Parafoveal Processing in Reading: On the Interplay between Verbal and Visual Factors

Denis Drieghe

Promotor: Prof. Dr. Wim Fias
Co-Promotor: Prof. Dr. Marc Brysbaert

Proefschrift ingediend tot het behalen van de academische graad van Doctor in de Psychologische Wetenschappen

2005
Parafoveal Processing in Reading: On the Interplay between Verbal and Visual Factors

Denis Drieghe

Promotor: Prof. Dr. Wim Fias
Co-Promotor: Prof. Dr. Marc Brysbaert

Proefschrift ingediend tot het behalen van de academische graad van Doctor in de Psychologische Wetenschappen

2005
“I’ve seen things you people wouldn’t believe...”
Roy Batty in Ridley Scott’s Blade Runner
CONTENTS

ACKNOWLEDGMENTS 9

CHAPTER 1 PARAFOVEAL PROCESSING IN READING: AN OVERVIEW 13

THE NAME OF THE GAME: PARAFOVEAL PROCESSING IN READING 17
WHAT'S AT STAKE? THE BIG QUESTIONS IN THE FIELD 21
  Is word processing in reading serial or parallel? 21
  Low-level oculomotor factors versus higher-level cognitive processes 22
  How much information is extracted from the right of the fixation? 24
THE PLAYERS: MODELS OF EYE MOVEMENT CONTROL IN READING 24
  The grandmaster: the E-Z Reader model 25
  Challenges for the E-Z Reader model 27
  The competitors 29
PARAFOVEAL PROCESSING IN READING: ON THE INTERPLAY BETWEEN VERBAL AND VISUAL FACTORS 31
  The board is being set: Explaining the title 31
  The pieces are moving: Research objectives 32
REFERENCES 34

CHAPTER 2 PARAFOVEAL-ON-FOVEAL EFFECTS ON EYE MOVEMENTS IN TEXT READING: DOES AN EXTRA SPACE MAKE A DIFFERENCE? 43

INTRODUCTION 44
EXPERIMENT 1 49
  METHOD 49
  RESULTS 51
  DISCUSSION 54
EXPERIMENT 2 55
  METHOD 56
  RESULTS 58
TOWARDS A NEW MODEL OF WORD SKIPPING IN READING 188

AVENUES FOR FUTURE RESEARCH 195

The inflated fixation duration prior to skipping 195

The ability to cancel a planned skip 195

Word length and launch site as determining factors in skipping behavior 196

Individual differences 196

Parafoveal-on-foveal effects 197

REFERENCES 198
ACKNOWLEDGMENTS

This doctoral thesis was accomplished while I was working as a Research Assistant of the Fund for Scientific Research – Flanders (F.W.O. – Vlaanderen) at the department of Experimental Psychology of Ghent University, Ghent, Belgium, and during a visit at the eye tracking lab of the University of Massachusetts in Amherst, Massachusetts, USA. I thank the aforementioned institutions for their support.

Looking back at the years it took me to prepare and finalize this thesis, one critical moment comes to mind that just has to be mentioned. It was somewhere at the end of my first year that Marc took me aside and told me he was leaving for London. It didn’t take me long to realize that this departure condemned me, being the sole member of the Ghent Eye Movement Team, to preparing a PhD on my own with no one to directly talk to about the details of my research, and no one to share the burden with of programming and maintaining the Eyelink system. I spent quite some time doubting the feasibility of such a venture, indeed, even reflecting the possibility of quitting the entire project. Finally, I decided to take my chances, go for it, and never look back. Two reasons pulled me through at that time. The first reason was a passion for my research. Secondly, I was coming to realize that Marc Brysbaert as a promotor, even though reduced to a mailbox, was still about as good a promotor as you can get. Therefore, the first person I would like to thank is Prof. dr. Marc Brysbaert, for teaching me to rely on my own capabilities, for being a ruthless critic whenever I needed one, and most importantly for being one of my most avid supporters. And if I ever would write an encyclopedia on great scientists, he would be in it, and he would be described as “The man who replies faster to his e-mails than his shadow”. I also would like to thank Prof. dr. Wim Fias for taking the rather ungrateful job as a pro forma promotor. Thanks to his signatures (and he
I really hate paperwork! I was granted the freedom of being able to do my research while answering to nobody but my mailbox. And that was just what I needed!

Many thanks go to Prof. dr. Dominiek Sandra and to Prof. dr. Ralph Radach who were both in my PhD guidance commission. I thank Dominiek for being, besides a good psycholinguist and a trusted referee for projects, also one of the nicest guys I’ve ever met. I thank Ralph for traveling all the way from Aachen to Ghent (even if it only took him about 40 minutes in his new Saab) just to listen to me, and for many helpful comments.

As mentioned earlier, part of the research that is presented here, was carried out in Amherst, Massachusetts. For this opportunity I would like to thank, in the first place, Prof. dr. Keith Rayner. At first, I think I was a bit impressed by the man’s astounding career. But, after playing an 8 hour Trivial Pursuit marathon with him on the way back from Niagara Falls to Amherst, I’ve learned to know him as a great guy to be around with and to talk about science, McDonalds (which he hates even more than I do) and the peculiar habit of Belgians to usually stay within a radius of 4.5 kilometers of the place where they were born. Many thanks to Sandy as well, I look back with great pleasure at our conversations concerning art and politics. I won’t name all the people of the Umass eye tracking lab, but I would like to thank them all for my pleasant stay there and mention a few: Carrick, Kiel (for his incredible story telling skills during our car trip to Washington), Chuck, Barb (the goddess of lab managers), Becca, Tim (who sucks big time at playing Diamond Mine on the computer), Slash, Jane, Mako, Adrian and Michael. I am really looking forward to returning to Amherst in the near future.

There are three more persons I have to mention who have added a certain flavor to my PhD whilst being abroad. They are my smoking buddies at conferences. Standing outside a conference building, smoking a cigarette with fellow researchers, creates a certain bond and that bond is particularly
powerful with the following three persons: Françoise Vitu from France, Raymond Bertram from Finland, and Steven Frisson from New York/UFSIA. All three of them I consider to be special friends.

But most of the time I was of course in Ghent at the Department of Experimental Psychology. Special thanks to all the psycholinguists at the Department, especially Timothy (my partner-in-crime for many of the experiments reported here) and Wouter (who was a great help while doing the lay-out of this thesis). And then there is a group of people that I have to thank for which I can find no better designation than the psychologica alcoholica team. These are the people that make sure that I don’t work too much and that almost every working day ends in a pub discussing science, gossip and movies. Thank you, Rob, Wim, Els, Michael, Free, and Trappist. And I have to mention especially Michael, Free and Trappist, for showing me how colleagues can become close friends. Thanks also to my former office mates for putting up with me and for the fun we had, Jan and Abdullah. I would also like to thank André for being a great help when I needed advice on administration, Lies and Christophe for technical support and Antoine for teaching me most valued (and funny) insights on what it means to work in our department.

A big “thank you” goes to two lifetime friends: Tom and Bart. Our weekly get-togethers were and are one of my most valued assets. They were always interested in what I had to say about my work, and for that I declare them slightly mad (and I am a psychologist, I can do that) and love them. I am indebted to Piet, Els, Siesel, Anneleen, and Michael for convincing me that my enthusiasm for eye movements is no reason to inflict it on them, but for showing me a good time all the same. And also thanks to Simon, David, Steve and Pieter for many pleasant walks and evenings.

This brings me to the final part of this, almost absurdly long, acknowledgments section. I thank my parents for the financial and emotional support they gave me all these years. I thank Tine for letting me use her
house and her internet connection when I wanted to retreat in isolation for a while, to write my model of word skipping. And of course, the last person I want to thank is the most important one. I thank Laure for being there every step of the way, for nurturing me in every possible way, for supporting me, for being what she is: my playmate-for-life.

Denis Drieghe

Ghent, January 29th, 2005
CHAPTER 1
PARAFOVEAL PROCESSING IN READING:
AN OVERVIEW

Science is a boys’ game
Anonymous

Contrary to our phenomenological impressions, our eyes do not move
smoothly across a line of text when we read. Eye movements in reading are
characterized by rapid jerk-like movements (saccades) and short periods of
steadiness (fixations). The first recorded observation that describes this
distinction goes back as far as 1879 and was made by Javal (cited in Huey,
1908). Saccades are necessary to direct the gaze to a new location, bringing
new information into the center of the visual field where acuity is best. They
occur several times each second and saccades typically move the eyes
forward about 7 to 9 character spaces, although there is considerable
variability (for reviews on these basic findings see Rayner, 1978; 1998).
Although most saccades in reading English or Dutch are made from left to
right, about 10 to 15 % of the saccades go back in the text. These eye
movements from right to left are called regressions. Research has shown that
during saccades vision is suppressed (e.g. Ishida & Ikeda, 1989; Wolverton
& Zola, 1983) indicating that only during eye fixations new information is
acquired by the processing system and consequently a mental representation
of what the text means can be constructed. Assuming that the word which is
fixated is recognized, it is tempting to assume that eye movements in reading
essentially consist of word-to-word movements. However, such a simple
sequence of motion is rarely observed in empirical data. Whereas some
words are fixated multiple times, other words are initially not fixated but are
immediately afterwards regressed to, or are not fixated at all. Indeed, the
rather loose relationship between eye movements and the layout of the text
made many of the first researchers believe that eye movements were
controlled by an autonomous oculomotor control center (e.g., Buswell, 1920; Erdman & Dodge, 1898; Huey, 1908). In this view, saccade sizes are relatively constant and are only influenced by the global difficulty of the materials being read. It wasn’t until the seventies that the idea of linguistic influences on eye movement behavior really began to be taken into account. Nowadays, the link between certain aspects of eye movements during reading (e.g. fixation times\(^1\)) and linguistic processing is so well-established, that the use of eye movement data is considered by the vast majority of researchers a vital tool for understanding the on-line operations involved in the reading process.

Before discussing those issues in the field of eye movements in reading that this dissertation will focus on, it is important to introduce a couple of terms and techniques that will be used and referred to frequently throughout the thesis. More concretely we will first deal with divisions that are being made in the visual field and methods often employed in eye movement research. By doing so we will touch upon a very important question: how much information can be obtained in a given eye fixation? What is the size of the perceptual span in reading?

Visual acuity is maximal in the center of the retina and rapidly decreases in the periphery. Fine visual discrimination can only be made within the fovea, which is the central 2° of vision. This is illustrated by the fact that it is very difficult to read using only non-foveal vision (Rayner & Bertera, 1979; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). The visual field is further divided in the parafovea (which extends 5° on either side of the

---

\(^1\) The most frequently reported fixation times in reading are: The gaze duration, which is the sum of the fixations from the moment the eyes land on the word/region of interest until the moment they move off again. The first fixation duration, which is the duration of the first fixation during first passage independent of the number of fixations that were made on the word/region. Finally, the total fixation duration, is the sum of all the fixations on the word/region of interest, regardless of whether a fixation happened during first passage or during re-reading.
fixation) and the *periphery* (the region beyond the parafovea). Accuracy is not nearly as good in the parafovea as in the fovea, and even worse in the periphery. Hence, we need to move our eyes on that part of the text that we want to see clearly.

The area and type of text that can be processed on a single fixation is called the perceptual span. The size of the perceptual span has been investigated using the *moving window* technique (McConkie & Rayner, 1975; Rayner, 1975). An example of this technique is shown in Figure 1b. Using this paradigm, the perceptual span is measured by looking at the minimal area or type of text that needs to be presented within the window around the fixation in order for the reading process not to be disturbed, in other words the cutoff is when there are no longer any differences in the eye movement pattern between the condition with and the condition without the moving window. Other techniques employed in this and other eye movement research are the *moving mask* technique (see Figure 1c) in which a mask obscures the text within the window whilst normal text is presented beyond the mask region, and the *boundary* technique (see Figure 1d). In this latter method, a single critical target word is replaced by another word when the reader crosses an invisible boundary location in the text (usually located before the target word). If a reader picked up information from the initially presented word, any inconsistency between this word and the word presented after the display change should appear in the fixation time on the target word.

Using the moving window technique McConkie and Rayner (1975; see also Den Buurman, Boersma, & Gerrissen, 1981; Rayner, 1986; Rayner & Bertera, 1979; Rayner, Well, Pollatsek, & Bertera, 1982) showed that the perceptual span extends up to fifteen characters to the right of the current fixation location. To the left the perceptual span extends only as far as three to four characters (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980). So the perceptual span is clearly asymmetric with more processing in the direction of the text that has not yet been fixated. The fact that the
Figure 1. Examples of the moving window, moving mask, and boundary paradigms. The first line shows a normal line of text with the fixation location marked by an asterisk. The next two lines show an example of two consecutive fixations with a window of 9 letter spaces (with other letters and spaces replaced by x's). The next two lines are an example of two consecutive fixations with a moving mask. The bottom two lines show an example with the boundary paradigm. The first line in the boundary example shows the text prior to the display change. When the reader's saccade crosses an invisible boundary location (the g in morning), the initially displayed stimulus (bottle) is replaced by the target word (coffee). The change occurs during the saccade so that the readers are not aware of the change. (Figure and caption reprinted from Starr and Rayner, 2001).

constituting variable for this asymmetry is reading direction is convincingly shown in research showing the opposite pattern for languages being read in the opposite direction (Pollatsek, Bolozky, Well, & Rayner, 1981). It is important to note here that even though the perceptual span extends up to fifteen characters to the right, the region that is used for extracting information to identify words is limited to about 7-8 letters (McConkie & Zola, 1984; Rayner et al., 1982; Underwood & McConkie, 1985). Beyond 7-8 letters only low-level information, such as word length and spacing, can be extracted.
THE NAME OF THE GAME: PARAFOVEAL PROCESSING IN READING

A large body of research has shown that, besides the essential foveal processing in reading, information from the word next to the currently fixated word is extracted and used in reading as well. This is also consistent with the size of the perceptual span, as described earlier. We will shortly review some of the studies concerning two phenomena that have clearly established the existence and the importance of parafoveal processing in reading: the parafoveal preview benefit and word skipping. Next we will examine the kind of information that can be extracted from the parafovea more closely. How detailed is this information? Does it entail semantic, phonological, or orthographic features of the word presented in the parafovea or is it limited to for instance the retention of visual features?

One of the most robust findings in research on eye movements in reading is that the preview of a word to the right of the fixation results in shorter fixations on that word when it is subsequently fixated. This parafoveal preview benefit was first demonstrated by Rayner (1975) and has been replicated quite a number of times (e.g. Blanchard, Pollatsek, & Rayner, 1989; Morris, Rayner, & Pollatsek, 1990; Rayner et al., 1982). It is most clearly shown by comparing conditions in which the parafoveal word is visible versus conditions in which, using the boundary paradigm, the parafoveal word is masked prior to the eyes landing on it. The size of the preview benefit (i.e. the difference between these conditions for the viewing times of the manipulated word when it is fixated) is typically in the order of 20 – 50 ms.

Another instance in which the importance of parafoveal processing is clearly shown is word skipping. A word is skipped when there is no fixation on the word during first passage. This is far from being a rare phenomenon; on average about one third of the words in a text are skipped (Rayner, 1998). One of the most conspicuous findings in word skipping is that it occurs more
frequently with short words than with long words (Brysbaert, Drieghe, & Vitu, 2005; Brysbaert, & Vitu, 1998; Rayner, 1979; Rayner & McConkie, 1976; Vitu, O’Regan, Inhoff, & Topolski, 1995). Word skipping also occurs more often when the previous fixation is close to the parafoveal word than when it is further away (Kerr, 1992; Rayner, Sereno, & Raney, 1996; Vitu et al., 1995). This is the so-called effect of launch-site (i.e. the position from which the saccade starts). However, besides low-level visuomotor factors, linguistic variables are also known to influence the probability of word skipping. Quite some studies have shown that words that are predictable from the preceding context are skipped more often than unpredictable words (Altaribba, Kroll, Sholl, & Rayner, 1996; Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981; Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987) and that short function words (e.g. “the”) are skipped more than content words (O’Regan, 1979; 1980; Gautier, O’Regan, & LaGargasson, 2000). A small effect of the frequency of the parafoveal word has also been observed in word skipping: High-frequency words are more likely to be skipped than low-frequency words, especially when the eyes are close to the parafoveal word on the fixation prior to skipping (Henderson & Ferreira, 1993; Radach & Kempe, 1993; Rayner & Fischer, 1996; Rayner et al., 1996). The effects of these lexical/linguistic variables on word skipping clearly show that some words that are skipped have been processed, at least up to a certain extent. Just how much parafoveal processing has occurred on the fixation prior to skipping the following word, is an issue of quite some debate in the field of eye movements in reading. We will return to this issue both in the Introduction and throughout the entire dissertation as it is truly at the core of this thesis.

Acknowledging the fact that parafoveal processing occurs, the question arises how detailed the information is that is being extracted from the parafovea and what kind of information is being integrated across saccades. McConkie and Zola (1979) examined the parafoveal preview benefit in an experiment in which from fixation to fixation all the letters changed from
uppercase to lowercase and vice versa (e.g. ChAnGe was replaced by cHaNgE). By applying this methodology they were able to establish that the parafoveal preview benefit was not due to retention of visual feature information, as the reading process was virtually not disrupted. If visual codes were important for integrating information across saccades, then the change of features between upper- and lowercase letters should have disrupted reading. From this study it was concluded that information must be integrated across saccades in an abstract form. Prior research (Rayner, 1978) had already established that orthographic information could be extracted from the parafovea. In a naming task, a word or a letter string was presented in the parafovea. When the participant made an eye movement to the stimulus in the parafovea, it was replaced by a target word and the task for the participant was to name this target word as fast as possible. The basic finding from this study, which was later replicated in normal reading experiments (e.g. Inhoff, 1990), was that the preview caused facilitation on the subsequent fixation and that if the first 3 letters of the initially presented stimulus were identical almost as much preview benefit was obtained as when the entire word was available for preview. So, readers are clearly able to obtain sub-lexical information from the parafovea such as partial word information.

At this point, it is important to try to integrate findings from a related research field in cognitive psychology. Visual word recognition is the area in psycholinguistics that is concerned with the processes related to recognizing a single word that is visually presented. Whereas in the seventies and eighties the field was dominated by models that assume a direct route from print to a full lexical recognition of the presented word (e.g. Coltheart, 1978), a shift occurred in the nineties based on accumulated evidence supporting models that assume an additional conversion from orthography to phonology (for a review see: Frost, 1998). In recent years, a strong phonological model of visual word recognition has been promoted, according to which the orthographic stimulus is first translated into a partial phonological code that makes access to stored word information. Once the
stored representation has been activated, additional information about the exact pronunciation and spelling becomes available (Drieghe & Brysbaert; 2002). Returning to the issue of the kind of information that is being extracted from parafoveal vision and taking this shift in models of visual word recognition into account, the importance of looking for potential extraction of phonological features becomes apparent. Pollatsek, Lesch, Morris and Rayner (1992) found that when a homophone of a target word (e.g. accept – except) was presented as a preview in the parafovea, facilitation was found on the processing of the target word when it was subsequently fixated. This facilitation was observed even though in a control condition a preview was presented that was equal to the homophone in terms of visual similarity with the target word. So, phonological information is clearly extracted from the word to the right by means of parafoveal processing.

Other candidates for the code conveying information across saccades have also been examined. No effects were found from morphological parafoveal information (Lima, 1987; Inhoff, 1989; but see Deutsch, Frost, Pollatsek, & Rayner, 2000) or from purely lexical parafoveal information (Lima & Inhoff, 1985). In this latter study the familiarity of the initial letter sequence of the words presented in the parafovea was manipulated. Quite surprisingly, semantic information extracted from the parafovea does not appear to boost word recognition either when the word is later fixated (or at least doesn’t do it fast enough). Rayner, Balota and Pollatsek (1986; see also Altaribba, Kambe, Pollatsek, & Rayner, 2001; Inhoff, 1982; Inhoff & Rayner, 1980) found no facilitation in the viewing times of a target word (e.g. tune) when a semantically related word (e.g. song) was presented as a preview in the parafovea, as compared to unrelated previews.

In this short overview of what is known about parafoveal processing in reading, we have consciously restricted ourselves to those observations and conclusions that almost all researchers in the field will agree upon. This puts us in a good position to leave the beaten path and move on to more slippery
ground. But before plunging into the details of current models of eye movements in reading we hold on to our relatively a-theoretical position a little longer to hover over the research field as a whole. What are the major questions in the eye movements in reading research and how does parafoveal processing always appear to be at the heart of the debate?

WHAT’S AT STAKE? THE BIG QUESTIONS IN THE FIELD

In 2001 an important review article on eye movements in reading was published by Matthew Starr and Keith Rayner in the Trends in Cognitive Sciences journal. They identified three controversial issues that needed to be resolved in future research. We will be as audacious as to steal their structure of organizing the quicksand areas in the field and deal with these three issues one by one.

IS WORD PROCESSING IN READING SERIAL OR PARALLEL?

In this Introduction we have reported quite some studies that established, in addition to foveal processing, the existence and importance of parafoveal processing during reading. One of the major questions in eye movements in reading research concerns the time course of these two types of processing. Does parafoveal processing only kick in when foveal processing has been concluded or do both types of processing occur simultaneously? In other words, do readers process information from more than one word at a time? Whereas nobody doubts that the letters within a word are processed in parallel the issue is more controversial when it comes to words. According to the highly influential E-Z Reader model (e.g. Reichle, Rayner, & Pollatsek, 2003), which we will describe in more detail in the next section, words are typically processed one at a time during reading. Word recognition is considered to be a serial process under the control of an attentional beam, with the word in the attentional beam being the only word that is being processed lexically. Attention is allocated serially during reading because
readers need to keep word order straight (Pollatsek & Rayner, 1999). A number of other models (e.g. SWIFT, Engbert, Longtin & Kliegl, 2002) assume that lexical processing is spatially distributed across words with a constant competition for processing resources between different words. The issue of whether word processing in reading is serial or parallel will be the main focus of Chapter 2.

LOW-LEVEL OCULOMOTOR FACTORS VERSUS HIGHER-LEVEL COGNITIVE PROCESSES

Are eye movements during reading controlled by low-level oculomotor strategies or are they controlled by moment-to-moment cognitive processes? This is a question that is embedded in the eye-mind lag issue: Are the acquisition of verbal/linguistic materials and the subsequent processing in the brain fast enough to influence the planning and execution of the next saccade? Or is the planning of a saccade based on a relatively dumb mechanism taking into account only low-level factors, such as word length? One of the nicest things about doing research in eye movements in reading - and this is a compliment to the researchers involved - is that when you look at the progress in the field, parts of this question and other questions have actually been answered and need to be reformulated. In 2001 (although its peak period can be situated in the first half of the nineties) this specific question was still mostly concerned with two options: are eye movements determined by low-level oculomotor factors and only obliquely associated with higher-level processing (the so-called oculomotor models, e.g. the strategy-tactics theory by O’Regan 1990; 1992), or is higher-level processing responsible for the lion’s share of the variance in eye movement data (i.e. processing models such as the E-Z Reader model). Nowadays most researchers will agree that this question has shifted to the relative importance of low-level versus high-level factors in specific phenomena. Too many studies have found influences of both low-level and high-level factors on eye movement behavior, a subtlety reflected in the architecture of almost all
current models of eye movements in reading. One distinction in which the relative contributions of low-level and high-level factors are clearly different is the decision of when to move the eyes versus the decision of where to move the eyes. Concerning the when decision, how long readers look at a word is mostly determined by the ease or difficulty associated with the processing of that word. Readers look longer at a low-frequency word than at a high-frequency word (e.g. Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner et al., 1996; Schilling, Rayner, & Chumbley, 1998, Vitu, 1991), and look less long at a word that is predictable from the preceding context (Balota et al., 1985; Binder, Pollatsek, & Rayner, 1999; Ehrlich & Rayner, 1981, Rayner & Well, 1996; Schustack et al., 1987; Zola, 1984) or a word that was acquired at an early age (Juhasz & Rayner, 2003, 2005).

While some low-level visual factors influence the decision of when to move the eyes, the main players in making that decision are clearly the linguistic properties of the words being read. The opposite seems to be true for the decision of where to move the eyes: low-level visual factors, such as word length and spacing between words (Rayner, Fischer, & Pollatsek, 1998), are the most important influences on saccade length and on the landing position in a word. For example, the saccade length is influenced by the word length of the currently fixated word and that of the word to the right of fixation (e.g. Blanchard et al., 1989; O’Regan, 1980; Rayner, 1979). A phenomenon in reading that eludes this convenient when/where dichotomy is word skipping. While word skipping is clearly closer to the question of where to move the eyes, influences of both low-level visual factors and high-level linguistic factors have been shown to affect skipping behavior (see above). Chapter 3 will focus on the question of the relative contribution of low-level

---

2 More concretely, whereas processing models have incorporated “low-level” explanations for certain phenomena (e.g. accounting for initial fixation locations in the E-Z Reader model, Reichle, Rayner, & Pollatsek, 1999), a purely oculomotor approach has been mostly abandoned (for an exception see Yang & McConkie, 2001).
and high-level factors in determining word skipping behavior, an issue also embedded in the following question.

**HOW MUCH INFORMATION IS EXTRACTED FROM THE RIGHT OF THE FIXATION?**

We have already summed up the basic findings on the types and amount of information that is being extracted during parafoveal processing, but a few questions remain. We will focus on the following question: How much information is extracted from a parafoveal word when the system decides to skip the word? Is the parafoveal word completely recognized by means of parafoveal processing or does the system make an educated guess taking into account only coarse information such as for instance word length? Both views are incorporated in models of eye movements in reading (respectively the E-Z Reader model and the EOVP model (Brysbaert & Vitu, 1998)) and it is our strong belief that answering this question will prove to be a cornerstone in deciding between various models. Chapter 4 will focus directly on the amount of processing of the parafoveal word prior to skipping.

**THE PLAYERS: MODELS OF EYE MOVEMENT CONTROL IN READING**

*Every eye tracker around the world used for reading research is oriented, either in a positive or in a negative sense, towards Amherst.*

*D. D.*

In recent years, research in eye movements in reading has been taken to the next level. A key development in the field is the emergence of computational models. In this section, we will describe the architecture of the most elaborate of those models at present, the E-Z Reader model (e.g. Reichle et al., 2003). This influential model was developed at the eye tracking lab of the University of Massachusetts at Amherst. We will concentrate on its
account of word skipping as this will be the focus of two of the empirical chapters in this dissertation. Afterwards we will discuss a few studies of which the results seem problematic for the E-Z Reader model. Finally, we will briefly present some of the competitors: Alternative models that offer a different explanation for eye movement control in reading as a whole or for specific phenomena in reading.

THE GRANDMASTER: THE E-Z READER MODEL

The E-Z Reader model of eye movement control in reading (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, 2003; Rayner, Reichle, & Pollatsek, 1998; 2000; 2005; Pollatsek, Reichle, & Rayner, 2003; 2005) is a quantitative model in which the core assumption is that cognitive processes associated with processing the fixated word are the engine behind eye movements in reading. As already mentioned, word recognition is considered to be a serial process under the control of an attentional beam, with the word in the attentional beam being the only word that is being processed. Two phases of word recognition are being distinguished. The termination of the first phase (identification of orthographic and phonological forms) cues the oculomotor system to begin programming a saccade to the next word. The termination of the second phase (full lexical identification) causes the attentional beam to shift to the next word. The shift of the attentional beam usually occurs before the eyes move to the next word and during the time that the attentional beam is on the parafoveal word (but the eyes are still on the previous word), parafoveal processing occurs. In this manner the E-Z Reader model accounts for the parafoveal preview benefit. Let us consider the case of a difficult foveal word (e.g. a low-frequency word). The end of the 1st phase of the word recognition of the foveal word triggers the start of programming a saccade to the parafoveal word. The time needed for programming a saccade is fairly constant, so the eyes land on the parafoveal word after a certain delay. If the foveal word is difficult, there is a longer 2nd phase of its word recognition
than there would be in the case of an easy foveal word, hence the shift of the attentional beam cued by the termination of this 2nd phase occurs later than for an easy word. Because the attentional beam spends less time in the parafoveal word prior to the arrival of the eyes, the model predicts that the preview benefit decreases as the processing difficulty of the foveal word increases, an observation first made by Henderson and Ferreira (1990).

In E-Z Reader, word skipping is based on the following sequence of events. If (a) the eyes are on word_{n}, (b) the attentional beam has shifted to word_{n+1}, and (c) if the first phase of word identification of word_{n+1} in the parafovea is rapid enough, the programming of the eye movement to word_{n+1} is cancelled and replaced by the programming of an eye movement to word_{n+2}. The saccade to word_{n+2} will be executed and as a consequence word_{n+1} will be skipped. It is important to note here that the second phase of the identification of word_{n+1} in the parafovea will usually be completed before the eyes actually move to word_{n+2}. In other words, this model predicts that in order for a word to be skipped, a significant amount of processing of the skipped word needs to have happened: the first phase of word recognition has been completed and the completion of a full lexical identification of that word has occurred or is imminent. Because of the above sequence of events, it is the processing ease of word_{n+1} that influences the likelihood of it being skipped and this ease is determined by word frequency and contextual predictability. In this manner the model can successfully predict the effects of predictability and frequency on word skipping. In addition to language variables, visual factors such as word length and launch site have a role in the E-Z Reader model as well. The model can account for the word length effect in word skipping because it assumes an inverse relation between the extraction of letter information and the distance of a letter from the center of the visual field. So the further away the eyes are from the target word, the more time will be needed to complete the first phase of the word recognition, and as a consequence the slimmer chances will be that the word will be skipped.
CHALLENGES FOR THE E-Z READER MODEL

As is apparent from the architecture of the E-Z Reader model described above, the model upholds a serial view on the time course issue of foveal and parafoveal processing. Only one word at a time is in the attentional beam and is being processed. In recent years this serial assumption has been questioned by a number of researchers who state that this assumption is in conflict with several studies that report so-called parafoveal-on-foveal effects. Parafoveal-on-foveal effects refer to the possibility that characteristics of the word to the right of the fixation influence the processing of the currently fixated word. It is generally assumed that the existence of such effects would be very damaging to the serial assumption of the E-Z Reader model. After all, in the E-Z Reader model parafoveal processing only begins after foveal processing has been concluded. How can parafoveal processing influence foveal processing if it is the termination of foveal processing that cues the shift of the attentional beam and subsequently cues parafoveal processing? In fact, a number of studies have indicated that the foveal viewing time can indeed be altered by the words presented in the parafovea (e.g. Inhoff, Radach, Starr, & Greenberg, 2000; Kennedy, 1998; 2000; Kennedy & Pynte, 2005; Kennedy, Pynte, & Ducrot, 2002; Schroyens, Vitu, Brysbaert, & d’Ydewalle, 1999; Starr & Inhoff, 2004; Underwood, Binns, & Walker, 2000; Vitu, Brysbaert, & Lancelin, 2004). However, there are methodological problems associated with some of these studies. More specifically, quite some of the studies that report parafoveal-on-foveal effects used tasks that closely resemble reading but doubts can be raised whether these tasks require the full range of psycholinguistic processes that occur during normal reading (for a discussion on the generalizability of these tasks, see: Rayner, White, Kambe, Miller, & Liversedge, 2003). In the Kennedy (1998) study, for instance, participants were presented with three words. The first word was either the word “looks” or the word “means”. In the first case the task for the participants was to decide whether the following two words were physically identical, in the second case they had to decide whether the adjacent words were synonyms. Among the findings
of this study a reduced gaze duration on the first word was reported in the case of long parafoveal word. This observation was taken as evidenced for a parallel processing account in which the system is sensitive to the rate at which sub-lexical parafoveal information can be acquired; a difficult second word (e.g. a long word) would attract an early saccade from the prior word. Criticisms on this task and others alike (e.g. the clothing search task, Schroyens et al., 1999) usually state that the task employed resembles more a variant of a visual search task than normal reading. Although not all of the studies use artificial tasks (e.g. Inhoff et al., 2000; Kennedy & Pynte, 2005) the validity of parafoveal-on-foveal effects is also endangered by an apparent lack of consistency: For example Underwood, Binns and Walker (2000) report longer fixations on the foveal word when the parafoveal word had an informative initial trigram, whereas Kennedy (1998; 2000) reports shorter fixations in these circumstances. Failures to replicate are also not uncommon, sometimes even within the same study (experiment 2 versus experiment 4 in Hyönä & Bertram, 2004). However, regardless of the lack of a model that can reliably predict the presence and direction of parafoveal-on-foveal effects, the accumulated evidence for these effects is quite substantial and growing, posing a serious threat for the E-Z Reader model.

A second challenge for the E-Z Reader model, closely linked to its serial assumption, is the explanation offered for the skipping of words during reading. As a result of the serial assumption in the time course issue of foveal and parafoveal processing the E-Z Reader model predicts that fixations on a word should be longer when the next word is skipped than when the next word is not skipped. This follows from the model because skipping results from the cancellation of the program to fixate word_{n+1} by the program to fixate word_{n+2}. Thus, a later program replaces an earlier program which should be reflected in a time cost. In fact, such an inflated fixation duration prior to skipping has been observed in several studies (Pollatsek, Rayner, & Balota, 1986; Pynte, Kennedy, & Ducrot, 2004; Rayner et al., 2004), but not in others (Engbert et al., 2002; Radach & Heller, 2000). Whereas such an inflated fixation duration prior to skipping is essential in a
serial approach, a model that upholds parallel processing of foveal and parafoveal processing does not necessarily predict such an effect to occur. As a result, the fixation duration prior to skipping has turned out to be one of the corner stones in the ongoing dispute on serial versus parallel processing of multiple words in reading and especially the explanation offered for word skipping by the E-Z Reader model.

THE COMPETITORS

While it is clearly beyond the scope of this Introduction to describe in detail the architecture of all the different models that have been proposed to account for eye movement control in reading, a very brief presentation of a few of the competitors of the E-Z Reader model is in order. It is important to note that in the choice of the models that are presented, we were not as such guided by the impact of a certain model in the field of eye movements in reading, but mainly by the importance of these models in the following chapters of this dissertation.

The discovery of parafoveal-on-foveal effects has been accompanied by the development of models of eye movement control embracing a parallel view on the time course issue of foveal and parafoveal word processing. The SWIFT model (Engbert et al., 2002; Kliegl & Engbert, 2003) for instance, while adopting quite a few of the architectural features of the E-Z Reader model, departs from it by assuming that lexical processing is spatially distributed across words and that a competition for processing resources between the different words is constantly going on; for example, a difficult word will use the majority of the resources leaving few resources for the processing of the other words. In SWIFT saccades are directed towards words that have the highest level of excitation, which occurs at intermediate amounts of lexical processing. Thus, word_{n+1} will be skipped if word_{n+2} has a higher level of excitation, and the model predicts that this occurs when word_{n+1} is more frequent, more predictable, and shorter. However, because
SWIFT does not assume that the next word is the default saccade target (as E-Z Reader does) there is no predicted “cost” in canceling a planned saccade to the next word (i.e. no inflated fixation duration prior to skipping).

Another suggestion about how processing of word\textsubscript{n+1} might affect the reading time of word\textsubscript{n} was made by Schiepers (1980). We will explain this model more in detail because it will be the starting point for Chapter 2. Schiepers started from the observation that in a perceptual identification task it takes on average 90 ms longer per degree of eccentricity to identify a word, arguably because it takes that much longer for the stimulus to activate the relevant representations in the brain. Given that one degree of visual angle roughly coincides with three letter positions\textsuperscript{3} (in the fonts used in most research) and that saccades usually are 7-9 letters long, Schiepers hypothesized that if word\textsubscript{n+1} was presented in foveal vision 210-270 ms after it had been presented in parafoveal vision, the parafoveal information from fixation \textsubscript{n} could be merged optimally with the foveal information on fixation \textsubscript{n+1}. By combining both sources of information, the speed of the activation of the word representation could be boosted as compared to when it was based on the foveal information alone. Schiepers argued that this might be the origin of the typical fixation durations of some 250 ms seen in text reading. When fixations are shorter or longer, part of the parafoveal preview benefit is lost, because the synchrony in the arrival of parafoveal and foveal information is less than perfect.

Finally, another model that we will briefly describe is the Extended Optimal Viewing Position model (Brysbaert & Vitu, 1998). This model is not as such a model on eye movement control in reading but is restricted to explaining the phenomenon of word skipping in reading. In 1998, Brysbaert and Vitu

\footnote{Research has shown that in reading, the numbers of letters are a more appropriate metric to use than degrees of eccentricity. The number of letter spaces crossed by saccades is relatively stable, independent of the visual angle (Morrison, & Rayner, 1981).}
(for an update see: Brysbaert et al., 2005) carried out a meta-analysis on all the studies that examined word skipping of target words of which the processing difficulty was manipulated (e.g. the predictability of the target word from the preceding context) while reporting the word lengths of the target words. They found that the vast majority of the variance in the word skipping data could be explained based on word length. In other words, to predict how often a word was skipped, it was more useful to know its word length than for instance its predictability. Starting from this finding Brysbaert and Vitu developed a small, elegant model that states that word skipping is based on an educated guess taking into account only rather coarse information such as the word length of the target word and the distance of the target word from the present fixation location. In this manner, the EOVP model is an alternative to the E-Z Reader model that states that a word is mainly skipped because it is recognized or because full recognition is imminent. Predictions derived from this model will be directly tested in Chapter 3.

**PARAFOVEAL PROCESSING IN READING: ON THE INTERPLAY BETWEEN VERBAL AND VISUAL FACTORS**

**THE BOARD IS BEING SET: EXPLAINING THE TITLE**

In this Introduction we have shown the existence and importance of parafoveal processing. We have discussed the major questions still present in the field of eye movements in reading, and showed how they all are closely linked to the issue of parafoveal processing. We have briefly described the current models that try to account for the various observations made on this topic. The current dissertation will focus on all three of the questions identified by Starr and Rayner (2001): We will examine the time course of foveal and parafoveal processing by testing a hypothesis directly derived from a parallel model, and we will try to find out how much of a parafoveal
word is processed when the system decides to skip the word. And while dealing with these two questions, we will try to answer how much of the observations we have made can be explained solely on the basis of low level factors (e.g. word length), on the basis of high level factors (e.g. predictability), or by a combination of both.

THE PIECES ARE MOVING: RESEARCH OBJECTIVES

In Chapter 2 we will directly test a hypothesis derived from the Schiepers model (1980). In this model, the benefit of seeing the upcoming word is due to the fact that the parafoveal information from fixation $n$ is combined with the foveal information from fixation $n+1$ to boost word recognition. We tested this assumption by adding an extra blank space between the foveal and the parafoveal word. According to the model, this should result in a 30 ms longer viewing time on word$_n$ because the system would have to wait longer for the incoming information from word$_{n+1}$. Such an observation would be incompatible with attention-based, sequential processing models, such as the E-Z Reader model, as the cue for starting to program the saccade to the parafoveal word should be independent of parafoveal processing.

In Chapter 3 we tested one of the assumptions of the EOVP model (Brysbaert & Vitu, 1998) that states that word skipping is mainly a function of the length of the upcoming word. If this is the case a short, unpredictable word should be skipped more often than a long, predictable word. We will also look for a potential interaction between the low-level factor word length and the high-level factor predictability: Will an unpredictable target word that has the same word length as a predictable target word be skipped more often, based on this congruency of word length, than a control condition?

In Chapter 4 we examine how much of the parafoveal word is processed on the prior fixation when the system decides to skip the following word. We
will do this by comparing the skipping rates of different previews using the boundary technique\(^4\). Will the system be able to pick up the difference between a predictable word and a non-word when this difference is limited to a single letter? Based on the E-Z Reader model we would expect so, for it states that a word is mainly skipped because it is \textit{fully} recognized by means of parafoveal processing. If this is the case the preview of a predictable word will be skipped more often than a preview that is identical to the predictable word with the exception of a single letter. And how often do readers skip a complete “garbage” word (e.g. \textit{f\(\text{hrrsos}\)}) that is presented as a preview? The EOVP model predicts this event to be not so rare because word length is the main player in determining skipping behavior in this model, whereas the E-Z Reader model predicts skipping probabilities close to zero. In Chapter 4 we will also examine how skipping rates of the parafoveal word are affected by the processing ease of the foveal word. If parafoveal processing is reduced in the case of high foveal load (e.g. a low-frequency word) and word skipping is based on the amount of parafoveal processing, we would expect to observe a reduced skipping ratio when the target word is preceded by a difficult word.

\[^4\text{Due to technical limitations of the eye tracking equipment available at Ghent University, this research was carried out during a five month research visit at the eye tracking lab of the University of Massachusetts at Amherst.}\]
REFERENCES


Schiepers (1980) proposed that in text reading, the currently fixated word and the next word are processed in parallel but with a time delay of 90 ms per degree of eccentricity. In his model, the benefit of seeing the upcoming word is due to the fact that the parafoveal information from fixation $n$ is combined with the foveal information from fixation $n+1$ to boost word recognition, at least when the fixation on word $n$ is of an optimal duration (between 210 and 270 ms). We tested this assumption by adding an extra blank space between the foveal and the parafoveal word. According to the model, this should result in a 30 ms longer processing time for the foveal word. However, reading time was shorter for a word followed by a double space than for a word followed by a single space. An effect of parafoveal word length was also observed with a longer word in the parafovea leading to shorter fixation times on the foveal word. Implications of these low-level parafoveal-on-foveal effects are discussed.

---

1 This paper is co-authored by Marc Brysbaert and Timothy Desmet.
INTRODUCTION

When people are reading, their eye movements are characterized by a sequence of saccades and fixations. The main purpose of the saccades is to bring new information into the center of the visual field, where visual acuity is highest. However, there is a large body of evidence that, in addition to foveal word processing, information from the word to the right of the fixation is extracted and used in reading as well (see Rayner, 1998 for a review). Two of the most important findings in this respect are the phenomenon of word skipping and the so-called parafoveal preview benefit. About one third of the words in a text are skipped during first-pass reading. This is particularly so for short words and words that lie close to the previous fixation location (i.e., when the saccade is launched from the second half of the word prior to the target word). There is also a smaller influence of the difficulty of the target word (see Brysbaert & Vitu, 1998; Brysbaert, Drieghe, & Vitu, in press, for a meta-analysis of the data). The parafoveal preview benefit refers to the finding that reading is slower when the letters of the word to the right of the currently fixated word are not visible than when they are visible (e.g. Blanchard, Pollatsek, & Rayner, 1989; Morris, Rayner, & Pollatsek, 1990; Rayner, 1975; Rayner, Well, Pollatsek, & Bertera, 1982). From these findings, it is clear that processing of parafoveal information plays a role in normal reading. There is, however, much more controversy over the question to what extent parafoveal information concerning word $n+1$ influences the fixation duration and gaze duration$^2$ of the currently fixated word $n$. This latter possibility is referred to as parafoveal-on-foveal effects and several suggestions of such effects have been made.

---

$^2$ The gaze duration is the sum of the fixations from the moment the eyes land on word $n$ to the moment they move off again.
A first way in which parafoveal processing of word $n+1$ might influence the gaze duration on word $n$, was proposed by Pollatsek, Rayner, and Balota (1986). They reported that the fixation duration was longer before a saccade that skipped the next word than before a saccade that was targeted at the next word. They interpreted this finding as evidence for the hypothesis that words were skipped as a result of a two-stage process. First, a saccade was programmed to word $n+1$, but if this word was recognized (or was likely to be recognized) before the saccade was initiated, the program could be cancelled and replaced by a new program for a saccade towards word $n+2$ (see Reichle, Rayner, & Pollatsek, 2003, for the latest update of this model of eye movement control). The cancellation of the original program and the replacement by a new one were the origin of the longer fixation duration on word $n$. Unfortunately, this finding is a bit controversial with some studies finding the effect and others that do not (e.g., Drieghe, Brysbaert, Desmet, & De Baecke, 2004; but see Drieghe, Rayner, & Pollatsek, submitted). A recent study suggests that longer fixations before a skipping saccade are observed only when long and difficult words are being skipped (Kliegl & Engbert, in press). When short and easy words are skipped, fixation durations actually tend to be shorter than when these words are fixated. Although the latter finding is a problem for most theories of eye movement control in reading, if it can be replicated it still is an example of how processing word $n+1$ may influence the gaze durations on word $n$.

Another suggestion of how parafoveal word $n+1$ might affect the gaze duration on word $n$ was made by Kennedy and colleagues (e.g., Kennedy, 1998; Kennedy, Murray, & Boissiere, 2004; Kennedy & Pynte, 2005). Kennedy (1998) reported that the gaze durations on word $n$ were shorter when word $n+1$ was a low-frequency word and when it was a long word. He interpreted this paradoxical parafoveal-on-foveal effect as evidence for a model of eye movement control (which has been referred to as the process monitoring hypothesis) in which word $n$ and word $n+1$ are processed in parallel (with some time delay depending on the length of word $n$) and in which the resources are allocated as a function of the difficulty of both
words. The harder word \( n+1 \) is to process, the stronger it pulls the eyes towards it, in order to optimize the extraction of visual information from the page of text. Again, however, the evidence for this parafoveal-on-foveal effect is not unequivocal, with some studies failing to report an effect of the difficulty of word \( n+1 \) on the gaze duration for word \( n \) (e.g., White & Liversedge, 2004), and others reporting a lengthening of the gaze duration for difficult parafoveal words (e.g., Hyönä & Bertram, 2004, Experiment 2; see Rayner & Juhasz, 2004, for a critical review of the evidence).

A final suggestion about how processing of word \( n+1 \) might affect the reading time of word \( n \) was made by Schiepers (1980). Schiepers started from the observation that in a perceptual identification task it takes on average 90 ms longer per degree of eccentricity to identify a word, arguably because it takes that much time for the stimulus to activate the relevant letter and word representations in the brain. Given that one degree of visual angle roughly coincides with three letter positions\(^3\) and that saccades usually are 7-9 letters long, Schiepers hypothesized that if word \( n+1 \) was presented in foveal vision 210-270 ms after it had been presented in parafoveal vision, the parafoveal information from fixation \( n \) could be merged with the foveal information on fixation \( n+1 \). By combining both sources of information, the activation of the word representation could be faster than if it were based on the foveal information alone. This, argued Schiepers, could be the origin of the typical fixation durations of some 250 ms seen in text reading. When fixations are shorter or longer, part of the parafoveal preview benefit is lost, because the synchrony in the arrival of parafoveal and foveal information is less than optimal.

The ideas of Schiepers (1980) were utilized by Schroyens, Vitu, Brysbaert, and d’Ydewalle (1999) to provide a neat explanation of a puzzling finding.

---

\(^3\) Nowadays we know that in reading the numbers of letters are a more appropriate metric to use than degrees of eccentricity. The number of letters crossed by saccades is relatively stable, independent of the visual angle (Morrison, & Rayner, 1981).
In their experiment, Schroyens et al. presented three alphabetic stimuli. The first one was a boundary stimulus, which either was a high-frequency word, a low frequency-word, or a homogeneous string of the letter z. There were two lengths of these boundary stimuli: 3 letters long (e.g., now, tic, zzz) and 5 letters long (e.g., first, vaunt, zzzzz). The second word was the target word and was a high-frequency or a low-frequency word of 7 letters (e.g., because, judaism). Finally, there was a third word with a length ranging from 4 to 8 letters. The task of the participants was to read the three stimuli and to indicate whether one of the words referred to an article of clothing (e.g., cap, skirt, trousers). The intriguing finding was that participants looked more than 20 ms longer at a zzzzz string than at a zzz string, even though there was no more information to be obtained from a 5-letter z-string than from a 3-letter z-string. Schroyens et al. ventured that the only reason for the longer gaze durations on zzzzz than on zzz was that in the former case the parafoveal word was on average one letter position further away from the fixation location. If fixation durations are partly determined by the need to synchronize the parafoveal information from the current fixation with the foveal information from the next fixation, then the oculomotor system had some 30 ms longer to wait before initiating the saccade.

Strong influences of word length on eye movement parameters have also been reported in studies that looked at the factors that govern eye movement control in text reading. Increases in word length are known to increase the probability of fixating a word (Brysbaert & Vitu, 1998; Rayner & McConkie, 1976) and of making a second fixation on that word (Vitu, O’Regan, Inhoff, & Topolski, 1995). Word length is also positively correlated with gaze duration, partly because of the increased tendency torefixate long words, but also partly due to increased fixation durations on long words (Calvo & Meseguer, 2002; Rayner & Fischer, 1996; Rayner, Sereno, & Raney, 1996). Interestingly, the issue of word length has never received much attention from researchers investigating visual word recognition with lexical decision and word naming. The prevailing wisdom (e.g., Balota, 1994, pp. 308-309; Harley, 2001, p. 148) seems to be that word
length does not have a strong effect on lexical decision and naming, as long
as words are controlled for frequency and lexical neighborhood, and as long
as the nonwords in the lexical decision task are properly chosen (Hudson &
Bergman, 1985). Because of these divergent views on the impact of word
length, it seemed worthwhile to us to explicitly test whether part of the word
length effect in text reading could be a result of the need to synchronize the
arrival of parafoveal and foveal information, as claimed by Schiepers (1980)
and recently endorsed by Schroyens et al. (1999) and Kennedy, Pynte, and
Ducrot (2002).

There is a very simple test of Schiepers’s conjecture. If the retinal
distance between the parafoveal and the foveal word affects the reading time
of the foveal word, then adding an extra space between both words should
result in a longer gaze duration on the foveal word. This extra time should be
in the order of 30 ms (as the parafoveal information has been shifted by one
third of a degree of visual angle). Prior studies using manipulations of the
spacing between words have concentrated primarily on the effects of
denying space information. This line of research has shown that reading
unspaced text is detrimental for the reading rate (for a review see Rayner &
Pollatsek, 1996) hence demonstrating the importance of the word
boundaries. Only a few studies have looked at the effects of double spacing,
and those that did so mostly used a letter search task (e.g. Jacobs, 1987;
Jacobs, & O’Regan, 1987). The study that comes closest to the current
experiment is a study by Rayner, Fischer, and Pollatsek (1998). In their
second experiment they used a so-called wide space condition. It consisted
of a blocked presentation of three blank spaces between the words. The task
was normal reading. The comparison between this spaced condition and
normal reading showed no significant differences, but the means strongly
suggested, contrary to the prediction from the Schiepers model, a reduction

4 The missing word length effect in visual word recognition is present even up to 9
letter words but is limited to skilled readers. Impaired and beginning readers show a
word length effect in smaller words (Nazir, 2000).
of the viewing times in the case of wide spacing. The only other studies we are aware of that used double spacing in normal or close to normal reading are Kolers, Duchnicky, and Ferguson (1981) and Heller and Müller (1983). Kolers and colleagues directly compared single and double spacing and reported no effects on individual fixations but a slightly lower number of fixations in the condition with the double spacing. In the study by Heller and Müller the distance between the words was varied between 1° and 7°. A larger distance between the pre-target and the target word resulted in longer saccades and prolonged fixation durations on the target, presumably because of a reduced parafoveal preview benefit.

**EXPERIMENT 1**

Whereas Rayner et al., Kolers et al., and Heller and Müller used a blocked presentation of the wide spacing, in our experiment we worked with normally spaced text that had an occasional extra blank space after target words of 5 letters. We chose this word length because we wanted to increase our chances of observing a single fixation on the target word (words that are shorter, are skipped too often; and words that are longer, are refixated too often). To ensure that the extra blank space would not draw too much attention, we used a large number of filler texts in the experiment.

**METHOD**

*Participants.* Participants were 40 first-year students at Ghent University, who participated for course credits. They all had normal, uncorrected vision and were native Dutch speakers.

*Apparatus.* Eye movements were recorded with a Senso-Motoric Instruments (SMI Eyelink) video-based pupil tracking system. Viewing was binocular but eye movements were recorded from the right eye only. A high speed video camera was used for recording. It was positioned underneath the
monitored eye and held in place by head-mounted gear. The system had a visual resolution of 20 seconds of arc. Fixation locations were sampled every 4 ms and these raw data were used to determine the different measures of oculomotor activity during reading. The display was placed at a distance of 69 cm from the participant’s eye, so that three characters coincided with 1° of visual angle. A chin rest was used to reduce head movements during the experiment.

Materials. We used the 36 text fragments created for the Drieghe et al. (2004) study. Each text fragment consisted of five lines of text. The original purpose of this stimulus set was to examine combined effects of word length (2 and 4 letter words) and predictability on word skipping, but this has no further relevance for the present study. The 5-letter words in the stimulus set served as the target words of the present experiment. All the targets were located in the middle portion of a line of a text and none was the last or penultimate word of a sentence. For each text, two variants were made according to a latin-square design, with half of the targets followed by one blank space, and the other half followed by two blank spaces. To increase the number of observations we allowed for two 5-letter words to serve as targets within the same text fragment. When this was the case, one variant always had one blank space after the first target and a double after the second target; for the other variant, the order was reversed. In total, there were 35 cases of words followed by a double space and 35 matched cases of words followed by a single space.

Procedure. Before the experiment started, participants were informed that the study was about the comprehension of short texts that were displayed on a computer screen. Text administration was self-paced. Participants stopped text presentation by pressing on a button. Each passage of text was presented as a whole. Participants were asked to read at their normal speed, and to

---

5 All materials are available from the first author upon request, denis.drieghe@UGent.be
answer any comprehension question that would follow the passage. On average, questions followed on one fourth of the trials. The participants had no difficulty answering these questions, which were simple true–false statements. They were correct 87% of the time. The initial calibration of the eye-tracking system generally took approximately 10 min and consisted of a standard nine-point grid. Following the initial calibration the participant was given 10 practice trials to become familiar with the procedure before reading the experimental text fragments. The 36 experimental text fragments were embedded in a pseudo-random order in 108 filler texts. Each participant was presented with one of the two possible variants of the critical text fragments according to a Latin square design. Participants completed a single session lasting about one hour, containing 144 text fragments to read.

RESULTS

Our primary dependent variable of interest is the single fixation duration on the target word. We will also report the gaze duration on the target word as well as the number of fixations on the target word. For the word after the target word, we will report the first fixation duration and gaze duration, as well as the properties of the saccade originating from the target word and landing on the following word. These latter measurements are reported to look at the effects the extra blank space has after the eyes have left the target word. 5.4% of the data were removed from the analyses because of track loss or because the fixation was shorter than 100 ms (see Morrison, 1984; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989, for justification). From this data set, the gaze duration and number of fixations on the target word were calculated. After these analyses, a supplementary reduction of the data set was done for calculating the other measurements, by selecting only

---

6 First fixation duration on the target word will not be reported because the target word was in the vast majority of the cases fixated only once (see analysis of fixation probability and number of fixations).
those trials in which there was a single fixation on the target word followed by a forward saccade. All in all, 1473 observations (of a total of 2800) were included in this reduced data set. All analyses were run over participants ($F_1$-analyses) and items ($F_2$-analyses).

**Fixation Times on the target**

A repeated measures ANOVA was carried out on the gaze durations on the target word, which are shown in Table 1. The gaze duration on the target word followed by a double blank space was shorter than when it was followed by a single blank space. This 8 ms effect was marginally significant by participants [$F_1(1,39) = 3.39, p < .10$] and was significant by items [$F_2(1,69) = 2.95, p = .05$].

The single fixation times on the target also revealed an effect opposite to what was expected. Instead of increasing the fixation duration, an extra blank space reduced the single fixation duration on the target word. A repeated measures ANOVA revealed that this 10 ms effect was significant both by participants [$F_1(1,39) = 5.61, p < .05$] and by items [$F_2(1,69) = 7.84, p < .01$]. The effect was not due to the fact that the target word was skipped less often in the two blank spaces condition than in the single blank space condition or to the fact that the target word was refixated more often in one of the conditions. This can be seen from the number of fixations on the target word, (.77 fixations single blank space versus .75 fixations in the double blank space condition, all $F$’s < 1) and the fixation probability of the target word (.72 in the single blank space condition versus a fixation probability of .70 in the double blank space condition, $F_1(1,39) = 1.23, p > .20$; $F_2(1,69)= 2.17, p < .10$), both shown in Table 1.
Table 1: Fixation time measures (in milliseconds), number of fixations and fixation probability as a function of number of blank spaces after the target.

<table>
<thead>
<tr>
<th></th>
<th>Number of blank spaces after the target word.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 space</td>
</tr>
<tr>
<td>Gaze duration Word N</td>
<td>236</td>
</tr>
<tr>
<td>Single fixation duration Word N</td>
<td>228</td>
</tr>
<tr>
<td>Number of fixations Word N</td>
<td>.77</td>
</tr>
<tr>
<td>Fixation probability Word N</td>
<td>.72</td>
</tr>
<tr>
<td>First fixation duration Word N+1</td>
<td>218</td>
</tr>
<tr>
<td>Gaze duration Word N+1</td>
<td>241</td>
</tr>
</tbody>
</table>

In our search for variables that moderated the reduction of the single fixation duration when the target word was followed by two blank spaces, we noticed that the reduction correlated with the length of word $n+1$ [$t(68) = 2.05, p < .05$, explaining 24% of the variance]. The reduction was larger for long parafoveal words than for short parafoveal words. For instance, it was 17 ms for a 4-letter word in the parafovea, whereas it amounted to 38 ms for an 8-letter word.

Fixation times on the word following the target.

As soon as the eyes landed on the word after the target word, the extra blank space manipulation no longer exerted an effect on the fixation times. The 6 msec difference in the first fixation duration was not significant [$F1 < 1; F2(1,61) = 3.39, p > .05$], nor was there any difference in the gaze duration [$F1 < 1; F2(1,61) = 1.45, p > .20$].

Characteristics of the saccade originating from the target.
As can be seen from Figure 1, the extra blank space caused a lengthening of the saccade out of the target word by 1.2 letter positions. This effect was significant both by participants \( F_1(1,39) = 40.87, p < .001 \) and by items \( F_2(1,61) = 50.30, p < .001 \). Because the lengthening fully compensated for the extra blank space, the average landing position on word \( n+1 \) was exactly the same in both conditions, regardless of the manipulation.

![Figure 1. Landing distribution of the saccade originating from the target word (in letter positions). The letter S indicates a blank space. Left hand curve is the one blank space condition, right hand curve is the two blank spaces condition.](image)

**DISCUSSION**

According to Schiepers’s (1980) model, foveal and parafoveal words are processed in parallel but with a time delay of 90 msec per degree of eccentricity. We hypothesized that adding an extra blank space to a word would result in the eyes staying for an extra 30 ms on this word before the synchrony became jeopardized. Therefore, inflated fixation durations on the word were predicted. What we found, however, was the complete opposite: Inserting an extra blank space after a target word did not result in longer fixations on the word, but in shorter fixations. This effect was marginally significant in the gaze durations on the target word, but was significant in the single fixation times. The direction of the effect and its size are highly comparable to the results obtained in a related study by Rayner, Fischer, and Pollatsek (1998). They reported on average 12 ms shorter fixation durations in their wide spacing condition. While there are some clear differences
between both studies (Rayner et al. used a blocked presentation and three blank spaces), it is reasonable to assume that the trend of an effect observed by Rayner et al. is the same effect we observe in the present experiment. Inserting an extra blank space between the words \( n \) and \( n+1 \) causes (a) a reduction of the viewing time on word \( n \), (b) a lengthening of the saccade from word \( n \) to word \( n+1 \) by one character position to compensate for the extra blank space, and (c) no spill-over effects when the eyes land on word \( n+1 \). We shall return to these findings in the general discussion.

A further (serendipitous) finding of the present experiment was that the reduction of the viewing times in the double blank space condition seemed to be modulated by the length of the parafoveal word. The difference between a single and a double blank space was larger for long words in the parafovea than for short words. However, before we speculate about the origin of this effect, it seemed appropriate to first try to replicate it in a proper experiment. After all, in Experiment 1 the length of the parafoveal word was not manipulated and, therefore, the parafoveal word lengths were unequally distributed.

EXPERIMENT 2

Experiment 2 replicated the first experiment but manipulated the parafoveal word length. Short parafoveal words were 4-letter words; long words were 8-letter words. In addition, we created mindless reading trials in which the words were replaced by z-strings. These meaningless stimuli allowed us to assess to what extent the effect of parafoveal word was due to language processing or to low-level oculomotor control processes.

\[\text{7 From the 70 target words, 15 were followed by a 2 letter-word, 16 by a 3 letter-word, 12 by a 4 letter-word, 3 by a 5 letter-word, 6 by a 6 letter-word, 5 by a 7 letter-word, 6 by an 8 letter-word, 2 by a 9 letter-word, 2 by a 10 letter-word, 2 by an 11 letter-word and one by a 16 letter-word.}\]
The task of z-reading, in which participants are asked to “fake” reading z-strings, is not new. In a study by Vitu et al. (1995) the task was used to compare the oculomotor behavior of readers reading normal text and readers scanning meaningless materials. Based on the similarity of the eye movement patterns in both conditions, they concluded that predetermined oculomotor strategies are an important determinant of eye movement control in reading. This conclusion was questioned by subsequent research. Rayner and Fischer (1996) reported many differences between text- and z-reading at a finer level of analysis, which they took as evidence for the hypothesis that eye movement control in reading is under immediate language control. Among the differences reported were increased fixation times and skipping rates in the z-string condition.

A comparison of text-reading and z-reading allowed us to determine whether the shorter fixation durations on a target word followed by a double blank space are due to the readability of the word (as a consequence of reduced lateral inhibition), or a low-level variable related to the lay-out of the different word blobs within the sentence. In addition, a comparison of text-reading and z-reading allowed us to see whether the effect of the length of the parafoveal word is language-inspired or whether it is due to a greater pulling of long word blobs (such as the global effect, proposed by Vitu, 1991).

METHOD

Participants. Thirty-two members of the Ghent University community participated in this experiment. All participants were native speakers of Dutch and had normal or corrected vision. They were paid 10€ for their participation.

Apparatus. The apparatus was the same as in Experiment 1.
Materials. We selected 30 text fragments from the 36 used in Experiment 1. These text fragments were altered to ensure that every text fragment featured four 5-letter words, two of which were followed by a 4-letter word, and two by an 8-letter word. The 5 letter-words served as the target words of the present experiment. All the target words were located in the middle part of a line of text and none was the last or penultimate word of a sentence. For each text fragment, two variants were made according to a latin-square design. Each variant had two instances of a double blank space, equally distributed over the short and long parafoveal words. In the alternate version, the single and double spaces were swapped. After the creation of the text fragments, we doubled the stimulus set by replacing all letters in the text fragments with the letter z, hence creating an extra 30 text fragments with 2 versions that mirrored all the properties of the original text fragments, with the exception of the letter identities. Overall, 120 text fragments were created.

Procedure. The procedure was the same as in Experiment 1 with a few exceptions. Participants were notified that 30 random trials would consist of z-strings and that they were to “fake” normal reading behavior. Z-string trials were also inserted in the practice trials. The 60 experimental fragments were embedded in a pseudo-random order in 82 filler fragments, which were all meaningful texts. On average, questions followed on one fourth of the text fragments. Participants had no trouble answering these questions. They were correct 96% of the time. Participants completed a single session lasting about 50 minutes, containing 142 fragments to read (112 texts and 30 z-strings).

8All materials are available from the first author upon request, denis.drieghe@UGent.be
RESULTS

Again, our primary dependent variable of interest was the single fixation duration on the target word. We will also report the gaze duration on the target word as well as the number of fixations on the target word and the fixation probability of the target word. To examine the effects of the extra blank space in the various conditions after the eyes left the target word, the first fixation and gaze duration on the following word will also be reported, together with the characteristics of the saccade originating from the target word. 3.0 % of the data were removed due to track loss or because the fixation was shorter than 100 ms. After the analyses of the gaze duration and the number of fixations on the target, an additional reduction of the data set was carried out, selecting those trials on which there was a single fixation on the target word followed by a forward saccade. For these analyses, 3906 observations of a total of 7680 were included in the data set. All analyses were run over participants (F1-analyses) and items (F2- analyses).

Fixation times on the target

In a repeated measures ANOVA of the gaze durations on the target word with letter identity (normal vs. z-strings), parafoveal word length (4 vs. 8 letter words) and the number of blank spaces after the target (1 vs. 2) as independent variables, there was a main effect of letter identity \[ F_1(1,31) = 9.57, p < .01; \] \[ F_2(1,58) = 238.63, p < .001 \]. The gaze durations in the z-string condition were clearly longer than in the text condition (by 58 ms on average; see Table 2).

When the analysis was restricted to the z-strings, there were no further significant effects: There was no main effect of word length [all F’s < 1], no main effect of the number of blank spaces \[ F_1 < 1; F_2 (1,58) = 1.13, p > .20 \], nor an interaction between these two factors [all F’s < 1]. The situation was different in the normal reading condition. There we obtained a clear main effect both of parafoveal word length \[ F_1(1,31) = 13.20, p < .01; \]
$F_2(1,59) = 5.76, p < .05$] and of number of blank spaces [$F_1(1,31) = 15.09, p < .001; F_2(1,59) = 16.94, p < .001$]. Gaze duration was on average 12 ms shorter when the target word was followed by a long word and it was also shorter when it was followed by a double blank space (on average 13 ms). There was no interaction between these two factors [all $F$’s < 1].

<table>
<thead>
<tr>
<th></th>
<th>Letters</th>
<th></th>
<th></th>
<th>Letters</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 letter-word in parafovea</td>
<td>8 letter-word in parafovea</td>
<td>4 letter-word in parafovea</td>
<td>8 letter-word in parafovea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 space</td>
<td>228</td>
<td>231</td>
<td>214</td>
<td>286</td>
<td>283</td>
<td>289</td>
</tr>
<tr>
<td>2 spaces</td>
<td>240</td>
<td>225</td>
<td>217</td>
<td>222</td>
<td>200</td>
<td>256</td>
</tr>
<tr>
<td>Gaze duration</td>
<td>228</td>
<td>217</td>
<td>221</td>
<td>200</td>
<td>256</td>
<td>262</td>
</tr>
<tr>
<td>Single fixation Duration</td>
<td>228</td>
<td>217</td>
<td>221</td>
<td>200</td>
<td>256</td>
<td>262</td>
</tr>
<tr>
<td>Number of Fixations</td>
<td>0.74</td>
<td>0.71</td>
<td>0.69</td>
<td>0.48</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>Fixation probability</td>
<td>0.67</td>
<td>0.63</td>
<td>0.62</td>
<td>0.41</td>
<td>0.50</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 2: Fixation time measures (in milliseconds) and number of fixations as a function of letter identity, parafoveal word length and number of blank spaces after the target.

A similar picture emerged in the analyses of single fixation durations. The fixations were substantially longer in the z-string condition than in the text reading condition (on average 47 ms; $F_1(1,31) = 11.56, p < .01; F_2(1,58) = 122.74, p < .001$), but when we restricted the analyses to the z-string data no further significant effects were observed: No main effect of word length [$F_1(1,31) = 1.83, p > .10; F_2 < 1$], no main effect of the number of blank spaces [all $F$’s < 1], nor an interaction between these two variables [$F_1(1,31) = 1.43, p > .20; F_2 < 1$]. In contrast, for text reading there was a significant
effect of parafoveal word length [$F(1,31) = 14.67, p < .001; F(1,59) = 6.90, p < .05$] and a significant effect of the number of spaces on the single fixation data [$F(1,31) = 26.72, p < .001; F(1,59) = 28.97, p < .001$]. Contrary to the gaze duration data, the interaction between these 2 factors was significant by participants [$F(1,31) = 4.42, p < .05$] and marginally significant by items [$F(1,59) = 3.97, p = .051$]. As in Experiment 1 single fixation durations were shorter before a double blank space and this effect was larger when the target word was followed by an 8 letter-word (21 ms) than when it was followed by a 4 letter-word (8 ms). Single fixation times were shorter when the target word was followed by a long word, and although this effect in the single space condition was rather small (4 ms), contrasts showed that it was significant by participants [$t(32) = 2.20, p < .05$] and marginally significant by items [$t(60) = 1.89, p > .05$].

**Number of fixations on the target and fixation probability of the target.**

In the analysis of the number of fixations on the target, a repeated measures ANOVA on all three factors showed a significant main effect of letter identity [$F(1,31) = 13.34, p < .001; F(1,59) = 147.16, p < .001$]. As shown in Table 2, the number of fixations were clearly lower for the z-string conditions (.53 vs .73 fixations). When analyzed separately, the z-string data showed no effect of parafoveal word length [all $F$'s < 1], but did show a significant effect of the number of blank spaces after the target word [$F(1,31) = 12.99, p < .01; F(1,59) = 27.56, p < .001$]. An extra blank space caused the z-string target word to have a higher number of fixations, with an average increase of .10 fixations. The interaction between the parafoveal word length and number of blank spaces was not significant [all $F$'s < 1]. The analysis of the data on normal reading showed a significant effect of parafoveal word length [$F(1,31) = 10.01, p < .01; F(1,59) = 6.14, p < .05$]. When the following word was an 8-letter word, the number of fixations on the target word was on average .06 lower. In normal reading there was no effect of the number of blank spaces after the target word [$F(1,31) = 2.15, p$
The fixation probabilities of the target, as shown in Table 2, show the exact same patterns as observed in the data on the number of fixations on the target. A repeated measures ANOVA on all three factors showed a significant main effect of letter identity \( F(1,31) = 25.76, p < .001; F(1,59) = 201.68, p < .001 \). The probability of fixating the target word was lower for the z-strings (.46 vs .65). When analyzed separately, there was no effect of parafoveal word length \( F(1,31) = 2.27, p > .10; F(1, 59) < 1 \), but there was an effect of the number of blank spaces after the target word \( F(1, 31) = 17.38, p < .001; F(2, 59) = 40.13, p < .001 \). An extra blank space caused the z-string target word to be fixated more often, with an average increase of .10 in fixation probability. The interaction between these 2 factors was not significant \( all F's < 1 \). In the normal reading data there was a significant effect of parafoveal word length \( F(1,31) = 11.10, p < .01; F(2,59) = 7.52, p < .01 \). When the following word was an 8-letter word, the probability of making a fixation on the target word was on average .06 lower. There was no effect of the number of blank spaces after the target word \( F(1,31) = 1.76, p > .10; F(2,59) = 1.05, p > .20 \) and there was no interaction between these two factors \( all F's < 1 \).

**Fixation times on the word following the target.**

When we restricted the data set to those cases in which a single fixation on the target word was followed by a fixation on the following word, we ended up with a large number of empty cells for the z-strings. A fixation on the next word followed in 20% of the trials only. Therefore we did not further analyze the data of the z-strings. A repeated measures ANOVA was carried out on the first fixation data in the normal reading condition, as shown in Table 3. The main effect of word length was marginally significant by participants \( F(1,28) = 3.45, p < .10 \) but not by items \( F(2 < 1 \). The effect of the number of blank spaces was marginally significant by participants...
[\(F_1(1,28) = 3.01, p < .10\)] and was significant by items [\(F_2(1,50) = 4.44, p < .05\)]. This was due to a significant difference between an 8 letter-word that followed a single blank space and an 8 letter-word that followed a double blank space [\(t_1(31) = -2.25, p < .05; t_2(58) = -2.14, p < .05\)], the latter showing a longer first fixation duration. The overall interaction between word length and the number of blank spaces was not significant [\(F_1(1,28) = 1.62, p > .20; F_2 < 1\)].

<table>
<thead>
<tr>
<th>Characteristics of the saccade originating from the target.</th>
</tr>
</thead>
</table>
| For the gaze duration data, there was a significant effect of word length by participants [\(F_1(1,28) = 6.13, p < .05; F_2(1,50) = 2.73, p > .10\)]. If the parafoveal word was an 8 letter-word gaze duration was on average 15 ms longer. There was no significant main effect of the number of blank spaces by participants [\(F_1(1,28) = 1.99, p > .10\)] but there was by items [\(F_2(1,50) = 5.33, p < .05\)]. After a double blank space gaze duration was on average 8 ms longer. There was no interaction between word length and the number of blank spaces [all \(F^\prime s < 1\)].

### Table 3: Fixation time measures (in milliseconds) on the word following the target word as a function of parafoveal word length and number of blank spaces after the target.

<table>
<thead>
<tr>
<th></th>
<th>4 letter-word in parafovea</th>
<th>8 letter-word in parafovea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 space</td>
<td>2 spaces</td>
</tr>
<tr>
<td>First fixation duration</td>
<td>212</td>
<td>215</td>
</tr>
<tr>
<td>Gaze duration</td>
<td>221</td>
<td>226</td>
</tr>
</tbody>
</table>

The data on the characteristics of the saccade originating from the target word and landing on the following word are shown in Figure 2. There was a significant effect of parafoveal word length on the saccade length [\(F_1(1,28) = 48.28, p < .001; F_2(1,50) = 56.64, p < .001\]: the saccade was 1.4 character
positions longer when landing into an 8 letter-word. The main effect of the number of blank spaces was also significant \( F(1,28) = 29.31, p < .001; 
F(1,50) = 56.32, p < .001 \). An extra blank space caused a lengthening of the saccade by 0.84 character positions. As in Experiment 1, the lengthening compensated for the extra blank space, making the average landing position on word \( n+1 \) almost identical. There was no interaction between parafoveal word length and the number of blank spaces after the target [all \( F \)'s < 1].

![Figure 2](image-url)

**DISCUSSION**

Experiment 2 replicated the finding of Experiment 1 that in text reading fixation durations on target words are shorter when the word is
followed by a double blank space than when it is followed by a single blank space. This effect was present both in the single fixation durations and in the gaze durations. Shorter fixation times were not observed in z-string reading, a finding that seems to support a reduced lateral masking interpretation. The z-string data replicated the basic findings that were reported for these materials before: Longer fixation times and more word skipping were observed than in normal text reading (e.g., Vitu et al., 1995; Rayner, & Fischer, 1996).

The data on the number of fixations on the target word and the fixation probability of the target word were also in line with those observed in Experiment 1: Adding a blank space between two words did not increase the probability of the first word being fixated. This contradicts predictions one of the authors previously made in the Extended Optimal Viewing Position model of word skipping (Brysbaert & Vitu, 1998). According to this model, word skipping for word \( n+1 \) depends on the length of word \( n+1 \) and the distance of word \( n+2 \) from the fixation location. Adding a space between word \( n+1 \) and word \( n+2 \) should increase the probability of fixating word \( n+1 \) (because word \( n+2 \) is farther away). Interestingly, this effect was observed when participants were reading meaningless z-strings: Chances of fixating a target word were 10% higher when there were two blank spaces after the word than when there was only one (see Table 2). So, whereas a double blank space had a significant influence on skipping rates obtained in the z-string data, in normal reading this manipulation had no effect on the skipping data. In short, this is a strong indication that not all word skipping in text reading is due to oculomotor factors.

Another dissociation between z-reading and text reading was found in the effect of the length of the parafoveal word. Whereas a long parafoveal word \( n+1 \) decreased the gaze duration on a 5-letter foveal word \( n \) and increased the likelihood of skipping the word \( n \), no such effect was observed for z-reading. This is a very interesting observation, because one of the interpretations of the parafoveal length effect has been that a long parafoveal
word pulls the landing position towards its center of gravity (i.e., the so-called global effect; Vitu, 1991; Gautier, O’Regan, & Le Gargasson, 2000). However, in that case we should have observed a similar effect in z-reading. The fact that the effect was not observed in z-reading is more in line with Kennedy’s (1998) conjecture that in text reading the eyes are pulled towards the region with the highest information (assuming that long words on average are more informative than short words). An alternative interpretation could be that z-reading, because of its longer fixations and saccades, is less influenced by the global effect than normal text reading.

The fixation times on the word after the target word showed the standard word length effect (Calvo & Meseguer, 2002; Rayner & Fischer, 1996; Rayner et al., 1996), and an effect of the number of blank spaces, mostly due to a longer fixation time on an 8-letter word preceded by a double blank space. The latter effect could be expected based on the reduced processing (shorter fixation durations, more skipping) of the previous word in the double-space condition. Both of these factors contribute to reducing the parafoveal preview benefit.

Finally, the data on the saccade originating from the target word and landing on the following word are also highly compatible with the data obtained in Experiment 1: The extra blank space was fully compensated by lengthening the saccade by approximately one character position, hence landing on the same site. The landing distributions in Figure 2 also show that it is not very meaningful to compare the average landing position for short and long parafoveal words, because in the former condition we clearly got a truncated distribution, with many saccades aimed at the word after the short parafoveal word.
GENERAL DISCUSSION

The possibility of parafoveal-on-foveal effects in eye movement control has become a major issue in recent research on eye movements in reading, because researchers see it as the critical test to determine whether the words in a line of text are processed one by one, or whether two or more words are being processed in parallel. According to the first view, the human visual attention system is able to limit word processing in text reading to one word at a time (i.e., there is an early selection of information). The most elaborate and detailed model of this type is the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2003; Rayner, Reichle, & Pollatsek, 1998; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, 2003). One of the core assumptions of the model is that attention covertly shifts from word to word. Only the word within the attentional beam is being processed, and the beam does not shift to the next word until full identification (or close to full identification) of the currently fixated word has been obtained. Words are processed serially because it is important for readers to keep the word order straight (Pollatsek & Rayner, 1999). The “leave-on-completion” assumption of the model can account for foveal-on-parafoveal effects (as reported by Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Schroyens et al., 1999), but does not predict parafoveal-on-foveal effects other than the extra time needed to replace a cancelled forward saccade to word \( n+1 \) by a new saccade to word \( n+2 \) (see the introduction).

However, as discussed in the introduction, in recent years a number of parafoveal-on-foveal effects have been published that seem to raise the possibility of parallel word processing in reading (Hyönä & Bertram, 2004; Inhoff, Radach, Starr, & Greenberg, 2000; Inhoff, Starr, & Shindler, 2000; Kennedy, 1998; 2000; Kennedy et al., 2002; Schroyens et al., 1999; Starr & Inhoff, 2004; Underwood, Binns, & Walker, 2000; Vitu et al., 2004). According to some (Radach, & Kennedy, 2004), the evidence now is so strong that we no longer have to question whether such effects exist but how
we can understand them, whereas others remain more cautious (Rayner & Juhasz, 2004; Rayner, White, Kambe, Miller, & Liversedge, 2003).

The discovery of parafoveal-on-foveal effects has been accompanied by the development of alternative models of eye movement control, all embracing a parallel view on foveal and parafoveal word processing (with late selection of information). The SWIFT model (Engbert, Longtin, & Kliegl, 2002; Kliegl & Engbert, 2003), for instance, adopts many of the architectural features of the E-Z Reader model, but departs from it by assuming a parallel, spatially distributed lexical processing. The Glenmore model (Reilly, & Radach, 2003) is an even more radical departure from the attention based, sequential processing models by replacing the entire concept of attention by a saliency map, based on the highly influential model by Findlay and Walker (1999). A similar view is defended in Yang and McConkie’s (2001) competition-inhibition model, which is also based on the Findlay and Walker model and which puts a very strong emphasis on non-cognitive factors to explain eye movements in reading.

A weakness of the available evidence on parafoveal-on-foveal effects, however, is that it has not yet been framed within a coherent model that allows researchers to predict which effect will be obtained when and why (Rayner & Juhasz, 2004). This is even more a problem because the effects are not always pointing in the same direction (see e.g., Hyönä & Bertram, 2004). In this paper, we set out to directly test a basic assumption of one coherent set of ideas that has been put forward and that recently has been referred to a number of times. According to Schiepers’s (1980) model, foveal and parafoveal words are processed in parallel but with a time delay of 90 msec per degree of eccentricity. In this model, the parafoveal preview benefit is not due to the fact that the attentional system already partly processed the parafoveal word by the time the eyes reach this word (as defended by E-Z Reader), but to the fact that the activation of word representations is boosted when the foveal information from fixation $n$ can be combined with the parafoveal information from fixation $n-1$. This
combination of information from different fixations critically depends on the synchrony with which the activation arrives in the relevant brain centers. Based on this assumption, we hypothesized that adding an extra blank space to a word would allow the eyes to stay for an extra 30 ms on this word before the synchrony became jeopardized. Therefore, an inflated fixation duration on the word was predicted. What we found, however, was the complete opposite: Inserting an extra blank space after a target word did not result in longer fixations on the word, but in shorter fixations.

Although we failed to find direct evidence for the Schiepers model, we did obtain evidence for parafoveal-on-foveal influences. There were three such influences. First, the fixation durations on the target words were not similar in the two-space condition as in the single-space condition; they were significantly shorter. Second, we found an effect of parafoveal word length with a longer word in the parafovea leading to shorter fixation durations and slightly less fixations on the prior word. And third, the effect of the double blank space was modulated by the length of the parafoveal word; the reduction in the single fixation time on the target due to the double blank space tended to be larger when the following word was an 8 letter-word than when it was a 4 letter-word. Interestingly, none of these effects were observed when we asked participants to mimic reading behavior when presented with z-strings. This strongly suggests that the effects we observed are not due to low-level oculomotor variables related to the length and the lay-out of the word blobs, otherwise we would have found the same effects in the z-scanning task.

We will start our discussion with the first finding, the reduced fixation duration prior to a double space. The fact that there was no similar effect in z-reading indicates that the origin of the effect is likely to be language related. The simplest explanation probably is reduced lateral masking of the letters in the double space condition, a phenomenon that would have no repercussions on the task of scanning z-strings. This processing advantage leads to faster word recognition with hardly any repercussions for the
Our second finding concerns the effect of parafoveal word length on viewing times: A long word $n+1$ in the parafovea leads to a shorter viewing time on word $n$. The effect of parafoveal word length was first reported by Kennedy (1998). In his experiment participants first viewed a fixation marker after which three words were presented on the screen. The first word was either the word looks or the word means. In the looks case participants had to indicate whether the two following words had the same spelling, in the means case participants had to indicate whether they had the same meaning. Kennedy concluded from his results that parafoveal word length acted to modify foveal inspection time, resulting in a shorter foveal fixation time in the case of a longer second word. A replication of the experiment using a task closer to normal reading (Kennedy, 2000, Experiment 2) also found this effect of parafoveal word length. In this task participants had to read strings of unrelated words, looking for rare occurrences of an article of clothing (see also Schroyens et al., 1999). Although this task was clearly closer to normal reading as compared to the previously used looks-means task, the generalizibility of the results to normal reading is still somewhat disputed (Rayner et al., 2003). An effect of parafoveal word length was also observed in a large data corpus of normal reading containing the eye movements of four German-speaking students reading the first two parts of Gulliver Travels (Radach, 1996). Kennedy (1998) further reported in this corpus an effect of parafoveal word length on the fixation durations of the foveal word: A long parafoveal word was associated with shorter single fixation durations on the foveal word. For a 5 to 8 letter foveal word for instance, the single fixation duration ranged from an average of 287 msec in the case of a 4 letter parafoveal word to an average of 274 msec in the case of 7 to 10 letter parafoveal word. Also in normal reading Hyönä and Bertram (2004, Experiment 2) reported a similar effect of parafoveal word length in Finnish. The parafoveal words they used consisted of a set of short (7-9
letters) and long (12-15 letters) compound words. In their experiment the targets preceding long compounds received a shorter gaze duration than those preceding short compounds. Hyönä and Bertram also interpreted this finding in terms of long parafoveal words attracting an early saccade towards them, but they were unable to replicate the finding in a follow-up experiment (2004, Experiment 4).

In a parallel processing model such as the one proposed by Kennedy (1998) the harder the word \( n+1 \) is to process, the stronger it pulls the eyes towards it, in order to optimize the extraction of visual information from the page of text. Such a mechanism could explain the effect parafoveal word length had on our fixation times on the target. The question remains however whether the attraction that the longer word in the parafovea exerts, finds it origin in processing difficulties associated with longer words. An alternative hypothesis comes to mind. The attraction of parafoveal word length could just be a consequence of a strategy that tries to distribute the fixation locations in the most efficient way, landing more on long words and skipping shorter words. If such a strategy exists, it is not inconceivable that it results in an attraction, a pulling force, if a very suitable candidate is close-by. An extra blank space prior to it could make the candidate stand out more, which would explain our third finding, why the parafoveal word length effect was larger in the double space condition than in the single space condition. The major difference between the mechanism described in the alternative hypothesis and the one proposed by Kennedy is that the alternative hypothesis does not assume that the parafoveal attraction is based on word processing in the parafovea. The only variable it requires is word length.

At this point, it is important to note that we see the explanation for the observed patterns in the data of the current study as a combination of two effects. The shorter fixation duration prior to a double blank space is due to a reduction of lateral inhibition, increasing the readability of the following word. The effect of parafoveal word length is explained by an attraction
exerted by long words resulting in a pulling force closely related to the ideas proposed by Kennedy (1998), although the present proposition downplays the original assumptions. Neither of these two influences can individually account for all the effects observed in the present study. A reduced fixation duration prior to a longer word can not be expected solely based on an reduced lateral inhibition hypothesis. Likewise, there is no reason to predict a reduced fixation duration prior to a double blank space based on the pulling force account. However, a double blank space could boost the saliency of a long word, resulting in the observed interaction between the double blank space manipulation and the effect of parafoveal word length.

Finally, it has to be acknowledged that the parafoveal-on-foveal effects unraveled in the present experiments, do not look very damaging for the serial assumption of the E-Z Reader model either. A distinction has to be made between the rather low-level parafoveal-on-foveal effects reported here and effects such as for instance the meaning of the word to the right of the fixation influencing the current fixation. Better visibility of a word due to less lateral interference is not incompatible with the principles underlying E-Z Reader. The same may be true for the effect of the length of the parafoveal word. Although E-Z Reader in our view underestimates the effect of word length in inter-word eye movement control (Brysbaert & Drieghe, 2003), in the latest version of the model (Reichle, Rayner, & Pollatsek, 2003) a pre-attention stage of processing has been incorporated allowing information about word length to be extracted prior to the shift of attention. While this recent adaptation was not specifically constructed for accounting for the effects reported above, it might offer an explanation for them (Rayner et al., 2003).

Indeed, one of the most striking results of the present experiment is the apparent ease with which the participants dealt with the breach in the spacing protocol. With the exception of the shortened fixation durations, the double blank space caused hardly any noticeable signs of changed eye movement behavior. Only in the condition in which a double blank space
preceded a long word did a clear effect of reduced parafoveal preview emerge. There was a swift adaptation of the outgoing saccade so that the landing position on the parafoveal word was the same in the double space condition as in the single space condition. This, incidentally, is a very clear demonstration of the fact that eye movements are determined by the visual lay-out of the text to be read, and are not selected at random from a distribution of possible saccade sizes (as has recently been suggested by McConkie (personal communication) whilst reviewing Brysbaert et al., in press). The participants were not aware of the space manipulation. About one third of them were asked after the experiments whether they had noticed anything unusual about the text fragments they had read, and none reported the occasional double spacing. A potential reason for not noticing the manipulation could be that it is altogether not such an uncommon phenomenon. We are not aware of any study reporting the frequency of unintended double spaces in normal texts, but from personal experience we can say that once one starts to pay attention to the phenomenon, an unintended double spacing in for instance e-mails does appear quite often. Another argument for the flexibility of readers to deal with changed spacing could be the common use of justified fonts, an option in most modern text editors, which also requires a swift adaptation from the reader in terms of adjusting to different letter sizes and spacing.

All in all, in what started as a direct test of a core assumption of the Schiepers (1980) model, our main conclusion must be that the model failed to make the correct prediction. On the basis of the present evidence, we cannot conclude that the fixation (and the gaze) duration on a word is the result of two forces: (1) the need to process the foveal word, and (2) the need to synchronize the parafoveal information from the current fixation with the foveal information from the next fixation. As a matter of fact, our results went reliably in the opposite direction. Therefore, we feel that Schiepers’s ideas can no longer be used as the basis for a parallel model of eye movement control in reading. What we did find was that an extra blank space speeds up the reading process, presumably due to a reduced lateral
masking. An effect of parafoveal word length was also reported, a long word leading to shorter fixation times and a fewer number of fixations on the previous word. This latter finding has been interpreted as a pulling force exerted by longer words, possibly resulting from a strategy to distribute fixations in text in the most efficient manner.
REFERENCES


Brysbaert, M., & Drieghe, D. (2003). Please stop using word frequency data that are likely to be word length effects in disguise. Behavioral and Brain Sciences, 26, 479.


CHAPTER 3
WORD SKIPPING IN READING: ON THE INTERPLAY OF LINGUISTIC AND VISUAL FACTORS.


An eye movement experiment is reported in which target words of two and four letters were presented in sentences that strongly raised the expectation of a particular word. There were three possible conditions: The expected word was present in the sentence, an unexpected word of the same length was present, or an unexpected word of a different length was present (all continuations were acceptable, but the latter two were difficult to predict). Our first purpose was to test one of the core assumptions of the Extended Optimal Viewing Position model of eye guidance in reading (Brysbaert & Vitu, 1998). This model states that word skipping is primarily a function of the length of the upcoming word. It leads to the prediction that an unpredicted two-letter word will be skipped more often than a predicted four-letter word, which is indeed what we observed. Our second aim was to determine if we could obtain an interaction between context predictability and parafoveal word length, by looking at what happens when the length of the parafoveal word does not agree with the length of the expected word. No such interaction was observed although the effects of both word length and predictability were substantial. These findings are interpreted as evidence for the hypothesis that visual and language-related factors independently affect word skipping.

1 This paper is co-authored by Marc Brysbaert, Timothy Desmet and Constantijn De Baecke.
2 We thank Keith Rayner, Françoise Vitu, Sarah White and an anonymous reviewer for the many helpful comments on earlier drafts of this paper.
INTRODUCTION

One of the current controversies in research on eye movements in reading concerns the kind of information extracted from parafoveal vision and the ways in which this information influences subsequent eye movements (Starr & Rayner, 2001). A large body of research indicates that in reading information is extracted from the word next to the currently fixated word. This parafoveal preview benefit is easily shown by comparing conditions in which the parafoveal word is visible with conditions in which it is masked until the reader’s eyes land on it (e.g. Blanchard, Pollatsek, & Rayner, 1989; Morris, Rayner, & Pollatsek, 1990; Rayner, 1975; Rayner, Well, Pollatsek, & Bertera, 1982; Schroyens, Vitu, Brysbaert, & d’Ydewalle, 1999). A major issue, however, is the extent to which the parafoveal information determines the length of the subsequent forward saccade out of the fixated word. In the present manuscript, we will deal with the probability of fixating the next parafoveal word only, although a very similar discussion exists about the landing position within the parafoveal word.

There are two main types of factors that can influence the length of a forward saccade: low-level, visuomotor variables and high-level, language-related variables. Visuomotor factors refer to the visual characteristics of the text and to limitations in the planning and execution of eye movements. The most prominent visual characteristics related to a text are the length of the foveal and the parafoveal word and the distance of the eyes from the parafoveal word (the so-called launch site). It is well-established that word skipping occurs more frequently with short words than with long words (Brysbaert, & Vitu, 1998; Rayner, 1979; Rayner & McConkie, 1976; Vitu, O’Regan, Inhoff, & Topolski, 1995), and that it occurs more often when the previous fixation was close to the parafoveal word than when it was far away (Kerr, 1992; Rayner, Sereno, & Raney, 1996; Vitu et al., 1995). The oculomotor limitations refer to the fact that in general there is some error
between the intended landing position and the actual landing position. This error has a systematic component (such as the tendency to undershoot far targets) and a random component (making the landing distribution over many trials looks like a normal distribution). Data from McConkie, Kerr, Reddix, and Zola (1988; see also McConkie, Kerr, & Dyre, 1994) nicely illustrate the importance of visuomotor factors in word skipping. McConkie et al. asked participants to read a long text, in order that they could analyse a large corpus of data. When the landing sites were plotted as a function of the launch site, there were systematic differences in the mean and the standard deviation of the Gaussian landing site distributions. Each time the saccade originated from a launch site one character position further away from the beginning of the target word, the mean of the landing position distribution was shifted leftward by about one third of a character and the variance of the distribution increased. To explain the shift of the landing distribution as a function of the launch site, McConkie et al. suggested the existence of a range effect. The oculomotor systems tends to undershoot targets at a large eccentricity and to overshoot targets at a small eccentricity. Another oculomotor limitation that has been proposed is the so-called global effect (Vitu, 1991; Gautier, O’Regan, & Le Gargasson, 2000). When the eyes make a saccade to a target word, the movement is influenced not only by the visual characteristics of the target word but also by the surrounding stimulus materials. This influence causes the eyes to be deviated away from the target, towards a cortically weighted centre of gravity defined by the global visual configuration surrounding the target. Applied to skipping this means that sometimes a parafoveal word will be skipped erroneously because the centre of gravity lies behind the word, or that it will be impossible to skip the word because the centre of gravity coincides with the word. Findings such as these indicate that any comprehensive theory of word skipping has to take into account involuntary word skipping or word fixation because of oculomotor error. Visuomotor factors also explain why word skipping is observed in studies where it is impossible to use language information from the parafoveal word because this word is masked until the eyes cross an
In addition to the low-level visuomotor factors, linguistic variables are also known to influence the probability of word skipping. A distinction can be made between characteristics at the word level (such as word frequency and visibility of the parafoveal word) and characteristics at the sentence or discourse level (such as word predictability from the context). The strongest language-related influence is the effect of contextual constraints: Words that are highly predictable from the preceding context are skipped more often than words that are not constrained (Altarriba, Kroll, Sholl, & Rayner, 1996; Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981; Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987). Contextual constraint is typically assessed with a sentence completion task in which subjects are given a sentence fragment up to the target word and are asked to write down the first word that comes to mind. Words that are produced by many participants are considered to be highly constrained. Effects of word-related variables on skipping have also been observed: High-frequency words are more likely to be skipped than low-frequency words, especially when the eyes are close to the target word on the fixation prior to the skipping (Henderson & Ferreira, 1993; Radach, & Kempe, 1993; Rayner, & Fischer, 1996; Rayner, Sereno, & Raney, 1996).

Brysbaert and Vitu (1998) were among the first to look at the effects of both types of variables together. A reason why visuomotor and linguistic variables had rarely been examined simultaneously until then is that both types of variables are highly correlated, making them difficult to disentangle. In general, high-frequency, familiar (i.e., easy) words tend to be shorter than low-frequency, unfamiliar words. To determine the relative importance of word length and word processing difficulty on word skipping probability, Brysbaert and Vitu ran a meta-analysis of all the studies that had manipulated the processing load of target words (by changing the word frequency or the word visibility in the parafovea) and that included
information about the length of the target words. Averaged over all the studies (and the different conditions within the studies), Brysbaert and Vitu found a consistent 4% difference in skipping rate between the easy and the difficult words, confirming the claim that language-related factors have an influence on the probability of word skipping. However, at the same time they observed a strong effect of word length. Skipping rate ranged from over 50% for two-letter target words to some 1% for nine-letter words, and this effect could be captured quite well with an exponential function. In a second analysis, they repeated the exercise for those studies that had manipulated the context predictability of the target words and had reported information about the length of the words. The effect of context predictability on skipping rate was slightly larger than that of the word-related variables and amounted to a 9% difference between predicted and unpredicted words. The impact of word length was exactly the same as in the word-level studies. Because of the importance of these findings for the rest of the present article, they are illustrated in Figure 1 (taken from Brysbaert & Vitu, 1998).

![Fig. 1. Skipping rate as a function of word length and contextual constraint (circle = predictable conditions; square = neutral conditions). Fitted curve based on nonlinear regression with $\exp(\text{word length})$ and contextual constraint as predictors. The upper curve represents the best fit for the predictable words; the lower curve is the best fit for the neutral words (reprinted from Brysbaert & Vitu, 1998).](image)

On the basis of these findings, Brysbaert and Vitu (1998) concluded that it was more informative to know the length of the parafoveal word in order to
predict word skipping probability than to know the ease of processing, in other worlds short words are skipped mainly because they are short and not because they are easy to process. To explain how word skipping could be based on the length of the parafoveal word, Brysbaert and Vitu turned to a finding they had reported before (Brysbaert, Vitu, & Schroyens, 1996). In this study, Brysbaert et al. (1996) had found that the probability of word recognition given a certain presentation duration and distance between the target word and the fixation location, could reasonably well be described by a Gaussian distribution that had the mode shifted slightly to the left of the word centre. Figure 2 depicts their finding. Brysbaert et al. (1996) called this effect, by analogy to the previously defined optimal viewing position (e.g. O’Regan & Jacobs, 1992; O’Regan, Levy-Schoen, Pynte, & Brugaillère, 1984), the Extended Optimal Viewing Position (EOVP) effect.

Fig. 2. Probability of word recognition as a function of presentation duration and word position relative to fixation location. Empirical data and best fitting Gaussian distributions (reprinted from Brysbaert et al., 1996).

Brysbaert and Vitu (1998) hypothesised that the EOVP information could be used by the eye guidance system to estimate the chances of identifying the upcoming parafoveal words within the time period of an average fixation
(200 – 220 ms), and to select the most appropriate parafoveal target word. The estimates are based upon (i) the length of the word blobs\(^3\) and the distance of the word blobs from the fixation location, and (ii) the standard deviation of the Gaussian EOVP curve. The latter depends on text difficulty and task demands, so that the spread of the Gaussian curve will be larger for easy texts and for cursory reading. The decision to skip a word is viewed here as an educated guess based on coarse visual information which becomes available rather early in the fixation.

Linguistic influences on skipping behaviour were included in the EOVP model as follows. Although the system disposes of an identification probability for the parafoveal word on the basis of word length and the eccentricity of the word, it still has to “decide” which word to pick as the most suitable parafoveal target word. This decision is taken by a system related to the discourse processing on the basis of partial information (i.e., before the word is fully recognised), or as Brysbaert and Vitu (1998, p. 142) phrased it:

“According to [other] theories, a word is skipped because it was recognized during the previous fixation. According to our view, a word is skipped because the language system estimates chances high enough that it (i.e. the parafoveal word) will be identified by the end of the current fixation or, at least, that bypassing the word will not hinder text understanding...”

Brysbaert and Vitu (1998) further assumed that occasionally the initial decision could be overruled by the processing rate of the parafoveal word. There are two ways in which this can happen. First, when the processing of the parafoveal word is easier than anticipated, a planned saccade to this word can be cancelled and replaced by a saccade to the next word. Alternatively, when processing of the parafoveal word turns out to be more difficult than

\(^3\) Brysbaert and Vitu (1998) use this term for the parafoveal word to stress the limited visibility of the words in parafoveal vision.
anticipated, a planned skip can be cancelled (or shortened) so that the parafoveal word is fixated after all. Brysbaert and Vitu ventured that the latter situation could be more frequent, as it may be easier to shorten the saccade of a planned skip than to lengthen the saccade of a planned fixation. New data on saccade generation during reading (Yang & McConkie, 2001) corroborate this idea that language processing influences eye movements not by directing each and every saccade but by interfering with an ongoing, more basic oculomotor process. Following Findlay and Walker (1999), Yang and McConkie (2001) hypothesised that saccades are initiated by a rhythmical sequence, on which cognitive interventions now and then exercise an inhibitory control. Although Yang and McConkie's competition – inhibition theory focused on fixation durations, they also claimed that processing-related inhibition not only affects the onset time of the next saccade but also its length.

Because interword eye movements are guided primarily on the basis of word length, the EOVP model of word skipping is closer to the oculomotor family of theories than to the cognitive-processing family of theories. So, how does the latter group deal with the issue of word skipping? We will only discuss the E–Z Reader model (Rayner, Reichle, & Pollatsek, 1998; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, in press), which at present is the most elaborated and detailed model. According to the E-Z Reader model, word recognition in reading is a serial process under the control of an attentional beam (i.e., only the word in the attention beam is processed). The eyes follow the shift of the attentional beam with a certain delay due to the eye movement programming time. In E-Z Reader, word skipping is based on the following sequence of events. The programming of an eye movement starts as soon as the word processing system reaches a stage from which word identification becomes likely.

---

4 This idea was originally formulated by Morrison (1984) and also features in the ideas of Henderson and Ferreira (1990).
Reichle and colleagues called this the familiarity check. At this stage, a word is not yet fully recognised, but the dynamics in the lexicon are such that it is likely to become so within a limited time period. When the familiarity check for the currently fixated word $n$ has occurred, the eye guidance system starts to program a saccade to the next word $n+1$. Visual attention and the eyes remain on the foveal word until it is completely processed. Upon full identification of the word $n$, attention shifts to word $n+1$ and the eyes are expected to follow as soon as the eye movement programming is completed. If, however, in the mean time the familiarity check of word $n+1$ happens and if the programming of the initial eye movement has not yet reached its final ballistic stage, then the eye movement program towards word $n+1$ can be cancelled and replaced by a new program to word $n+2$. In this situation, skipping of word $n+1$ will take place. Because of the above sequence of events, it is the processing ease of word $n+1$ that influences the likelihood of it being skipped and this ease is determined by word frequency and contextual predictability. In addition to language variables, visual factors such as word length and launch site have a role in the E-Z Reader model as well, because they limit the amount of early visual processing (in ms) that is completed during a fixation. The E-Z Reader model assumes an inverse relation between the extraction of letter information and the distance of the letter from the centre of the visual field hence. Hence, a longer word on the right of the foveal word will decrease the probability that the programmed eye movement to word $n+1$ is replaced by an eye movement to word $n+2$.

So, both the EOVP model and the E-Z Reader model incorporate visual as well as language-related factors to calculate the probability that a parafoveal

---

5 To test the viability of the E-Z Reader model, Reichle et al. (1998) compared the results of the model on a set of sentences with human eye movement data that had been gathered before. For each word of the sentences context predictability was assessed with the use of the sentence completion task. A similar approach was followed to test another recent computational model of eye guidance in reading, the so-called SWIFT model (Engbert, Longtin, & Kliegl, 2002).
word will be skipped. However, there are two main differences. First, EOVP puts the primacy on the visual variables of word length and launch site (see Figure 1), whereas in the mathematical elaboration E-Z Reader sees a primary role for word frequency and context predictability. The second difference between EOVP and E-Z Reader is that in the latter model word skipping nearly always occurs because the parafoveal word was processed. In the EOVP model, the difference between easy and difficult words is hypothesised to be due more to cancelled skips for difficult words than to extra skips for easy words. The latter agrees with the finding that word skipping is often followed by a short fixation duration and a regressive eye movement (Brysbaert & Vitu, 1998; Vitu & McConkie, 2000), as if the oculomotor system by the time of the skip received a signal that the initial educated guess was wrong.

The present experiment was designed to further test what happens when the variables word length and contextual constraint are manipulated simultaneously. Given that the effect of contextual constraint on the probability of word skipping is stronger than the effect of word frequency, it seems worthwhile to verify whether word length in this situation still plays the basic role, as claimed by the EOVP model, or whether contextual constraints (which build up over several words and can be fed to the language system before the target word is encountered) will be playing the leading part, as claimed by the E-Z Reader model. In Brysbaert and Vitu's (1998) meta-analysis (Figure 1), the word length variable was largely a between-experiments variable (i.e., the data for the different word lengths came from different articles), whereas the contextual constraints variable was a repeated measure (i.e., was obtained within a single experiment with the same participants reading high- and low predictable words in a random sequence). Needless to say, this opens the possibility that some of the word length effect was due to confounded variables.

A second, related question is whether cross-talk is possible between the EOVP-based information system and the language processor. Basically,
what the EOVP system provides is information about the length of the upcoming word blob. There is, however, one situation in which this information could be very useful to the language system, namely when contextual constraints well in advance predict the length of the upcoming word. Suppose you get the following sentence (from Rayner et al., 2001) “Most cowboys know how to ride a … if necessary.” Virtually all (American) readers expect the five-letter word “horse” to appear after the sequence "to ride a". Fewer expect the five-letter word “llama” or the twelve-letter word “hippopotamus”. However, the difference between both unexpected continuations is that the first (llama) agrees with the predicted word length, whereas the second (hippopotamus) clearly violates the expectation. Will the system be fooled by the fact that the length-preserving intruder strongly resembles the expected word blob, or will it make no difference whether the intruder is of a different length as the expected word?

To investigate our research questions, we constructed two sets of sentences. The first set consisted of sentences in which a 2-letter word was strongly expected on the basis of the preceding sentence context, and we presented either this 2-letter word, or an unexpected but still acceptable 2-letter word, or an unexpected but acceptable 4-letter word. We also constructed a second set of sentences in which a 4-letter word was expected from the context and we presented this word, an intruder of the same length, or an intruder of 2 letters. The combination of these 6 conditions allowed us to investigate the effects of word length and predictability separately (by comparing the skipping rates for the different lengths and predictability ratios while the other variable was held constant). It also allowed us to see whether an unexpected 2- or 4-letter word is more likely to be fixated when in addition to being unpredictable from the context, it also violates the length expectations (we could achieve this by comparing the skipping rates for unexpected 2-letter words when a 2-letter word was expected vs. when a 4-letter word was expected, and by comparing the skipping rates for unexpected 4-letter words when a 4-letter word was expected vs. when a 2-letter word was expected).
To obtain a significant effect of contextual constraints on skipping behaviour, it is necessary to ensure that the contextual manipulation is powerful enough (Hyönä, 1993; Rayner, & Well, 1996). Therefore, we took great care to construct stimulus materials that very strongly raised the expectation of a particular word. In addition, we worked with short words that could easily be identified in the parafovea, and with a kind of stimuli that allowed predictions as low as the lexical level\(^6\). This kind of optimal stimuli is provided by the separable verbs in Dutch. These verbs consist of two parts that can be written quite far apart from one another. For example, the verb *opvallen* (to attract attention), consists of two parts: a particle (*op*) and a non-separable verb (*vallen*), which in itself has another meaning (i.e., *to fall*). Because the meaning of separable verbs can differ from that of the non-separable verb part, it is generally assumed that the lexicon temporarily has to keep several different word forms activated (Kempen, 1995; Schreuder, 1990). Otherwise, it is difficult to understand why Dutch-speaking people have no difficulties understanding the sentence "Door zijn oranje haarbos viel de punker op." (literally: Because of his orange hair fell the punker up [in reality: Because of his orange hair the punk attracted a lot of attention.]).

\(^6\) Hyönä (1993) did not find a significant difference in skipping rate due to contextual constraints. He attributed this to the possibility that the difference in sentence completion ratios in his stimulus materials (32 % low predictable versus 65 % highly predictable) was not large enough. Recently, however, an alternative explanation has been put forward (Calvo, & Meseguer, 2002; Calvo, Meseguer, & Carreiras, 2001). Calvo et al. suggest that a distinction should be made between contextual predictability based on associative priming among the various words in the sentence (e.g. the word *wedding* priming the word *cake*), and contextual predictability based on more elaborative inferences at the discourse level. The former could affect early processing stages such as skipping probability, whereas the latter only affects later processing stages. Calvo et al. suggest that the stimuli used by Hyönä may have belonged to the latter category.

\(^7\) *viel* [fell] is the simple past singular tense of the verb *vallen* [to fall].
The particles of separable Dutch verbs were used in our experiment as the target words and the non-separable verb always preceded the target by several words. This gave us two advantages. The first is of a practical nature; it is relatively easy to construct sentences in Dutch for which the particle is strongly expected from the context. The second reason is that the lexical node representing the separable verb is likely to be activated during the fixation prior to potential skipping, thus maximizing the opportunity for linguistic influences to encourage skipping. Because there is some evidence that context effects are stronger when the contextual constraints have been built up for several sentences (Ehrlich & Rayner, 1981) we also chose to use text passages that filled five lines of text on the screen. The target appeared about halfway along the fourth line. An example sentence with a 2 letter predictable word is shown in table 1, one with a 4 letter predictable word is shown in table 2.

Table 1. Example sentence 2 letter predictable word

<table>
<thead>
<tr>
<th>Preceding context:</th>
<th>Following context:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hanneke begon haar stage als verpleegster. Na een lange en zware opleiding was ze blij dat ze eindelijk eens kon ervaren hoe het er allemaal praktisch aan toe gaat. Een vroedvrouw begeleidde haar bij de eerste taak. Hanneke maakte het bed</td>
<td>Daarna werd de vroedvrouw weggeroepen en stond ze er alleen voor.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Continuations:</th>
<th>Translation:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 letter word predictable</td>
<td>op volgens de instructies van de vroedvrouw.</td>
</tr>
<tr>
<td>2 letter word neutral</td>
<td>na volgens de instructies van de vroedvrouw.</td>
</tr>
<tr>
<td>4 letter word neutral</td>
<td>vast volgens de instructies van de vroedvrouw.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Continuations:</th>
<th>Translation:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 letter word predictable</td>
<td>made the bed according to the midwife's instructions.</td>
</tr>
<tr>
<td>2 letter word neutral</td>
<td>imitated the bed according to the midwife's instructions.</td>
</tr>
<tr>
<td>4 letter word neutral</td>
<td>fastened the bed according to the midwife's instructions.</td>
</tr>
</tbody>
</table>

Table 1. Example sentence 2 letter predictable word

---

8 Keep in mind that the term “target word” refers to the two- or four-letter particle parts of the separable verbs.
PRECEIVING CONTEXT:
Op een vrije namiddag besloot ik iets aan mijn kookkunst te doen. Om het eenvoudig te houden wilde ik beginnen met het bereiden van pudding. Een tijdje nadat ik de melk op het vuur had gezet, hoorde ik mijn moeder roepen: Haast je of de melk kookt

CONTINUATIONS
12 letter word predictable: over en dan is de pudding niet lekker meer.
2 letter word neutral: in en dan is de pudding niet lekker meer.

FOLLOWING CONTEXT:
Het was echter al te laat en ik kon opnieuw beginnen.

TRANSLATION:
On a free afternoon I decided to do something about my cookery skills. To keep it simple I started with preparing a pudding. A while after I had put the milk on the fire, I heard my mother calling: Hurry or the milk boils

CONTINUATIONS:
4 letter word predictable: over and the pudding won’t taste good anymore.
2 letter word neutral: down and the pudding won’t taste good anymore.

FOLLOWING CONTEXT:
However, it was already too late and I could start over again.

Table 2. Example sentence 4 letter predictable word

EXPERIMENT

METHOD

Participants. Fifty-seven first-year university students of Ghent University participated in the experiment which involved two sessions of one hour each. They all had normal uncorrected vision and were native speakers of Dutch.

Apparatus. Eye movements were recorded by a Senso-Motoric Instruments (SMI Eyelink) video-based pupil tracking system. Viewing was binocular but eye movements were recorded from the right eye only. A high speed video camera was used for recording. It was positioned underneath the monitored eye and held in place by head-mounted gear. The system has a visual resolution of 20 seconds of arc. Fixation locations were sampled every 4 ms and these raw data were used to determine the different measures of
oculomotor activity during reading. The display was 69 cm from the subject’s eye and three characters equalled 1° of visual angle. A chin rest was used to reduce head movements during the experiment.

**Materials.** Thirty-six target sentences with a preceding context were created (see Appendix). Half of the target sentences were built from a strongly expected 2-letter word; half from a strongly expected 4-letter word. For each sentence, two additional variants were made by replacing the expected target word by an acceptable but unexpected word of the same length and an acceptable but unexpected word of the complementary length. To validate our stimulus materials, a group of 40 first-year students who did not participate in the eye-tracking experiment were presented with the sentence frames up to the target word and asked to produce the next word in the sentence. For the target sentences with an expected two-letter target word, the sentence completion ratios were .82, .003, and .003 for the expected 2-letter word, the unexpected 2-letter word, and the unexpected 4-letter word respectively. Virtually the same completion ratios were obtained for the sentences with an expected four-letter word: .82, .004, and .006 for the expected 4-letter word, the unexpected 4-letter word and the unexpected 2-letter word. In the selection of the target words, we matched the conditions as closely as possible on the frequency of the separable verbs. The mean frequency per million for the two 2-letter and the one 4-letter targets according to the Celex database for Dutch (Baayen, Piepenbrock, & Van Rijn, 1993) were respectively 28.4, 14.8, and 7.4. For the target sentences with two possible 4-letter target words and one 2-letter word the mean log frequency per million were respectively 21.3, 14.8, and 20.2. Because we were not able to fully match the conditions on frequency this factor will be considered in the data-analysis. The sentences could have alternative endings, depending on the target word used. This was done to insure that the sentences as a whole remained plausible. O’Regan (1990, 1992) hypothesized that readers adopt a global reading strategy (e.g. careful or risky reading) that influences the fixation times and saccade lengths. It is not inconceivable that an implausible sentence ending could lead to a more
careful strategy on the following trials, thus influencing skipping behaviour. Precautions were also taken to insure that the alternative endings all had the same continuations in terms of word lengths and spacing for at least 10 characters following the target word. If not, this would affect the global visual configuration surrounding the target (cf. the global effect).

Procedure. Before the experiment started, participants were informed that the study was about reading comprehension of short texts, which would be displayed on a screen. Text administration was self-paced. The passages of the text were presented as a whole. Participants indicated when they had finished reading the text passage by pressing a button. They were told to read at their normal rate and that periodically they would be asked to answer a comprehension question about the passages. This was done on one-fourth of the trials. The participants had no difficulty answering the questions; the questions were simple true – false statements, and the participants were correct 90% of the time. The initial calibration of the eye-tracking system generally required approximately 10 min and consisted of a standard nine-point grid. Following the initial calibration the participant was given 10 practice trials to become familiar with the procedure before reading the experimental sentences. The 36 experimental sentences were embedded in a pseudo-random order in 220 filler sentences. Each participant was presented one of the three possible conditions per sentence according to a Latin square design. Participants completed two one-hour sessions, each session containing 128 trials.

RESULTS

Our primary dependent variable of interest was the probability of skipping the target word. We also calculated fixation times on the target words to check whether our patterns of data were compatible with prior findings. Because the word length of the target words was very short (i.e. 2 letters and 4 letters) there were virtually no double first-pass fixations on the target
words. Only in 5.6% of the cases was a fixation on the target word in first-pass followed by a re-fixation. Therefore, only single fixation durations (i.e., the duration of the fixation on the target given that there was only one fixation on the target) will be reported, together with the total fixation times (including regressions on the target). The region used for computing target word fixations consisted of the letters of the target word together with the space in front of the target word and only first-pass skipping was considered. The target word was presented halfway through the fourth line. However, for two stimuli with a predictable 2-letter word the targets were presented closer to the beginning of the fourth line. This caused a lot of participants to directly fixate on the target or past the target, when coming from the third line and making a return sweep to the fourth line (92% of the cases for these two stimuli). Because we wanted to control our data for launch site these 2 stimuli were excluded from the analyses. 7.7% more data were eliminated due to four possible reasons: (i) track loss, (ii) the participants first fixation in the 4th line was directly on the target or past the target, (iii) the participant had not read all five lines when he/she pressed the button to indicate that they had reached the end of the text, and (iv) the fixation was shorter than 100 ms (see Morrison (1984) and Rayner, Sereno, Morris, Schmauder, & Clifton (1989) for justification). Because a Latin square design was used with relatively few observations in the different cells, the group variable was included in all analyses reported below. If this is not done, the power of the design may be deflated because of random fluctuations between the participants or between the stimuli allocated to the different cells (Brysbaert & Mitchell, 1996; Pollatsek & Well, 1995). All analyses were run over participants (F1-analyses) and stimulus materials (F2-analyses).

**Skipping probability**

The skipping probabilities for the different conditions are shown in Table 3. Separate ANOVAs were run for the two different sentence sets. For the sentences with the expected 2-letter words, our manipulation of predictability was significant (F1(2,108) = 23.54, MSe = 0.04, p < .001;
F2(2,26) = 17.34, MSe = 0.01, p < .001). This effect, however, was largely due to the difference between the unexpected 4-letter words on the one hand and the 2-letter words on the other hand. Contrasts showed that the difference between the predicted and the unpredicted 2-letter words did not reach significance in a two-tailed test (2 pr vs. 2 unpr: t1 = 1.53, p > .10; t2 < 1, n.s.; 2 pr vs. 4 unpr: t1 = 5.78, p < .001; t2 = 7.99, p < .001; 2 unpr vs. 4 unpr: t1 = 5.16, p < .001; t2 = 3.68, p < .001; all the reported p values were Bonferroni adjusted).

For the sentence group with the expected 4-letter words, the effect of target type was significant as well (F1(2,108) = 26.76, MSe = 0.03, p < .001; F2(2,30) = 12.61, MSe = 0.02, p < .001). The difference in skipping rate for the expected 4-letter words and the unexpected 4-letter words reached the significance level for the F1 analysis but not for the analysis over stimuli (4 pr vs. 4 unpr: t1 = 2.30, p < .05; t2 = 1.22, p > .10; 4 pr vs. 2 unpr: t1 = -4.09, p < .001; t2 = -3.38, p < .01; 4 unpr vs. 2 unpr: t1 = 7.72, p < .001; t2 = -4.23, p < .001).

**Sentences with expected 2-letter words**

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Word Length</th>
<th>Skipping Probability</th>
<th>Single Fixation Duration</th>
<th>Total Fixation Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td>2</td>
<td>.79</td>
<td>219</td>
<td>236</td>
</tr>
<tr>
<td>Neutral</td>
<td>2</td>
<td>.74</td>
<td>232</td>
<td>385</td>
</tr>
<tr>
<td>Neutral</td>
<td>4</td>
<td>.56</td>
<td>235</td>
<td>413</td>
</tr>
</tbody>
</table>
Sentences with expected 4-letter words

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Word Length</th>
<th>Skipping Probability</th>
<th>Single Fixation Duration</th>
<th>Total Fixation Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td>4</td>
<td>.55</td>
<td>230</td>
<td>283</td>
</tr>
<tr>
<td>Neutral</td>
<td>4</td>
<td>.46</td>
<td>239</td>
<td>329</td>
</tr>
<tr>
<td>Neutral</td>
<td>2</td>
<td>.71</td>
<td>230</td>
<td>289</td>
</tr>
</tbody>
</table>

Table 3. Skipping probability and fixation time measures (in milliseconds) on the target word as a function of contextual constraint and target word length.

The skipping rates for unexpected 2-letter words were very similar when they replaced an expected 2-letter word (74%) and when they replaced an expected 4-letter word (71%). In a one-way between groups ANOVA run on the stimuli data they did not differ significantly from each other (F(1,32) < 1. n.s.). Although there seemed to be a bigger difference for the unexpected 4-letter words when they substituted an expected 4-letter word (46% skipping) than when they substituted an expected 2-letter word (56% skipping), the ANOVA showed that these two values did not differ significantly from each other either (F(2,32) = 1.16, MSE=0.04, p >.10).

To further examine the relative effects of word length and word predictability, we ran a multiple linear regression analysis on the skipping rates for all the 34 x 3 = 102 test sentences. Four predictors were entered: word length, sentence completion rate, log frequency per million, and average launch site. Frequency was entered in the analyses because in our stimulus selection it was not possible to completely match the frequency in all the conditions. Because word skipping depends on launch site, we entered
this variable as a predictor as well. The results were quite straightforward: Word length emerged as the only significant variable \((t(97) = 6.43, p < .001)\) explaining 31% of the variance. Word predictability was not significant \((t(97) = 1.41, p > .10)\), nor was word frequency \((-1 < t(97) < 1, \text{n.s.})\). Finally, average launch site per sentence failed to predict average word skipping rate too \((-1 < t(97) < 1, \text{n.s.})\), although there was a clear difference in launch site between the sentences in which a word was skipped and the sentences in which the word was fixated, as shown in Table 4. Launch sites were more than two letter positions further away when the target word was fixated than when it was skipped, both for the conditions with the expected 2-letter words \((F1(1,129) = 74.03, \text{MSe} = 7.25, p < .001, F2(1,47) = 27.39, \text{MSe} = 5.42, p < .001)\) and for the conditions with the expected 4-letter words \((F1(1,152)=108.18, \text{MSe}=6.95, p < .001, F2(1,53)=87.86, \text{MSe}=2.71, p < .001)\).

<table>
<thead>
<tr>
<th>Word length Constraint</th>
<th>Launch site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skipping</td>
</tr>
<tr>
<td>2 predicted</td>
<td>5.82</td>
</tr>
<tr>
<td>2 neutral</td>
<td>5.29</td>
</tr>
<tr>
<td>4 neutral</td>
<td>5.63</td>
</tr>
<tr>
<td>4 predicted</td>
<td>5.07</td>
</tr>
<tr>
<td>4 neutral</td>
<td>5.19</td>
</tr>
<tr>
<td>2 neutral</td>
<td>5.19</td>
</tr>
</tbody>
</table>

Table 4. Mean launch site in function of target type and skipping

**Fixation times**
Table 3 also displays the single fixation durations and the total fixation durations for the different sentence types. For the sentences with expected 2-letter words, the effect of target type was not significant for the single fixation durations ($F_1(2,38) = 1.07$, $MSe = 4125$, $p > .10$; $F_2(2,20) = 2.51$, $MSe = 1563$, $p > .10$). A different pattern was found for the total fixation durations: The main effect of target type was significant ($F_1(2,42) = 9.19$, $MSe = 21857$, $p < .001$; $F_2(2,20) = 21.32$, $MSe = 6162$, $p < .001$) and the predictable 2-letter words differed significantly from the neutral 2-letter words and neutral 4-letter words but the latter two were not significantly different from each other ($2\text{ pr vs. } 2\text{ unpr: } t_1 = -3.41, p < .01; t_2 = -2.80, p < .01; 2\text{ pr vs. } 4\text{ unpr: } t_1 = -7.03, p < .001; t_2 = -10.84, p < .001; 2\text{ unpr vs. } 4\text{ unpr: } t_1 < 1, \text{n.s.}; t_2 < 1, \text{n.s.}$).

For the single fixation durations of the expected 4-letter words the effect of target type was not significant ($F_1(2,78) < 1, \text{n.s.}; F_2(2,28) < 1, \text{n.s.}$). For total fixation durations, the main effect of target type was significant over participants ($F_1(2,80) = 4.16$, $MSe = 10345$, $p < .05$) and over stimuli ($F_2(2,28) = 4.85$, $MSe = 4242$, $p < .05$). The neutral 4 letter word target differed significantly from the predictable 4 letter word but the other comparisons did not differ significantly from each other ($4\text{ pr vs. } 4\text{ unpr: } t_1 = -2.27, p < .01; t_2 = -3.28, p < .05; 4\text{ pr vs. } 2\text{ unpr: } t_1 < 1, \text{n.s.}; t_2 < 1, \text{n.s.}; 4\text{ unpr vs. } 2\text{ unpr: } t_1 = 2.04, p > .05; t_2 = 2.06, p > .05$).

**Eye movements after the target word**

The difference in total fixation times between sentences with expected and unexpected words is a first indication that the former type of sentence indeed induced the processing difficulties we aimed at. To obtain more information about the extent to which the unexpected sentence continuations disturbed the ongoing processing, we looked at regression probabilities and spill-over effects. Table 5 shows the percentage of immediate regressions to the target word both when this word had been skipped and when it had been fixated. First, it is clear that regressions were more likely in sentences with
unexpected words (18%) than in sentences with expected words (10%). In addition, the data in Table 5 replicated Vitu and McConkie's (2000) finding that immediate regressions are more likely after a word has been skipped (19%) than after it has been fixated (11%). Finally, regressions to 2-letter words (12%) were less likely than regressions to 4-letter words (18%). The effects of these three variables can easily be estimated by using them as predictors in a multiple regression analysis with the probabilities of regressive eye movements listed in Table 5 as dependent variable.

<table>
<thead>
<tr>
<th>Word length Constraint</th>
<th>Probability of regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skipping</td>
</tr>
<tr>
<td>2 predicted</td>
<td>.125</td>
</tr>
<tr>
<td>2 neutral</td>
<td>.228</td>
</tr>
<tr>
<td>4 neutral</td>
<td>.275</td>
</tr>
<tr>
<td>4 predicted</td>
<td>.153</td>
</tr>
<tr>
<td>4 neutral</td>
<td>.208</td>
</tr>
<tr>
<td>2 neutral</td>
<td>.166</td>
</tr>
</tbody>
</table>

Table 5. Probability of regression to the target in function of target type and skipping

To look at the total disruption caused by the unexpected continuations, we also calculated the Cumulative Region Reading Time (CRRT) for the two-word region after the target. The CRRT (Brysbaert & Mitchell, 1996) is the sum of all fixations from the moment the eyes cross the front boundary of the region to the moment they cross the back boundary⁹. It includes the first-

⁹ Measurements exist that are similar or the same as the CRRT, eg. the Regression Path Reading Time. For a discussion on the benefits of such measurements see Liversedge, Paterson, & Pickering (1998).
pass gaze duration of the region, all the fixation durations of the regressive movement (if there is one), and the rereading of the critical region after the regression. CRRTs as a function of the different sentence types are shown in Table 6. These data clearly illustrate the processing difficulties induced by the unexpected words.

<table>
<thead>
<tr>
<th>Word length Constraint</th>
<th>CRRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 predicted</td>
<td>473</td>
</tr>
<tr>
<td>2 neutral</td>
<td>689</td>
</tr>
<tr>
<td>4 neutral</td>
<td>630</td>
</tr>
<tr>
<td>4 predicted</td>
<td>440</td>
</tr>
<tr>
<td>4 neutral</td>
<td>601</td>
</tr>
<tr>
<td>2 neutral</td>
<td>625</td>
</tr>
</tbody>
</table>

Table 6. Cumulative Region Reading Time for the two-word region past the target in function of target type (in ms).

**GENERAL DISCUSSION**

The first goal of the current study was to test one of the core assumptions of the Extended Optimal Viewing Position (EOVP) model of eye guidance in reading (Brysbaert & Vitu, 1998). This model states that the most important factor in word skipping is the length of the upcoming word, even when words are highly expected on the basis of the sentence context. Because the model has been formulated in mathematical terms, it is possible to predict the effects due to word length and context predictability quantitatively. More specifically, percentage skipping rate for predictable words is assumed to be equal to $100(e^{-0.34 \text{ word length}}) + 5$; and that of unpredictable words to $100(e^{-0.34 \text{ word length}}) - 5$. Thus, for 2-letter words, the expected skipping rates are 56% and 46% for predictable and unpredictable words respectively; for 4-letter words, they are 31% and 21%. These can be compared with the data obtained in the present experiment (see Table 3). Expected 2-letter words
resulted in a skipping rate of 79%, unexpected 2-letter words in a skipping rate of 72% (averaged across the two conditions with unexpected 2-letter words). Expected 4-letter words induced a skipping rate of 55% for the predictable words and 51% for the unpredictable. Although the skipping rates in our experiment were considerably higher than expected on the basis of the EOVP-model (possibly because the sentences were not presented in isolation but as part of a text passage), the model successfully predicted the 75%-52% = 23% difference between 2-letter words and 4-letter words (25% difference predicted). In addition, the impact of the word length variable was more pronounced than that of word predictability, both in absolute terms and according to the regression analysis. An unpredictable 2-letter word was more likely to be skipped (72%) than a predictable 4-letter word (55%). To our knowledge, this is the first study that simultaneously manipulated word length and word predictability, and that confirmed Brysbaert and Vitu's (1998) predictions about word length in a repeated measures design.

In the predictable 4-letter word conditions a difference of 9% in skipping rate due to contextual constraint was observed. This effect of predictability was reduced to a 5% difference in skipping rate in the 2-letter word conditions. The fact that the effect was smaller in the 2-letter word conditions may be due to ceiling effects: the 79 % skipping in the predictable 2 letter word condition is among the largest thus far observed. As such, we believe that the 9% difference due to contextual constraint in the 4-letter word conditions is the more accurate measure of the contextual constraint effect. This 9% difference is consistent with prior research (see Figure 1; and Brysbaert & Vitu, 1998). 10

10 The only effect of contextual constraint that clearly deviates in size from the present results is the 23% difference reported by Vonk (1984), but this finding has been questioned on methodological grounds (Brysbaert & Vitu, 1998).
Although the 9% difference is in line with previous studies and presents further evidence for the importance of linguistic variables in determining the probability of word skipping, it is clear that a difference of such a small size causes problems for a purely language-based model of eye movement control in reading. Remember that we had an 80% difference in completion scores between the expected and the unexpected words. This was the highest difference we could achieve and is in line with the differences that were used in other studies with length-matched stimulus materials. In addition, we worked with short words that are known to be skipped frequently (i.e., because they are often identified in parafoveal vision according to the E-Z Reader model). Still, we were not able to increase the difference in skipping rate between the expected and the unexpected words. Apparently, a 9% difference is among the largest one can obtain with contextual constraints. This is double the 4% effect Brysbaert and Vitu (1998) revealed for word frequency, but a way short of the effects attributed to word frequency and contextual constraints in the E-Z Reader model. In our view, the importance of word length should be more directly considered in the model. Our bet is that a very large part of the frequency and contextual constraint effects that are currently presented as evidence in favour of the E-Z Reader model, will actually turn out to be word length effects in disguise, certainly as far as word skipping is concerned (see Calvo & Meseguer, 2002, for effects of word length on other first-pass reading time measures).

Another difference between the E-Z Reader model and the EOVP model is that in the former all word skips consist of a cancelled saccade to word \( n + 1 \), which is replaced by a saccade to word \( n + 2 \). This replacement process is assumed to require extra time, so that fixation durations prior to skipped words on average will be longer than fixation durations prior to fixated words (Reichle et al., 1998). Although Hogaboam (1993) and Pollatsek, Rayner, and Balota (1986) reported empirical evidence for an inflated fixation duration prior to skipping, the effect has subsequently been questioned by McConkie, Kerr, and Dyre (1994) and Radach and Heller (2000) who failed to find a reliable difference between the fixation duration
distributions before skipped and unskipped words. However, the latter two studies were based on correlational data (i.e., large corpora of eye movements from free text reading) rather than on experimentally manipulated data. Therefore, we decided to check whether our better controlled data would confirm this prediction of the E-Z Reader model. For each sentence, we computed the average fixation duration before the critical saccade as a function of whether or not the target word was skipped. As can be seen in Table 7, no evidence whatsoever was found for an inflated fixation duration prior to skipping, rather the reverse. This is in line with the EOVP model which states that cancelled saccades are replaced with both longer and shorter alternatives, depending on the reason why the saccade was cancelled. The idea that target modification does not necessarily imply a delay of the saccade can also be found in the SWIFT model (Engbert, Longtin, & Kliegl, 2002). This recent computational model of eye movement control in reading focuses on a more spatially distributed view of lexical processing but also uses, like the E-Z Reader model, word frequency and predictability instead of word length as determinants of processing difficulty for predicting skipping rates.

<table>
<thead>
<tr>
<th>Word length Constraint</th>
<th>Prior fixation duration</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skipping</td>
<td>No Skipping</td>
</tr>
<tr>
<td>2 predicted</td>
<td>206</td>
<td>203</td>
</tr>
<tr>
<td>2 neutral</td>
<td>202</td>
<td>193</td>
</tr>
<tr>
<td>4 neutral</td>
<td>198</td>
<td>224</td>
</tr>
<tr>
<td>4 predicted</td>
<td>204</td>
<td>217</td>
</tr>
<tr>
<td>4 neutral</td>
<td>205</td>
<td>216</td>
</tr>
<tr>
<td>2 neutral</td>
<td>207</td>
<td>218</td>
</tr>
</tbody>
</table>

Table 7. Fixation durations prior to the target in function of target type and skipping (in milliseconds).
Finally, despite the main effects of contextual constraint and word length, there was no evidence for the interaction we looked for. We hypothesised in the framework of the EOVP model that a mismatch between the length of the expected parafoveal word and the length of the actually presented word might be picked up by the language processor in order to prevent a word skip over an unexpected word. If a word of two letters is strongly expected on the basis of the previous context, then information about the length of the parafoveal word cannot be used to make a decision between this word and a 2-letter intruder, but it can be used to decide between this word and a 4-letter intruder. Or formulated otherwise, an unexpected 2-letter word should be more likely to be skipped in a sentence where a 2-letter word is expected than in a sentence where a 4-letter word is expected. Although it is always difficult to interpret a null-effect, certainly when the effect goes in the expected direction (as is the case for the 2-letter words: 74% vs 71%; see Column 1 of Table 3), the findings of the 4-letter words seem to go enough in the opposite direction (46% vs. 56%) not to expect too much of a lack of power of the present experiment.

So, we failed to find evidence for the idea that the eye guidance system is sensitive to the match between global visual information extracted from the parafovea and word-length anticipations based on the meaning of the text. Such cross-talk between low-level visual information and high-level language information was first suggested by Hochberg (1975) and also features in Clark and O'Regan's (1999) ideas about parafoveal word recognition. Inhoff, Radach, Eiter, and Juhasz (in press) recently called this view that knowledge of word length assumes linguistic function in the word recognition process the word length constraint hypothesis.

Our failure to find evidence for the parafoveal word length constraint hypothesis in eye guidance might be explained by assuming that word length can influence word recognition only if it is backed up by a minimum of incoming sensory (orthographic) information. This would be in line with Clark and O'Regan (1999) who ventured that word length *in combination*
with information about the first and the last letter of the word determines parafoveal word recognition. Such a view would also agree with Marslen-Wilson's (1989) ideas of context influences on auditory word recognition. He claimed that such influences do not pre-activate the next word but facilitate the processing of matching bottom-up information, so that the word is recognised faster. According to this view, the mismatching word "na" in Table 1 would not be mistaken for the expected word "op" on the basis of the word length because there is no minimum of matching sensory information entering the system. Skipping data which are compatible with this minimum of incoming information hypothesis were reported by Balota et al. (1985). They presented predictable words in the parafovea. The target words were either correctly spelled or had a misspelling in their last letters so that they effectively became non-words. No difference in skipping rate was found between the predictable words and the non-words that were based on these predictable words and that shared the same onset.

So, our results leave open the possibility that a combination of parafoveal word length information and text-based expectancies may be used in eye guidance, when there is enough matching bottom-up information (from the first letters or the outer letters). However, other recent evidence reported by Inhoff et al. (in press) tempers the enthusiasm about this possibility. Inhoff et al. manipulated the length information of parafoveal target words by either presenting the correct word or by replacing one middle letter of the word with a blank space so that it seemed as if there were two short words coming up in the parafovea (e.g. subject vs. sub ect). When the eyes crossed the invisible border in front of the target word, the display changed and the preview was replaced by the intended target word. In a first experiment, Inhoff et al. manipulated word length information and orthographic information about the parafoveal word. In the second experiment, the word length preview was manipulated for high-frequency and low-frequency words. In Experiment 1, strictly additive effects of word length preview and orthographic preview were found. No extra word processing advantage was observed when both the length and the orthographic properties of the
preview matched the target word than could be expected on the basis of the individual main effects. Similar results were obtained in Experiment 2, where the effect of word length preview did not differ for high- and low frequency words. These findings led Inhoff et al. to conclude that spatial and linguistic information are controlled by functionally autonomous processing systems.

On the basis of our findings and those of Inhoff et al. (in press), it seems to us that the most likely interpretation within the frame of the EOVP model is that word length information and linguistic information independently influence the skipping decision. Word length is used in the very beginning of the fixation to obtain a rough estimate of the chances of recognising the parafoveal word by the end of the fixation, and this estimate is used to decide whether or not to program a saccade to this word. The initial decision can be overruled on the basis of the incoming linguistic information, but this latter decision does not take into account the word length information on which the original decision was based. These are two functionally independent subsystems.

Interestingly, whereas our context manipulation had a clear effect on the likelihood of skipping the target word, it failed to return a reliable effect on the duration of the first (and only) fixation on the target word, if this word was fixated. Looking at Table 3, we see an effect of some 11 ms between expected and unexpected words of the same length, together with a 6 ms difference between unexpected 2-letter and 4-letter words. This is in line with other evidence (e.g., McConkie & Dyre, 2000; Schroyens, Vitu, Brysbaert, & d’Ydewalle, 1999) that fixation durations are not always fully determined by the on-going processing of the fixated word. Apparently, the eye guidance system is sometimes programming one saccade ahead, meaning that processing difficulties have to be dealt with by means of regressions (see the total fixation durations in Table 3 and the regression probabilities in Table 5) or longer fixations on subsequent words (the so-called spill-over effect). As Schroyens et al. (1999) argued, this may be due
to a tension that exists between the need to fully process the foveal word and
the urge to move the eyes fast enough to take optimal advantage of the
parafoveal preview. On the other hand, it must not be forgotten that because
of the high skipping rates associated with our short target words, the average
fixation times were based on very small numbers of observations, so that we
may not have been able to pick up the full size of the effect.

To summarize, the results of our study indicate a strong word length effect
on skipping behaviour as well as an effect of contextual constraint. These
effects combined with the finding that expected word length did not
influence skipping behaviour adds further evidence to the view that verbal
and visual factors on word skipping are controlled by two functionally
different subsystems.
REFERENCES


APPENDIX

The sentences used in the experiment. For each sentence, the three possible continuations are given, followed by their sentence completion ratio. The first 18 sentences include an expected 2-letter word; the last 18 sentence groups include an expected 4-letter word.

1. De meid van de pastoor moest normaal gezien niet zoveel werken, maar die dag was duidelijk een uitzondering. ‘s Avonds was ze enorm vermoeid. Het was een lange dag geweest voor haar en de meid slaakte dan ook een diepe zucht. Zij rolde het tapijt
   
   *(op)* toen het vijf uur was en ze naar huis mocht. 67.5 %
   *(na)* toen het van de trap viel en in de kelder belandde. 0 %
   *(open)* toen het besmeurd was doordat ze erop morste. 2.5 %
   Gelukkig moest ze van de pastoor niet iedere week de tapijten reinigen.

2. Hanneke begon haar stage als verpleegster. Na een lange en zware opleiding was ze blij dat ze eindelijk eens kon ervaren hoe het er allemaal praktisch aan toe gaat. Een vroedvrouw begeleide haar bij haar eerste taak. Hanneke maakte het bed
   
   *(op)* volgens de instructies van de vroedvrouw. 90 %
   *(na)* volgens de instructies van de vroedvrouw. 0 %
   *(vast)* volgens de instructies van de vroedvrouw. 0%
   Daarna werd de vroedvrouw weggeroepen en stond ze er alleen voor.

3. De politiepatrouille kreeg een oproep binnen dat er ergens een winkeldief-stal gepleegd werd. Toen de patrouille aankwam was de dief net weggevlucht. Ze zette meteen de achtervolging in en zat de dief algauw op de hielen. De patrouille haalde de wagen van de dief
   
   *(in)* met een gevaarlijk manoeuvre. 80 %
   *(op)* met een gevaarlijk manoeuvre. 2.5 %
   *(open)* met een gevaarlijke techniek. 0%
   Gelukkig waren de agenten hierin getraind en gebeurden er geen ongelukken.

4. Een bloemenverkoper ging iedere maandag naar Brussel om bloemen in te kopen. Toevallig moest een van zijn vaste klanten ook in Brussel zijn. De verkoper zei dat hij de klant om negen uur een lift kon geven. Hij haalde de klant
   
   *(op)* om geen tijd te verliezen. Hij wist dat 72.5 %
   *(in)* om haar voor te kunnen zijn. Hij wist dat 0 %
   *(door)* op zijn lijstje met kopers. Hij wist dat 0%
   de klant zelf niet zo stipt was en vreesde dat deze te laat zou komen.
5. In een chalet in de Ardennen ontstond brand door een weggegooide sigaret. Vijf minuten na het ontstaan van de brand werd een ketting gevormd voor het doorgeven van emmers water. Frank zette de emmer voor de voeten van Els. Zij nam de emmer op en schaterde het uit. De emmer bleek immers leeg te zijn. 65%

Men had de emmer zo ruw doorgegeven dat al het water eruit gevallen was.


\[
\text{in dat de dokter voorgeschreven heeft, maar spuwt het onmiddellijk terug uit en weigert er nog van te drinken. } \quad 70\% \\
\text{af dat de dokter voorgeschreven heeft, maar laat het onmiddellijk terug vallen en weigert er nog van te drinken. } \quad 0\% \\
\text{waar dat de dokter voorgeschreven heeft, maar laat het onmiddellijk terug uit en weigert er nog van te drinken. } \quad 0\%
\]

7. In het leger van de Pruisen was een bevel een bevel. De soldaten wisten dat ze een opdracht van de officier nooit konden weigeren. Deze had immers het recht om iemand die niet gehoorzaamde ter plaatse dood te schieten. De officier droeg een soldaat

\[
\text{op een mitrailleur schoon te maken. Hoewel dit een van de saaiste werkjes was, gehoorzaamde deze onmiddellijk. } \quad 67.5\% \\
\text{af tot de diepste kelder van het huis. De soldaat had geweigerd een van de saaiere werkjes uit te voeren. } \quad 0\% \\
\text{voor als verzetsheld tijdens de oorlog. Hoewel hij een van de saaiste werkjes had, gehoorzaamde hij altijd. } \quad 0\%
\]

8. Mijn vrouw is een echte slaapkop. In het weekend slaapt ze vaak tot elf of twaalf uur. Maar ook op weekdagen overslaapt ze zich regelmatig. Zelf ben ik iemand die doorgaans geen problemen heeft om wakker te worden. 's Morgens sta ik meestal vroeg

\[
\text{op voor de kinderen. Intussen bereid ik hun ontbijt en lees ik de krant. Daarna breng ik ze met de wagen naar school. } \quad 100\% \\
\text{in voor de kinderen. Intussen bereid ik hun ideeën. Intussen bereid ik hun open voor ideeën. Intussen bereid ik hun } \quad 0\% \\
\]


\[
\text{af om meer in de wacht te kunnen slepen. } \quad 72.5\%
\]


10. De week voor vaderdag ging een jongetje op zoek naar een geschikt cadeau. Uiteindelijk koos hij voor een gestreepte vlinderdas en vroeg aan de kassa om deze mooi te verpakken. Er was nog net genoeg geschenkverpakking. De verkoopster pakte het cadeautje in dat de jongen gekocht had voor vaderdag. Daarna overhandigde ze hem de nogal dure rekening. Ze gaf het pas terug nadat de rekening werd betaald.

11. De chauffeur van de baas van een grote firma moest reeds om vijf uur in de morgen opstaan om naar de luchthaven van Zaventem te rijden. Die dag was er immers een grote vergadering en de chauffeur haalde zijn baas op om naar de vergadering te gaan. Gelukkig mocht hij daarna terug naar huis en kreeg hij de rest van de dag vrijaf.


13. Het baasje van een hond merkte op een morgen dat deze wel heel erg rustig was. Toen daarin na een week geen enkele verandering kwam, ging hij naar de dierenarts. De diagnose bij de hond was terminale kanker aan de nieren. Zijn baasje wou zijn lijden niet onnodig rekken en maakte het dier af met een pijnloos prikje. Enkele uren later overleed de hond.
14. Na vele brieven en telefoontjes van theaterliefhebbers en van het publiek besloot de theaterdirecteur om Wachten Op Godot van Beckett nog maar eens op te voeren. Die avond was het zover! Iedereen had ermaar uitgekeken. Na veel uitstel voerde het gezelschap eindelijk weer het toneelstuk van de bekende schrijver op omdat iedereen het wou zien. 80 %
af omdat iedereen het slecht vond. 0 %
door omdat iedereen het goed vond. 2.5 %

15. De kunstenaar had me beloofd dat ik een schilderij van hem mocht kopen voor een zacht prijsje. We spraken af dat ik het op een avond in het weekend zou komen halen. Het doek was in fel oranje geschilderd en hing aan de muur in de grijze kamer. Het viel direct op toen ik het licht aanstak. Ik vertelde 95 % voor verder onderzoek. Ik vertelde 0 % toen ik het weg nam. Ik vertelde 0 %
de kunstenaar dat ik toch liever een van zijn andere werken wou zien.

16. Een grapjas wou eens de boel op stelten zetten tijdens de nieuwjaars- receptie voor de professoren van een universiteit. Hij had zich zeer goed voorbereid: hij had zijn haren geel gekleurd en droeg versleten kleren. Tussen de deftige mensen viel hij snel op als voorbeeld hoe het niet moest. 90 % in als voorbeeld hoe het niet moest. 0 % dood als gevolg van grote verveling. 0 %

17. Griet liep al de hele week nogal zenuwachtig. De spanning bereikte een hoogtepunt op vrijdag. Er was die avond namelijk een groot galabal. Griet trok zich na zeven uur terug op haar kamer en nam haar make-up. Ze maakte zich op als ballerina en ging naar het bal. Het bal bleek echter als protest tegen het dure bal. Het bal bleek echter 82.5 % als gastvrouw van het bal. Toch bleek het echter 0 % een gigantische flop te zijn wegens de slordigheid van de organisatoren.

18. Omdat het al zo lang geleden was dat ze iets van zich had laten horen, besloot moeder een brief te schijven naar een jeugdvriendin. Ze had echter maar weinig inspiratie en de brief werd pas zeer laat 's nachts voltooid. Ze nam de brief en kleefde er een postzegel op om hem te versturen. 100 %

19. Op een vrije namiddag besloot ik iets aan mijn kookkunst te doen. Om het eenvoudig te houden wilde ik beginnen met het bereiden van pudding. Een tijdje
nadat ik de melk op het vuur gezet had, hoorde ik mijn moeder roepen: Haast je of de melk kookt

over en dan is de pudding niet lekker meer.  95 %
door en dan is de pudding niet lekker meer.  0 %
in en dan is de pot goed voor de vuilnisbak.  0 %
Het was echter al te laat en ik kon opnieuw beginnen.

20. Sinds het aantreden van de nieuwe koning was er bij de bevolking heel wat protest te horen. Dit stak steeds duidelijker de kop op. Op het marktplein van de hoofdstad begon een revolutie te ontstaan. Deze zou echter niet lang duren. De soldaten sloegen de opstand terwijl de vadsige koning feest vierde in zijn paleis en dronk op de behaalde overwinning.

21. Wanneer haar moeder haar zakgeld overhandigde hoefde ze er bij Edith niet aan toe te voegen dat ze spaarzaam moest zijn. Edith zorgde ervoor nooit schulden te maken. Zij was een meisje dat met geld kon omgaan. Na haar inkopen hield Edith altijd nog vijftig frank vrij en kocht zich een lekker ijsje met het overgebleven geld.

22. De alerte ondernemer specialiseerde zich in het opkopen van failliete bedrijven, die toch nog rendabel kunnen gemaakt worden. Het nieuws dat de textielfabriek de boeken neergelegd had verspreidde zich als een vuurtje. De ondernemer nam het failliete bedrijf over en verdiende er een goede cent aan. Daarna keek hij alweer uit naar een nieuwe kans.

23. Tijdens de griepepidemie in het voorjaar voelde ik mij op een morgen ook maar mottig. Toen ik me ’s avonds nog niet beter voelde, maande mijn vrouw me aan om een dokter te raadplegen, die me terstond een medicijn aanraadde. De dokter schreef mij het medicijn voor op een bladje. Dat gaf ik aan mijn vrouw die er onmiddellijk mee naar de apotheek ging.

24. Nu de overheidssubsidies waren verhoogd, kon de gemeente werk maken van de vernieuwing van de Stationsstraat. Toen de werken voltooid waren, werd de straat
feestelijk ingehuldigd. Onder luid gejuich van de omwonenden knipte de
burgemeester het lint

\textit{door} en huldigde zo de vernieuwde straat 95%
\textit{stuk} en huldigde zo de vernieuwde straat 0%
\textit{af} en huldigde zo de vernieuwde straat 0%
in. Daarna werd een reuzebarbecue georganiseerd voor alle inwoners.

25. Toen Joris van zijn werk naar huis reed, kwam hij in een file terecht op een
plaats waar er doorgaans vlot verkeer is. Een automobilist had een fietser omver
gereden. Bovendien had de chauffeur vluchtmisdrijf gepleegd. Hoewel hij de fietser
eraan had reed de automobilist

\textit{door} alsof er niets 65%
\textit{rond} alsof er niets 0%
\textit{om} alsof er niets 0%

eer was. Gelukkig had iemand zijn nummerplaat genoteerd.

26. De buren konden door hun raam binnenkijken in het bureau van de bookhouder.
Deze was volop bezig het geld te tellen. Hierbij concentreerde hij uitermate
omdat hij geen bankbriefje mocht missen. Bij het tellen sloeg de boekhouder echter
een bankbriefje

\textit{over} en raakte in de war. De buren 80%
\textit{gade} en raakte in de war. De buren 0%
\textit{om} en raakte in de war. De buren 0%
zagen dat hij geïrriteerd opnieuw begon met het geld te tellen.

27. Een seriemoordenaar ging voornamelijk op zoek naar blonde slachtoffers.
Zonder het zelf te weten had Marian, die blond is, onnoemelijk veel geluk. Hoewel
zij het volgende slachtoffer op zijn lijstje was sloeg de seriemoordenaar haar

\textit{over} en vermoordde de volgende op de lijst. Marian 50%
\textit{gade} en vermoordde de volgende op de lijst. Marian 0%
\textit{af} en vermoordde de volgende op de lijst. Marian 0%
heeft nooit geweten dat ze ook op zijn lijstje stond.

28. De ouders van Els vroegen haar om een spaarrekening te openen bij een andere
bank, omdat deze meer rente opbracht. Om te kunnen beginnen had Els geld nodig
op haar nieuwe spaarrekening. De bankbediende schreef duizend frank van haar
zichtrekening

\textit{over} als start van de nieuwe rekening. 85.5%
\textit{voor} als remedie tegen eenzaamheid. 0%
\textit{op} als start van de nieuwe rekening. 2.5%
Na enige jaren was dit bedrag natuurlijk heel wat hoger geworden.

29. Op het strand van Calais hadden de reinigingsdiens op een dag heel wat werk.
Overal lagen blikjes en eetkasten, weggegooid door toeristen. De dag ervoor had
immers een veelbesproken recordpoging plaatsgevonden. Een atleet zwom het kanaal

over maar vroeg zich achteraf af wat hij hiermee 92.5 %

zon over maar vroeg zich achteraf af wat hij hiermee 0 %
in maar vroeg zich achteraf af wat hij hiermee 0 %
wou bewijzen. Hij besloot om het nooit meer over te doen.

30. In de klas van Tom is iedereen stiekem verliefd op een heel knap meisje. Bovendien is ze zeer intelligent en behaalt ze goede schoolresultaten. Alsof dit nog niet genoeg is heeft ze ook nog eens een prachtige stem. Het meisje is wel heel verlegen. Wanneer zij aangesproken wordt slaat zij altijd de ogen neer en er komt een rode blos op haar wangen. 85 %

31. Na zijn ontslag bij de firma waar hij al jaren werkte, zag vader Janssens zich genoodzaakt om een onderbetaalde job aan te nemen. Op zijn leeftijd is het immers niet zo evident een goedbetaalde positie te vinden. Het gezin Janssens had het financieel heel moeilijk. Ze kwamen amper rond het verlies van hun dochter die stierf bij een ongeluk. 80 %

32. Na het begin van het nieuwe schooljaar kwam ook de ouderraad voor de eerste keer samen. Alle ouders van de leerlingen konden zich kandidaat stellen om in deze raad te zetelen. De ouders waren compleet onbekenden voor elkaar. Na het etentje stelde iedereen zich voor met naam en voornaam. Daarna 82.5 %

33. Vroeger aten alle leden van ons gezin steeds samen. Dit gold zowel voor het ontbijt, het middagmaal als het avondmaal. Voor de laatstgenoemde maaltijd was het zelfs traditie dat het gezin uit één grote schotel at. Iedereen nam er een stuk uit en gaf de schotel door aan de volgende. 92.5 %

Meestal werd er heel wat afgelachen tijdens zo'n avonden.

34. Na de inbraak in een juwelierszaak werd de buurt hermetisch afgesloten door enkele rijkswachters. Omdat er nogal wat schade was en omdat het een juwelierszaak betrof diende men ook een expert op te roepen. Met behulp van de expert stelden de rijkswachters de schade
35. Elise liep thuis altijd gestrest rond tijdens de examens. Als ze niet op haar kamer was om te studeren, maakte ze ons het leven zuur met haar humeur. Ze was ook bang om op het laatste moment iets te vergeten. Voor het examen nam ze haar cursus vluchtig door uit schrik iets te zullen. Ze besefte echter dat dit niet veel meer zou uitmaken.

36. Gisteren moesten we verplicht aanwezig zijn op de voordracht van een saaie professor. Het eerste onderwerp sprak niet echt tot de verbeelding en tot overmaat van ramp behandelde hij dit enorm langdradig. Na het uitmelken van het eerste onderwerp ging hij snel over op een ander onderwerp. Maar ook dit kon blijkbaar niemand bekoren.
Two experiments examined word skipping in reading. In Experiment 1, skipping rates were higher for a preview of a predictable word than for a visually similar nonword, indicating full recognition in parafoveal vision. In Experiment 2, foveal load was manipulated by varying the frequency of the word preceding either a 3-letter target word or a misspelled preview. Again, a correct preview increased skipping; a high foveal load reduced skipping rate, but there was no interaction. Also, the pattern of effects in fixation times was the same as in the skipping data. Experiment 2 also showed significant skipping of nonwords similar to the target word, indicating skipping based on partial information.
INTRODUCTION

How long readers look at a word is primarily determined by the ease or difficulty associated with the processing of that word. A very robust finding in research on eye movements in reading is that readers look longer at a low-frequency word than at a high-frequency word (e.g. Inhoff & Rayner, 1986; Rayner, & Duffy, 1986; Rayner, Sereno, & Raney, 1996; Schilling, Rayner, & Chumbley, 1998, Vitu, 1991). Other variables that also reflect the ease of processing, such as predictability of the word from the preceding context (Balota, Pollatsek, & Rayner, 1985; Binder, Pollatsek, & Rayner, 1999; Ehrlich, & Rayner, 1981, Rayner, & Well, 1996; Schustack, Ehrlich, & Rayner, 1987; Zola, 1984) or the age at which the word was acquired (Juhasz & Rayner, 2003, 2005), have also been shown to affect how long a word is looked at. While some low-level visual factors influence the decision of when to move the eyes, a strong case can be made that the linguistic properties of the words are the main determiners of that decision. The opposite seems to be true for the decision of where to move the eyes: low-level visual factors, such as word length and spacing between words (Rayner, Fischer, & Pollatsek, 1998), are the most important influences on saccade length and on the landing position in a word. For example, the length of a saccade is influenced by the length of the currently fixated word and the length of the word to the right of fixation (e.g. Blanchard, Pollatsek, & Rayner, 1989; O’Regan, 1980; Rayner, 1979), and readers tend to make their first fixation about halfway between the beginning and the middle of a word (Deutsch & Rayner, 1999; McConkie, Kerr, Reddix, & Zola, 1988; Pollatsek & Rayner, 1982; Rayner, 1979).

3 The effect of word length on the fixation time on a word is hard to classify in that it almost certainly influences how difficult a word is to identify, but it also is likely to have effects that are related to eye movement control.
A phenomenon in reading that eludes this convenient when/where dichotomy is word skipping (the phenomenon that readers do not fixate on each word in the text). To be precise, about 30% of the words in a text do not receive a direct fixation during reading (Rayner, 1998). While word skipping is clearly closer to the question of where to move the eyes, influences of both low-level visual factors and high-level linguistic factors have been shown to affect skipping behavior. One of the most robust findings in word skipping is that short words are skipped more often than long words (Brysbaert & Vitu, 1998; Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Rayner, 1979; Rayner & McConkie, 1976; Vitu, O’Regan, Inhoff, & Topolski, 1995). But it has also been shown that words that are predictable from the preceding context are skipped more often than words that are not predictable (Altaribba, Kroll, Sholl, & Rayner, 1996; Balota, et al., 1985; Drieghe et al., 2004; Ehrlich & Rayner, 1981; Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987) and that high-frequency words are more likely to be skipped than low-frequency words (even when their lengths are matched), especially when the eyes are close to the target word on the preceding fixation (Henderson & Ferreira, 1993; Radach & Kempe, 1993; Rayner & Fischer, 1996; Rayner et al., 1996). So, clearly both visual and lexical/linguistic variables affect whether a word is skipped. But arguably the most convincing piece of evidence that word skipping is not easily placed in the classic when/where dichotomy is that even though predictability has a clear effect on the skipping rates, it has no effect on the position of the landing site in cases where the word is actually fixated (Rayner et al., 2001; Vonk, Radach, & van Rijn, 2000). This indicates that there is a distinction between the mechanisms that determine the saccade target (which word to fixate) and the ones that determine the actual landing site (where to fixate in the word), a distinction we believe should be present in the architecture of any comprehensive model of eye movements in reading.

Returning to the effects of predictability and frequency on word skipping, these effects clearly show that some words that are skipped have been
identified, at least to a certain extent. However, the extent to which a word that is skipped was processed during the prior fixation remains an issue of some debate in the literature (e.g. Radach & Kennedy, 2004; Rayner & Juhasz, 2004; Reichle, Rayner, & Pollatsek, 2003), and views on this matter differ rather dramatically. At one extreme, a word is skipped based on an “educated guess”, taking only coarse information about the target word into account (Brysbaert, Drieghe, & Vitu, 2005; Brysbaert & Vitu, 1998; Drieghe et al., 2004). At the other extreme, a word is mainly skipped because it was recognized in parafoveal vision on the prior fixation (e.g. Reichle et al., 2003). Thus, while a broad consensus exists among researchers in the field on the determinants of the where/when decision, the debate on word skipping continues. Moreover, the debate is enlivened by data on word skipping that have proven hard to simulate by models that do a fairly good job in simulating fixation duration data (Kliegl, & Engbert, 2004; Rayner, Ashby, Pollatsek, & Reichle, 2004).

Before we turn to the issues on word skipping that the current study will address, we will outline a model of reading which will help to frame these questions. The E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, 2003; Rayner, Reichle, & Pollatsek, 1998; 2000; 2005; Pollatsek, Reichle, & Rayner, 2003; 2005) is a quantitative model in which the core assumption is that cognitive processes associated with processing the fixated word serve as the engine behind forward eye movements in reading4. Word recognition is considered to be a serial process under the control of an attentional beam, with the word in the attentional beam being the only word that is being processed lexically. In addition, the model posits two phases of word recognition. The termination of the first phase, which could be identified with the identification of the

---

4 The model is serial in that it posits that only one word is processed at a time, but letters within a word are assumed to be processed in parallel (possibly with the exception of long polymorphemic words).
orthographic and phonological forms, cues the oculomotor system to begin programming a saccade to the next word. The termination of the second phase, which entails full lexical identification, causes the attentional beam to shift to the next word. Given the parameters of the model, the shift of the attentional beam usually occurs before the eyes move to the next word, and during the time that the attentional beam is on the next word (but the eyes are still on the previous word), parafoveal processing occurs. This mechanism is how the E-Z Reader model accounts for the fact that information is extracted from the word next to the currently fixated word during reading. This parafoveal preview benefit can be seen most clearly from the fixation time on a word that was presented in parafoveal vision on the prior fixation, as compared to when it was masked in parafoveal vision (e.g. Blanchard, et al., 1989; Morris, Rayner, & Pollatsek, 1990; Rayner, 1975; Rayner, Well, Pollatsek, & Bertera, 1982; Schroyens, Vitu, Brysbaert, & d’Ydewalle, 1999). Moreover, because the model posits that the gap in time between the eye movement signal to fixate the next word and the attention shift decreases as a function of the difficulty of processing the currently fixated word, it predicts that this preview benefit decreases as processing difficulty increases.

The model primarily predicts skipping by the following mechanism: If (a) the eyes are on word$_n$, (b) the attentional beam has shifted to word$_{n+1}$, and (c) if the first phase of word identification of word$_{n+1}$ in the parafovea is rapid enough, the programming of the eye movement to word$_{n+1}$ is cancelled and replaced by the programming of an eye movement to word$_{n+2}$. (The second phase of the identification of word$_{n+1}$ in the parafovea should usually

---

More precisely, the end of the 1st phase of the word recognition of word$_n$ cues the oculomotor system to start programming a saccade to word$_{n+1}$. The amount of time needed for the programming of a saccade is fairly constant, so the eyes land on the next word following a certain delay after the end of the 1st phase. If word$_n$ is difficult, there is a longer 2nd phase of the word recognition processes than for an easy word, hence the shift of the attentional beam, caused by the termination of this 2nd phase, occurs later.
complete before the eyes move to word_{n+2}) So, while the attentional beam goes to every word in the text, the eyes do not necessarily fixate each word. In the model, the amount of processing needed to complete the first phase of word recognition is related directly to the frequency and the predictability of the word. In this manner the model can successfully predict the effects of predictability and frequency on word skipping. The model can also account for the word length effect in word skipping because it assumes an inverse relation between the extraction of letter information and the distance of a letter from the center of the visual field. So the further away the eyes are from the target word, the more time will be needed to complete the first phase of the word recognition, and as a consequence of that the slimmer chances will be that the word will be skipped. This mechanism accounts both for the well-documented word length effect in reading (McConkie et al., 1988) as well as the effect of launch site (a word close-by will also be skipped more often independent of word length). What chiefly distinguishes this model from models that embrace a more low-level approach to explain word skipping (Brysbaert et al., 2005; Brysbaert & Vitu, 1998; Drieghe et al., 2004), is that in order for a word to be skipped, a significant amount of processing of the skipped word needs to have happened: the first phase of word recognition has been completed and the completion of full lexical identification of that word has occurred or is imminent.

One phenomenon that is predicted by the E-Z Reader model is that fixations on a word should be longer when the next word is skipped than when the next word is not skipped (all else being equal). This follows from the model because skipping results from the cancellation of the program to fixate word_{n+1} by the program to fixate word_{n+2}. Thus, a later program replaces an earlier program. In fact, such an inflated fixation duration has been observed in several studies (Pollatsek, Rayner, & Balota, 1986; Pynte, Kennedy, & Ducrot, 2004; Rayner et al., 2004), but not in others (Drieghe et al., 2004; Engbert et al., 2002; Radach & Heller, 2000). As a result, trying to explain this phenomenon has been viewed by some as an important arena for
understanding word skipping, and eye movements in reading more generally. We will return to this issue below.

A second phenomenon that is not directly related to word skipping but is related to the amount of processing that parafoveal words receive is the so-called parafoveal-on-foveal effect. This effect refers to the possibility that the processing of parafoveal information from a word not only aids later foveal processing and sometimes leads to skipping a word, but that it also can affect the processing time on the prior word (other than by the mechanism discussed in the prior paragraph). This phenomenon has been a primary reason why some researchers have rejected the serial processing assumption of the E-Z Reader model and proposed parallel processing of foveal and parafoveal words. In fact, a number of studies have indicated that the foveal viewing time can be altered by the words presented in the parafovea (e.g. Inhoff, Radach, Starr, & Greenberg, 2000; Kennedy, 1998; 2000; Kennedy, Pynte, & Ducrot, 2002; Schroyens, et al., 1999; Starr & Inhoff, 2004; Underwood, Binns, & Walker, 2000; Vitu, Brysbaert, & Lancelin, 2004). However, there are methodological problems associated with some of these studies, as well as failures to obtain consistent effects across experiments. We think a fair summary is that it is clear that an unusual beginning of the word_{n+1} can produce longer fixations on word_{n} (Inhoff, Starr, & Shindler, 2000; Underwood, et al., 2000), but that it is far less clear that the meaning of word_{n+1} influences the fixation time on word_{n} (for a review see Rayner, White, Kambe, Miller & Liversedge, 2003). In a recent study, Kennedy and Pynte (2005) used a large corpus of eye movement data and claimed to find further evidence of the meaning of the word to the right of fixation influencing the current fixation (particularly when word_{n+1} was a short word). However, there are problems with corpus analyses in that there is no control over difficulty levels associated with the location in the text from which two consecutive words are culled. However, as we will elaborate below, parafoveal-on-foveal effects are not necessarily inconsistent with a serial processing model such as the E-Z Reader model because the model predicts that not all saccades land on the intended word.
In fact, there are abundant data indicating that eye movements, like other motor movements, have variability and usually do not land exactly on their target (McConkie et al., 1988). In particular, it is quite reasonable from the quantitative data accumulated by McConkie et al. and others to conclude that it is not rare for saccades to fall short of the targeted word so that word_n is fixated even though word_{n+1} was the intended target and is the attended word (see also, Rayner, Warren, Juhasz, & Liversedge, 2004). Whether E-Z Reader (or a serial model) can predict these effects quantitatively, however, is an open question.

Largely spurred by these two phenomena, several models have appeared that have argued for parallel processing of foveal and parafoveal words, notably SWIFT (Engbert, Longtin, & Kliegl, 2002) and Glenmore (Reilly & Radach, 2003). That is, lexical processing in these models is spatially distributed across words and a competition for processing resources between the different words is constantly going on; for example a difficult word will use most of the resources leaving few resources for the processing of the other words. In the SWIFT model for example, lexical processing is distributed over a four-word attentional gradient, and (contrary to the E-Z Reader model), lexical processing is not the engine behind eye movements during reading to the same extent as in the E-Z Reader model. Instead, saccades are initiated after a variable (random) time interval to maintain a preferred mean rate of eye movements. Obviously, such parallel processing models do have the capability of predicting parafoveal-on-foveal effects. Whether they give an adequate explanation of such effects we think is also an open question, as well as whether they can account for when such effects do not occur.

How do these models explain skipping? Again using SWIFT as an example, saccades are directed towards words that have the highest level of excitation, which occurs at intermediate amounts of lexical processing. (That is, the default saccade target is not the following word, as in the E-Z Reader model.) Thus, word_{n+1} will be skipped if word_{n+2} has a higher level of excitation, and the model successfully predicts that this occurs when word_{n+1}...
is more frequent, more predictable, and shorter. However, because SWIFT does not assume that the next word is the default saccade target, there is no predicted “cost” in canceling a planned saccade to the next word, as it is the case in the E-Z Reader model. Kliegl and Engbert (2004) attempted to resolve the inconsistency in the literature we discussed earlier on whether there is a cost in skipping on the fixation time on the prior word. In a study using a large corpus (where other factors were controlled post-hoc), they found that fixations before skipped words were shorter before short or highly frequent words and longer before long or low frequency words. However, this issue is complex as assessing this effect depends on essentially correlational analyses. That is, whether the reader skipped a word or not was determined by the reader, so that one never can achieve the same amount of stimulus or participant control over the situations in which readers skip and the situations in which they don’t, as in for instance fixation times.

The current study examines word skipping using the E-Z Reader model as its focus, as we don’t think that the phenomena discussed above are fatal to E-Z Reader’s explanation of skipping or other parafoveal phenomena in reading. In particular, we wish to highlight the two major assumptions that E-Z Reader makes to explain word skipping. First, the model states that a word is skipped because it is recognized (processed up though the first stage) on the prior fixation by means of parafoveal processing. Second, it states that some skipping will occur because of saccadic error. However, for now, we will focus on the first mechanism. In particular, there appear to be two prior studies whose results seem somewhat at odds with the assumption that skipping results from a fairly full analysis of the parafoveal word.

The first study (Balota et al., 1985) examined the skipping of misspelled parafoveal words in sentences such as: “Since the wedding was today, the baker rushed the wedding cake to the reception”, where the target word (italicized) was quite predictable. This study used the eye-contingent display change paradigm, the boundary paradigm (Rayner, 1975), in which a preview stimulus was replaced by either a predictable target word cake or by
an unpredictable target word *pies* when the reader crossed the invisible boundary located before the target. Of major interest for the present purposes is how often various preview stimuli were skipped when a certain word was fairly predictable. In fact, when the predictable word *cake* was in the parafovea, it was skipped 11% of the time, whereas the non-predictable (but sensible) word *pies* was only skipped 2% of the time. (A non-word that was visually dissimilar to the predictable word *picz* and a word that was semantically anomalous in the sentence frame *bomb* also had low skipping rates.) However, a non-word that was visually similar to the predictable word *cahc* was skipped 11% of the time. This study was one of the first to show that a predictable word is skipped more often than an unpredictable word in the same location, and thus that skipping was due to the word that was actually there rather than simply due to guessing that the word was likely to be there. However, there are a few features of this study we would like to address.

First, as noted above, Balota et al. reported no difference between a predictable word (*cake*) and a non-word preview that was visually similar to the predictable word (*cahc*). This led them to conclude that the decision to skip the target word was not based on a full analysis of the parafoveal word. This conclusion, however, is somewhat at odds with the E-Z Reader model we presented earlier. That is, if a word is skipped because it is almost fully recognized, how can there be no difference at all between a predictable word and a visually similar non-word? However, it is not inconceivable that when a word that is skipped from a far launch site, it is skipped based on coarse information. The system would accommodate for this sub-optimal processing by compensating for it on the fixation after the skipping. This latter assumption would be compatible with the findings of Binder, Pollatsek, and Rayner (1999) who reported that readers often still attend to a word after it is skipped (plausibly when a saccade overshot the target word) and with the data reported by Reichle et al. (1998) that the duration of a fixation after a skip is also inflated. Thus, perhaps this lack of difference in the Balota et al. experiment was because a majority of the skips were from a
reasonably distant launch site. Unfortunately, Balota et al. did not report skipping rates as a function of different launch sites.

The first experiment reported here is essentially a replication of the Balota et al. study, but an important difference is that we also examined the skipping data as a function of launch site. In order to create a sensitive test of whether there would be a difference between the predictable word and the visually similar non-word preview at close launch sites, we increased the visual similarity by reducing the difference to a single letter. In addition, there is the question about whether there is a difference between a preview of an unpredictable word (pies) and a preview of a non-word (picz) derived from the unpredictable word that is both visually dissimilar to the predictable word (cake). That is, analogous to the question about the predictable word, does it make a difference in skipping rate that one is a word and one is not? The original Balota et al. data are not diagnostic, because there were likely to have been floor effects. To amend this, the original study was replicated in Experiment 1 but all the words longer than 6 letters were taken out of the stimulus set. Because short words are skipped more often than long words, this should increase the overall skipping rates and thus make floor effects less likely. Finally, we also added an extra condition in which a preview was presented that was an unpronounceable, orthographic illegal non-word. This condition was added to determine whether the visually dissimilar condition constituted the lower boundary of skipping behavior.

Experiment 1 focused on the question of how much processing of a parafoveal word is necessary to modulate skipping. This question was also addressed in Experiment 2, but the focus in Experiment 2 was on the question of whether word skipping is modulated by foveal load. Henderson and Ferreira (1990) showed that when foveal load is increased (e.g. a low-frequency word prior to the target word) the parafoveal preview benefit is reduced (see also Kennison & Clifton, 1995; Schroyens et al., 1999; White, Rayner, & Liversedge, 2005). Models such as the E-Z Reader model explain this phenomenon by stating that because the processing of the word $n$ takes
longer, the time window for parafoveal processing to occur between the arrival of the attentional beam and the actual arrival of the eyes on word $n+1$ will be smaller, hence less processing will have occurred. Because the E-Z Reader model relies heavily on parafoveal processing in explaining skipping behavior, the model would clearly predict that a higher foveal load will be accompanied by a lower skipping rate of the following word. If this is observed, then it would be another piece of data indicating that word skipping is importantly determined by word processing, contrary to other “where to move the eyes” decisions. This is even more so the case since previous research has shown that foveal load has only a small effect on the saccade length originating from the target word (e.g., Rayner, Ashby et al., 2004).

The second study that seemed problematic for the E-Z Reader model’s account of word skipping was by White (2004), as she reported finding no effect of foveal load on the skipping of the following word. She used five to six-letter foveal words and four-letter target words (i.e., the words whose skipping rates were assessed). The preview of the target word was either correct or misspelled. While there was a main effect of preview (the correct previews were skipped more often than the incorrect previews), no other significant effects were observed with the exception of an incorrect preview being skipped less often when it was preceded by a low frequency word. These findings can clearly not be accounted for by the mechanisms incorporated in the E-Z Reader model. However, because the overall skipping rates in this study were rather low, the lack of an effect of foveal load could have been due to a lack of power. As a result, we decided to use shorter target words. In Experiment 2, we employed three letter target words (for which the preview was either correct or misspelled) which were preceded by either a high-frequency or a low-frequency five letter word.
EXPERIMENT 1

The primary question explored in Experiment 1 was whether the findings of Balota et al. (1985), indicating that there is no difference in skipping rate between a predictable word and a nonword that was visually similar to it would still be true if one examined situations in which the prior fixation was close to the target region (i.e., at a close launch site). In addition, to make the test more sensitive, we used shorter stimuli than were used than in the original study to avoid floor effects.

If skipping is merely based on a crude estimate of whether a predictable word was present, the skipping rates should be about the same when the predictable word and a nonword visually similar to it are present in the parafovea and those skipping rates should be higher than the other conditions in which the preview of the target word is orthographically different from the predictable target word. Moreover, if the preview is orthographically different from the target word, it shouldn’t matter whether it is a word or nonword and/or whether the word fits in with the sentence. We expected that this might be the pattern when the launch site is far from the target word region. However, we thought that at close launch sites, there would be a more complete analysis of the preview stimulus, and thus that skipping rates would be at least sensitive to whether the preview was the predictable word or the nonword that was visually similar to it. It was less clear whether skipping rates would be at all influenced by the lexicality or sensibility of the preview if it wasn’t a candidate for the predictable word.

METHOD

Participants. Twenty-four members of the University of Massachusetts community participated in this experiment. All were native speakers of American English and had 20/20 vision or contacts. They were either given extra credit in a Psychology course or paid $8 for their participation.
Apparatus. Participants were seated 61 cm from a 15-inch NEC MultiSync FGE color monitor. All sentences were displayed on a single line with a maximum length of 80 characters. At this distance, 3.8 character positions equaled 1 degree of visual angle. An eye contingent boundary technique was used (Rayner, 1975) in which display changes occurred within 5 ms of detection of when an invisible “boundary” was crossed; the boundary was between the last letter of the prior word and the space preceding the target word. Eye movements were recorded using a Fourward Technologies Dual Purkinje Eyetracker (Generation V) interfaced with a Pentium computer. Although reading took place binocularly, eye movements were recorded only from the participants’ right eye, sampling the eye’s position every millisecond.

Materials. The sentences were selected from the materials used by Balota et al. (1985). In the original study, 96 sentence frames were used. Two separate norming procedures were used to assess the predictability of the predictable and unpredictable (but not anomalous) words. In the first procedure, 20 participants were presented with the sentences up to and including the target word. Their task was to indicate, on a 5-point scale, how well the base word fit into the sentence (1 = the word did not fit very well; 5 = the word fit very well). The mean ratings for the predictable and unpredictable words were 4.47 and 2.32, respectively. In the second procedure, 20 participants who did not participate in the first norming study, were given the sentence frame up to, but not including, the target word and were asked to generate the next word in the sentence. The predictable words were generated 64% of the time, whereas the unpredictable words were generated less than 1% of the time. Target words ranged in length from 4 to 8 characters, with a mean of 5.2 characters. For the current experiment we removed the 7 and the 8 letter words from the Balota et al. stimulus set, maintaining 84 sentence frames from the original 96 sentence frames. The average word length of the reduced stimulus set was 4.7 characters.
For each sentence, the target word was always the predictable word, and there were six possible previews that were either taken from the Balota et al. study or adapted given the criteria below for constructing the nonword previews. (An example is shown in Table 1.) The preview was either the predictable \((P)\) word (e.g. liver), an unpredictable \((U)\) word (e.g. heart), or a word that was semantically anomalous \((SA)\) in the sentence frame (e.g. files). The materials for these three conditions came from the original study. The visually similar condition \((VS)\) was formed by altering the penultimate letter of the predictable word, creating a non-word (e.g. livor). If the penultimate letter was an ascender or a descender, this letter was replaced by respectively an ascender or descender. The same procedure was used to make the preview for the visually dissimilar \((VD)\) condition (e.g. heant, which is visually dissimilar to the predictable word) where the base word was the unpredictable word. Finally the condition, which for the sake of convenience will be called the orthographically illegal \((OI)\) condition, was a non-word whose first three letters always constituted a unpronounceable combination that does not appear in the English language as letters at the beginning of a word (e.g. frhos). The previews were always the same length as the target. The average word frequency, based on the Francis and Kučera (1982) norms were 58.2 per million for the predictable words, 58.1 per million for the unpredictable words, and 88.9 per million for the semantically anomalous words. As each of the 84 sentence frames was read only once by a participant, there were 14 sentences per condition per participant. The 84 experimental sentences were embedded in a pseudorandom order in 60 filler sentences.

**Procedure.** When a participant arrived for the experiment, a bite bar was prepared, which served to eliminate head movements. Participants were given a general description of the experimental procedure and were asked to read sentences on the monitor as their eye movements were monitored. They were also told that they would be asked questions about the sentences and were instructed to read for comprehension. The initial calibration of the eye-tracking system required about 5 minutes. Each participant read 10 practice
sentences to become familiar with the procedure. Prior to the presentation of each sentence, a series of five boxes appeared on the monitor, extending from the first to the last character position of an 80-character sentence. During this calibration check, participants looked at each box so that the experimenter could verify that the eye position was accurately recorded. If the calibration was not accurate, the participant was recalibrated. If the calibration was accurate, the participant looked at the first box and the experimenter displayed the sentence. Questions about the meaning of the sentence were asked after 25% of the trials and participants had little difficulty answering the questions (accuracy 96%). The experiment lasted about 35 minutes.

| 1. Predictable word.          |
| The doctor told Fred that his drinking would damage his *liver* very quickly. |
| 2. Unpredictable word.        |
| The doctor told Fred that his drinking would damage his *heart* very quickly. |
| 3. Semantically anomalous word.|
| The doctor told Fred that his drinking would damage his *files* very quickly. |
| 4. Visually similar non-word.  |
| The doctor told Fred that his drinking would damage his *livor* very quickly. |
| 5. Visually dissimilar non-word.|
| The doctor told Fred that his drinking would damage his *heant* very quickly. |
| 6. Orthographic illegal word. |
| The doctor told Fred that his drinking would damage his *frhos* very quickly. |

Note: The stimuli shown in italics indicate the preview for each condition prior to the eyes' crossing of the display change boundary. The preview was always replaced by the predictable word after the boundary had been crossed.

Table 1. An example sentence from Experiment 1 with each of the 6 parafoveal preview conditions.

RESULTS

Our primary interest in this experiment was the probability of skipping the target word during the first pass through the text (not taking regressions into account). In addition to the overall skipping probability, we examined the skipping probability conditional on the distance of the launch sites from the target word. For the cut-off point between a close and a far launch site we chose 5 character positions, since this allowed an approximately even...
division of the data (45% of the saccades, regardless of whether the target word was skipped or fixated, were launched from 5 or fewer character positions from the space in front of the target word).

We also calculated the fixation times on the target word. When the target word was fixated, in 91.7% of the cases it received a single fixation. Therefore, we will restrict the fixation duration analyses to when there was a single fixation on the target word. Since our materials in this experiment were identical up to the target word, we were also able to examine the fixation duration of the last fixation prior to either skipping of or landing on the target word. And finally, although not the focus of the present article, we also examined the fixation duration prior to the target word regardless of whether it was skipped or not as our incorrect previews constitute an interesting situation for examining potential parafoveal-on-foveal effects. Fixation durations of less than 100 ms and more than 1200 ms on the target word were removed from the analyses. Trials on which the eye-tracker lost track of the eye position were also excluded from the analyses, as well as trials in which the eyes triggered the boundary but remained on the word before the target (usually the last letter of this word). Finally, when the fixation duration was greater than three standard deviations from the mean for that participant in that condition, it was also removed (this was the case for 1 observation). As a result, about 16% of the trials were excluded from the analyses, and these trials were approximately equally distributed across conditions. A series of repeated measures analyses of variance (ANOVAs) were undertaken with participants (F1) and items (F2) as random variables.

Skipping the target word. The skipping rates associated with Experiment 1 are shown in Table 2. The effect of preview on the skipping rates of the

---

6 On some occasions, the Dual Purkinje Eye-tracker will register a saccade that crosses the boundary (triggering the display change), but the eye then (within a few milliseconds) “hooks” back to land on a character prior to the boundary location.

7 Deleted cells were treated as missing cases.
target word during first pass reading was close to significant, $F_1(5,115) = 2.13, p = .067, F_2(5,415) = 1.962, p = .083$. Contrasts showed that this was mostly due to the 5% difference in skipping rate between the predictable word preview and the average of the other five conditions, $F_1(1,23) = 6.18, p < .05; F_2(1,83) = 5.16, p < .05$. There also appeared to be a difference between the U, SA, and VS condition on the one hand and the VD and OI condition on the other; however, contrasts showed that this was not significant, $F_1(1,23) = 3.76, p > .05; F_2(1,83) = 1.82, p > .10$. In addition, there was no longer any significant effect of preview when the predictable condition was removed from the analysis ($F$s < 1).

<table>
<thead>
<tr>
<th>Preview Conditions</th>
<th>Skip critical word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
</tr>
<tr>
<td>Predictable word</td>
<td>.20</td>
</tr>
<tr>
<td>Unpredictable word</td>
<td>.16</td>
</tr>
<tr>
<td>Sem. Anomalous word</td>
<td>.16</td>
</tr>
<tr>
<td>Visually similar non-word</td>
<td>.16</td>
</tr>
<tr>
<td>Visually dissimilar non-word</td>
<td>.13</td>
</tr>
<tr>
<td>Orth. Illegal non-word</td>
<td>.12</td>
</tr>
</tbody>
</table>

Table 2. Probability of skipping the target word during first pass for all the data (All), saccades launched from 5 or less character positions (45 % of the data), saccades launched from 6 or more characters (55 % of the data) and for all the data not restricted to first pass (All + regr).

When we restricted the data set to saccades launched from five or fewer character positions from the target word, the effect of preview on the
skipping rates was significant, $F_1(5, 115) = 3.35, p < .01$; $F_2(5, 415) = 3.14, p < .01$. A predictable word was skipped 14% more often than the average of the other conditions, $F_1(1, 23) = 250.05, p < .001$; $F_2(1, 83) = 16.30, p < .001$, and there was no significant difference among the other five conditions ($F$s < 1). Finally, when the skip was launched from six or more character positions from the target word, there was virtually no difference among the conditions, with skipping rates being low in all conditions ($F$s < 1). Table 2 also includes the overall skipping probability. In this measure, if the word was initially skipped but then regressed back to, it isn’t counted as a skip. This presumably should index whether the reader realized there was something wrong after they skipped the target word. In the original Balota et al. study there was no difference in this measure between the P and the VS condition. The effect of preview on these skipping rates was significant, $F_1(5, 115) = 3.66, p < .01$; $F_2(5, 415) = 3.46, p < .01$. Contrasts showed that this variance was again mainly due to the 8% difference between the predictable word and the other conditions, $F_1(1, 23) = 11.71, p < .01$; $F_2(1, 83) = 9.42, p < .01$. There was no significant difference among the other five conditions, $F_1(4, 92) = 1.06, p > .20$; $F_2(4, 332) = 1.32, p > .20$. Although the overall skipping appears to be a bit smaller for non-word previews than for word previews, this effect was not significant, $F_1(1, 23) = 2.24, p > .10$; $F_2(1, 83) = 1.52, p > .20$.

**Fixation times.** Fixation times in Experiment 1 are shown in Table 3. As mentioned above, since the vast majority of gaze durations on the target consisted of a single fixation, we shall only report the single fixation duration, the mean fixation duration when there was a single fixation on the target. The effect of preview on the single fixation duration times was significant, $F_1(5, 115) = 11.42, p < .001$; $F_2(5, 415) = 5.81, p < .001$. Contrasts showed that three groups could be distinguished: the fixation times

---

8 Single fixation durations and gaze durations show exactly the same pattern.
in the predictable condition were about 20 ms less than those in the U and VS conditions, which were in turn about 20 ms less than those in a 3rd group containing the SA, the VD and the OI conditions. However, some of the comparisons were no longer significant in the analysis across stimuli after the p values were Bonferroni adjusted (P versus U & VS, $F_1(1,23) = 8.65, p < .01$, $F_2(1,83) = 4.49, p < .12$; U & VS versus S, VD & OI, $F_1(1,23) = 6.30, p < .05$; $F_2(1,83) = 3.44, p < .20$; P versus S, VD, & OI, $F_1(1,23) = 61.62, p < .001$, $F_2(1,83) = 25.70, p < .001$).

<table>
<thead>
<tr>
<th>Target word</th>
<th>Single fixation</th>
<th>Prior to the target word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Overall</td>
</tr>
<tr>
<td>Predictable word</td>
<td>262</td>
<td>225</td>
</tr>
<tr>
<td>Unpredictable word</td>
<td>284</td>
<td>236</td>
</tr>
<tr>
<td>Sem. Anomalous word</td>
<td>305</td>
<td>236</td>
</tr>
<tr>
<td>Visually similar non-word</td>
<td>279</td>
<td>239</td>
</tr>
<tr>
<td>Visually dissimilar non-word</td>
<td>301</td>
<td>239</td>
</tr>
<tr>
<td>Orth. Illegal non-word</td>
<td>305</td>
<td>242</td>
</tr>
</tbody>
</table>

Table 3. Fixation times (in ms) of the single fixation duration on the target word, of the last fixation prior to the target word, the last fixation prior to the target word restricted to fixations at 3 or less character positions from the target word, last fixation restricted to more than three character positions away from the target word, last fixation prior to skipping the target word and the last fixation prior to landing on the target word.

The effect of preview on fixation durations prior to the target word was not significant (all $F$s $< 1$) (shown in Table 3, columns 2), regardless of whether the target word was skipped or fixated. An examination of the means, however, indicates that the overall fixation time in the predictable word
condition was somewhat shorter than in the other conditions. To make sure we did not miss a potential parafoveal-on-foveal effect on these viewing times, we examined the fixation durations of the fixations that were very close to the target word (three or fewer character positions, 42.4% of the data) or further away. These fixation times are shown in Table 3 (column 3 and 4). The effect of preview was not significant (all $F$s < 1) but contrasts showed that the 20-25 ms difference between the fixation duration in the predictable condition and the other conditions was significant for the analysis across participants, $F(1,20) = 6.06$, $p < .05$, but not for the analysis across items, $F_2 < 1$. There were clearly no reliable differences among the conditions when the launch site was at least three characters from the target word (all $F$s < 1).

Turning to the question of whether there was an inflated fixation duration prior to skipping, when we compared the last fixation duration on the prior word conditional on whether the target word was skipped (shown in Table 3, columns 5 and 6), we did indeed find that this fixation time was about 34 ms greater when the target word was skipped than when it was fixated, $F(1,8) = 15.17$, $p < .001$.  

DISCUSSION

We had anticipated replicating the finding by Balota et al. (1985) that there was no significant overall difference in skipping rates between a predictable word and the nonword that was visually similar to it, but that we would find a difference when the prior fixation was at a suitably close launch site. Contrary to our expectations, we found that there was a difference between the predictable words and the visually similar nonwords in both the overall

---

9 No $F_2$ analyses are reported for this comparison due to a high number of missing cells in the fixation duration prior to skipping matrix. The high number of missing cells also made any further analysis for these data unwarranted.
analysis and the analysis restricted to a close launch site, even though the visual similarity between the P and VS condition was higher than in the original study. In fact, we found little difference in skipping among the conditions in which the preview was not the predictable word. One possibility for the discrepancy between the present study and the original study may be the viewing conditions of the experiments. First, the quality of monitors has improved (with many more pixels per character) in the 20 years between the original study and the present study. This may have made extraction of parafoveal information more efficient in the present study. Second, in the original study, three characters equaled 1 degree of visual angle, whereas in the present study 3.8 characters equaled 1 degree of visual angle. This closer packing of the information in the present study may have also made extraction of parafoveal information more efficient. Thus, the original lack of finding a difference between the predictable and the visually similar conditions may have been due to poorer parafoveal viewing conditions. The present findings are therefore consistent with the assumption that the decision to skip the target word was based on a full analysis of the parafoveal word, as the difference between the P and the VS condition was a single letter. In contrast, the skipping rates of the VS condition were not different of those of the VD condition even though the orthographic difference was large. It should also be noted that virtually the entire skipping effect occurred when the launch site was close, again indicating that the effect was likely to be due to fairly full processing of the skipped word.

Another interesting finding in the skipping data is the lack of difference between the non-predictable conditions. In the original Balota et al. (1985) study, there were also no significant differences between the N, VD and SA conditions, but this could have been due to floor effects. The fact that there is no difference in skipping between a neutral word and a semantically anomalous word in the present study could be expected based on previous research (Altarriba, Kambe, Pollatsek, & Rayner, 2001; see also Rayner, Balota, & Pollatsek, 1986) showing that semantically related words are not
skipped more often than semantically unrelated words. In the E-Z Reader model this finding is explained by stating that the decision to skip a word is instigated by the completion of the first phase of word identification of the target word. Whereas predictability appears to boost performance in this phase, the extraction of semantic features from the parafoveal word apparently does not (or at least doesn’t do it fast enough), explaining the lack of an effect of semantic inhibition in the SA condition. Indeed, the mechanism explaining skipping behavior incorporated in the E-Z Reader model does not make any differential predictions on skipping behavior in terms of inhibition. Rayner and Well (1996) showed that to obtain an effect of predictability on skipping you need a large enough difference between the predictable and the unpredictable target words in terms of sentence completion ratio. A medium constraint target word resulted in faster viewing times when the target word was actually fixated but did not differ from an unpredictable word in terms of skipping. The system apparently decides to cancel the planned saccade to the target word only when the speed of the first phase of word recognition of the target word is boosted a lot. Further evidence for a restriction of this mechanism to the condition with the predictable preview is provided by the strength of the word length effect in the current experiment. As noted previously, a considerable amount of prior research has clearly demonstrated that short words are skipped more often than long words, presumably due to reduced visibility in the case of long words. If the skipping rates are higher in the predictable condition due to enhanced word recognition of the target word, then the low-level visual effect of word length would be relatively smaller as compared to the other conditions that do not have this influence. We ran a simple regression analysis on the skipping data of the six conditions with word length as a predictor. Word length was not a significant predictor for the skipping rates in the predictable condition (P: $F(1,82) = 1.33, p > .20$) but was in all the other conditions (U: $F(1,82) = 23.77, p < .001$; SA: $F(1,82) = 3.91, p = .05$; VS: $F(1,82) = 9.17, p < .01$; VD: $F(1,82) = 7.30, p < .01$; OI: $F(1,82) = 17.68, p < .001$). Even though we have no doubt that, given a very large data
set, the effect of word length in the predictable condition would also become a significant predictor we cautiously take this analysis as an indication that there is a less pronounced word length effect in the predictable condition. This would be due to the strong presence of another influence, the modulation of skipping data by enhanced word recognition of the target word due to its predictability from the preceding context. Taking all these arguments into account, the fact that the 5 non-predictable conditions do not differ from each other is in agreement with the E-Z Reader model. However, what is incompatible with the model is the high rate of skipping “garbage” words and nonwords, an issue we will address in more detail in the General Discussion.

The single fixation durations on the target showed a pattern of $P < U$ and $VS < SA, VD, and OI$. That the predictable word should receive the shortest fixation times is not surprising, since it was the only condition in which the preview matched the target word after the boundary change had occurred. That some inhibition could be expected from the $S, VD, and OI$ condition is also not surprising. The fact that the target word in the $VS$ condition was read faster was undoubtedly due to orthographic overlap with the target word after the reader had landed upon it, however this orthographic overlap was not strong enough for the visually similar non-word to be read significantly faster ($5 \text{ ms}$) than the unpredictable word, perhaps due to some inhibition from the non-wordness of the $VS$ preview attenuating the orthographic overlap advantage.

The fixation durations prior to the target word were very interesting. We did find that the fixation duration prior to skipping a word was inflated, adding further evidence for the existence of this effect (Pollatsek et al., 1986). Furthermore, we found that when the eyes were very close to the target word (three or fewer character positions), the fixation durations in the five non-predictable conditions were longer than the fixation durations in the predictable preview condition. We will also defer discussion of these two effects to the General Discussion.
EXPERIMENT 2

The primary goal of Experiment 2 was to explore whether there is an influence of foveal load on the skipping of the following word. Previous research (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; White et al., 2005) demonstrated that a high foveal load leads to reduced parafoveal preview benefit. In the E-Z Reader model, a word is skipped because it is recognized in parafoveal vision. Thus, it follows that the chances of recognizing the parafoveal word would be reduced when the amount of parafoveal preview is reduced. Accordingly, we expected that a high foveal load would lead to a reduced skipping rate of the following word. As we noted earlier, a prior study (White, 2004) did not find an effect of foveal difficulty of word_n on skipping rates of word_{n+1}. However, there may have been a power problem as the skipping rates in this study were quite low, probably because 5-6 letter foveal words and 4 letter target words were used. As a result, we used shorter words for both the foveal words whose difficulty is being manipulated (5 letters) and for the ensuing target words that are being examined for skipping probabilities (3 letters). Note that the same prediction is made by parallel models that assume a constant competition for processing resources between the different words; a difficult word will use most of the resources leaving few resources for the processing of the other words. Because less parafoveal processing has occurred, the next word will become a more attractive candidate to program a saccade to, and thus it is less likely to be skipped.

METHOD

Participants. Twenty members of the University of Massachusetts community participated in this experiment. All participants were native speakers of American English and had 20/20 vision or contact lenses. They were either given extra credit in a Psychology course or paid $8 for their participation.
Apparatus. The apparatus was the same as in Experiment 1.

Materials. 32 sentence frames were created so that the word prior to the target word was either a low- or high-frequency adjective (see Appendix). The mean frequencies, as assessed in the Francis and Kučera norms (1982), were 5 counts per million for the low-frequency adjectives and 270 counts per million for the high frequency adjectives. Word length was controlled: the word prior to the target word was always a 5 letter word adjective, and the target word was always a 3 letter noun. The mean frequency of the three letter target words was 135 counts per million. Two possible previews were created: a correct preview and a misspelled preview. In the misspelled condition, the middle letter of the 3 letter noun was always replaced by the letter x. The combination of the two possible adjectives preceding the target (low- and high-frequency) and the two possible previews (correct and incorrect) produced a 2 x 2 design of which an example is given in Table 4. As each of the 32 sentence frames was read only once by a participant, there were 8 sentences per condition per participant. The 32 experimental sentences were embedded in a pseudorandom order in 112 filler sentences. The boundary was set, as in Experiment 1, between the last letter of the prior word and the space preceding the target word.

To ensure that any differences that were found between the skipping of a target word preceded by either a low- or high-frequency adjective were not due to differences in predictability, we assessed how predictable the target words were in the two frequency conditions. Fourteen participants who did not participate in the eye-tracking study were given the sentence frame up to and including one of the two possible preceding adjectives, and were asked to generate the next word in the sentence. In fact, there was virtually no difference in predictability between the two conditions: the target word was generated 6.25% of the time given the sentence frame with a high-frequency adjective and 6.70% of the time given the sentence frame with a low-frequency adjective.
Table 4. An example sentence of experiment 2 with each of the 4 parafoveal preview conditions.

**Procedure.** The procedure was the same as in Experiment 1.

**RESULTS**

Our primary interest in this experiment was the probability of skipping the target word during the first pass through the text. We will also report the fixation times on the word prior to the target word, to confirm that our frequency manipulation was effective, and fixation times on the target word. The latter is interesting in terms of replicating the basic finding of Henderson and Ferreira (1990) that the parafoveal preview benefit is reduced in the case of high foveal load. Contrary to Experiment 1, we will not report an analysis of the fixation duration on the word prior to the target word as a function of skipping or landing, or as a means to look for potential parafoveal-on-foveal effects. Because the word prior to the target word was not identical in every condition, both the suitability of this design and its statistical power to examine these effects is considerably reduced.

As the target word was very short, it is not surprising that in the vast majority of the cases when there was a fixation on the target word, only one fixation occurred (97.5%). Therefore, as in Experiment 1, only single fixation durations will be reported for the target word. Since one of the manipulations was foveal load, it is of course essential that the word prior to the target word was fixated. Therefore we will restrict all our analyses to those trials in which there was a single fixation on the adjective preceding
the target word. A single fixation on the adjective was the most frequent event when the word was fixated (94.6%), but more importantly, single fixation duration is the cleanest measure to use in this situation, as a second fixation would allow two opportunities to get a parafoveal preview, making the analysis unnecessarily complicated.

Target fixation durations of less than 100 ms and more than 1200 ms were removed from the analyses, as well as trials on which the eye-tracker lost track of the eye position. We also removed trials in which the eyes triggered the boundary but remained on the word before the target (usually on its last letter). Finally, when the fixation duration was greater than three standard deviations from the mean for that participant in that condition, it was also removed (this was the case for 1 observation). All in all, the reported analyses were carried out on 572 trials, or 74% of all the trials. Because a Latin square design was used with relatively few observations in the different cells, the counterbalancing group variable was included in all analyses reported below to increase the power of the design (Pollatsek & Well, 1995). A series of 2 (foveal load) x 2 (preview) repeated measures analyses of variance (ANOVAs) were undertaken with participants (F1) and items (F2) as random variables.

**Skipping the target word.** The skipping probabilities in Experiment 2 are shown in Table 5. The 12% overall difference between the correct and incorrect preview conditions was significant, $F(1,16) = 7.16$, $p < .05$; $F(2,27) = 5.79$, $p < .05$, and the 8% overall difference between conditions with high and low foveal load was significant across participants, $F(1,16) = 7.14$, $p < .05$, and marginally significant across items, $F(2,27) = 3.46$, $p = .07$. Although there was a greater effect of the correctness of the preview in the high foveal load conditions than in the low foveal load conditions, the interaction of correctness by foveal load was not close to significant, $F < 1$.

---

10 Deleted cells were treated as missing cases.
Post-hoc t-tests revealed that there was no significant difference in skipping between a correct and an incorrect preview in the case of a low foveal load, $t1(19) = -1.21, p > .20$; $t2 < 1$, nor was there between a high foveal load and a low foveal load in the case of a correct preview, all $t$s < 1. These results indicate that most of the variance in the skipping data can be accounted for by the difference in skipping rate between an incorrect preview with a high foveal load, and the other three conditions.

<table>
<thead>
<tr>
<th></th>
<th>Correct Preview</th>
<th>Incorrect Preview</th>
</tr>
</thead>
<tbody>
<tr>
<td>High foveal load -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low frequent word</td>
<td>.40 (.28)</td>
<td>.25 (.14)</td>
</tr>
<tr>
<td>Low foveal load -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High frequent word</td>
<td>.45 (.37)</td>
<td>.37 (.24)</td>
</tr>
</tbody>
</table>

Table 5. Skipping probabilities as a function of preview and foveal load. Skipping probabilities taking regressions into account are shown in parenthesis.

Fixation times. The fixation times are shown in Table 6. For single fixation durations on the adjective prior to the target word there was a 27 ms effect of foveal load (i.e., the frequency of the adjective), $F1(1,16) = 7.01, p < .05$; $F2(1,27) = 4.81, p < .05$, but there was no effect of the correctness of the preview of the following noun (all $F$s < 1), nor any interaction between these two factors (all $F$s < 1). The 15 ms effect of correctness of the preview in the low foveal load condition was also not significant, $t1(19) = -1.34, p = .19$; $t2(31) < 1$.

For the single fixation duration on the target word itself both the 39 ms advantage when the preview was correct, $F1(1,13) = 4.78, p < .05$; $F2(1,21) = 11.89, p < .01$, and the 29 ms advantage when the foveal load was low were significant, $F1(1,13) = 5.47, p < .05$; $F2(1,21) = 5.43, p < .05$ (this latter effect could be a frequency spillover effect). The very small (3 ms) interaction between these two factors in the predicted direction was not close to significant (all $F$s < 1).
Chapter 4

Single fixation times

Prior word

Target

Correct Preview Incorrect Preview Correct Preview Incorrect Preview

High foveal load - Low frequent word

315
310
296
333

Low foveal load - High frequent word

278
293
266
306

Table 6. Fixation times (ms) on the word prior to the target word and on the target word.

Discussion

Even though our frequency manipulation on the word prior to the target word was effective, we did not replicate the basic finding of Henderson and Ferreira (1990; see also Kennison & Clifton, 1995; Schroyens et al. 1999; White et al., 2005) that the parafoveal preview benefit is reduced in the case of a high foveal load; there was only a small and insignificant interaction between the foveal load and the preview condition on the single fixation duration on the target word. What we did find were large main effects of foveal load and preview condition on the single fixation times: an average 39 ms effect of preview and an average 29 ms effect of foveal load.

A possible explanation for this discrepancy could lie in the difference in how the incorrect preview was implemented in the studies. In the Henderson and Ferreira (1990) study, both a visually similar parafoveal preview and a visually dissimilar parafoveal preview were used besides the correct preview, and whereas the similar preview maintained the first three letters of the correct preview, the dissimilar preview consisted of a random string of
letters. The effects reported by Henderson and Ferreira were entirely due to the difference in preview benefit observed between the correct and the visually similar condition on the one hand and the dissimilar condition on the other hand. Based on this analysis, it is possible that our incorrect preview was not dissimilar enough to elicit the interaction because only the middle letter was changed in the incorrect preview condition. However, that explanation seems unlikely because we did obtain large preview effects (i.e., differences between the correct and incorrect preview) on the fixation times on the target word. In addition, a large spill-over effect from the frequency manipulation on the prior word was also observed, indicating that our foveal load manipulation was powerful. So while our experiment did not replicate the interaction effect of preview and foveal load reported by Henderson and Ferreira (1990), we did obtain substantial preview and foveal load effects. Spillover has been explained within the E-Z Reader model as one of the consequences of a reduced parafoveal preview. A large foveal load will reduce the amount of parafoveal processing that can be done before the eyes arrive on the next word, causing a longer fixation after a difficult word because more processing is still left to be done in order to identify the newly fixated word. So while our fixation times on the target did not replicate a reduced parafoveal preview, as it is traditionally assessed by comparing the fixation time when there was a correct preview versus an incorrect preview, there were indications that our foveal load manipulation was effective in reducing the amount of parafoveal processing. We will explore alternative explanations for the essentially additive pattern of data we obtained in the General Discussion.

Turning to the skipping rates, effects of foveal load and preview were also observed, but mostly because the skipping rate of the incorrect preview with a high foveal load was considerably lower than the other conditions. We

11 A random string of letters also constituted the incorrect preview in the Kennison and Clifton and White et al. studies, whereas Schroyens, et al. (1999) used a random permutation of pixels. No visually similar condition was used in these studies.
should note that our pattern of skipping rates is similar to that of White (2004) except that she did not obtain any significant effect of foveal load on skipping rates. That is, she also found that the lowest skipping rates were associated with a high foveal load and an incorrect preview. Although the patterns of effects observed in the single fixation times and the skipping rates seem to suggest a common underlying cause (i.e. reduced parafoveal processing), the fact that the high foveal load and incorrect preview condition stands out in the skipping data leads us to believe that the story may be more complicated for saccade target selection. We will also discuss this issue further in the General Discussion.

**GENERAL DISCUSSION**

We will first discuss what we think we have learned about the causes of skipping, then discuss the effects of skipping on processing neighboring words, and finally touch on related issues, such as the effects of our manipulations on the fixation durations on the target word and neighboring words.

The first issue is why words are skipped. Our data raise two issues: (a) the causes for the differences in skipping rate between our conditions and (b) why letter strings are skipped even when they are nonwords or words that are anomalous in the sentence context. Clearly, the fact that we obtained differences in skipping rate for target words as a function of the letters that were there (with the length of the letter string held constant) indicates that processing of the word to the right of fixation is influencing the frequency with which it is skipped. Moreover, this assertion is relatively uncontroversial. However, the extent of the processing of the skipped word that is causing these differences in skipping rates remains an issue of some debate in the literature (e.g. Radach & Kennedy, 2004; Rayner & Juhasz, 2004; Reichle, Rayner, & Pollatsek, 2003). For example, the E-Z Reader model (Reichle et al., 1998, 2003) posits that one of the primary mechanisms
for skipping a word is when a word (word_{n+1}) is recognized very quickly in parafoveal vision. This very rapid recognition produces a program to fixate the following word (word_{n+2}), which occurs early enough to cancel the program to fixate word_{n+1}. In contrast, other models (e.g. Brysbaert et al. 2005; Brysbaert & Vitu, 1998; Drieghe et al., 2004) assume that skipping is based on coarser visual information.

A major motivation for Experiment 1 was that Balota et al. (1985) found no difference between the skipping rates for a predictable word and for a nonword that was visually similar to the predictable word – a finding at odds with the assumption that full processing of the parafoveal word was a major cause of skipping. As we argued earlier, however, such a pattern might occur when the launch site for a typical saccade comes from some distance from the skipped word, and thus the difference between the predictable word and a nonword that is orthographically similar to it might not be discriminated by the visual system. As a result, we attempted to replicate the experiment, but examining carefully the locations of where saccades were launched that either did or didn’t result in skipping the target word. Our findings were (a) that there was a large difference in skipping probability between the predictable word and the visually similar nonword from near launch sites, but (b) almost no difference in skipping probability between these conditions from far launch sites.

This, of course, raises the question of why there was a difference between the two experiments. One possibility is that virtually all of the skipping in the Balota et al. study was from far launch sites; however, this seems improbable. Instead, we think that the most likely reason for the difference between the two experiments is that the parafoveal information was more difficult to extract in the original Balota et al. experiment, largely because the font in their display system was not nearly as legible as the fonts in current computer display systems and possibly because our words, on average were somewhat shorter than in Balota et al. If one only pays attention to significant results, this would be the end of the story: predictable
words are skipped more than anything else in the same location. However, there is a suggestion that the words in the other two parafoveal word conditions and the nonwords in the visually similar condition were skipped a bit more (3-4%) than the other two nonword conditions. In terms of the E-Z reader model, this could be explained by assuming (a) that the words in the other two conditions were identified rapidly a small fraction of the time without the aid of predictability and (b) that the visually similar nonword was occasionally misidentified as the predictable word. However, it is an open question as to whether this could really be predicted by the model quantitatively.

In sum, the results of Experiment 1 indicate that the differences in skipping rates between the conditions is based chiefly on a complete identification of the word in the parafovea, consistent with the E-Z Reader model. Moreover, we think we have made a strong test of this because the difference between the predictable word and the visually similar non-word was a single letter. This leaves open the question of why there was over a 10% skipping rate even for nonwords that presumably didn’t look like any word. The two mechanisms posited by E-Z Reader seem the most plausible. First, there is quite a bit of evidence that indicates that there is error in saccadic programming, such that there is a non-trivial probability that the word targeted is not the word fixated. Thus, there is a reasonable probability that a saccade intended for the target word or nonword overshot the word and resulted in a skip. There is some confirmation for this hypothesis when one looks at the regressions back to the target word. When the preview was the predictable word, the rate was only 2%, whereas it varied from 3-7% in the other conditions, suggesting that the preview was processed, intended to be fixated, and then there was a “double-take”. A second mechanism posited by E-Z Reader 8 (Pollatsek et al., 2005; Rayner et al., 2005) is that some skipping is based solely on predictability. That is, a decision is made to skip the following word based on no visual information other than that the parafoveal string is approximately the length of the predicted word. Whether
these two mechanisms can predict a 12% skipping rate for orthographically illegal strings of the correct length is an open question.

Are the results of Experiment 2 compatible with these conclusions? As indicated earlier, the major results of Experiment 2 were: (a) that a correct preview of the target word was skipped more than a visually similar nonword; (b) that skipping rates were higher when the word before the target word was higher frequency; (c) but that most of the above two differences were due to skipping rates being lower for nonword previews preceded by low frequency words than in the other three conditions. The first finding of a difference in skipping rates between the correct word and the nonword is quite compatible with Experiment 1. As the target words were short and relatively frequent, identification times for these words could have been short enough to produce increased skipping, even without being predictable. Similarly, making the prior word lower frequency would delay the beginning of parafoveal processing (according to the E-Z Reader model) and thus reduce the amount of skipping.

There are two problematic aspects to the data, however: the pattern of the interaction mentioned in point (c) above; and the fact that skipping rates were still 25% in the worst condition (a nonword following a low-frequency word). Let’s consider the second phenomenon first. Are these skipping rates abnormally high? First, the target words are all short reasonably high frequency nouns. As a result, one would expect them to be identified reasonably quickly, especially as the prior words were also reasonably short (about 5 letters) and thus the fixation prior to the target word was likely to be pretty close to it. However, this doesn’t explain why visually similar nonwords would also be skipped. The simplest explanation is that some of the time, the nonword is misidentified as the word and that, in these cases, the misidentification doesn’t slow processing all that much. This explanation, however, appears to run into trouble because when fixation times on the target word were examined there was a healthy (35-40 ms) difference in fixation time between when the correct and incorrect preview
were presented. This might not be a problem, though, if one considers the cases when the target word was skipped as those trials when the incorrect (middle) letter was misidentified and the cases when the word was fixated as those trials when the incorrect letter was correctly identified. Clearly, this is a speculative post-hoc explanation that would need to be verified somehow.

To interpret the size of the skipping rates obtained in Experiment 2, it is also important to note that the analysis was restricted to instances in which the reader fixated on the 5-letter adjective preceding the target word (a 3-letter noun). Taking into account the fact that the average saccade length reported in the literature is 8 character positions and that the perceptual span for letter identification expands 7 to 8 letters to the right of the fixation (see Rayner, 1998 for a review of studies examining these factors), it is safe to say our target word was in an area of high visibility and that skipping the word would not entail executing an especially large saccade. Keeping this in mind, the selection of the fixation data carried out in Experiment 2 has another important consequence. Prior research (McConkie et al., 1988) has established the existence of a so-called range effect; the oculomotor system tends to undershoot targets at a large eccentricity and to overshoot targets at a small eccentricity. By restricting our analyses to those instances on which there was a fixation on the 5 letter adjective preceding the target word, we are also increasing the chances of involuntary overshooting of the target (as compared to studies that do not select close-by launch sites). This latter phenomenon would also explain some of the high skipping rates regardless of condition.

The pattern of the interaction, however, seems harder to understand. According to the E-Z Reader model, skipping occurs only when identification of word_{n+1} is really rapid, and thus we had expected the opposite interaction: where skipping rates would be higher when the preview

---

12 In Experiment 2 the average saccade length based on the three saccades prior to encountering the target word was 7.5 character positions.
was correct and the prior word was high frequency than those in the other three conditions. It seems more difficult to explain why there should be a bigger difference in skipping rates between the correct and incorrect preview conditions when the prior word is low frequency, and thus presumably less processing of the parafoveal word is taking place. It might seem that a possible explanation is that fixation times on the prior word are 30 ms longer when that word is lower frequency and thus allows more processing of the target word. However, according to the E-Z Reader model, the lower the frequency of a word, the less time there is to process the next one parafoveally because the signal for an eye movement precedes the attention shift to the next word, and furthermore the gap between these two events increases with decreasing frequency of the word. Although it is possible that a more parallel encoding model might be able to explain this interaction by pointing to the increased fixation time on the target word, it is far from clear that it can. That is, the issue is not how much total time there is to process the parafoveal word, but how much time there is before the signal to skip the word.

We also wondered whether this strange interaction could be due to fast readers and slow readers each having a different pattern than the overall pattern, with the overall pattern being the result of averaging the two different patterns. Another possibility we considered was that the interaction was largely due to fast readers having developed a strategy that allows them to skip words more frequently. That is, given that fast readers make fewer fixations and have longer saccades (Rayner, 1998) they may adopt a strategy of skipping short words by default and only canceling such saccades when everything points in the direction of a long saccade being inappropriate. To examine this issue in more detail, we split our participants into two groups containing the 10 fastest readers and the 10 slowest readers, based on the overall reading speed; the 10 fastest readers had an average total viewing time of the sentences in the experiment that was shorter than 3005 ms. As can be seen in Table 7, there was little evidence for a difference between the groups in terms of the pattern of the interaction, even though the 10 fastest
readers skipped 37% more, on average, and made 17% more regressions back to the target word after skipping. Thus, we view the pattern of interaction in the skipping data as a problem we still haven’t solved.

<table>
<thead>
<tr>
<th></th>
<th>10 fastest readers</th>
<th>10 slowest readers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct Preview</td>
<td>Incorrect Preview</td>
</tr>
<tr>
<td>High foveal load -</td>
<td>.57 (.36)</td>
<td>.38 (.18)</td>
</tr>
<tr>
<td>Low frequent word</td>
<td>.23 (.19)</td>
<td>.12 (.09)</td>
</tr>
<tr>
<td>Low foveal load -</td>
<td>.63 (.49)</td>
<td>.50 (.32)</td>
</tr>
<tr>
<td>High frequent word</td>
<td>.27 (.24)</td>
<td>.23 (.17)</td>
</tr>
</tbody>
</table>

Table 7. Skipping probabilities of the 10 fastest and slowest readers as a function of preview and foveal load. Skipping probabilities taking regressions into account are shown in parenthesis.

To quickly summarize the above, in spite of a couple of aspects that are not easy to explain, the overall pattern of skipping data is consistent with the following principles. First, a reasonable amount of skipping is explained by the parafoveal word being easy to identify fully, either on the basis of it being predictable, short, and/or high frequency. Second, there is a residue of skipping that seems to be explained by mistargeting of saccades, which would lead to more skipping for shorter words. Third, it appears that some additional skipping might be explained when a string that is similar to either a frequent and/or predictable word is misidentified as that word. Fourth, some additional skipping might be due merely to guessing that the next stimulus is a predictable word if the parafoveal string is approximately the right length.\(^{13}\)

\(^{13}\) Although prior research (Drieghe et al., 2004) has shown that this effect is difficult to obtain in the complete absence of any orthographic overlap with the predictable target word.
Now let us move on to other phenomena related to skipping. The first is how skipping word_{n+1} affects fixation times on the prior word. As indicated in the introduction, the phenomenon of inflated fixations prior to word skipping has been considered a cornerstone in the discussion of the time course of foveal and parafoveal processing. The E-Z Reader model, which posits serial processing of words, predicts a cost associated with the canceling of the saccade to word_{n+1} and the replacement by a saccade to word_{n+2}. Parallel models, such as SWIFT (Engbert et al., 2002) or Glenmore (Reilly & Radach, 2003), do not assume such a cost associated with the skipping of a word. In Experiment 1, we observed a large (34 ms) inflation of fixation time on the prior word if the target word was skipped as did Pollatsek et al. (1986) and Rayner et al. (2004). Admittedly there are results to the contrary (e.g., Kliegl & Engbert, 2004). However, as we indicated earlier, the comparison is complicated as any such comparison is correlational because the reader, and not the experimenter, decides when a word is skipped. One such correlational artifact that could work to produce these inflated times prior to skipping is that when a fixation is longer, it gives the reader more time to process the next word and hence skip it. However, we think that the most plausible artifacts of the correlational structure of this comparison would work against finding this cost due to skipping and could explain why null results are sometimes found. That is, words are more likely to be skipped by good or motivated readers (or readers paying close attention at that moment) and such readers are also more likely to produce shorter fixation times. We should point out that the existence of inflated fixation durations prior to skipping is not necessarily threatening to parallel models for the reason indicated above: longer fixation durations on the foveal word (e.g. due to random variations in fixation times) will allow increased parafoveal preview, and as a consequence more skipping.

The second finding is that the prior fixation durations in the five non-predictable conditions were longer than the prior fixation duration in the predictable preview condition in Experiment 1. These results are compatible with parafoveal-on-foveal effects. However, the effect was localized to trials
when the prior fixation was on the last three characters prior to the beginning of the target word. Such an effect was first reported by Rayner (1975). He found that the fixation durations at the launch site were longer when the following letter string was a nonword than when the launch site was three or fewer character positions away from the beginning of the target word (similar to our finding in Experiment 1). Rayner, Warren et al. (2004) also reported what could be assumed to be a parafoveal-on-foveal effect in a study dealing with plausibility. When word_{n+1} was anomalous, Rayner, Warren et al. (2004) found that readers’ fixations were longer in the three character region preceding the target word. They attributed the finding to (a few) mislocalized saccades (undershoots). It is interesting in the present experiment (as well as Rayner, 1975) that not only were the fixations longer in the three character region preceding the target word for all the other conditions than when the predictable word was in the parafovea, but the fixations on the target word region were also longer for these other conditions. This is not surprising, because when readers undershoot a target word, the duration of the mislocated fixation should plausibly be increased (because the reader is really processing the target word) and they would then fixate directly on the target word (to confirm what they have already read). But, of course, in our experiment (as well as Rayner, 1975), a display change occurs between the two fixations and the difference between the pre-display change word and the post-display change is also likely to lead to longer fixations.

To summarize this last discussion, our experiments were not really designed to examine fixation durations, and hence not designed to test these predictions of serial and parallel models of attention in reading. However, we think there is nothing in the data to indicate that the serial processing assumption of the E-Z Reader model is wrong, and the large cost of skipping on the fixation time of the prior word observed in Experiment 1 is quite compatible with such an assumption and might be hard to parallel models to predict quantitatively.
The other issue our experiments addressed is the relationship between the difficulty of processing a word and the amount of parafoveal processing that occurs on the next word. In our discussion above, we examined this issue with respect to skipping rates, and found that the pattern of results in Experiment 2 was different from that predicted by the E-Z Reader model and also at odds with the findings reported by Henderson and Ferreira (1990; see also Kennison & Clifton, 1995; Schroyens et al., 1999; White et al., 2005). That is, in the prior findings, there was a greater benefit from a correct preview (and/or cost from an incorrect preview) when processing the prior word was easy (e.g., high frequency). As indicated earlier, in the skipping data, we observed the opposite interaction, but in the fixation times on the target word (which was the focus of the earlier studies), we observed additive effects of whether the preview was correct or not and whether the prior word was high or low frequency. There are two differences between our stimuli and those in the prior experiments that may explain the difference. First, as indicated earlier, our incorrect preview was less distorted than the previews previously used in research examining the effects of foveal load on parafoveal processing (Henderson and Ferreira, 1990; Kennison & Clifton, 1995; Schroyens et al., 1999; White et al., 2005), as there was only the change of a single internal letter. However, the difference in pattern between Experiment 2 and these other studies can’t merely be due to insensitivity of the present design, as there was almost a 40 ms main effect due to the correctness of the preview in the fixation time data. Second, our target words were only three letters, and there may be something different about how these short words are processed that explains the pattern of effects, although the data do not offer any obvious clues for what would account for the difference in pattern.

The prediction of the E-Z Reader for the interactive pattern in which there is a bigger preview effect when the prior word is high frequency is based on the assumption that the second stage of lexical processing, L2 is solely a function of word frequency. However, this is undoubtedly an oversimplification and other factors are likely to come into play. One
possibility is that competition among possible lexical entries is another factor influencing the later aspects of word identification. In an activation-verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982), for example, such a competition among lexical entries is explicitly the second stage of lexical access. Thus, one possibility for three letter words is that the frequency of the word is only a minor determinant of the time for later stages of word identification, and that factors such as neighborhood size and whether there is a higher frequency neighbor play a more important role in these later stages of lexical identification and thus in the amount of preview benefit. This process might also be modulated by a rapid identification of the “outside” letters relative to identification of the middle letter due to less lateral inhibition of these letters.

In conclusion, we have found strong evidence that word skipping is usually based on a full identification of the word in the parafovea, consistent with the mechanisms described in the E-Z Reader model. Our first experiment showed that even when a preview of a word is different from the preview of a predictable word by only one single letter, this manipulation is already sufficient to cause a difference in skipping behavior. The fact that in the second experiment our frequency manipulation on the prior word led to comparable patterns in the fixation times and skipping data of the following word adds further evidence to the argument that both factors are influenced by a common phenomenon, the amount of parafoveal processing. However, we did not find a reduced parafoveal preview benefit in the case of high foveal load as reported by Henderson and Ferreira (1990), possibly due to parafoveal processing reaching ceiling values or due to there being something different about how three-letter words are accessed. Inconsistent with earlier E-Z Reader models and with most other models of eye movements in reading (e.g. the SWIFT model, Engbert, et al., 2002), we
found a non-trivial amount of skipping of “garbage” words\textsuperscript{14}. This was especially true when the eyes were close to the target word and the target word was very short. It was by the use of shorter stimuli in both experiments, as compared to most previous research on word skipping, that this finding was established. These data indicate that sometimes a reader prefers to skip a word based on more coarse information, presumably by relying on extra processing that will be done on the fixation after the target. Whether this “guessing” mechanism constitutes a real default or whether our reported results were due to individual strategies (e.g., those of fast readers) will have to be examined in future research. In both cases, models of eye movements in reading will have to take into account an amount of skipping based on an incomplete identification.

\textsuperscript{14} E-Z Reader 8 (see Pollatsek et al., 2005; Rayner et al., 2005, incorporates a “guessing” mechanism from predictability that is consistent with such a phenomenon)
REFERENCES


White, S. J. (2004, May). Foveal processing load does not modulate the probability of skipping the following word: Implications for models of eye movement control in reading. Paper presented at the Sixth European Workshop on Language Comprehension, St Pierre d’Oléron, France.


APPENDIX

The 32 sentences used in experiment 2. The 3 letter target word is always preceded by respectively the low-frequency adjective and the high frequency adjective.

1. The artist painted a lilac/brown sky which clashed with the orange flowers.
2. One day last week, a nasty/small man came in who made a fuss about the high prices.
3. Lying on the table was a crisp/whole nut waiting to be eaten.
4. The farmer had a lot of trouble with a moody/quiet pig that wouldn't eat its food.
5. The secret ingredient in the cocktail was a tasty/fresh egg added at the last minute.
6. There was still a humid/small rag on the floor that the cleaners forgot to pick up.
7. The cafe was noted for its foamy/green tea that wasn't available elsewhere.
8. The box contained a hairy/small pup that was going to be a birthday present.
9. It was so hot outside that they spent the day in a shady/small hut next to the sea.
10. In southern Italy they put spicy/fresh oil on your pizza whether you want it or not.
11. We walked in the dense/heavy fog for more than three hours before we returned home.
12. When she went to the theater, she wore a fancy/black hat that went well with her eyes.
13. On most days, there was a weird/young guy who didn't seem to belong in the bar.
14. The room was filled with stale/fresh air because the windows were broken.
15. The traveling circus featured a queer/great act involving two clowns and a huge duck.
16. They spent their vacation in a surprisingly roomy/large hut before they went back.
17. In their hurry, they bought a shaky/small bed that was broken after two weeks.
18. Reaching into his pocket, he found a shiny/brown key he showed to the children.
19. The actress spoke a soggy/wrong bit of dialogue, after which some people left.
20. The joke was too much for his frail/great ego and he left in a hurry.
21. As she got older, the plump/heavy cow could no longer support her own weight.
22. In the late evening, there were no sober/happy men left except for the driver.
23. He once bought a mauve/green car which surprised everyone because he was so dull.
24. To give to the needy is considered a pious/moral act although few people do it often.
25. As she didn't want to get into the messy/empty bus she took a taxi.
26. The salesman was a jolly/happy guy who was always in for a joke.
27. The house was ruled by a cocky/black dog that killed the hamster yesterday.
28. The house was haunted by a timid/small cat that always ran away when she saw someone.
29. Nobody could have guessed that the agile/young boy actually hated athletics.
30. There was no trouble hearing the noisy/brown owl preparing for his nightly escapades.
31. For the wedding, she was wearing a silky/white hat that belonged to her grandmother.
32. After lunch he always had a cup of yummy/green tea regardless of the weather.
CHAPTER 5
GENERAL CONCLUSIONS

The aim of the eye movement research presented in this doctoral dissertation was to further investigate the phenomenon of parafoveal processing in reading. In this final chapter the main empirical findings of this thesis are summarized and the implications for current theories of eye movement control in reading are discussed. We will present a new model of word skipping in reading in which we will try to integrate all the findings reported in this thesis on this topic. The chapter is concluded with some directions for future experimental investigations of parafoveal processing in reading.
RESEARCH OVERVIEW AND THEORETICAL IMPLICATIONS

As mentioned in the Introduction the research presented in this doctoral thesis deals with issues concerning all three major questions in the field of eye movements in reading (Starr & Rayner, 2001). Whereas Chapter 2 focused on the question whether readers process information from more than one word at a time, Chapter 3 and Chapter 4 mainly focused on how much processing of the word to the right has occurred when the system decides to skip the following word. We will deal with both issues separately in summarizing the main findings and discussing the theoretical implications. While doing so we will constantly deal with the third big question in the field: the relative contribution of low-level oculomotor factors versus high-level cognitive processes in determining eye movements in reading.

IS WORD PROCESSING IN READING SERIAL OR PARALLEL?

The question of whether readers process two or more words in parallel or are always limited to processing one word at a time has been a hot topic in the field ever since the discovery of so-called parafoveal-on-foveal effects. These effects refer to the possibility that the processing of the parafoveal word can affect the processing time on the prior word. Such a finding would be considered most damaging to serial models such as the E-Z Reader model (e.g. Reichle, Rayner, & Pollatsek, 2003). Whereas quite a few studies report evidence of such effects (e.g. Hyönä & Bertram, 2004; Kennedy, 1998; 2000, Kennedy & Pynte, 2005; Schroyens, Vitu, Brysbaert, & d’Ydewalle, 1999), there are methodological problems associated with some of these studies, as well as failures to obtain consistent effects across experiments. As a result, a coherent model that allows researchers to predict which effect will appear when and why is still lacking (Rayner & Juhasz, 2004). This is why
we decided in Chapter 2 to directly test a parallel model that had not yet received that much attention in the literature.

According to Schiepers (1980) the currently fixated word and the next word are processed in parallel but with a time delay of 90 ms per degree of eccentricity. In his model, the benefit of seeing the upcoming word is due to the fact that the parafoveal information from fixation $n$ is combined with the foveal information from fixation $n+1$ to boost word recognition, at least when the fixation on word $n$ is of an optimal duration (between 210 and 270 ms). When fixations are shorter or longer, part of the parafoveal preview benefit is lost, because the synchrony in the arrival of parafoveal and foveal information is less than optimal. Another reason why we wanted to test this specific model was because it gave a plausible explanation for a puzzling finding reported by Schroyens et al. (1999). In their experiment participants looked 20 ms longer at a *zzzz* string than at a *zz* string. Schroyens ventured that this effect could be due to the parafoveal word being further away in the case of the *zzzz* string. The need to synchronize would have cued the system to wait longer (for the incoming parafoveal information) before initiating the saccade. In Chapter 2 we tested this in normal reading by adding an extra blank space between the foveal and the parafoveal word. Contrary to our expectations, this manipulation did not cause the eyes to stay longer on the foveal word. We observed a reduced fixation duration prior to a double blank space. A follow-up experiment showed that this effect was not present when we asked participants to mimic eye movements in reading while scanning *z*-strings. So, clearly the phenomenon was linked to the specific task of reading. The most likely explanation of this reduced fixation duration is that the double blank space caused a reduced lateral masking of the surrounding letters, increasing the readability of the two words that are separated by the double blank space.

A further, serendipitous finding in Chapter 2 was that there was a significant effect of parafoveal word length on the viewing times of the foveal word: a longer parafoveal word was associated with a shorter viewing time on the
prior word. We explained this effect by an attraction exerted by long words resulting in a pulling force closely related to the ideas proposed by Kennedy (1998). In the Kennedy (1998) model, the harder word$_{n+1}$ is to process, the stronger it pulls the eyes towards it, in order to optimize the extraction of visual information from the page of text. The question remains however whether the attraction that a longer word in the parafovea exerts, does indeed find its origin in the processing difficulties associated with longer words. We presented an alternative mechanism that downplays the assumptions of the Kennedy (1998) model. The attraction of parafoveal word length could just be a consequence of a strategy that tries to distribute the fixation locations in the most efficient way, landing more on long words and skipping shorter words. If such a strategy exists, it is not inconceivable that it results in an attraction, a pulling force, when a very suitable candidate is close-by. An extra blank space prior to it could make the candidate stand out more, which would explain another finding of Chapter 2; why the parafoveal word length effect was larger in the double blank space condition than in the single space condition. The major difference between the mechanism proposed in Chapter 2 and the one proposed by Kennedy is that our hypothesis does not assume that the parafoveal attraction$^1$ is based on word processing in the parafovea. The only information it requires is the low-level variable word length.

What are the theoretical implications of these findings for some of the current models of eye movement control in reading? Needless to say, our observations put the Schiepers (1980) model back in the freezer. We also have to acknowledge that the parafoveal-on-foveal effects unraveled in the experiments reported in Chapter 2 do not look very damaging for the serial assumption of the E-Z Reader model. A distinction has to be made between the rather low-level parafoveal-on-foveal effects reported here and effects such as for instance the meaning of the word to the right of the fixation

$^1$ Just for the record, the original title we had in mind for Chapter 2 was: “The attraction of parafoveal word length in reading: Size does matter.”
influencing the current fixation. Better visibility of a word due to less lateral interference is not incompatible with the principles underlying E-Z Reader. The same may be true for the effect of the length of the parafoveal word, although it remains an open question whether the model can quantitatively predict such an effect. The Kennedy (1998) model is of course by no means threatened by the findings reported in Chapter 2, we merely pointed out that to account for these specific findings a more low-level approach could do the job as well.

A parafoveal-on-foveal effect of a higher-level nature was observed in Experiment 1 of Chapter 4: fixation durations were longer on the prior word when there was a non-predictable preview compared to when the predictable word was presented as preview. However, this effect was limited to those instances where the eyes were on the last three characters of the prior word. Parafoveal-on-foveal effects are not necessarily inconsistent with a serial processing model such as the E-Z Reader model on the condition that the effect is restricted to when the position of the eye fixation is very close to the parafoveal word. The reason why E-Z Reader can accommodate for such effects is because it incorporates saccadic error. Data collected by McConkie, Kerr, Reddix, and Zola (1988) convincingly showed that eye movements, like other motor movements, have variability and usually do not land exactly on their target. Applied to Chapter 4 this would mean that sometimes saccades fall short of the targeted word so that word_{n} is fixated even though word_{n+1} was the intended target and is the attended word (see also Rayner, Warren, Juhasz, & Liversedge, 2004). When readers undershoot a target word, the duration of the mislocated fixation should plausibly be increased (because they process a non-predictable preview) and they would then fixate directly on the target word (to confirm what they have already read). Needless to say, when readers undershoot a target word they will usually land on the very last or penultimate letter of the prior word.

In summary, we have provided evidence for both a low-level parafoveal-on-foveal effect (i.e. the attraction of parafoveal word length) and a higher-level
parafoveal-on-foveal effect (i.e. the processing of a non-word preview affecting foveal processing) but have to admit that neither one of these two effects poses a serious threat to a serial attention-shift model such as E-Z Reader. Chapter 2 shows a limited low-level parallel processing (the 4 ms effect in normal single spaced reading is of course pretty small), whereas the effects in Chapter 4 could be a consequence of saccadic error.

One final remark needs to be made with regard to the experiments reported in Chapter 2. Besides the findings we already discussed, another lesson can be learned from the observations made in this study. The results of the second experiment in Chapter 2, in which we asked participants to fake eye movements during reading while scanning z-strings, constitute a warning for generalizing results obtained from artificial tasks to normal reading. If nothing else, the differential influence of the double blank space manipulation and the parafoveal word length manipulation on the fixation probabilities in both tasks indicate that different processes are at play. The fact that the extra blank space reduced the skipping of the foveal word in the scanning of z-strings, but had no effect on the skipping rates in normal reading, also shows that during normal reading not all word skipping is due to oculomotor factors; a statement which brings us to the second section of this Research Overview.

**HOW MUCH OF THE PARAFOVEAL WORD IS PROCESSED PRIOR TO SKIPPING?**

Views on the extent to which a word that is skipped was processed during the prior fixation differ rather dramatically. At one extreme, a word is skipped based on an “educated guess”, taking into account only coarse information about the target word (the EOVP model, Brysbaert & Vitu, 1998). At the other extreme, a word is mainly skipped because it was recognized in parafoveal vision on the prior fixation (the E-Z Reader model, Reichle et al., 2003). It will undoubtedly have surprised the unknowing
reader that both models have served as starting points and theoretical frames for empirical chapters in this dissertation, respectively the EOVP model for Chapter 3 and the E-Z Reader model for Chapter 4. Besides practical reasons (Chapter 3 was carried out in Ghent under the long-distance guidance of Marc Brysbaert, Chapter 4 was carried out in Amherst under the guidance of Keith Rayner, a few offices further down the hall), there is also another reason: both models give very feasible accounts for explaining skipping behavior while taking all possible factors into account. The EOVP model states that low-level factors are the main players in determining skipping behavior and that linguistic factors only slightly modulate the skipping probabilities. The E-Z Reader model puts the emphasis on the amount of lexical processing of the parafoveal word. However, the latter model also incorporates a word length effect, in terms of less optimal extraction of information due to reduced visibility, and takes saccadic error into account. So it is perhaps more accurate to weaken the opening statement of this paragraph somewhat and look upon these two theories as two kinds of spectacles through which we look at the same play, while wondering which model has the most efficient casting director – a theoretical enriching experience, to say the least!

In Chapter 3 we wanted to examine two predictions derived from the EOVP model. Brysbaert and Vitu (1998, for an update see Brysbaert, Drieghe & Vitu, 2005) carried out a meta-analysis on the data of all the studies that examined skipping rates of target words of which the processing difficulty was manipulated and that reported the word lengths of the target words. They found that the vast majority of the variance in the word skipping data could be explained based on word length. In other words, to predict how often a word was skipped, it was more useful to know its word length than for instance its predictability. However, in their analysis, the word length variable was largely a between-experiments variable (i.e., the data for the different word lengths came from different articles), whereas the contextual constraints variable was a repeated measure (i.e., was obtained within a single experiment with the same participants reading high- and low
predictable words in a random sequence). Needless to say, this opens the possibility that some of the word length effect was due to confounded variables. So, the first prediction of the EOVP model that we wanted to test in an experiment especially designed for this purpose was to look whether an unpredictable short word would be skipped more often than a predictable long word. This was exactly what we observed.

A second prediction was also derived from the EOVP model. The question here was whether there was any cross-talk possible between the EOVP system and a linguistic factor such as the predictability of the target word. We examined whether an unpredictable word that shared the same word length with the predictable target word was skipped more often, based on this word length congruency, than a control condition (an unpredictable word that was different from the target word in word length). No such cross-talk was observed and we take this as evidence for the statement that visual and language-related factors independently affect word skipping. The EOVP model was subsequently fine-tuned to this observation: Word length is used in the very beginning of the fixation to obtain a rough estimate of the chances of recognising the parafoveal word by the end of the fixation, and this estimate is used to decide whether or not to program a saccade to this word. The initial decision can be overruled on the basis of the incoming linguistic information, but this latter decision does not take into account the word length information on which the original decision was based. These are two functionally independent subsystems.

To what extent are the findings reported in Chapter 3 compatible with the skipping mechanism proposed by the E-Z Reader model (Reichle et al., 2003)? In this model, word skips consist of a cancelled saccade to word_{n+1}, which is replaced by a saccade to word_{n+2}. This replacement process is assumed to require extra time, so that fixation durations prior to skipped words on average will be longer than fixation durations prior to fixated words. However, such an inflated fixation duration prior to skipping was not observed in Chapter 3. Another finding that is somewhat at odds with the
assumptions of the E-Z Reader model is related to the size of the predictability effect on word skipping: The predictable target word was skipped 9% more often than an unpredictable target word. Even though this is a strong effect, the differences in sentence completion ratio for the unpredictable and the predictable target words were among the largest ever reported, and the size of the effect still comes a long way short from the whopping word length effects (23%) we observed. As mentioned earlier, the E-Z Reader model does incorporate a word length effect by assuming less optimal extraction of information due to reduced visibility in the case of a long word. However, it remains to be seen whether the E-Z Reader model can accommodate for such a huge word length effect, as the difference in word lengths used in the experiment was limited to 2 letters (2 letter words versus 4 letter words). Hence, the question remains whether the E-Z Reader model has cast the right factor as the main player in determining skipping behaviour (see also Brysbaert & Drieghe, 2003).

Chapter 3 approached the problem of how much processing has occurred on the fixation prior to skipping by trying to determine how much of word skipping can be explained based on a low-level factor such as word length. In Chapter 4 we took the opposite approach to this issue. We started from the assumption that a word was mainly skipped because it was recognized on the prior fixation and looked at two studies that seemed to be in conflict with this assumption.

The first experiment reported in Chapter 4 was essentially a replication of the Balota, Pollatsek and Rayner (1985) study in which the skipping probabilities of a predictable target word (cake) and a visually similar non-word (cahc) were compared. Balota et al. reported no differences in skipping probabilities between these two conditions. This led them to conclude that the decision to skip the target word was not based on a full analysis of the parafoveal word. This conclusion, however, is somewhat at odds with the E-Z Reader model we presented earlier. That is, if a word is skipped because it is almost fully recognized, how can there be no difference at all between a
predictable word and a visually similar non-word? We hypothesized that such a pattern might occur when the launch site for a typical saccade in the original study came from some distance from the skipped word, and thus the difference between the predictable word and the orthographically similar non-word might not have been discriminated by the visual system. Therefore, in Chapter 4 we attempted to replicate the experiment, but examined carefully the launch sites of the saccades that either did or did not result in the skipping of the target word. Our findings were (a) that there was a large difference in skipping probability between the predictable word and the visually similar non-word from close launch sites, but (b) almost no difference in skipping probability between these conditions from far launch sites. The discrepancy between the present findings and those reported by Balota et al. was explained in terms of poorer viewing conditions in the original study, making the parafoveal information more difficult to extract. The font in their display system was not nearly as legible as the fonts in current computer display systems. It should also be noted that virtually the entire effect of predictability occurred when the launch site was close, indicating that the effect was likely to be due to fairly full processing of the skipped word.

The fact that there were no significant differences in the 5 non-predictable conditions is also in line with the assumptions of the E-Z Reader model. The lack of difference in skipping between a neutral word and a semantically anomalous word could be expected based on previous research (Altarriba, Kambe, Pollatsek, & Rayner, 2001; Rayner, Balota, & Pollatsek, 1986) showing that semantically related words are not skipped more often than semantically unrelated words. In the E-Z Reader model this finding is explained by stating that the decision to skip a word is instigated by the completion of the first phase of word identification of the target word. Whereas predictability appears to boost performance in this phase, the extraction of semantic features from the parafoveal word apparently does not (or at least does not do it fast enough), explaining the lack of an effect of semantic inhibition. Likewise, the mechanism explaining skipping behavior
incorporated in the E-Z Reader model does not make any differential predictions on skipping behavior in terms of inhibition. The system apparently decides to cancel the planned saccade to the target word only when the speed of the first phase of word recognition of the target word is sufficiently boosted.

Contrary to Chapter 3, we did observe a substantial inflated fixation duration prior to skipping (34 ms), adding further support for the existence of this effect (e.g. Pollatsek, Rayner, & Balota, 1986). At this point, we will first discuss the findings of Experiment 2 of Chapter 4 and then turn to the issue of the high rate of skipping non-word previews which was observed in both experiments reported in Chapter 4. Besides the Balota et al. (1985) paper, the second study of which the outcome seems at odds with the assumptions of the E-Z Reader model was reported by White (2004). This study focused on the question whether there is an influence of foveal load on the skipping of the following word. Previous research (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; White, Rayner, & Liversedge, 2005) demonstrated that a high foveal load leads to a reduced parafoveal preview benefit. In the E-Z Reader model, a word is skipped because it is recognized in parafoveal vision. Thus, it follows that the chances of recognizing the parafoveal word would be reduced when the amount of parafoveal preview is reduced. However, White (2004) did not find an effect of the foveal difficulty of word \(n\) on the skipping rates of word \(n+1\). There may have been a power problem in this study as the skipping rates were quite low, probably due to the use of 5 to 6 letter foveal words and 4 letter target words. Therefore, in the second experiment of Chapter 4 we used shorter words for both the foveal words whose difficulty was manipulated (5 letters) and for the ensuing target words that were being examined for skipping probabilities (3 letters). Even though our frequency manipulation on the word prior to the target word was effective, we did not replicate the basic finding of Henderson and Ferreira (1990; see also Kennison & Clifton, 1995; Schroyens et al. 1999; White et al., 2005) that the parafoveal preview benefit is reduced in the case of a high foveal load. This was presumably due to our
preview manipulation being less distorted than the previews used in prior research, or due to differences in how a three letter word is processed as compared to longer words. What we did find in the skipping rates was a main effect of foveal load and preview, but mostly because the skipping rate of the incorrect preview with a high foveal load was considerably lower than the other conditions. We should note that our pattern of skipping rates is similar to that of White (2004) except that she did not obtain any significant effect of foveal load on skipping rates. A similar pattern of main effects was also observed in the single fixation times on the target: an average 39 ms effect of preview and an average 29 ms effect of foveal load. Although the patterns of effects observed in the single fixation times and the skipping rates seem to suggest a common underlying cause (i.e. reduced parafoveal processing), the fact that the high foveal load and incorrect preview condition stands out in the skipping data leads us to believe that the story may be more complicated for saccade target selection. The truth is, even though we explored a few possible explanations (e.g. the strategy used by a fast reader) we still regard the unexpected pattern in the skipping data as a jigsaw puzzle yet to be solved. The fact that we replicated the (non-significant) pattern found by White (2004), only adds further importance to this observation.

Another interesting finding made both in the first and the second experiment of Chapter 4 was the high skipping rate of non-word previews. This finding is of course incompatible with the word skipping mechanism embedded in the E-Z Reader model. The model states that a word is mainly skipped because it was recognized in parafoveal vision. How could this apply to the preview of a non-word? Three possible explanations were presented of which the combination could account for these problematic findings: (a) as already mentioned the E-Z Reader model incorporates saccadic error, explaining why the skipping rate for a non-word preview is above zero. And we expect saccadic error to be larger in these specific analyses. Prior research (McConkie et al., 1988) has established the existence of a so-called range effect; the oculomotor system tends to undershoot targets at a large
eccentricity and to overshoot targets at a small eccentricity. It is important to note that by restricting our analyses to those instances when the eyes were very close to the target word, we are also increasing the chances of involuntary overshooting of the target (as compared to studies that do not select close-by launch sites). This phenomenon could explain some of the high skipping rates regardless of condition. (b) Some additional skipping might be explained when a string that is similar to either a frequent and/or predictable word is misidentified as that word. In other words, skipping can be based on an incomplete identification, an observation that is compatible with the original Balota et al. (1985) paper. After all, even if the viewing conditions were relatively bad in that study, skipping did take place. In all likelihood some of that word skipping was based on a partial analysis of the parafoveal word. (c) Finally, some skipping might be merely due to guessing that the next stimulus is a predictable word if the parafoveal string is approximately the right length. Even though this latter hypothesis could explain some of the skipping of the non-word previews in Chapter 4, based on the findings reported in Chapter 2 we regard this hypothesis as somewhat unlikely.

Where does this leave us with regard to the theoretical implications of the findings reported in Chapter 3 and Chapter 4 on the issue of word skipping? We believe it to be a fair summary that even though these two chapters provide an amplitude of new data and new insights, they also constitute a tough nut to crack in terms of integrating all the findings in one coherent view on word skipping. Let’s summarize. We have shown that word length is in all likelihood the main player in determining skipping behavior and that visual and language-related factors independently affect word skipping (Chapter 3). In Chapter 4 we observed that skipping can be based on a full analysis of the parafoveal word; the predictable word was skipped more often, even though the difference with the visually similar word was limited to a single letter. However, in Chapter 4 we also observed a skipping rate of non-word previews that was just too high to be merely due to saccadic error, implicating that the system can sometimes skip based on an incomplete
identification of the parafoveal preview. No further differences were reported between the skipping rates of the non-predictable previews. Whereas in Chapter 4 we observed an inflated fixation duration prior to skipping, no traces of such a phenomenon were found in Chapter 3. And finally, there is still the unexplained pattern of the skipping data in the foveal load experiment: the effects were mostly due to a reduced skipping of a non-word preview that was preceded by low-frequency word. Neither of the two theoretical frames from which we started (i.e. the EOVP model and the E-Z Reader model) can accommodate for the combination of these findings. Therefore we choose not to sum up which findings are compatible with these two models and which are not, but instead we choose to present a new model of word skipping in reading that tries to accommodate for all these findings.

**TOWARDS A NEW MODEL OF WORD SKIPPING IN READING**

Whereas the proposed new model incorporates features of both the EOVP model and the E-Z Reader model, it also departs from them on a number of fundamental characteristics. One of these characteristics, and perhaps the most important one, deals with the issue of the *default saccade target*. Before taking a stand on this issue, we will discuss a study that offers interesting insights on this topic. Reilly and O’Regan (1998) tried to simulate the empirical findings of McConkie, et al. (1998) on landing site positions with four different computer models. The first computer model simulated a word-by-word reading strategy, in which the target word was always the first word to the right of the fixation location. The second computer model was based on a strategy in which the target word was the longest word in the 20-letter window to the right of the fixation location. The strategy of the third computer program consisted of skipping words as a function of their frequency. Finally, the fourth computer program was based on an implementation of Morrison’s (1984) sequential attention model, the predecessor of the E-Z Reader model. Reilly and O’Regan (1998) observed that the second computer program yielded the best
The empirical data were best predicted by a strategy that consisted of simply choosing the longest word in the parafovea, without any further identification of the words involved. Bearing these findings in mind we turn to one of the core assumptions of the new model.

The mechanism for the default saccade target that will be incorporated in the present model will rely heavily on the concept of preferred saccade length. We believe that every individual reading a piece of text develops a preferred saccade length. Influencing factors on this variable are text difficulty, task demands and individual differences. A difficult text will induce a smaller preferred saccade length as compared to an easy text. Likewise, reading for comprehension will induce a different saccade length than for instance proof reading where we expect a shorter preferred saccade length. Finally, a slow reader or a beginning reader will also have a shorter preferred saccade length (see Rayner, 1998 for an overview on related findings).

The word located at the preferred saccade length² from the current fixation location will be the initially preferred saccade target. However, a large word in the parafovea will exert a certain attraction, not only in terms of attracting an early saccade from the currently fixated word but also in terms of making it stand out more as an ideal candidate for being the saccade target. This latter phenomenon is closely linked to earlier observations made on the so-called global effect (Vitu, 1991; Gautier, O’Regan, & Le Gargasson, 2000); when the eyes make a saccade to a target word, the movement is influenced not only by the visual characteristics of the target word but also by the surrounding stimulus materials. This attenuation of the saccade target is also related to the findings reported both in Chapter 2 and by Reilly and O’Regan.

² Even though the preferred saccade length incorporates no less than three factors (individual differences, task demands and text difficulty), establishing the preferred saccade length for an individual should be relatively easy. By calculating his/her average saccade length in the filler items presented in the experiment, a good approximation of the size can already be obtained.
(1998). More concretely, say that the preferred saccade length would lead to a saccade landing between two words. One of the ways in which the attraction of parafoveal word length can come into play is by making the word with the largest word length the saccade target in this situation. Even though the present model is only concerned with word skipping, it is of course important to note that after the saccade target has been set, the ultimate saccade length will also be attenuated for the desire of the system to land on a position within the word that is more or less optimal for word recognition (i.e. *the preferred viewing position*, Rayner, 1979). However, we will keep the discussion focused on the issue of saccade targets as this will be crucial in our account of skipping behavior.

Summarizing, every reader develops a preferred saccade length during the reading of a text. This preferred saccade length is attenuated by an attraction exerted by long words in the parafovea and results in an *intended saccade target*. We use the word “intended” to fully acknowledge the impact of saccadic error.

What does this mean in terms of which word will usually be the intended saccade target during reading? Let’s take the example of a participant reading a text who has developed a preferred saccade length of 8 characters. Taking this distance into account, as well as average word lengths, the intended saccade target will usually be word_{n+1}. The intended saccade target will be word_{n+2} when either (a) word_{n+1} is short (e.g. function words will frequently be in this scenario), (b) the eyes are close to word_{n+1}, (c) if word_{n+2} is a long word and word_{n+1} is reasonably short, the attraction of parafoveal word length will kick in, or (d) combinations of the previous possibilities.

---

3 We are fully aware that using saccade length as a determining factor in word skipping is a way of incorporating both word length and launch site in one single parameter. As such, our explanation of the choice of the saccade target is much closer to the mechanism described by the EOVP model than for instance the E-Z Reader model.
The information necessary to determine the intended saccade target comes available very soon during the fixation as it only requires low-level information such as parafoveal word length and spacing. With the intended saccade target established, we have reached the second phase of the word skipping model: the three possible scenarios we believe to occur in word skipping in reading, and the frequency of occurrence of these scenarios. The first scenario we will describe is by far the default mechanism when a word is skipped.

1) The intended saccade target is word_{n+2}. The saccade is carried out and as a result word_{n+1} is skipped. If the word processing of word_{n+1} has not yet been completed, the system will compensate by either fixating longer on word_{n+2} while processing word_{n+1} (Binder, Pollatsek & Rayner, 1999) or by making a regression to word_{n+1} (Vitu & McConkie, 2000). This way eye movement control and language processing are never more than one step out of phase.

2) The intended saccade target is word_{n+1}. However, the parafoveal processing of word_{n+1} is going really well in terms of the recognition process of word_{n+1}. The system decides to invest extra effort into the parafoveal processing of word_{n+1}, stays longer on word_{n}, and replaces the intended saccade to word_{n+1} by a saccade to word_{n+2}. As a result word_{n+1} is skipped and we observe an inflated fixation duration on word_{n} as compared to when the system does not decide to skip word_{n+1}.

3) The intended saccade target is word_{n+2}. However, the processing of word_{n} is difficult and the incoming information concerning word_{n+1} is also not very promising. The system decides to cancel the intended skip of word_{n+1} and replaces the saccade program by a saccade to word_{n+1}. 


It is important to put a lot of emphasis on the frequency of occurrence of these three scenarios. The first scenario is clearly the default in word skipping. The system prefers to carry out a saccade to the intended saccade target. The second scenario is not so frequent, even in a specially designed experiment using extremely well predictable target words, the resulting effect is rather limited (e.g. 9 % in Chapter 3). The third scenario is quite rare; readers prefer to correct an incomplete identification after the skip has been executed. Apparently, it is more economical for a reader sometimes to make a ‘wrong’ forward saccade that is immediately corrected, than to adopt a cautious strategy of targeting the next word (see also Brysbaert, et al., 2005). A canceling of an intended skip is reserved for rather extreme situations (e.g. the high foveal load – incorrect preview condition in Chapter 4). Indeed, one can wonder whether a significant effect of scenario 3 can be obtained in normal reading without the use of artificial previews.

How does this model, which we have baptized the *Quidditch* model of word skipping, stand on the time course issue of foveal and parafoveal processing? Quidditch is a parallel model in the sense that the parafoveal word length information is processed simultaneously with the processing of the foveal word. The model assumes some lexical processing in the parafovea that occurs at the same time as the foveal processing but at this point we choose to be rather conservative with regard to the amount of parallel lexical processing. As a consequence, it remains to be seen whether the model

---

4 We named the model after Quidditch, the premier ‘sport’ of the wizarding world described in the Harry Potter books by J. K. Rowling. Without going into details, every one of the three scenarios described above is associated with one of the three balls used in Quidditch: the basic process of deciding to skip a word based on the intended saccade target (getting the *Quaffle* through one of the hoops), the more intensive process of an inflated fixation duration which is rewarded by a longer saccade length (catching the *Snitch*), and finally the rather painful process of canceling an intended skip (being hit by a *Bludger*).

5 The million dollar question here is: How detailed is the information that is being extracted from the parafovea while at the same time the foveal word is being
could also be up to a certain extent compatible with a serial model that allows parallel processing of low-level features in the parafovea. For now, we consider the Quidditch model to be a model with parallel processing of low-level features and with a limited amount of parallel lexical processing.

How does this model account for all the findings reported in Chapter 3 and Chapter 4? Quidditch is consistent with the notion of word length being the main player in determining skipping behavior. The mechanism we described earlier of how the system comes to the intended saccade target is clearly primarily determined by parafoveal word length. Even though the intended saccade target can be overruled on the basis of incoming linguistic information, this latter decision does not have to take into account the word length information on which the intended saccade target was based. In this sense, word length information and linguistic information are dealt with by functionally independent subsystems (Chapter 3). Because Quidditch relies so heavily on the notion of preferred saccade length, it makes differential predictions for experiments in terms of the word lengths of the target words being used and whether the analysis is restricted to close-by launch sites. If you use very short words, chances are high that you will be mostly confronted with the intended saccade target being word_{n+2}. As a result, scenario 2 will be rather rare and you will not find a significant inflated fixation duration prior to skipping (Chapter 3). If on the other hand the target words are a bit longer, an inflated fixation duration can be observed (Chapter 4). In Experiment 1 of Chapter 4 the conditions were ideal for observing the 2nd scenario of word skipping.

Quidditch also explains why an inflated fixation duration prior to skipping has not been observed in corpus studies that did not control for the size of the saccade length that would entail word skipping (Radach & Heller, 2000). If scenario 1 is by far the most frequent scenario, it is impossible to get a...
significant inflated fixation duration because this effect, resulting from scenario 2, will have drowned in the far more frequent presence of scenario 1. Quidditch is also the very first account that is compatible with the puzzling data reported by Kliegl and Engbert (2005). In an extensive corpus study they observed fixation durations prior to skipping a short word that were shorter than prior to skipping a longer word. This is a prediction that can be directly derived from the architecture of Quidditch: skipping a short word will usually be scenario 1, whereas the skipping of a longer word will more frequently result from the 2nd scenario. Because the 2nd scenario is associated with an inflated fixation duration prior to skipping, Quidditch predicts longer fixation times prior to skipping long words.

The rather high rate of skipping non-word previews can also be explained by Quidditch. The use of short words (especially in Experiment 2 of Chapter 4) will have caused the intended saccade target to be quite frequently word_{n+2}. The pattern of effects in the skipping data of Experiment 2 of Chapter 4 (reduced skipping of an incorrect preview after a low frequency word), are compatible with scenario 3 of Quidditch. The system does have an emergency stop but it really hates to use it. This is illustrated by the fact that our different non-predictable previews had no differential impact on the skipping rates of Experiment 1 in Chapter 4: an illegal preview was not skipped less than an unpredictable word, presumably because there was no low-frequency word prior to it, making the manipulation not powerful enough to elicit scenario 3.

Whereas the model already does a good job in explaining both the findings of the present doctoral thesis and those by Kliegl and Engbert (2005), we cannot call Quidditch an adequate model of word skipping in reading before certain criteria have been met. One of the ways to differentiate between a good and a rather poor model is the extent to which the model provides predictions that can be immediately tested in an experiment, leading to either a confirmation or a falsification of the model. This is a challenge we will deal with in the final section of this doctoral dissertation.
AVENUES FOR FUTURE RESEARCH

THE INFLATED FIXATION DURATION PRIOR TO SKIPPING

Quidditch leads to a few very simple predictions. We will start with the inflated fixation duration prior to skipping. Some research has reported the existence of such an effect (e.g. Pollatsek, et al., 1986; Chapter 4), whereas others found no traces of an inflated fixation duration prior to skipping (e.g. Radach & Heller, 2000; Chapter 3). According to Quidditch one of the determining factors of the presence or absence of this effect is the length of the saccade that would entail the skipping of the following word. If the skipping of word$_{n+1}$ requires quite a long saccade, an inflated fixation duration prior to skipping should be observed because chances are high that the intended saccade target was word$_{n+1}$, but the system cancelled the saccade and replaced it by a saccade to word$_{n+2}$ (scenario 2). As the required length of the saccade diminishes, so should the presence of the effect because word$_{n+2}$ will more often be the intended saccade target, hence no inflated fixation duration is expected because there was no planned saccade to word$_{n+1}$ cancelled and replaced.

THE ABILITY TO CANCEL A PLANNED SKIP

In Experiment 2 of Chapter 4 we observed a reduced skipping when an incorrect preview of word$_{n+1}$ was preceded by a low frequency word$_n$. We took this as evidence for a system that can cancel a planned skip but is reluctant to do so. This reluctance was illustrated by the lack of differences in the skipping of non-predictable previews in the first experiment of Chapter 4. Whereas these data formed the main reason to incorporate the third scenario in the Quidditch model, questions can be raised as to whether the canceling of a planned skip can be observed in normal reading (i.e. without the use of incorrect previews). By presenting a short low frequency word$_n$ followed by a short low or high frequency target word$_{n+1}$, the
necessity of this third scenario in word skipping can be established. From the present model we would predict that, given a fixation on the word $n$ and the fact that both word $n$ and word $n+1$ are short words, the intended saccade target will frequently be word $n+2$. Taking into account the reluctance of the model to cancel a planned skip, the low frequency word $n+1$ will still be skipped quite often. This prediction of only a limited reduction of skipping behavior in such a situation is incompatible with most other models of eye movement control in reading. Some reduction of the skipping rates of the low frequency target word $n+1$, as compared to skipping the high frequency target word $n+1$, will be observed due to scenario 3. Should this not be the case, this would establish that scenario 3 does not feature in normal reading.

WORD LENGTH AND LAUNCH SITE AS DETERMINING FACTORS IN SKIPPING BEHAVIOR

By using the saccade length as the determining factor in deciding which word will be the intended saccade target, we incorporate both the effects of launch site and parafoveal word length. As far as we know, no research has been conducted to examine the relative importance of launch site and parafoveal word length in skipping. In other words, if you control word length and look at the effect of launch site on word skipping, and vice versa, would one of the factors have a larger impact on determining skipping behavior?

INDIVIDUAL DIFFERENCES

While it is clear that very strong individual differences exist in eye movements in reading, not much research has been conducted on this topic. This is especially the case for word skipping (for an exception see: Kennison & Clifton, 1995). However, from experience in running word skipping experiments we can say that personal differences in word skipping are rather dramatic. Whereas some readers skip almost every other word, other readers
do hardly ever skip a word. Quidditch predicts that for those latter readers, who are usually the slowest readers, those instances in which they skip will be mostly due to a correction of the intended saccade target as described in scenario 2, and as a result a strong inflated fixation duration prior to skipping is expected in these readers as compared to fast readers.

**Parafoveal-on-foveal effects**

The E-Z Reader model (Reichle et al., 2003) can account for parafoveal-on-foveal effects if these effects are limited to those instances when the eyes are very close to the parafoveal word (i.e. the last or penultimate letter of the foveal word). The reason why the E-Z Reader model can accommodate for such effects is because it incorporates saccadic error. When readers undershoot a target word, the duration of the mislocated fixation could plausibly be influenced (because they process the preview) and they would then fixate directly on the target word (to confirm what they have already read). In our opinion, this is a hypothesis that still needs to be adequately confirmed. More precisely, one would expect the effect to be present almost solely on the last letter and one would expect the effect to originate from a relatively small subset of the data on the last letter (saccadic error that causes the eyes to fixate the wrong word is not thought to occur so often). There is no reason why a parallel model would expect the effect to be due to only a subset of the data. This can easily be tested given a large enough dataset, allowing a sufficient amount of data per character position on the foveal word.
REFERENCES


Brysbaert, M., & Drieghe, D. (2003). Please stop using word frequency data that are likely to be word length effects in disguise. *Behavioral and Brain Sciences*, 26, 479.


White, S. J. (2004, May). Foveal processing load does not modulate the probability of skipping the following word: Implications for models of eye movement control in reading. Paper presented at the Sixth European Workshop on Language Comprehension, St Pierre d’Oléron, France.
