "Subjects… can read thoroughly or superficially, and the eye movements vary accordingly" (p. 140). He went one step further in adding the format of reading materials as a potentially important variable. In his own empirical work, Heller was the first to quantify influences of reading intention on oculomotor behavior. In his reading studies, the intention of the reader was implemented by instructing participants to read (1) silently for comprehension, (2) with internal sub-vocalization, (3) quietly mumbling for comprehension, or (4) aloud. In this design, both reading aloud and for comprehension lead to increases in mean fixation durations and fixation frequencies.

Hendriks and Kolk (1997) examined the influence of silent reading for meaning or only sub-vocally pronouncing the words on the eye movement pattern of adults while they were reading a meaningful text. Whereas the pronunciation condition just requires phonological recoding and not mandatory semantic recoding, the silent condition can only be fulfilled using semantic activation. Therefore, the pronunciation task should activate mainly the sub-lexical route and the reading-for-meaning task the lexical route. It was actually found that the pronunciation task was characterized by a larger number of fixations, longer fixation duration and a lower vergence velocity. These observations corresponds to a more sub-lexical reading strategy because smaller processing units are improper for lexical access but increase at the same time the processing speed of phonological information (Hendriks 1996). Strategic control is reflected in the ability to define the word processing area and depending on this, one of the routes is chosen.

Rayner, Sereno, and Raney (1996) also examined the influence of strategic control on reading in general and eye movement behavior in particular. They observed that the word frequency effect changes as a function of the reading task. If the participants were asked to read a text for comprehension, the effect was very strong, whereas it vanished when they had to search for target words in the text. This finding may imply that word frequency effects can be related to the processing depth of the material and may in turn also vary with the reading task. Since a visual search task demands less processing depth, a more superficial reading strategy is sufficient and the frequency effect may no longer be observed. In contrast, reading for comprehension requires a higher processing depth and more elaboration that leads to a more careful reading strategy thus enhancing the frequency effect. Presumably, readers adapt their reading behavior to the respective requirements of the reading task and by doing this can save additional resources for other cognitive processes.

In Greenberg, Inhoff, and Weger (2006), a comparison was made between reading tasks performed with and without the additional requirement of detecting target letters. The goal of this study was to determine if eye movement measures were affected by the additional requirement of having to perform a letter detection task. While the detection task gave rise to longer fixations and fewer word skips, it did not modulate the standard effects of word class (function word vs. content word) and text

predictability. The overall oculomotor pattern suggested that the task does not substantially disrupt lexical processing. This appears, on the surface, to be at variance with Rayner et al. (1996), who did find a modulation of the word frequency effect with their word (as opposed to letter) detection task. Whatever the reason for the difference in results, it serves to highlight the inherent sensitivity of the reading process to the nature of task demands.

Modes of oculomotor influence

Although it appears straightforward to imagine global adjustments in reading behavior, it is not at all clear how such modulations might be implemented by the oculomotor control system. The simplest idea would be that the adjustments are made directly on the level of global parameters such as saccade length and fixation duration. If reading gets more careful, more fixations with longer durations should be made and saccade amplitudes should decrease.

However, this form of modulation appears unlikely given our present state of knowledge about the local fixations patterns associated with word processing in reading. To pick just one example, it is now clear that an increase in the difficulty of reading will usually lead to more refixations, which in terms of average parameters would lower both mean fixation duration saccade amplitude (Inhoff & Radach, 1998). There are now detailed quantitative descriptions of the local metrical properties of oculomotor eye movement behavior. Importantly, it has been shown that landing positions of initial progressive inter-word saccades form a truncated normal distribution with maxima about halfway between the beginning and the center of the target word. This phenomenon has been termed the "preferred viewing position" (Rayner, 1979) and, following McConkie, Kerr, Reddix, and Zola (1988), a substantial body of work has been conducted to describe in detail the metrics of landing site distributions within words. If strategic adjustments to visuomotor control are made in response to top-down factors, a reasonable prediction is to assume that conditions associated with more linguistic processing should lead to a leftward shift of landing positions. This would serve to accommodate the increased need for refixations on the same word, which would profit from a more even distribution of fixation positions across the word.

A similar line of reasoning can be developed with respect to the optimal viewing position (O'Regan, 1990). It has been shown in several single-word recognition tasks that there is an optimal fixation position at or just left of the word center where word processing performance is maximal (O'Regan & Jacobs, 1992). In the case of continuous reading, an analogous form of optimality is assumed to be the frequency of immediately refixating the same word as function of the position where the word is first fixated. This refixation curve is u-shaped with a minimum close to the word center (McConkie, Kerr, Reddix, Zola, & Jacobs, 1989). In a discussion of his strategy and tactics model, O'Regan (1992) makes a clear prediction about the implementation of "careful" versus "superficial" reading in terms of refixations. He argues that readers designate a region around the word center to be acceptable for processing the word with a single fixation. If a saccade lands outside of this region, a refixation is programmed automatically. In a situation of more "risky" reading, this region would be extended, leading to more single fixation cases and a reduction in the steepness of the refixation curve. Conversely, if reading is more careful, the region for optimally placed fixations would decrease resulting in a steeper refixation curve. Although the "strategy and tactics" theory has been quite influential in theoretical discussions, the present work is the first attempt to test one of its most clear-cut predictions.

The present study

The foregoing sections have reviewed approaches to quantifying and manipulating top-down effects in reading that range from ecologically valid reading situations studied within a correlational analysis framework to factorial designs that impose somewhat artificial task demands on the reader. Our goal is to strike a balance between these two extremes: we want as naturalistic a reading task as possible, but we also want the power of a factorial experimental design.

Much of the methodological variation in current reading research using eye movements can be described along two dimensions: format of reading materials and tasks given to ensure reading for comprehension. From an informal survey of reading studies using eye movement methodology selected from the reference sections of Rayner (1998) and several chapters of the recent volume edited by Kennedy, Radach, Heller, & Pynte (2000), the following is clear: text materials for typical studies were either presented as single sentences (about 2/3) or in the form of brief passages (about 1/3), while studies using large amounts of integrated discourse were quite rare. Looking at reading tasks, we found a majority of experiments using comprehension check questions ($n = 48$), followed by the tasks of paraphrasing content ($n = 22$), semantic consistency judgements ($n = 6$), forced choice word recognition tasks ($n = 5$), summarizing passage content ($n = 3$), and finding a headline for a passage ($n = 1$).

Based on these results we decided to maximize the potential benefit of our work by examining the two main dimensions in a factorial design. The first factor "format of reading materials" was implemented by presenting simple declarative sentences either as single items in a quasi-random order or embedded in a coherent corpus of text, representing an essay about life and culture of the Inuit people. The second factor, "Reading Task", was implemented as a variation intended to evoke different levels of linguistic processing along a continuum of careful/deep versus risky/superficial reading. To this end two different item types were utilized: Comprehension questions targeted either simple (location, object) or complex semantic relations within the target sentence. In contrast, in multiple-choice items participants were simply asked to decide which one out of four words had been in the sentence or passage just read. Importantly, the actual reading instruction was held constant across all conditions, so that all results can be attributed to the reading situation a participant got involved in.

Method

Participants

Thirty-six undergraduate students from RWTH Aachen University participated in the experiment for course credit. All had normal or corrected to normal vision and reported no history of reading disabilities.

Materials

The total corpus of text consisted of 7,570 words, representing an essay about life and culture of the Inuit people (adopted from a book by Jeier, 1977). This text corpus was segmented into 108 six-line passages each including two target sentences. Depending on the condition, target sentences were either presented as two out of six lines within a passage or as single items in a fixed random order. Within passages (pages), critical sentences never occupied the first or last line. Table 1 gives examples for these conditions.

Target words were embedded in one-line declarative sentences with a line width maximum of 82 characters, avoiding the two sentence beginning and ending positions. These target words were intended to be representative of the range of common word length and frequencies in German narrative texts. The purpose of varying both factors orthogonally was to avoid the well-known problem of correlation between word length and word frequency in coherent text (e.g., Kliegl, Olson, & Davidson, 1982; Kliegl, Grabner, Rolfs, & Engbert, 2004). This was accomplished by varying word length in three steps including short (4–5 letters), medium (7–8 letters) and long (10–11 letters) words. To vary word frequency, nouns were

Table 1 Examples of the variation of format and reading task

Examples of two single sentences:

 Man fand in der Siedlung neben einem alten *Schuh* auch Tonlampen und Kochgefäße.

 (In the settlement, in addition to an old *shoe*, exciter lamps and cookware were found).

 Eine ganze *Reihe* von Theorien über das Leben vor 2000 Jahren wurden bestätigt.

 (A whole *range* of theories about life 2,000 years ago were verified.)

Example for a passage including the two sentences:

 Dabei handelt es sich um wichtige Ausgrabungen, die in Nukleet gemacht worden sind.

 Man fand in der Siedlung neben einem alten *Schuh* auch Tonlampen und Kochgefäße.

 Sie enthalten neben ihrem zum Teil ästhetischen Wert auch wichtige Informationen.

 Eine ganze *Reihe* von Theorien über das Leben vor 2000 Jahren wurden bestätigt.

 Die Funde waren zum großen Teil aus Schiefer und wurden mit mannigfaltigen

 Steinwerkzeugen bearbeitet. Giddings nannte diese frühe Kultur Norton-Kultur.

Example for multiple-choice verification task:

 Man fand in der Siedlung neben einem alten Schuh auch Tonlampen und Kochgefäße.

 (In the settlement, in addition to an old shoe, exciter lamps and cookware were found).

 Which word was in the sentence/passage?

 Schlitten *Siedlung fand* Eisscholle

 (sledge) (settlement) (found) (ice floe)

Example for comprehension question task:

 Question: Welche Gegenstände wurden gefunden?

 (Which objects were found?)

 Correct answer: Schuh, Tonlampen und Kochgefäße.

 (shoe, exciter lamps and cookware)

Target words are printed in italics

sampled form the CELEX data base (Baayen, Piepenbrock, & van Rijn, 1993) in three frequency ranges, between 0 and less than 1 per million, between 10 and less than 100 per million and equal to or greater than 100 per million.

Computation of statistical word frequency for our sample of target words was supplemented with a rating of word familiarity on a 7-point scale from "unknown" to "ubiquitous" (see Juhasz & Rayner, 2003 and Williams & Morris, 2004, for recent discussions of word familiarity effects). This lead to the exclusion of some potential target words with mismatching frequency and familiarity values from a larger sample of candidate words. Measures were taken to ensure that the variation in word frequency would primarily reflect the degree of processing difficulty on the lexical level. One such measure was to control for morphological complexity according to CELEX (see Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003; Pollatsek, Hyönä, & Bertram, 2000 for evidence of effects of morpheme processing on reading time measures). A further step was to exclude words with extreme values in overall orthographic regularity, as indicated by mean positional bigram frequencies. Table 2 shows mean values of word frequency, familiarity and morphological complexity for the cells of the 3 (word length) × 3 (word frequency) design.

The position of target words within the current line of text was controlled such that they never occupied the first two or last two positions. They were also evenly distributed between the left, central and right part of the line for all cells of the 3 × 3 experimental design (see Vitu, Kapoula, Lancelin, & Lavigne, 2004, for an analysis of line position effects). The word preceding the target was an adjective of 6–10 characters in length. The word-length range for the

adjectives was selected on the basis of analysing a large corpus of reading data (Radach & McConkie, 1998) to maximize the proportion of cases with one fixation on the word before the target.

Apparatus

All text was presented on a 21″ EyeQ CRT monitor at a pixel resolution of 1,024 × 768 in fixed-width courier font. At a viewing distance of 71 cm, each letter subtended approximately 1/3 of a degree of visual angle. Saccadic eye movements were recorded using an SR Research Eyelink video-based eye tracking system, running at 250 Hz. Viewing was binocular but eye movements were recorded only from the right eye.

Design

The present study included two experimental sessions taking place on different days, usually within the same week. Participants were randomly assigned to one of two reading task groups. Within each group (verification vs. comprehension), the order of format (sentences vs. passages) was counterbalanced between two sessions. Experimental sessions consisted of a training block with eight practice trials followed by the experimental stimuli.

Procedure

Participants were tested individually. At the beginning of the experimental session a calibration routine was initiated when the participant pressed the space bar of a keyboard. During calibration, readers in the sentence reading

Table 2 Word frequency (top panel) and word familiarity ratings (center panel) for the cells of the 3 × 3 target word design with each cell including 24 words. For each frequency range, means in both frequency and familiarity for short, medium and long words are not significantly different ($p > .1$). Bottom panel shows the number of morphological components (according to CELEX, Baayen et al. 1993). For each of the word length ranges, means for high, medium and low frequency word are not significantly different ($p > .1$)

	4–5 letter words	7–8 letter words	10–11 letter words
Mean word frequency (per million)			
Low frequency	0.49	0.53	0.51
Medium frequency	6.41	6.26	6.35
High frequency	174.87	129.58	132.77
Word familiarity rating ($n = 20$)			
Low frequency	4.64 (0.67)	4.47 (0.90)	4.54 (0.61)
Medium frequency	3.44 (0.78)	3.62 (0.58)	3.40 (0.59)
High frequency	2.38 (0.54)	2.23 (0.36)	2.33 (0.35)
Number of morphological components			
Low frequency	1.04	1.54	1.67
Medium frequency	1.13	1.58	1.79
High frequency	1.13	1.50	1.79

conditions were asked to fixate a sequence of three fixation markers as they appeared in fixed order at the horizontal midline of the screen. In the paragraph reading condition nine fixation markers were positioned in a two-dimensional array. Calibration was immediately followed by a validation routine that determined the stability and accuracy of the initial measurement. Successful calibration was followed by the presentation of a fixation marker, consisting of a plus sign, shown at the left side of the screen. A second pressing of the space bar replaced the fixation marker with text which remained visible until the sentence or passage was read, which was signaled by the reader with a second space bar pressing. Questions or multiple choice items (see Table 1) were presented after each passage (passage reading) or at unpredictable intervals after sentence presentation (sentence reading). Subjects were asked to respond orally and to press the space bar after responding. This self-paced sentence reading procedure was used throughout the experiment. Each experimental session was usually completed in 45–60 min. Participants in all conditions were asked to read the sentence or paragraph one time silently for comprehension so that they would be able to respond to questions. No further explanations were given so that readers were led to accommodate the present reading situation on the basis of the type of questions and formats they encountered during the following training trials.

Data analysis

A word was considered fixated when a fixation fell on one of its constituent letters or the blank space preceding it. Fixation durations of less than 70 ms and of more than 1,000 ms were removed from analyses. In analyses of target words, fixations with durations above 3SD of the cell mean were also eliminated. Excluded were also trials in which the first fixation on the target word was not preceded by a progressive saccade with a length of 20 characters or less. Together with blinks or track losses, these restrictions resulted in the rejection of about 5.2% of all observations. In the analyses of these data, initial fixation durations were defined as the duration of the first fixation on the word, irrespective of whether the target was subsequently refixated. Gaze durations included the time spent viewing the target word during first-pass reading, including the time spent refixating it, but excluding saccade durations. Total viewing time was the summed duration of all fixations made on the word. Landing positions of incoming initial saccades were rounded to a tenth of a character before averaging; with the space preceding the target word coded as position 0.1–0.9 (see Inhoff & Radach, 1998, Inhoff & Weger, 2003 and Rayner, 1998, for discussions of oculomotor measures). Depending on the conditions in question, eye movement parameters were subjected to repeated

measures or mixed analyses of variance using subject (F1) variability in the computation of error terms.

Results

Comprehension scores

Word verification scores (based on 54 multiple-choice items) in the verification task were near perfect for both sentence and passage reading, although numerically there was a slight advantage for sentence reading ($M = 98.8\%$, SD $= 1.6$) as compared to passage reading ($M = 96.1\%$, SD $= 2.7$), $t(17) = 3.71$, $p = .002$. The near ceiling performance is due to the fact that two out of four alternatives were correct, reflecting an attempt to design this condition to be minimally demanding. The format difference was far more pronounced when participants responded to questions in the comprehension task ($M = 76.54\%$, SD $= 10.54$ vs. $M = 49.85\%$, SD $= 11.68$, respectively), $t(17) = 9.60$, $p < .001$. Here, only a perfect answer resulted in a full point, while answers that came close were counted as half. The difference in performance between format conditions is not surprising, given the fact that during passage reading more information needs to be processed and held in memory before a response can be made. Note that the scores from the verification and comprehension tasks (which were varied between participants) cannot be directly compared due to the different item formats and scoring procedures.

Participants also completed a German reading comprehension test, assessing memory for detail, discourse level comprehension and pronomial inferences (LVT, Hacker, Handrick & Veres, 2002). The results revealed virtually the same text comprehension performance for both groups of subjects that were randomly assigned to the verification and comprehension groups, $t(34) < 1$. According to normative data, the overall text comprehension score ($M = 16.5$) corresponds to a level of "very good text comprehension".

Top-down effects on word viewing time measures

Table 3 reports word-based viewing time and fixation frequency measures for all four cells in the design, using observations for word length 4–9. Statistical analyses were conducted using mixed two-way ANOVAs based on subject means with the within-subject variable format (sentence reading vs. passage reading) and the between-subject variable task (verification vs. comprehension) as independent variables. The pattern of results reported below was also confirmed in separate ANOVAs for all word length 4–9. In addition, analyses were repeated

Table 3 Viewing time and fixation frequency measures for the four conditions of the format x task design. Standard errors are shown in parentheses. Values are based on pooled observations for word length 4–9 ($n = 36$)

	Sentence verification	Passage verification	Sentence comprehension	Passage comprehension
First fixation duration	210 (66)	207 (67)	228 (77)	225 (76)
Single fixation duration	212 (65)	210 (66)	232 (76)	228 (76)
Gaze duration	247 (109)	239 (102)	281 (130)	263 (120)
Total viewing time	283 (145)	307 (173)	371 (231)	393 (245)
Fixations in first pass	1.31 (0.58)	1.27 (0.56)	1.38 (0.63)	1.29 (0.57)
Number of passes	1.17 (0.41)	1.30 (0.58)	1.35 (0.67)	1.54 (0.93)
Total number of fixations	1.40 (0.66)	1.51 (0.80)	1.64 (0.89)	1.80 (1.39)

without target words, again leading to the same pattern of results.

Fixation durations were significantly longer for the comprehension task as compared to the verification task, $F(1,34) = 7.21$, $p = .011$, $\eta_p^2 = 0.18$. There was no significant effect of format, $F(1,34) = 1.51$, $p > .10$, nor a significant interaction of task and format, $F < 1$. Similar results emerged for single fixation durations, which were significantly longer for the comprehension task as compared to the verification task, $F(1,34) = 7.18$, $p = .011$, $\eta_p^2 = 0.18$. There was no significant effect of format, $F(1,34) = 2.52$, $p > .10$, nor a significant interaction of task and format, $F < 1$.

Gaze durations were significantly longer for the comprehension task as compared to the verification task, $F(1,34) = 8.09$, $p = .007$, $\eta_p^2 = 0.19$. Unexpectedly, they were also significantly longer for sentence reading as compared to passage reading, $F(1,34) = 11.67$, $p = .002$, $\eta_p^2 = 0.26$. Finally, the effect of format was marginally larger in the comprehension condition, $F(1,34) = 3.09$, $p = .088$, $\eta_p^2 = 0.08$. Total viewing times followed a different pattern, being significantly longer for the comprehension task as compared to the verification task, $F(1,34) = 12.05$ $p = .001$ $\eta_p^2 = 0.26$. Furthermore, in line with our expectation they were now also significantly longer for passage reading as compared to sentence reading, $F(1,34) = 6.92$, $p = .013$, $\eta_p^2 = 0.17$. Note that this effect is a remarkable reversal from the effect for the gaze durations, which are part and parcel of total viewing times. There was no significant interaction, $F < 1$.

The data sample used in Table 3 and the statistical analyses reported above also served as the base for Fig. 1,

Fig. 1 Decomposition of word-based viewing times into time spent for initial fixation durations, refixations on the same word and re-reading during later passes. Four separate panels show data for two levels of format (sentence vs. passage) and task (verification vs. comprehension)
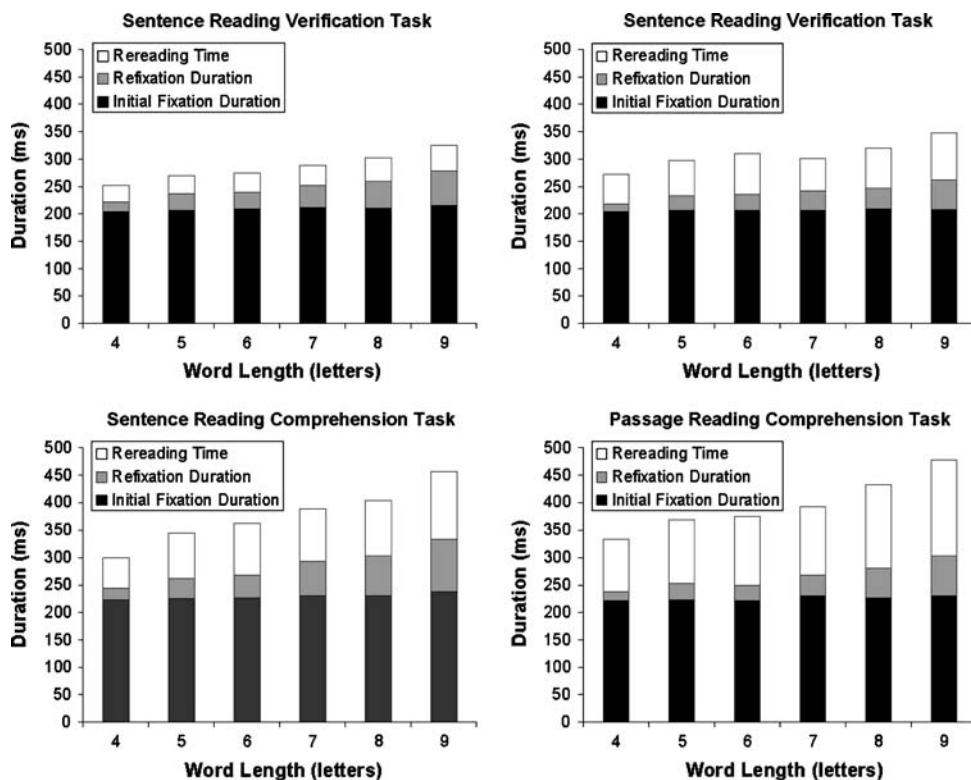
presenting a decomposition of word-viewing time in terms of initial fixation duration, time spent refixating the same word and time used for re-reading the same word during later passes. Note that fixation duration and refixation time are equivalent to gaze duration and that the sum of fixation duration, refixation time and re-reading time is identical to total reading time. The figure provides an effective way to visualize the proportion of time spent for the different components of word-viewing time over the range of word length studied, making the rather dramatic reversal of the reading format effect quite apparent (Figs. 2, 3).

Top-down effects on spatial eye movement measures

Statistical analyses were also computed for spatial eye movement parameters, again using mixed two-way ANOVAs with the within-subject variable format (sentence reading vs. passage reading) and the between-subject variable task (verification vs. comprehension) as independent variables. Results for initial saccade landing positions indicated no significant difference for task, $F(1,34) = 1.03$, $p > .10$. However, incoming saccades landed further into the word for the passage reading condition as compared to sentence reading, $F(1,34) = 20.47$, $p < .001$, $\eta_p^2 = 0.38$. There was no significant interaction, $F < 1$. Similarly, the

amplitude of incoming progressive saccades were not affected by task, $F < 1$, but were larger in the passage reading condition as compared to sentence reading, $F(1,34) = 56.10$, $p < .001$, $\eta_p^2 = 0.62$. There was no significant interaction, $F(1,34) = 2.12$ $p > .10$. Looking at the starting positions of these saccades, launch distance (relative to the beginning of the target word) was not affected by task, $F < 1$, but was located further away from the word beginning for passage reading as compared to sentence reading, $F(1,34) = 14.22$, $p = .001$, $\eta_p^2 = 0.30$. There was no significant interaction, $F < 1$. These results further illustrate the remarkable difference between the reading of the same sentence as a single item vs. in the context if a passage: In sentence reading saccades come from locations that are closer to the target word. Everything else equal, this should lead to a rightward shift of saccade landing positions (e.g., Radach & McConkie, 1998). Instead, these landing positions are located significantly further to the left.

Landing site distributions

A more detailed quantitative analysis of landing sites was performed, again using the data set from Table 3 (see
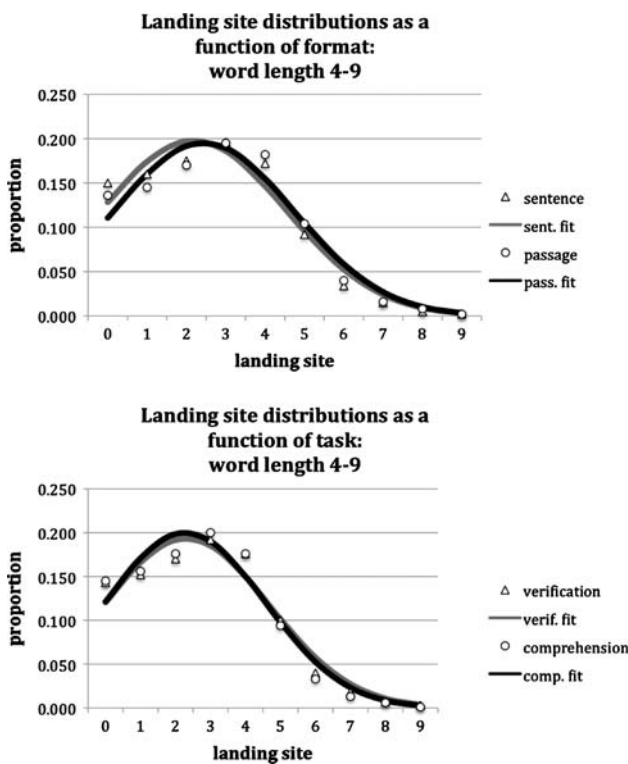

Fig. 2 Plots of location-based landing site distributions and Gaussian curve fits for words of length 4 through 9. Data are broken down by two levels of format (sentence vs. passage, *upper panel*) and task (verification vs. comprehension, *lower panel*)
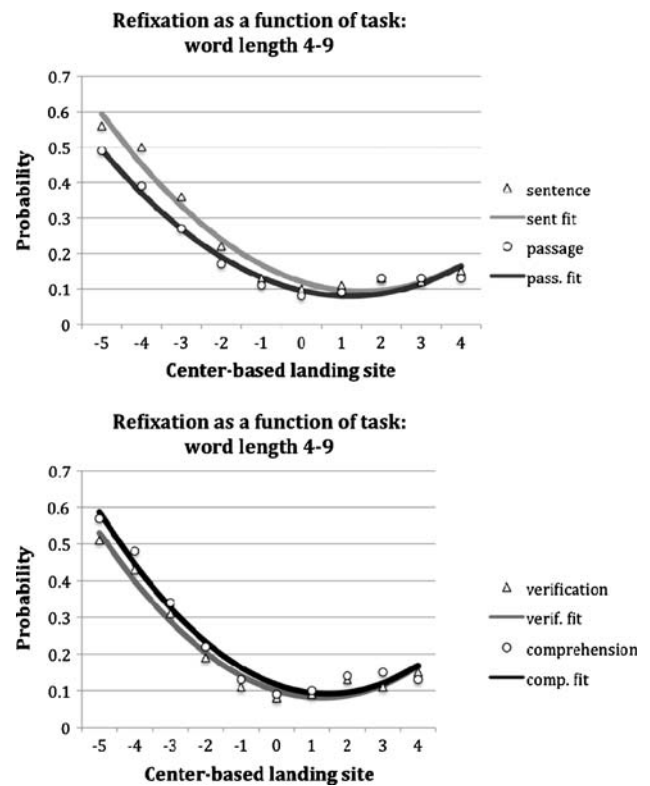

Fig. 3 Plots of location-based refixation probabilities and polynomial curve fits for words of length 4 through 9. The refixation data is normalized for word length, with zero indicating the word center. Data are broken down by two levels of format (sentence vs. passage, *upper panel*) and task (verification vs. comprehension, *lower panel*)

Table 4 Parameters from normal curves fitted to landing site distributions associated with the 2×2 experimental conditions

$y = ae^{-\frac{(b-x)^2}{2c^2}}$	Format		Task	
	Sentence	Passage	Verification	Comprehension
a	0.198	0.195	0.193	0.199
b	2.180	2.442	2.306	2.275
c	2.344	2.292	2.373	2.276
Corr.	0.99	0.98	0.98	0.98
SE	0.028	0.034	0.018	0.019

Table 5 Parameters for the refixation curves associated with the 2×2 experimental conditions

$y = a + bx + cx^2$	Format		Task	
	Sentence	Passage	Verification	Comprehension
a	0.120	0.095	0.010	0.117
b	−0.036	−0.026	−0.029	−0.035
c	0.012	0.011	0.011	0.012
Corr.	0.99	0.99	0.99	0.99
SE	0.034	0.026	0.028	0.034

Fig. 2). Following the seminal work of McConkie et al. (1988), the proportions of fixations landing on each letter in a given word were fitted with a Gaussian curve of the form: $y = ae^{-\frac{(b-x)^2}{2c^2}}$. In this analysis, x is the fixation location on the word and y is the proportion of fixations on that location. The parameters a, b, and c are give in Table 4.
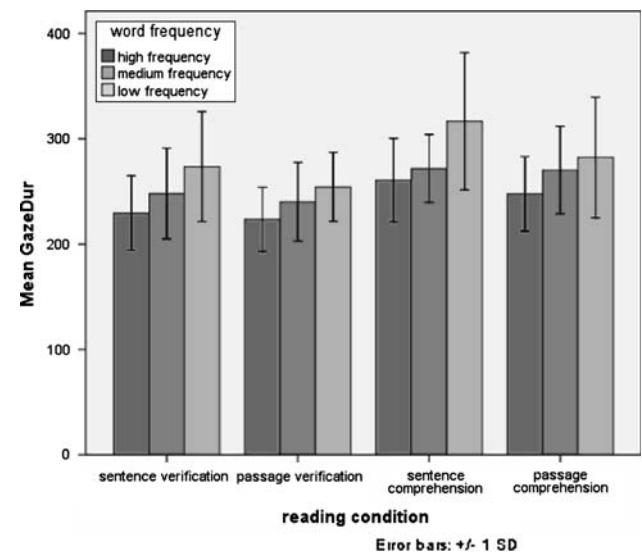
Assuming McConkie et al. are correct in proposing an underlying Gaussian mechanism for word targeting, the variation in curve fit parameters can give a clue to any modulation occurring of the mechanism in response to either reading task demands or text format. The results of fitting Gaussians to the landing site distributions showed most variation in the central tendency parameter (b), with the largest variation for the format condition with the mean landing position for passages occurring further into the word than that for sentences. In sum, a variation in the format (sentence vs. passage) serves to shift the distribution horizontally, without changing other parameters.

Refixations

Looking at the frequency of immediately refixating a word, there was no significant task effect, $F(1,34) = 1.86$, $p > .10$. However, refixations were more frequent during sentence reading as compared to passage reading, $F(1,34) = 20.26$, $p < .001$, $\eta_p^2 = 0.37$. There was no significant interaction, $F(1,34) = 4.01$ $p > .05$. Again, a more detailed analysis was performed using a curve fitting technique. Similar to the approach of McConkie et al. (1989), the u-shaped refixation curves for the 2×2 conditions of text format and task difficulty were fitted with quadratic polynomials of the form: $y = a + bx + cx^2$ where x is the first fixation location in the word and y is the probability of a refixation being launched from there (see Fig. 3). Table 5 shows the set of fitted parameters and measures of their goodness of fit. As can be seen, the largest parameter difference between the curves for each treatment level is for the vertical offset parameter (a). This difference is substantially more pronounced for the task than the format factor. It is apparent that the consequence

Fig. 4 Gaze duration as a function of word frequency. Data are presented for two levels of format (sentence vs. passage) and task (verification vs. comprehension)

of the format condition is a vertical shift of the distribution with virtually no effect on its horizontal position and little change in the steepness of the curve. Note that a variation in steepness was the specific prediction of the strategy and tactics theory of eye movement control (O'Regan, 1992).

Top-down modulation of word frequency effects

Figure 4 presents gaze duration as a function of word frequency, with results provided separately for the two levels of format (sentence vs. passage) and task (verification vs. comprehension). It appears from the figure that the word frequency effects may be more pronounced in sentence reading as opposed to the passage condition. Indeed Table 6 reports the difference between the low and high frequency conditions for initial fixation duration, single fixation duration, gaze duration and total viewing time. It is quite apparent that both task and format appear to modulate the influence of word frequency.

Table 6 Size of frequency effects as expressed in the difference between high and low frequency conditions. Data are presented for two levels of format (sentence vs. passage) and task (verification vs. comprehension)

	Sentence verification	Passage verification	Sentence comprehension	Passage comprehension
Difference in first fixation duration	17 (22)	25 (22)	21 (22)	17 (24)
Difference in single fixation duration	26 (25)	31 (25)	35 (17)	24 (25)
Difference in gaze duration	44 (29)	31 (22)	56 (37)	35 (43)
Difference in total viewing time	59 (36)	40 (55)	86 (65)	52 (83)

Statistical analyses were conducted using mixed $3 \times 2 \times 2$ ANOVAs based on subject means for target words with the within-subject variables frequency (low, medium and high) and format (sentence reading vs. passage reading) and the between-subject variable task (verification vs. comprehension) as independent variables. As could be expected, fixation durations were strongly affected by word frequency ($F(1,34) = 29.61$, $p < .001$, $\eta_p^2 = 1.0$). In addition there was a significant task effect, ($F(1,34) = 4.58$, $p < .05$, $\eta_p^2 = 0.55$) while the effect of format and all interactions were not significant. Similarly, results for single fixations included a large frequency effect $F(1,34) = 54.46$ $p < .001$, $\eta_p^2 = 1.0$, an effect of task $F(1,34) = 5.27$, $p < .05$, $\eta_p^2 = 0.61$ but no effect of format $F(1,34) = 2.67$ $p > .10$, $\eta_p^2 = 0.36$. The three-way interaction between frequency, format and task did not approach significance $F(1,34) = 1.53$, $p > .10$, $\eta_p^2 = 0.32$.

Most interesting are the results for gaze durations. Again, the influence of word frequency was very large $F(2,68) = 49.29$, $p < .001$, $\eta_p^2 = 1.0$ and there were now also significant effects of both format $F(1,34) = 10.57$, $p < .05$, $\eta_p^2 = 0.89$ and task $F(1,34) = 5.85$ $p < .05$, $\eta_p^2 = 0.65$. Importantly, there is also a significant interaction between word frequency and format ($F(1,34) = 7.60$, $p < .05$, $\eta_p^2 = 0.94$), confirming that indeed the frequency effect was modulated by whether a target words first-pass reading occurs in a single sentence or a passage of text. Looking at total viewing durations, in addition to a strong frequency effect $F(2,68) = 35.89$, $p < .001$, $\eta_p^2 = 1.0$, there was also a task effect ($F(1,34) = 12.75$, $p < .001$, $\eta_p^2 = 0.93$, an effect of format $F(1,34) = 7.02$, $p < .05$, $\eta_p^2 = 0.73$, and again an interaction between frequency and format ($F(2,68) = 5.97$, $p < .05$, $\eta_p^2 = 0.87$.

General discussion

The present study is the first to examine the effects of two top-down factors on reading using virtually identical target sentences and words and also identical reading instructions. The present work quantified the extent to which dynamic reading behavior as expressed in eye movements is influenced by the type of task induced via application of verification items, presumably triggering a more superficial reading style versus comprehension questions, presumably inducing a more deliberate reading with deeper cognitive processing. Moreover, taking together the various pieces of evidence presented above, it appears that there is a marked difference in how readers approach the tasks of reading when confronted with single sentences versus passages as part of integrated nonfictional text.

Effects on temporal measures

When readers had to answer detailed questions on what they had read (comprehension), as opposed to responding to multiple-choice questions (validation), word-viewing times as measured by a range of indices were significantly shorter. This finding is unsurprising, since the more demanding comprehension task required more careful reading with an overall increase in the number word re-fixations. The effect of text format (sentence vs. passage) on viewing times was, however, a little more complicated. Overall, the total viewing time for words was significantly greater for passages as opposed to single line sentences. However, readers' first-pass viewing times (initial fixation and first-pass gaze duration) were shorter for passages compared to sentences. This suggests that readers of passages perform a quick first pass over the text followed by a re-reading. What is particularly striking about this finding is that a global factor such as reading format has such a direct and significant effect on local word-viewing time parameters. This provides a timely reminder that readers dynamically adapt in unexpected ways to what might on the surface appear to be relatively innocuous changes to reading conditions.

Landing sites

Landing site distributions showed differences primarily for format conditions. In this case, the initial landing site for words in passages was shifted further into the word in comparison to the sentence condition. One possible explanation for this is that in first-pass passage reading

(which tended to be faster with fewer refixations), readers maximized information acquisition by landing nearer to the word center. In contrast, if a reader has the global strategy of more careful reading, she will make more refixations. In this case, a more effective targeting strategy is to land nearer to the word beginning for the first fixation in order to maximize information acquisition from a subsequent refixation.

In the case of task difficulty, there was a slight tendency for the mean landing site in the easier, validation task to be further into the target word. Again, this is consistent with a strategy to make more refixations in the case of a complex reading task, where a leftward shift in landing site optimizes information pick-up for refixations. The findings on saccade landing positions can also be seen as an interesting analog to studies showing that there are small but significant effects of local processing on saccade targeting (e.g., Hyönä, 1995; Radach, Heller, & Inhoff 2004; White & Liversedge, 2006).

Refixations

As discussed above, the analysis of refixation curves showed some subtle variations across the different experimental manipulation. In the case of the task condition, the more demanding task of answering comprehension questions caused a consistent elevation in the tendency to refixate across all fixation locations. In the text format condition, there was a combination of a small general increase in the probability to refixate combined with a slightly increased tendency for refixations to be triggered nearer the beginning of the word in the simpler, verification condition. This is fully consistent with the strategic adaptions suggested above, assuming that the elevation of the curve is a function of cognitive processing demands (McConkie et al. 1989). It is interesting to note that the minimal increase in the steepness of the refixation curve in the case of more "careful" reading may or may not be seen as support for the tactics part of O'Regan's strategy and tactics theory. In any case it should be acknowledged that this hypothesis had a stimulating influence on the present work.

Frequency effects

There was a marked reduction in the size of the frequency effect when target words were read as part of an integrated passage. There was also a reduction in the size of the frequency effect as a function of task difficulty: the validation task gave rise to smaller effect sizes than the comprehension task. Overall, these effects were more apparent in the gaze duration and total reading time measures than for initial or single fixation measures.

While the differential frequency effects for task can be accounted for by a deeper level of processing in the comprehension condition, the decrease in the effect for paragraph reading may be due to the temporal demands for higher order post-lexical processing. The aggregated viewing times for this form of reading would consequently be less sensitive to lexical access effects, since these processes had already been carried out during the first pass.

Modeling implications

What implications do these results have for current models of eye-movement control in reading? There are now several computational models that capture key components of the dynamics of normal reading like the E-Z reader model by Reichle, Rayner and Pollatsek (2003), the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005), the split fovea model (McDonald, Carpenter, & Schillcock, 2005), the Mr. Chips model (Legge et al. 2002) and the Glenmore model (Reilly & Radach, 2006). All of these models are successful in that they simulate key aspects of "normal reading". However, it is left unspecified where exactly on the continuum between skimming and very carefully scrutinizing the scope of "normal reading" lies and how variations along this dimension may affect the dynamics of the models. This present generation of models can neither readily accommodate task difficulty effects, nor effects arising from text format as found in the present study. The format effect in particular poses a challenge to all current models, since they would have to be able to account for the more search-like pattern of movements present in the paragraph reading condition. This would also mean taking account of the two-dimensional spatial distribution of information in a text as well as some features of its discourse level structure. These aspects of reading have been largely ignored to-date by the modeling community. The task-level effects are perhaps more easily accommodated within current models, since they seem to involve the adjustment of a more global reading parameter that could be implemented as a threshold or a perceptual span adjustment. For example, in the Glenmore model, the saccade triggering mechanism is in the form of a global activity measure that crosses a threshold. Raising this threshold can be used to simulate more careful reading. However, accounting for the extensive re-reading in the paragraph format condition is still beyond the scope of Glenmore and, indeed, any other model.

Summary

In our selective review of literature at the beginning of this paper, we encountered a striking pattern of complex and sometimes paradoxical findings: word-viewing behavior is

less influenced by context for highly skilled readers than for average readers (Ashby et al. 2005); a word search task eliminates lexical effects (Rayner et al. 1996), whereas a letter search task appears not to do so (Greenberg et al. 2006); reading to pronounce (i.e., less lexical reading) increases the number and duration of fixations, even for skilled readers (Hendriks and Kolk 1997); and frequency of word $n + 1$ affects viewing times of word n for readers of English but not of French (Pynte & Kennedy, 2006). What these results tell us, and this is something researchers from the comprehension research "culture" have known for some time, is that the notion of "normal" reading is a convenient construct, if not an illusion (albeit a useful one). When we study reading in an ecologically plausible context, a complex array of factors impinge upon the low-level processes of information acquisition and visuomotor control. The present work has served to identify two major sources of such influence. Presenting a sentence in the context of a passage quite dramatically modulates many aspects of the reading process, all within the boundaries of natural reading (see Kliegl, Nuthmann, & Engbert, 2006; Rayner, Pollatsek, Drieghe, Slattery & Reichle, 2007, for a controversial discussion). It is well possible that the format effects described in the present paper account for some of the differences between research traditions that lean towards more cognitive versus visuomotor theories of reading, as the latter tended to rely on corpus work with smaller frequency effects (O'Regan et al. 1994; Rayner et al. 1996).

Thirteen years ago, Churchland, Ramachandran and Sejnowski (1995) argued in a seminal paper entitled "A Critique of Pure Vision" that vision researchers needed to radically alter their theoretical view of how the visual perceptual system functioned. In their view, vision did not comprise a set of relatively independent, mutually agnostic and more or less sequential processes. They made the case that vision is a dynamically interactive system where high-level factors routinely and directly affect low-level processes. We have no reason to believe that the processes involved in reading behave any differently.

Finally, we mentioned earlier that the success of computational models might have led to something of a consolidation of the division between the comprehension-based and eye movement-based research cultures. There is an alternative and more optimistic outlook. Our modeling and data acquisition technologies are now of a level of sophistication that it is more feasible than ever to design computer models better able to capture faithfully the striking dynamics of the whole reading process. We believe that the results described in this paper will contribute to creating the necessary empirical push for such an undertaking.

References

Aaronson, D., & Ferres, S. (1984). Reading strategies for children and adults: some empirical evidence. *Journal of Verbal Learning and Verbal Behavior, 23*, 89–220.

Aaronson, D., & Ferres, S. (1986). Reading strategies for children and adults: A quantitative model. *Psychological Review, 93*, 89–112.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*, 285–311.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology, 58*, 1065–1086

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The Celex Lexical database. (CD-ROM). Linguistic Data Consortium. University of Pennsylvania, Philadelphia, PA.

Bisanz, G. L., Das, J. P., Varnhagen, C. K., & Henderson, H. R. (1992). Structural components of reading time and recall for sentences in narratives: exploring changes with age and reading ability. *Journal of Educational Psychology, 84*, 103–114.

Butcher, K. R., & Kintsch, W. (2003). Text comprehension and discourse processing. In F. Healy & R. W. Proctor (Eds.), *Handbook of psychology: Experimental psychology* (Vol. 4, pp. 575–595). New York: Wiley.

Churchland, P., Ramachandran, V., & Sejnowski, T. (1995). A critique of pure vision. In C. Koch, & J. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 22–60). Cambridge, MA: MIT Press.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel (Ed.), *Eye movements: A window on mind and brain* (pp. 341–372). Amsterdam: Elsevier.

Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory and Language, 50*, 268–288.

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

Garrod, S. C., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language, 42*, 526–544.

Greenberg, S. N., Inhoff, A. W., & Weger, U. W. (2006). The impact of letter detection on eye movement patterns during reading: Reconsidering lexical analysis in connected text as a function of task. *Quarterly Journal of Experimental Psychology, 59*, 987–995.

Haberlandt, K., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General, 114*, 357–374.

Hacker, W., Handrick, S. & Veres, T. (2002). Leseverständnistest (Test of Reading Comprehension). Technical report and test manual. Institute of Psychology. Technical University of Dresden.

Haenggi, D., & Perfetti, C. A. (1994). Processing components of college level reading comprehension. *Discourse Processes, 17*, 83–104.

Heller, D. (1982). Eye movements in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 139–154). Amsterdam: North Holland.

Hendriks, A. (1996). Vergence eye movements during fixations in reading. *Acta Psychologica, 92*, 131–151.

Hendriks, A. W., & Kolk, H. H. J. (1997). Strategic control in developmental dyslexia. *Cognitive Neuropsychology, 14*, 321–366.

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 68–81.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1–11.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29–54). Oxford: Elsevier.

Inhoff, A., & Weger, U. (2003). Advancing the methodological middle ground. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 335–344). Oxford: Elsevier.

Jeier, T. (1977). *Die Eskimos: Geschichte und Schicksal der Jäger im hohen Norden. (The Eskimos: History and destiny of the hunters in the far north)*. Düsseldorf: Econ Verlag.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1312–1318.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words. evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology, 94*, 223–244.

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

Kennedy, A., Radach, R., Heller, D. & Pynte, J. (Eds.) (2000). *Reading as a perceptual process.* Oxford: Elsevier.

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

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 136*, 520–529.

Kliegl, R., Olsen, R. K., & Davidson, B. J. (1982). Regression analyses as a tool for studying reading processes. *Memory & Cognition, 10*, 287–296.

Legge, G. E., Hoven, T. A., Mansfeld, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Research, 42*, 2219–2234.

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

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28*, 1107–1118.

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

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Reviews of oculomotor research, vol. 4: Eye movements and their role in visual and cognitive processes* (pp. 395–453). Amsterdam: Elsevier.

O'Regan, J. K. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 333–354). New York: Springer.

O'Regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effects in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 185–197.

O'Regan, J. K., Vitu, F., Radach, R., & Kerr, P. (1994). Effects of local processing and oculomotor factors in eye movement guidance in reading. In J. Ygge & G. Lennerstrand (Eds.), *Eye movements in reading*. New York: Pergamon.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 820–833.

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research, 46*, 3786–3801.

Radach, R., Heller, D., & Inhoff, A. W. (2004). Does orthographic information modulate saccade amplitudes in reading? *European Journal of Cognitive Psychology, 16*, 27–51.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for the future. *European Journal of Cognitive Psychology, 16*, 3–26.

Radach, R., & McConkie, G. W. (1998). Determinants of fixation positions in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 77–100). Oxford: Elsevier.

Radach, R., Reilly, R., & Inhoff, A. W. (2007). Models of oculomotor control in reading: towards a theoretical foundation of current debates. In R. van Gompel, M. Fischer, W. Murray & R. Hill (Eds.), *Eye movements: A window on mind and brain*. Oxford: Elsevier.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241–255.

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

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1188–1200.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Journal of Cognitive Systems Research, 7*, 34–55.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–526.

Stanovich, K. E. (1980). Towards an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32–71.

Stanovich, K. (2003). Understanding the styles of science in the study of reading. *Scientific Studies of Reading, 7*, 105–126.

Tinker, M. (1958). Recent studies of eye movements in reading. *Psychological Bulletin, 55*, 215–231.

Van Gompel, R., Fischer, M., Murray, W., & Hill, R. (Eds.) (2007). *Eye movements: A window on mind and brain*. Oxford: Elsevier.

Vitu, F., Kapoula, Z., Lancelin, D., & Lavigne, F. (2004). Eye movements in reading isolated words: Evidence for strong biases towards the center of the screen. *Vision Research, 44*, 321–338.

White, S. J., & Liversedge, S. P. (2006). Linguistic and nonlinguistic influences on the eyes' landing positions during reading. *Quarterly Journal of Experimental Psychology, 59*, 760–782.

Williams, R. S., & Morris, R. K. (2004). Eye movements, word familiarity and vocabulary acquisition. *European Journal of Cognitive Psychology, 16*, 312–339.