A Special Issue of the

European Journal of Cognitive Psychology

EYE MOVEMENTS AND INFORMATION PROCESSING DURING READING

GUEST EDITORS

Ralph Radach

Alan Kennedy

Keith Rayner

Psychology Press
Taylor & Francis Group

A Special Issue of
The European Journal of Cognitive Psychology

Eye Movements and Information Processing During Reading

Edited by

Ralph Radach
University of Aachen, Germany
Alan Kennedy
University of Dundee, UK
and
Keith Rayner
University of Massachusetts, Amherst, USA

Psychology Press
Taylor & Francis Group

HOVE AND NEW YORK

Published in 2004
by Psychology Press Ltd
27 Church Road, Hove, East Sussex BN3 2FA
http://www.psypress.co.uk

Simultaneously published in the USA and Canada
by Taylor & Francis Inc
29 West 35^{th} Street, New York, NY 10001

This edition published in the Taylor & Francis e-Library, 2005.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Psychology Press is part of the Taylor & Francis Group

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 0-203-48334-0 Master e-book ISBN

ISBN 0-203-59551-3 (Adobe eReader Format)
ISBN 1-84169-956-X (Print Edition)
ISSN 0954-1446

Cover design by Hybert Design, Waltham St Lawrence, Berkshire, UK

Contents*

* This book is also a special issue of the *European Journal of Cognitive Psychology* and forms Issues 1 and 2 of Volume 16 (2004).

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 1–2

Preface

The measurement and analysis of eye movements is one of the most powerful ways to study the workings of the human mind. This Special Issue on eye movements and information processing in reading presents a state of the art overview of experimental research based on this methodology. The origin of this collection of articles dates back to two back-to-back symposia on "Attention and Eye Movement Control in Reading" and "Visual and Cognitive Information Processing in Reading" at the 2001 conference of the European Society for Cognitive Psychology (ESCOP) in Edinburgh. These symposia were organised by Ralph Radach and Alan Kennedy. As a result of stimulating discussions and positive feedback from participants, Radach and Kennedy decided to pursue the possibility of publishing a set of articles dealing with eye movements in reading as a Special Issue. At this point in the process, Keith Rayner was approached about being involved with the project. So as to have a uniform editing policy, it was agreed that Rayner would handle the editing chores (with Radach serving as editor of the papers that Rayner was an author of). All articles submitted for the Special Issue were subsequently thoroughly reviewed by the peer review process. Thanks go to all those who served as reviewers for these articles.

We believe that eye movements provide a unique opportunity to examine principles of human information processing in a well-structured visual environment while people engage in a natural cognitive task. At the same time oculomotor measures can be used as a tool to develop and test psycholinguistic hypotheses on the processing of written language. The articles in this Special Issue all contribute to both aspects, addressing issues that dominate current debates in the field.

Seen from the angle of visual information processing, a major theme in the present articles is the role played by parafoveal information for different types and levels of processing and for oculomotor control. This includes effects of visual and linguistic word properties on the selection of words for fixation and the specification of saccade amplitudes. Clearly the most controversial issue in this context is the allocation of attention, with positions ranging from a sequentially moving spotlight to a gradient of spatially distributed processing. Related to this is the issue of serial vs. parallel word processing and the fundamental question regarding to what extent the duration of reading fixations is related to lexical processing. Taking a psycholinguistic perspective, the topics addressed include several levels of language processing from orthography to pragmatic information in sentence reading. New approaches are taken to study morphologically complex words and to reveal the complex nature of the popular but delusive concept of word frequency. All articles reflect current theoretical discussions centred around the ongoing development of computational models of the reading process and contribute to the empirical base of these discussions.

Together, the present collection of articles, supplemented with an introduction to the field and a commentary on major issues, presents a comprehensive and up-to-date view on a dynamically developing research area. We also think that it reflects the excitement with respect to new findings that currently exists in the field. Most of the contributions are from European-based researchers, and two of the authors (Inhoff and Rayner) from the United States have strong ties to Europe (as they were born in Germany and England, respectively).

We trust that this Special Issue will be attractive to readers of the *European Journal of Cognitive Psychology,* especially for those interested in basic and applied psycholinguistics, attention and visual perception, motor control, and the modelling of complex cognitive processes.

RALPH RADACH
ALAN KENNEDY
KEITH RAYNER

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000303

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 3–26

Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research

Ralph Radach

University of Aachen, Germany

Alan Kennedy

University of Dundee, UK

The study of eye movements has become a well established and widely used methodology in experimental reading research. This Introduction provides a survey of some key methodological issues, followed by a discussion of major trends in the development of theories and models of eye movement control in fluent reading. Among the issues to be considered in future research are problems of methodology, a stronger grounding in basic research, integration with the neighbouring area of research on single word recognition, more systematic approaches to model evaluation and comparison, and more work on individual variation and effects of task demands in reading.

Our aim in this Introduction is to provide a brief survey over the field of eye movements and information processing in normal reading. We will start with a discussion of some important methodological issues, intended primarily to help readers from other areas understand the terminology and get a feel for the kind of data that form the base of work in the field. We will then concentrate on what has been a dominant theme of much recent work, the development of theories and models of eye movement control in fluent reading. This discussion will provide a basis for identifying issues that may have been neglected in prior research and might consequently be considered as topics for the future.

Experimental reading research is an innovative and rapidly expanding field of scientific research (see Rayner, 1998; Kennedy, Radach, Heller, & Pynte, 2000; Hyönä, Deubel, & Radach, 2003; for recent discussions). Within the area, thestudy of eye movements has played a pivotal role, partly because they are aninherent behavioural manifestation of the reading process in action (and henceprovide the researcher with an excellent nonintrusive methodology), but alsobecause they have proved to be a powerful way of studying the workings of thehuman mind. There are at least three different, albeit overlapping, theoreticalperspectives from which the topic can be approached.

First, from the viewpoint of perception and motor control, reading can be viewed as a task where visual processing and sensorimotor control take place in a highly structured visual environment. The text page contains letters, words, lines of text, and paragraphs, forming a hierarchy of visual objects. However, it is an array of visual information significantly less complex than any realistic scene of natural objects. Reading thus allows for the examination of basic issues in perception and motor control in an ecologically valid situation, while subjects are engaged in a complex visual-cognitive task. In reality, this first approach is more a theoretical possibility than a routine research strategy. One of the relatively few, but influential, examples is the linking of eye movement control in reading to a perceptual centre of gravity phenomenon carried out by O'Regan (1990). Another key import from basic oculomotor research that has had a significant influence on theory building is the principle of temporal overlap in the programming of successive saccades introduced into reading research by Morrison (1984).

A second theoretical perspective, rooted in the mainstream of cognitive psychology, sees reading as a special case of a complex process of information acquisition, functioning at roughly the same level as the process of understanding a pictorial scene. A multitude of processing modules need to work together on several levels to turn the arrays of combined visual features comprising letters into words and concepts that eventually form a mental representation of text. Of central importance in this process are the coordination of processes, the timescale in which they unfold, and their interaction with underlying memory systems. Examples here might be the issue of serial vs. parallel processing of linguistic information at the letter and word level, or the question of whether certain types of information are processed in a modular or an interactive way (Pollatsek & Rayner, 1989). Eye movement research from this perspective has led in the development of several realistic computational models of reading that combine mechanisms of oculomotor control and linguistic processing to account for a wide

Correspondence should be addressed to Dr. R.Radach, Institute of Psychology, University of Aachen, Jaegerstr. 17, 52056 Aachen, Germany. Email: ralph@psych.rwth-aachen.deth-aachen.de ralph@psych.rwth-aachen.de
The authors are extremely grateful to Albrecht Inhoff, Wayne Murray, and Keith Rayner for many insightful comments and suggestions. This paper was prepared with the assistance of Grant No R000223650 from the UK Economic and Social Research Council to Alan Kennedy.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000295

range of empirical phenomena (see Reichle, Rayner, & Pollatsek, in press, for a detailed review).

Finally, eye movement measures have played a central role as a tool to develop and test psycholinguistic hypotheses about the processing of written language. In the recent past this has been the most prominent, and arguably the most fruitful, approach in the literature. There have been literally hundreds of studies exploiting the potential of temporal measures like fixation and gaze durations as online indicators of processing load, and spatial measures like fixation position and saccade amplitudes as indices of the direction and sequence of processing.

There are strong interrelations between these theoretical perspectives. For example, visuomotor research can make transparent the methodological strengths and weaknesses of eye movement measures in psycholinguistic research, but equally psycholinguistically motivated research has raised questions that demand examination in terms of their visuomotor dimensions (see, for example, the issue of long-range regressive saccades triggered by higher order linguistic processing; Kennedy, 2003). The papers assembled in the present Special Issue of the *European Journal of Cognitive Psychology* contribute, both in isolation and in different forms of interrelation, to all three of these perspectives.

SOME BASIC METHODOLOGICAL ISSUES

During reading, we move our eyes in a sequence of very fast, relatively well coordinated, movements known as saccades. These movements are interrupted by fixations, periods of relative stability in the position of the visual axis, during which visual information can be extracted. It is interesting to note that many basic facts about eye movements in general have been established in the context of reading research and that the present division into more visuomotor vs. more cognitive ways of looking at eye movements did not exist in an earlier research tradition (Dodge, 1903; Dodge & Cline, 1901; see also Wade, Tatler, & Heller, in press).

Research on the metrics of saccades in reading has shown that fixations are positioned in a very systematic, word-based, fashion. There are systematic relations between the duration and number of fixations on a given segment of text and its visual and linguistic processing. Perhaps one of the basic findings in this context is that the information acquired during a given fixation can influence the duration of that fixation as well as the amplitude of the outgoing saccade. This is referred to as *direct* control of eye movements (Rayner & Pollatsek, 1981).

The eye movement data reported in this Special Issue are the result of a relatively complex process of data reduction and analysis that obviously starts with the collection and storage of raw data. At this stage, issues like measurement accuracy and calibration play a significant role (McConkie, 1981). At the next stage, eye movement data need to be segmented and classified, including the determination of basic oculomotor events like saccades and fixations. Here, definitions of, and decisions on, minimal saccade amplitude and fixation duration and the handling of eye blinks play a critical role. At the next level of the process, data must be aggregated into units of analyses that can serve as the basis for statistical analyses. This stage of data analysis contains a number of choice points where the decisions taken can have significant consequences. In recent years, discussion of these methodological issues in their own right has started to gain momentum, but no generally agreed standards have yet emerged (see Inhoff & Radach, 1998; Inhoff & Weger, 2003; Murray, 2000; Rayner, 1998; Vonk & Cozijn, 2003). There is, however, a reasonable consensus (as is evidenced in this Special Issue) on the appropriate measures to describe saccades and fixations relative to critical words in text. Assuming that many readers are not familiar with the terminology and conventions of the field, Tables 1 and 2 present definitions of some of the more important spatial and temporal eye movement measures.

It is evident from Table 1 that word-based spatial eye movement measures reflect three aspects of oculomotor behaviour: the fact that a word is, or is not, fixated; the position fixated; and the amplitude of incoming and outgoing saccades. There are many possible combinations of these properties. For example, it has proved theoretically fruitful to distinguish between the amplitudes of saccades into and out of words and those within them (inter—vs. intraword movements) and to consider separately whether these saccades go from left to right (progressive saccades, when reading English and other left-to-right orthographies) or are directed against the normal reading direction (regressive saccades). Similar refinements are possible in the case of the temporal measures shown in Table 2, for example, distinguishing the first and second of two fixations from cases where exactly one fixation is made on a critical word. Each of these measures is, to some degree, pertinent to a particular theoretical question, making the measure of mean fixation duration (e.g., per condition or cell in an experimental design) somewhat problematic. For example, effects that manifest themselves in the number of fixations, may become obscured in the measure of their average duration (Blanchard, 1985; see Radach & Heller, 2000, for further discussion of relations between spatial and temporal aspects of eye movement control).

A useful notion for the description of fixation patterns is the concept of a "pass", denoting the first encounter with some defined segment of text, the equivalent gaze durations being referred to as first pass gaze durations. A complication emerges when a segment of text is read more than once. In this case, the respective gaze durations are referred to as second pass gaze durations, although it is possible that a particular word may receive its first fixation during such a pass, and in this case it is clearly necessary to explain whether the relevant data are allocated to the first or second pass.

Typically, the methodology used in eye movement studies of fluent reading involves the construction of experimental sentences or short passages of text that

TABLE 1

Definitions of the more commonly used word-based spatial eye movement measures

Parameter	*Definition*
Saccade amplitude, saccade length, saccade extent	Distance, in character positions, between the mean position of two successive fixations
Fixation probability, inverse measure: skipping rate	Relative frequency with which a word is fixated at least once; inverse measure: frequency of "word skipping"
Fixation position, fixation location	Position within a word (in characters) where a fixation is located, the empty space between words coded as zero
Launch distance, launch site	Distance in characters between the location of the prior fixation and the beginning (or centre) of the current word
Fixation frequency	Mean absolute number of fixations per word, for the current pass (defined as first, second, etc. encounter with specified text)
Refixation probability/frequency	Relative frequency of making at least one additional fixation before leaving a word

The relevant metric is in terms of letters rather than degrees of visual angle.

TABLE 2

Definitions of the more commonly used word-based temporal eye movement parameters

Parameter	*Definition*
Initial/first fixation duration	Duration of the first fixation within a word, irrespective of whether more fixations follow
Refixation duration	Summed duration of additional fixations within the current pass prior to an exit from the word
Gaze duration	Summed duration of all fixations before leaving the word (within the current pass)
Re-reading time	Summed duration of all fixations made after leaving the word for the first time
Total reading time, total fixation time	Summed duration of all fixations made on the critical word

Aggregated viewing time measures are usually computed without taking into account the durations of intraword saccades.

include critical target words. Relevant independent variables can then be systematically manipulated and the consequential effects on eye movement measures observed. An alternative, quasiexperimental approach is the recording of eye movements while corpora of natural text are being read. Kliegl, Olson, and Davidson (1983) refer to the technique of analysing this kind of data as orthogonal sampling. The corpus can be analysed with respect to the eye movements on particular words with specific properties related to some target variable, with other variables controlled. McConkie, Kerr, Reddix, and Zola (1988) used this technique in their seminal analysis of saccade landing site distributions, systematically sampling distributions for many combinations of word length and launch distance relative to the target word centre. This method has weaknesses, because it is always possible that particular outcomes have been mediated by variables not included in the sampling scheme, or arise as a result of hidden interactions. But it can be very useful for exploratory analyses and for the generation of hypotheses to be subsequently tested in more controlled experiments (as was the case with the McConkie et al. data). On the other hand, given that laboratory studies (perhaps particularly those involving single words examined in isolation) are somewhat remote from normal reading, a strong argument for the generality of a result obtained with highly selected stimulus material can be mounted if traces of the particular effect also occur in a normal text reading corpus.

A critical problem in the analysis of eye movement data is that there are at least three different ways to respond to processing difficulty: The eyes can remain on the critical word (or critical region) until the problem has been resolved; they can proceed with a progressive saccade (possibly related to an increase in viewing time measures); or they can execute a regressive saccade to reread segments of text. To handle complexities like this, several authors have proposed hybrid eye movement parameters that aggregate patterns of fixations over several words. One such measure has been termed "total pass" or "regression path" reading time (Kennedy, Murray, Jennings, & Reid, 1989; Konieczny, 1996; Liversedge, Paterson, & Pickering, 1998; Murray, 2000), defined as the sum of all fixations from the first fixation within the critical region until an exit from the region via its right (see also Rayner & Duffy, 1986, and Duffy, Morris, & Rayner, 1988, who used essentially the same procedure). A similar approach is the cumulative region reading time analysis proposed by Brysbaert and Mitchell (1996). Hybrid measures like this are assumed to reflect the time needed to notice and repair a higher level processing problem. Recently, there have been attempts to develop measures that go beyond the scope of sentence processing to describe phenomena such as shift of topic within a text or the detection of global semantic inconsistencies (Hyönä, Lorch, & Rinck, 2003).

The final and most important step in the analysis of eye movements is their interpretation within the context of a given theoretical framework. Until the mid1980s, interpretation of eye movement data was guided by the *eye–mind* and *immediacy* assumptions proposed by Just and Carpenter (1980). These suggested that processing coincides with, and is bounded by, the position fixated at any point in time, and that this processing starts at the point of fixation and continues until all possible analyses, up the semantic/thematic level, are completed. Numerous subsequent studies have shown that the spatial and temporal relationship between eye movements and information processing cannot be captured by these simple principles (see Rayner, 1998, for a comprehensive review).

First, while fixating a particular word, there is a substantial amount of preprocessing of the next word. Second, processing may spill over from one word to another: Encountering a difficult word can have consequences for the viewing time on a subsequent word (Rayner & Duffy, 1986). The eye and the mind are not tightly synchronised: The mind is sometimes a bit ahead of the eyes, but can also lag a little behind. Nonetheless, the relation between fixation positions and durations and local processing is strong enough to produce reliable effects when sampled over groups of participants and items and in this sense eye movement measures provide an extremely sensitive index of local processing load.

There is, of course, a limitation inherent in any essentially nonintrusive measure and this certainly applies to eye movement recordings. The experimenter is not in control of the specific pattern of data that arise in response to a given experimental manipulation. Further, and in contrast to many research paradigms used to study words in isolation, there is no straightforward way to determine how much visual information is available during a fixation, and precisely what type of information is being extracted. To some extent, this loss of experimental control can be compensated for using the technique of eyemovement-contingent display changes first introduced by McConkie and Rayner (1975) and Rayner (1975). This technique capitalises on the fact that little or no visual processing normally takes place during saccades. Movements of the eyes can be used to trigger changes in the display being processed. It is possible in this way to achieve control over the availability of (as yet uninspected) visual and linguistic information. For example, in the *boundary technique,* crossing an invisible boundary causes a change (during the ongoing saccade) in the text about to be fixated. Such changes (e.g., unmasking a previously masked item) usually go unnoticed by readers, but have profound consequences for processing. O’Regan (1990) has raised the question of whether perceptual consequences of display changes may influence results obtained with eye movement contingent display techniques. Such objections have been resolved to a large degree by direct demonstrations that the speed with which a display change is implemented has no effect on eye movements and fixations made after a saccade has crossed the critical boundary (Briihl & Inhoff, 1995; Inhoff, Starr, Liu, & Wang, 1998).

Display changes can also be implemented during a fixation. In this case, the changes are capable of being noticed by readers and masking techniques can be applied in combination with linguistic variations. As an example, McConkie, Underwood, Zola, and Wolverton (1985), briefly masked a whole line of text and replaced a word with another, visually similar, word at certain intervals after fixation onset. Readers were then asked to report which word they had seen. Similarly, the *fast priming paradigm* was developed by Sereno and Rayner (1992) to examine the time course of information uptake during a fixation. In this case, the movement of the eyes onto a target word causes the brief presentation of a prime stimulus that is replaced after a short time interval with a target word. Viewing time measures on the target are used to index the use of linguistic information conveyed by the prime. Inhoff, Connine, and Radach (2002) recently developed the *contingent speech technique,* where a visual display change is combined with the auditory presentation of a companion word immediately before, during, or after a critical word is fixated. The method provides data on the time course of phonological or semantic representations in working memory as reflected in gaze durations during the course of reading a number of words after the auditory stimulus was presented.

THEORIES AND MODELS OF EYE MOVEMENT CONTROL IN READING

In this section the discussion will focus on the development of detailed explanations of eye movement control. How are saccades generated during reading? How do linguistic processes and the visuomotor machinery interact to produce the observed oculomotor phenomena? These general questions represent an underlying theoretical preoccupation of the past two decades, namely, the identification of factors responsible for the *"where?"* and *"when?"* of eye movement control in reading.

With respect to the *"where?"* issue, there is now a consensus that the positioning of fixations is word-based, such that a key aspect is the decision as to which word is to be selected as the target for the next fixation. Fortunately, there are only a few alternatives to be considered in this decision, as the vast majority of saccades departing from a specific word either land on the same word again (refixations on word N), go to the immediately following word (N+l) or to the word beyond that (N+2). Interword regressions, bringing the eyes to positions *left* of the current word boundary, also predominantly land on words N-1 or N-2 (Radach & McConkie, 1998; Vitu & McConkie, 2000). Word targeting decisions are primarily based on low level information like word length and saccade launch distance, but cognitive factors such as word frequency and predictability also play a significant role (Brysbaert & Vitu, 1998; Drieghe, Brysbaert, Desmet, & De Baecke, 2004). Somewhat independent of the selection of which word to fixate is the specification of the precise saccadic amplitude needed to bring the eyes to the selected target. There are several lines of evidence suggesting that saccades (including refixation saccades) are aimed at the centre of the selected target word, which can be related to the fact that locations near the word centre appear to be optimal both for saccade targeting and foveal word processing. It is also largely agreed that, owing to visuomotor constraints, the eyes are systematically deviated from this target location, creating the well known difference between the "optimal" and the "preferred" saccade landing position, the latter being located about halfway between the beginning and the centre of the target word (McConkie et al., 1988; O'Regan, 1990; Rayner, 1979; see Inhoff & Radach, 2002, for a recent review).

The temporal aspect of eye movement control *("when?")* primarily concerns the question as to when a given saccade is triggered or, more precisely, the time course of processing events and control decisions occurring during a fixation. A closely related question is what information is used within this time frame to guide the eyes. As noted above, there is no doubt that control decisions can be based on information acquired online during the ongoing fixation. Effects of information acquired earlier in the text on viewing time measures can be significant when contextual constraint, either in terms of local word associations (Zola, 1984) or more global context, is very strong (Schustack, Ehrlich, & Rayner, 1987). A common way to determine the strength of contextual constraint is via measures of the predictability of the next word on the basis of all prior words in a sentence (e.g., Rayner & Well, 1996) or a longer passage of text (e.g., Vonk, Radach, & van Reijn, 2000).

Attempts to account for the detailed time course of oculomotor control during reading (e.g., Pynte, Kennedy, & Murray, 1991) have recently been buttressed by converging technologies. For example, Sereno, Rayner, and Posner (1998) found in an ERP study of single word processing that the evoked potential responses for low frequency and high frequency words start to diverge at about 130 ms. This represents an estimate of the lower bound of the time to achieve lexical access. At the other end of the timescale, results using the double step paradigm suggest that the amplitude of a saccade can be modified no later than 70–90 ms before the end of the current fixation (Deubel, O'Regan, & Radach, 2000). McConkie et al. (1985) propose a visual stimulus influence deadline of 80–100ms before the onset of the impending saccade. With these kinds of constraint in mind, Sereno and Rayner (2000b) conclude that the interval during which lexical processing can conceivably influence when the eyes are to move is severely limited. It follows that lexical processing must be sufficiently advanced within the first 100 to 150ms of a fixation in order to "intelligently" trigger the next eye movement (p. 79).

The issues sketched above have been at the core of a controversy that has dominated discussions in the field for more than a decade. One position in this debate has been that eye movements in reading are largely controlled by lexical processing, with lexical access determining both the position of fixations, their duration, and the probability that a word will be skipped. Typical of this position is the attention-based sequential processing model introduced by Morrison (1984) and reformulated by Henderson and Ferreira (1990) and Rayner and Pollatsek (1989). In the Morrison model, lexical access on a word causes a shift of visual attention to the next word, followed after a certain latency by a saccade. If the next word is easy to process, a second attention shift can take place (the eyes remaining on the first word). In this way, a process of cancellation and reprogramming of the ensuing saccade could lead to the next word being skipped. The Morrison model successfully accounted for some basic phenomena, in particular the existence of parafoveal processing, but suffered limitations such as the lack of a mechanism to handle refixations and an inability to account for the fact that parafoveal processing is modulated by current foveal difficulty. The existence of visuomotor constraints was acknowledged in this model, but they did not play a significant role in its instantiation.

The alternative position has been to claim that low-level visual processing and oculomotor factors are mainly responsible for the positioning of the eyes over a line of text and that it is these factors that have the stronger influence on viewing time measures. This view found its clearest expression in the Strategy— Tactics theory of O'Regan and colleagues (O'Regan, 1990, 1992; O'Regan & Lévy-Schoen, 1987), although work by McConkie et al. (1988) can also be seen in this tradition. As a more recent computational model by Reilly and O'Regan (1998) has shown, simple heuristics like "fixate the largest word within a window of 20 letter spaces" can indeed give a reasonable account of the where aspects of eye movement control. Similarly, the Mr.Chips model of Legge, Klitz, and Tjan (1997) proposed that the eyes are primarily guided by basic visual heuristics. However, models restricted to low-level oculomotor control plainly fail to handle the complexities of temporal control (Rayner, Sereno, & Raney, 1996).

With the benefit of hindsight this controversy (like many others) can be seen as little more than a question of emphasis. The two different approaches in reality define the extremes of a continuum and it is likely that any successful model will have to accommodate both "cognitive/attentional" and "visual/oculomotor" aspects. Indeed, O'Regan, Vitu, Radach, and Kerr (1994) acknowledge the limitations of an exclusively low-level view and suggest that a successful theory will need to combine elements of both traditions. And, equally, Rayner et al. (1996), while certainly giving precedence to the role of lexical processing, acknowledge the part played by visuomotor factors in arguing for some kind of "hybrid model". It may be more productive to view those eye movement patterns which are determined by purely visuomotor processing as providing a "carrier signal" which can be modulated by cognitive influences (see also Deubel et al., 2000; Yang & McConkie, 2001).

Reichle, Pollatsek, Fisher, and Rayner (1998) presented the first realistic algorithmic attention shift model of eye movement control in reading in the form of their E-Z Reader model. Although undeniably in the attention-shift tradition, its latest instantiation (Reichle et al., in press) also draws to some extent on lowlevel oculomotor control processes. The model has been used to fit the corpus of reading data obtained by Schilling, Rayner, and Chumbley (1998) and successfully accounts for a wide range of effects, including variation in viewing time as a function of word frequency, the effects of predictability, short-duration fixations, word-skipping, and spillover effects. Overall, the E-Z Reader model has set a standard against which alternative models will have to be evaluated.

Critical discussion of attention-based sequential control models has taken two main routes. The first points to evidence that more then one word may be processed in parallel rather than in a strictly sequential fashion (Hyönä & Bertram, 2004; Inhoff, Radach, Starr, & Greenberg, 2000; Kennedy, 1995, 1998, 2000; Kennedy, Pynte, & Ducrot, 2002; Pynte, Kennedy, & Ducrot, 2004; Starr

& Inhoff, 2004; Underwood, Binns, & Walker, 2000). While the parafoveal preview effect is entirely consistent with the operation of a sequential control model, properties of a word in the parafovea should not influence concurrent foveal processing and demonstrations that they do are clearly embarrassing. The research agenda has now shifted from a focus on *whether* such parallel processing occurs, to determining its precise nature and, in particular, whether such processing cross-talk is restricted to sublexical properties (Reichle et al., in press).

A second source of critical debate has centred on the question as to whether properties of the sequential model, such as lexical processing, attention shifts, and saccade programming can plausibly operate in a serial manner, given the temporal constraints outlined above. For example, for a word to be skipped the following sequence of events needs to occur before any decision to modify the ongoing saccade program can be made: (1) complete lexical processing of the current word, (2) a shift of attention to the to-be-skipped word, and (3) initial lexical processing (at least) of this newly attended word. This may take place under certain, rather unusual, circumstances (e.g., when the next word is short and very easy to process), but doubts can be raised as to whether such a process could plausibly be the default mechanism responsible for all cases of word skipping (Deubel et al., 2000).

In response to such critiques, a number of alternative models have been developed. The SWIFT model (Engbert, Longtin, & Kliegl, 2002; Kliegl & Engbert, 2003), has some features of the E-Z Reader architecture but departs in one important respect in proposing that lexical processing is "spatially distributed". The Glenmore model by Reilly and Radach (2003) has some similarities with SWIFT, but represents an even more radical departure from attention based sequential processing models. The model appeals to the theoretical framework provided by Findlay and Walker (1999), in which the spatial aspect of saccade generation is seen in terms of activation and competitive inhibition within a spatial salience map. It also incorporates a connectionist letter and word processing module that, together with low-level processing, codetermines spatial and temporal control decisions. A very detailed theory of saccade timing that also combines visuomotor and cognitive influences has been developed by Yang and McConkie (2001, 2004).

ISSUES FOR THE FUTURE

Thirty-five years ago it is unlikely that a discussion on the future of reading research would have even mentioned eye movements. Now, their measurement is accepted as an indispensible technology and an understanding of their control processes as a critical step in moving towards a complete theory of that complex cognitive activity. In the face of such a pace of development, it is patently idle to guess at the detail of the direction future research may take, but it is, nonetheless, possible to identify important topics that may have received insufficient attention to date.

It is clear that scientific work in our field will include the continuation of fruitful streams of empirical work, some of which will be discussed in detail below. There can also be no doubt that the development of computational models of information processing and eye movement control in reading will be a central part of the research agenda in the immediate future. New models will be proposed and it will become commonplace to discuss empirical data in close relation to these computational models, in particular as models are developed that are capable of dealing with higher level linguistic processing at a syntactic, semantic, and thematic level (see Rayner, 1998, Table 2, for an overview on these effects).

Methodological problems

It will have become clear from the discussions to this point that research on eye movements and information processing in reading has a solid theoretical and methodological grounding. Nonetheless, there is still room for improvement in the methodological apparatus used. For example, Inhoff and Radach (1998) reported the results of an informal survey to which 32 researchers using oculomotor measures responded. It was clear that a number of important issues have remained either unaddressed or unresolved. All the researchers believed that there is a need for increased discussion of measurement-related and methodological issues. More strikingly, two thirds of them considered the definition of functional oculomotor events (e.g., criteria for defining saccades and fixations) to be controversial. More than one third wished to appraise the specification and interpretation of extant measures. We are at a point where uncertainty over the specification and interpretation of oculomotor measures is limiting the appraisal, comparison, and exchange of empirical findings.

In the last couple of years there has been some progress towards broader discussions of these methodological problems (e.g., Inhoff & Weger, 2003; Murray, 2000; Rayner, 1998), but there is still a lack of empirical work on core issues. As an example, Irwin (1998) has shown that lexical processing continues during saccades, but virtually all studies still report gaze durations excluding saccade durations. Rayner (1998) discussed this problem and argued that, at least when words are the unit of measure, the inclusion of saccade durations has very little effect. It may, however, make more of a difference when larger units of analysis are used. Vonk and Cozijn (2003) were the first to address this question directly by comparing results of a sentence reading experiment with and without the inclusion of intraword saccade durations. The results were reassuring in so far as they indicated that a similar pattern of results emerged in both analyses, but there were, nonetheless, also subtle differences in gaze durations and changes in the significance of results. Another issue that calls for discussion and standardisation is the plethora of measures proposed for the analysis of complex interword eye movement patterns (see Murray, 2000, for a critical discussion). We believe that problems of this kind need more focused attention and that the development of agreed standards will benefit the development of the field.

Grounding of reading research in work on basic visuomotor functions

There is a rich literature on saccade generation in the traditions of oculomotor physiology and systems theory, which has produced explicit and virtually complete theoretical frameworks and models. This neurobiological evidence has recently been reviewed by Munoz (2002; see also Carpenter, 2000, for a brief summary) and van Gisbergen and van Opstal (1989) have provided a detailed discussion of formal biophysical saccade control models. Given this background of affairs it is odd that experimental reading research has remained largely disconnected from the neurobiology relevant either to visual processing or to oculomotor control. A recent attempt to remedy this situation can be found in Reichle et al. (in press), who seek to relate current evidence from neuroscience to the framework of their E-Z Reader model. However, there are still lively debates on the identification of brain areas responsible for even the most basic aspects of processing (e.g., the recognition of visual word forms) within the neuroscience community and it is difficult at present to derive predictions for experimental outcomes or modelling constraints from this literature. Something related in part to the fact that current functional brain image technologies cannot provide online information in a time frame that can be related to the known time course of the subprocesses of fluent reading (Sereno & Rayner, 2000b).

Somewhat similar problems arise in attempts to link relevant aspects of fundamental research in perception to eye movement control in reading. Perhaps the most glaring example is the concept of “attention”, which has played a central role in theories of eye movement control for over a century (Allport, 1993; Helmholtz, 1910). Attempts to define this concept can easily become bogged down in a maze of conflicting terms, circular arguments, and contradictory evidence. Both visual selection for the purpose of preparing eye movements and visual selection for the purpose of object (or word) recognition may be termed “attention”. Sometimes these are equated; sometimes they are strictly separated; and sometimes they are seen as related by a common mechanism, referred to as “attention” (Schneider & Deubel, 2002). Unsurprisingly, in view of this, opinions about the relation between eye movements and attention range from obligatory coupling to complete independence, depending on definitions, theoretical assumptions, and research paradigms (see Radach, Inhoff, & Heller, 2002, for a discussion). The popular view that attention *moves* in terms of an adjustable spotlight has been challenged and contrasted with activity-distribution models (LaBerge, Carlson, Williams, & Bunney, 1997). If a movement is assumed to take place, the question of how long it takes to prepare and execute such a shift of attention receives answers on a spectrum of opinion ranging from an “attentional dwell time” of around 50 ms (Duncan, Ward, & Shapiro, 1994; Treisman & Gelade, 1980) to claims that “visual attention moves no faster than the eyes” (Ward, 2002).

A more systematic approach to the problem of relating eye movements in reading to basic research on visuomotor functions could take the same form as the multiple task strategy suggested below with respect to research on word recognition. One successful example here is the principle of temporal overlap in the programming of successive saccades, derived from the double step paradigm (Becker & Jürgens, 1979). It formed an important element of Morrison’s (1984) model and provided the starting point for the later distinction between labile and nonlabile stages of saccade preparation in reading (Reichle et al., 1998). In addition to this multiple task approach, a second research strategy may be to address basic issues from within the task of normal reading. Promising lines of research that attempt to establish direct links from phenomena observed in a reading situation to basic perceptual or oculomotor mechanisms are the intriguing studies by Reingold and Stampe (2000, 2003) on the issue of saccadic inhibition evoked via gaze contingent display changes and the current work relating microsaccades in continuous reading to covert attention shifts (Engbert & Kliegl, 2003; Hafed & Clark, 2002).

It seems obvious that the starting point in any attempt to establish a productive relation between reading research and basic work on visuomotor func-tions must be a theory of saccade generation. Such a comprehensive theoretical framework, with clear relevance for a fuller understanding of reading, has recently been proposed by Findlay and Walker (1999). They suggest that the selection of saccade targets is accomplished via parallel processing and competitive inhibition within a two-dimensional “salience map” (in effect, fines-sing the problem of “attention”). The trigger for the execution of a saccade arises from a dynamic interrelation between a move centre, implementing the selection mechanism, and an independent fixate centre. This theoretical conception is rooted in a large body of work from different areas of basic research and although there are discussions on many details of modelling, the framework as such is widely accepted. More importantly, the theory explicitly allows for direct and indirect cognitive influences on saccade generation, opening a route for the development of submodels for the specific case of reading. The Competition/Interaction theory (Yang & McConkie, 2001), the SWIFT model (Engbert et al, 2002; Kliegl & Engbert, 2003), and the Glenmore model (Reilly & Radach, 2003) have all either included or implemented elements of the theoretical framework suggested by Findlay and Walker.

Integration with work on single word recognition

During the last two decades, eye movement research has contributed greatly to our understanding of how words are processed. But in many respects eye movement measurement has been seen primarily as a powerful addition to existing chronometric methods. This has, for example improved our understanding of the role played by phonology in visual word recognition (Pollatsek, Rayner, & Lee, 2000); morphological effects in complex word processing (Andrews, Miller, & Rayner, 2004; Inhoff & Radach, 2002; Pollatsek, Hyönä, & Bertram, 2000); or the role played by neighbourhood (Pollatsek, Perea, & Binder, 1999) and spelling to sound regularity (Sereno & Rayner, 2000a) effects. Nonetheless, the relationship between the well-established tradition of work on single word processing in cognitive psychology and equivalent work on word processing in fluent reading has remained wary and uncertain. The two subfields of reading research have maintained a state of (more or less) peaceful coexistence rather than productive cooperation. This is not to say there has been no interaction, but research questions, hypotheses, and theoretical ideas typically have been transferred from the area of word recognition into the field of research on continuous reading, with virtually no examples of transfer in the opposite direction.

Over the last decade there have been several attempts at the systematic comparison of eye movement data and results obtained using single word recognition paradigms, such as naming and lexical decision time (e.g., Folk & Morris, 1995; Grainger, O'Regan, Jacobs, & Segui, 1992; Inhoff, Briihl, & Schwartz, 1996; Juhasz, Starr, Inhoff, & Placke, 2003; Perea & Pollatsek, 1998). As an example, Schilling et al. (1998) examined word recognition performance in both naming and lexical decision and compared these to eye movement measures of viewing duration in normal reading. The same set of target words and the same group of participants was used. In all three methods a similar pattern of results emerged. Studies like this are an important first step towards collaboration between both research traditions, but it does not necessarily follow from a correspondence in the pattern of results obtained with different paradigms that the outcomes reflect the operation of the same underlying mechanisms. For example, facilitatory effects of orthographic neighbourhood density in both naming and lexical decision may, in fact, originate from very different processing mechanisms (Grainger & Jacobs, 1996). As a consequence, Grainger (2000) suggested: "the multi-task approach must be supplemented with a clear theoretical analysis of the mechanisms that determine performance in a given task, how these mechanisms relate to those hypothesized to be operational in normal reading outside of the laboratory, and how they relate to the mechanisms hypothesized to be operational in other tasks used to study word recognition and reading" (p.154). We believe that this approach of modelling "functional overlap" (Grainger, 2003; Jacobs & Grainger, 1994) will be important to achieve progress in the collaboration of both subfields of reading research.

Viewing this issue from the other side, current computational models of the reading process are relatively well specified with respect to eye movement control, but are clearly underspecified with respect to the core process of letter and word recognition. This becomes apparent when comparing the word processing modules of putative models of the reading process, such as E-Z Reader, SWIFT, or Glenmore, and the sophistication of current models of word recognition, such as the activation-verification model (Paap, Chun, & Vonnahme, 1999), the DRC (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), or MROM (Grainger & Jacobs, 1996; Jacobs, Rey, Ziegler, & Grainger, 1998). Of course, these models are, sometimes to an embarrassing degree, also models of performance in particular word recognition tasks. Given they are freed from the demands of syntactic and thematic processing and need not involve themselves with the oculomotor dynamics characteristic of reading it is unsurprising that it has, so far, proved extremely difficult to demonstrate their range of applicability to normal reading. It simply remains an article of faith that the subprocesses of (largely lexical) processing accounted for by these models are engaged in normal fluent reading. Perhaps, as an alternative to plugging more and more sophisticated word processors into eye movement control models, existing word processing models might accept as a necessary constraint modules for eye movement control. In this case, the move towards convergence of both areas would become a two-way road, including attempts to fine tune word processing models for the task of normal continuous reading.

Evaluation and comparison of reading models

There is no shortage of computation models of fluent reading. However, a gap is opening up between the computational sophistication of the models and an equivalent sophistication in techniques for their evaluation and comparison. Although the concept of "goodness-of-fit" itself may be debated, the primary criterion in assessing the success of any model is its fit to as many phenomena as possible. If this remains the primary criterion a situation will shortly arise where several, and possibly vastly different, models will justifiably claim to have achieved their goal. The more models differ in terms of underlying theoretical assumptions, architecture and principles of implementation, the more difficult will be their comparison. Jacobs (2000) discusses in some detail a set of general principles for model evaluation, basically arriving at the conclusion that existing computational models of reading cannot be compared because of their intrinsic heterogeneity. He makes two important exclusions. The first is the principle of "nested modelling" in which the E-Z Reader model provides an excellent example. Here, increasingly complex variants include older versions as special cases and thus represent a better approximation to reality by accommodating a larger empirical content. Unfortunately, this approach to model evaluation must be restricted to comparisons within the same family of models. A second, much more general, exclusion is what Jacobs refers to as the strong inference approach, where alternative models can be compared with respect to identical sets of data and evaluated on identical criteria.

A necessary precondition for this second approach is that the models are indeed comparable, or can be made comparable, and that accepted criteria for comparisons can be established (Jacobs & Grainger, 1994). A critical step in meeting this objective is the development of a common database to be used for the parameterisation and testing of alternative models. Some progress in this direction has already been made by Reichle et al. (1998), who made available the data set collected by Schilling et al. (1998), allowing Engbert and Kliegl (2001) and Engbert et al. (2002) to test their own models against this common baseline. The outcome achieved a high degree of comparability. Similar approaches involving common data sets can be found in the fields of machine learning, data mining, and, most importantly, language acquisition (MacWhinney, 1995). In the case of reading research, such a body of benchmark data should be extended to include a wider range of languages and scripts (Kliegl, Grabner, Rolfs, & Engbert, 2004). In the longer term it may also become possible to compare different computational models within a common implementation framework (Schmidt & Fayad, 1997).

Individual variations in reading and effects of task demands

There is a body of literature on both intra- and interindividual variation in the reading process, but in relation to the total amount of empirical work in the field the proportion is surprisingly small. In his review on the eye movement work of the last two decades, Rayner (1998) refers to the “classical” issues of reading skill, developmental changes in eye movement control, speed reading, and eye movements in dyslexia. However, little is known about the origins of such differences and about how individual variation in basic cognitive functions affect reading and, possibly, vice versa. The problem can be illustrated using the example of individual variation in working memory. Both Kennison and Clifton (1995) and Osaka and Osaka (2002) selected groups of participants on the basis of their performance in working memory span and found significant differences in several measures of oculomotor reading behaviour. Research like this is needed to better understand both the role of underlying cognitive component processes for reading performance and the role of reading in the composition of individual intellectual abilities.

Intraindividual factors, such as reading intention, motivation, or global strategy are widely assumed to affect comprehension, but little is known in detail of the way reading processes may be modulated, if at all, by such topdown factors (Heller, 1982; Tinker, 1958). This problem obviously has an important methodological dimension. At present, it is widely believed that it does not matter all that much whether a reader is asked to respond to a simple word verification task or answer difficult comprehension questions after fin-ishing a sentence or paragraph. Similarly, the differences in eye movement behaviour, if any, in reading sentences, compared to paragraphs or much longer segments of connected text (such as whole books) remain unquantified and it remains an open question whether text processing within new media environments has any specific impact on the reading process. Factors such as the defined reading task and the format of material are likely to affect reading speed and level of linguistic processing on a scale from "superficial" to "careful" reading. But it is also reasonable to hypothesise that they modulate the visuomotor and cognitive microprocesses of reading, e.g., in terms of local fixation patterns and word processing times. Radach, Huestegge, and Heller (2001) have data that point in this direction, showing, e.g., that word frequency effects are smaller and initial fixation positions shifted to the right when reading the same material in coherent text compared to single sentences presented in random order. It is particularly important in this context to know the exact nature of any global top-down influence on the inner workings of eye movement control (see O'Regan's, 1992, hypothetical distinction of "careful" vs. "risky" reading). If such factors turn out to be of importance, they will also need to be incorporated as modulating elements in models of the reading process. Knowing the scale and impact of these varieties of the reading process is at the core of defining the content and purpose of experimental reading research as a whole.

There can be no doubt that work on eye movement control has reached a level of sophistication where applied problems, particularly with respect to individual differences and task demands, can fruitfully be addressed. But the research community has, as yet, taken little advantage of these possibilities. As a single example, it is surprising that the issue of "speed reading" has hardly attracted a single scientific study during the last 15 years that would meet the methodological standards raised above (see Rayner & Pollatsek, 1989, for a discussion of earlier work). Equally, the community of experimental reading researchers is patently underrepresented in the vast literature on developmental dyslexia and other forms of impaired reading. There have been virtually no studies attempting to analyse word-based viewing time measures and local fixation patterns in dyslexic readers (see Hyönä & Olson, 1995, for a significant exception). We believe that the progress documented in this introductory paper, together with the range of contributions in this Special Issue, suggest the time is ripe to bring a new level of theoretical and methodological sophistication to bear on this and a number of other applied questions.

PrEview proof published online September 2003

REFERENCES

Allport, D.A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D.E.Meyer & S.M.Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 183–218). Cambridge, MA: MIT Press.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*(1/2), 285–311.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19,* 967–984.

Blanchard, H.E. (1985). A comparison of some processing measures based on eye movements. *Acta Psychologica, 58,* 1–15.

Briihl, D., & Inhoff, A.W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and of exterior letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 55–67.

Brysbaert, M., & Mitchell, D.C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *Quarterly Journal of Experimental Psychology, 49A,* 664–695.

Brysbaert, M, & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–148). Amsterdam: Elsevier.

Carpenter, R.H. (2000). The neural control of looking. *Current Biology, 10,* 291–293.

Coltheart, M.,Rastle, K.,Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Deubel, H., O'Regan, K., & Radach, R. (2000). Attention, information processing and eye movement control. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 355–374). Oxford, UK: Elsevier.

Dodge, R. (1903). Five types of eye-movement in the horizontal meridian plane of the field of regard. *American Journal of Physiology, VIII,* 307–329.

Dodge, R., & Cline, T.S. (1901). The angle velocity of eye-movements. *Psychological Review, 8,* 145–157.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology, 16*(1/2), 79–103.

Duffy, S.A., Morris, R.K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Duncan, J., Ward, W., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369,* 313–315.

Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics, 85,* 77–87.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43,* 1035–1045.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636.

Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22,* 661–721.

Folk, J.R., & Morris, R. (1995). Multiple lexical codes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1412–1429.

Grainger, J. (2000). From print to meaning via words? In A.Kennedy, R.Radach, D.Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 147–161). Oxford, UK: Elsevier.

Grainger, J. (2003). Moving eyes and reading words: How can a computational model combine the two? In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 457–470). Amsterdam: Elsevier Science.

Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518–565.

Grainger, J., O'Regan, K., Jacobs, A., & Segui, J. (1992). Neighbourhood frequency effects and letter visibility in visual word recognition. *Perception and Psychophysics, 51,* 49–56.

Hafed, Z. M, & Clark, J.J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Resesarch, 42,* 2533–2545.

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

Helmholtz, H. (1910/2000). *Handbuch der physiolgischen Optik* (Vol. 3, 3rd ed.). A.Gullstrand, J. von Kries, & W.Nagel (Eds.), *Helmholtz's Treatise on Physiological Optics* (Vol. 3. Trans. J.P. C.Southall, New York: Optical Society of America, 1925). (Reprinted in 2000, Bristol: Thoemmes Press. Original work published 1910).

Henderson, J.M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 411–429.

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology, 16*(1/2), 104–127.

Hyönä, J., Lorch, R.F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam: Elsevier Science.

Hyönä, J., Munoz, D., Heide, W., & Radach, R. (Eds.). (2002). The brain's eye: Neurobiological and clinical aspects of oculomotor research. In *Progress in brain research, Vol. 140.* Oxford, UK: Elsevier Science.

Hyönä, J., & Olson, R.K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1430–1440.

Hyönä, J., Radach, R., & Deubel, H. (Eds.). (2003). *The mind's eye: Cognitive and applied aspects of eye movement research.* Amsterdam: Elsevier Science.

Inhoff, A., & Radach, R. (2002). The biology of reading: Use of spatial segmentation in the reading of complex words. *Comments on Modern Biology: Part C. Comments on Theoretical Biology, 7,* 121–138.

Inhoff, A.W., Briihl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming and delayed naming tasks. *Memory and Cognition, 24,* 466–476.

Inhoff, A.W., Connine, C., & Radach, R. (2002). A contingent speech technique in eye movement research on reading. *Behavior Research Methods Instruments and Computers, 34,* 471–480.

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

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 221–246). Oxford, UK: Elsevier.

Inhoff, A., Starr, M., Liu, W., & Wang, J. (1998). Eye-movement-contingent display changes are not compromised by flicker and phosphor persistance. *Psychonomic Bulletin and Review, 5*(1), 101–106.

Inhoff, A.W., & Weger, U. (2003). Enhancing the methodological middle ground. In J.Hyönä, R. Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 335–343). Amsterdam: Elsevier Science.

Irwin, D. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology, 36,* 1–27.

Jacobs, A.M. (2000). Five questions about cognitive models and some answers from three models of reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 721–732). Oxford, UK: Elsevier.

Jacobs, A.M., & Grainger, J. (1994). Models of visual word recognition: Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6)*,* 1311–1334.

Jacobs, A.M., Rey, A., Ziegler, J.C., & Grainger, J. (1998). MROM-P: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In J.Grainger & A.M.Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 147–188). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

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

Kennedy, A. (1995). *The influence of parafoveal words on foveal inspection time*. Paper presented at the AMLaP-95 Conference, Edinburgh, UK.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–223). Oxford, UK: Elsevier.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kennedy, A., Brooks, R., Flynn, L.-A., & Prophet, C. (2003). The reader's spatial code. In J.Hyönä, R. Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 193–212). Amsterdam: Elsevier.

Kennedy, A., Murray, W.S., Jennings, F., & Reid, C. (1989). Comments on the generality of the principle of minimal attachment. *Language and Cognitive Processes, 4,* 51–76.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A*(4)*,* 1307–1337.

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

Kennison, S.M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*(1), 68–81.

Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In J.Hyönä, R.Radach, & H.Deubel, (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 391–411). Amsterdam: Elsevier Science.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1/2), 262–284.

Kliegl, R., Olson, R.K., & Davidson, B.J. (1983). On problems of unconfounding perceptual and language processes. In K.Rayner (Ed.), *Eye movements in reading and perceptual and language processes* (pp. 333–343). New York: Academic Press.

Konieczny, L. (1996). *Human sentence processing: A semantics-oriented parsing approach.* (IIGReports 3/96). Universität Freiburg, Institut für Informatik und Gesellschaft.

LaBerge, D., Carlson, R.L., Williams, J.K., & Bunney, B.G. (1997). Shifting attention in visual space: Tests of moving-spotlight models versus an activity-distribution model. *Journal of Experimental Psychology: Human Perception and Performance, 23*(5), 1380–1392.

Legge, G.E., Klitz, T.S., & Tjan, B.S. (1997). Mr.Chips: An ideal observer model of reading. *Psychological Review, 104,* 524–553.

Liversedge, S.P., Paterson, K.B., & Pickering, M.J. (1998). Eye movements and measures of reading time. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55–76). Oxford, UK: Elsevier.

MacWhinney, B. (1995). *The CHILDES Project: Tools for analyzing talk* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McConkie, G.W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods and Instrumentation, 13,* 97–106.

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

McConkie, G.W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.

McConkie, G.W., Underwood, N.R., Zola, D., & Wolverton, G.S. (1985). Some temporal characteristics of processing during reading. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 168–186.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Munoz, D.P. (2002). Saccadic eye movements: Overview of neural circuitry. In J.Hyönä, D.Munoz, W.Heide, & R.Radach (Eds.), *The brain 's eyes: Neurobiological and clinical aspects of oculomotor research* (pp. 89–96). Oxford: Elsevier.

Murray, W.S. (2000). Sentence processing: Issues and measures. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 649–664). Oxford, UK: Elsevier.

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

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

O'Regan, J.K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 363–383). Hove, UK: Lawrence Erlbaum Associates Ltd.

O'Regan, J.K., Vitu, F., Radach, R., & Kerr, P.W. (1994). Effects of local processing and oculomotor factors in eye movement guidance in reading. In J.Ygge & G.Lennerstrand (Eds.), *Eye movements in reading* (pp. 329–348). Oxford, UK: Pergamon Press.

Osaka, N., & Osaka, M. (2002). Individual differences in working memory during reading with and without parafoveal information: A moving-window study. *American Journal of Psychology, 115,* 501–513.

Paap, K.R., Chun, E., & Vonnahme, P. (1999). Discrete threshold versus continuous strength models of perceptual recognition. *Canadian Journal of Experimental Psychology, 53,* 277–293.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 767–779.

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

Pollatsek, A., Perea, M., & Binder, K.S. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1142–1158.

Pollatsek, A., & Rayner, K. (1989). Reading. In M.I.Posner (Ed.), *Foundations of cognitive science.* Cambridge, MA: MIT Press.

Pollatsek, A., Rayner, R., & Lee, H.W. (2000). Phonological coding in word perception and reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 399–426). Oxford, UK: Elsevier.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1/2)*,* 178–202.

Pynte, J., Kennedy, A., & Murray, W.S. (1991). Within-word inspection strategies in continuous reading: The time-course of perceptual, lexical and contextual processes. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 458–470.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 165–192). Oxford, UK: Elsevier.

Radach, R., Huestegge, L., & Heller, D. (2001, Sept.). *The role of global top down factors for eye movement control in reading.* Paper presented at the conference of the European Society for Cognitive Psychology, Edinburgh, UK.

Radach, R., Inhoff, A.W., & Heller, D. (2002). The role of attention and spatial selection in fluent reading. In E.Witruk, A.Friederici, & T.Lachmann (Eds.), *Basic mechanisms of language and language disorders* (pp. 137–154). Dordrecht, The Netherlands: Kluwer.

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

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psycholog, 7,* 65–81.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–22.

Rayner, K., & Duffy, S.D. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity and lexical ambiguity. *Memory and Cognition, 14,* 191–201.

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology, 33A,* 351–373.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

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

Rayner, K., & Well, A.D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review, 3,* 504–509.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

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

Reilly, R., & Radach, R. (2003). Glenmore: An interactive activation model of eye movement control in reading. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429–455). Amsterdam: Elsevier Science.

Reilly, R.G., & O'Regan, J.K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research, 38*(2), 303–317.

Reingold, E.M., & Stampe, D.M. (2000). Saccadic inhibition and gaze contingent research paradigms. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 119–146). Oxford, UK: Elsevier.

Reingold, E. M, & Stampe, D.M. (2003). Modulation of saccadic inhibition by cognitive or attentional factors: Implications for models of reading. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 347–360). Amsterdam: Elsevier.

Schilling, H.E., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26,* 1270–1281.

Schmidt, D.C., & Fayad, M. (1997). Object-oriented application frameworks. *Communications of the Association for Computing Machinery, 40,* 32–38.

Schneider, W.X., & Deubel, H. (2002). Selection-for-perception and selection-for-spatial-motoraction are coupled by visual attention: A review of recent findings and new evidence for stimulus driven saccade control. In W.Prinz & B.Hommel (Eds.), *Attention and performance XIX: Common mechanisms in perception and action.* Oxford, UK: Oxford University Press.

Schustack, M.W., Ehrlich, S.F., & Rayner, K. (1987). Local and global sources of contextual facilitation in reading. *Journal of Memory and Language, 26,* 322–340.

Sereno, S.C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 173–184.

Sereno, S.C., & Rayner, K. (2000a). Spelling-sound regularity effects on eye fixations in reading. *Perception and Psychophysics, 62*(2), 402–409.

Sereno, S.C., & Rayner, K. (2000b). The when and where of reading in the brain. *Brain and Cognition, 42,* 78–81.

Sereno, S.C., Rayner, K., & Posner, M.I. (1998). Establishing a timeline of processing during reading: Evidence from eye movements and event-related potentials. *Neuro Report, 9,* 2195–2200.

Starr, M., & Inhoff, A. (2004). Attention allocation to the right and left of a fixated word: Use of orthographic information from multiple words during reading. *European Journal of Cognitive Psychology, 16*(1/2), 203–225.

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

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp.247–268). Oxford, UK: Elsevier.

Van Gisbergen, J.A., & Van Opstal, A.J. (1989). The neurobiology of saccadic eye movements: Models. *Reviews of Oculomotor Research, 3,* 69–101.

Vitu, F., & McConkie, G.W. (2000). Regressive saccades and word perception in adult reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 301–326). Oxford, UK: Elsevier.

Vonk, W., & Cozijn, R. (2003). On the treatment of saccades in and regressions in eye movement measures of reading time. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 291–311). Amsterdam: Elsevier Science.

Vonk, W., Radach, R., & van Reijn, H. (2000). Eye guidance and the saliency of word beginnings in reading text. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 269–300). Oxford, UK: Elsevier.

Wade, N., Tatler, B.W., & Heller, D. (in press) Dodge-ing the issue: Dodge, Javal, and the discovery of "saccades" in eye movement research. *Perception.*

Ward, R. (2002). Attention moves no faster than the eyes. In W.Prinz & B.Hommel (Eds.), *Attention and performance XIX: Common mechanisms in perception and action*. Oxford, UK: Oxford University Press.

Yang, S.-N., & McConkie, G. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research, 41*(25–26), 3567–3585.

Yang, S.N., & McConkie, G.W. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology, 16*(1/2), 226–261.

Zola, D. (1984). Redundancy and word perception during reading. *Perception and Psychophysics, 36,* 277–284.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 27–51

Orthographic regularity gradually modulates saccade amplitudes in reading

Ralph Radach

University of Aachen, Germany

Albrecht Inhoff

SUNY Binghamton, New York, USA

Dieter Heller

University of Aachen, Germany

The present research tested the hypothesis that variations in the orthographic regularity of word beginnings influence landing positions and amplitudes of interword saccades in continuous reading. Participants were asked to read sentences including target words of low, medium, and high frequency of initial quadrigrams that were either single-root nouns or noun-noun compounds. Saccades landed further into words with more regular beginnings, irrespective of whether the target was a compound word or not. Critically, the orthographic landing site effect was graded, suggesting that orthographic information continuously modulates saccade amplitude before and after the decision to move has been made.

Fluent reading can be seen as a dual task where the main process is the extraction of meaning from print and the secondary task is the planning and execution of adequate oculomotor behaviour. Consequently, information beyond a fixated word is used for two distinct purposes: First, it can be used to initiate letter and word processing before a word is fixated and, second, it can be used to guide the eye movement programming system. In classic research both of these aspects have been related to the concept of the perceptual span. It is the *total perceptual span* within which word length information critical for saccade programming can be acquired, while letter and word information can only be extracted from a smaller region around the current fixation, which includes the fixated word and generally the next (parafoveal) word in the text. This range is occasionally referred to as a *letter identification span* (see Rayner, 1998, for a recent review).

Parafoveal letter and word processing is often quantified in terms of a parafoveal preview benefit, determined as the difference in fixation or gaze duration on a target word between conditions with versus without useful parafoveal target previews. Key results have shown that readers acquire useful orthographic and phonological information from a parafoveal word. Yet, as shown by McConkie and Zola (1987), there is a rather sharp drop-off in letter identification performance from the centre of fixation toward the parafovea. Related to this is the fact that more useful parafoveal letter information is extracted from the beginning letters of a to-be-fixated word in comparison to centre or ending letters (Briihl & Inhoff, 1995; Inhoff, 1989; Rayner, Well, Pollatsek, & Bertera, 1982).

The second aspect of parafoveal information processing in reading is related to its use in the control of eye movements. There is broad consensus that the visual configuration around the current fixation position provides the primary base for the decision which word should be the target for the next saccade (O'Regan, 1990; Rayner, Sereno, & Raney, 1996). For example, McConkie, Kerr, Reddix, and Zola (1988) proposed that text is parsed by routines of low level vision into an array of low spatial frequency word-objects that the eye movement control system can select as targets for saccades. The choices that can be made during this targeting process are quite limited. The most likely alternatives are the execution of a refixation on the same word, a saccade to the adjacent parafoveal word or a saccade that "skips" this word and goes further to the right. Empirically, this selectivity is reflected in the multimodality of landing site distributions including cases of refixations, interword saccades, and "word skipping" (Inhoff & Radach, 2002; Radach & McConkie, 1998). There is a substantial body of literature indicating that saccade targeting decisions are codetermined by low-level visual and linguistic processing. One example is the discussion on the relative contributions of both components to the probability of fixating a word (see Brysbaert & Vitu, 1998, for a review).

In addition to selecting the next to-be-fixated word, the spatial parameters for a saccade aiming at this word need to be specified. It is widely assumed that the intended target position of these saccades is the centre of the selected word. Two major lines of evidence support this view. One includes results from research on

---

Correspondence should be addressed to R.Radach, Institute of Psychology, University of Aachen, Jaegerstr. 17, 52056 Aachen, Germany. Email: ralph@psych.rwth-aachen.de The authors are grateful to Holger Olawsky for his support in the preparation of the materials and for doing most of the experimental work for this study. Christian Vorstius provided valuable assistance in data analysis and in the preparation of this manuscript. We also wish to thank Jukka Hyönä, Keith Rayner, and Sarah White for very helpful comments on an earlier version of this paper.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000222

the processing of briefly presented single words (Brysbaert, Vitu, & Schroyens, 1996; O'Regan, Lévy-Schoen, Pynte, & Brugaillére, 1984; Nazir, Heller, & Sußmann, 1992; O'Regan & Jacobs, 1992; Stevens & Grainger, 2003), indicating the existence of an optimal viewing position at or slightly left of the word centre. The counterpart of these effects for the case of continuous reading can be seen in the fact that the probability of refixating the current word follows a u-shaped function with a minimum at locations close to its centre. This phenomenon was first demonstrated by McConkie, Kerr, Reddix, Zola, and Jacobs (1989) and has been replicated in numerous reading studies (see Vitu, McConkie, Kerr, & O'Regan, 2001, for a recent discussion of issues related to the optimal viewing position in text reading).

However, if the centre of the to-be-fixated word is the optimal target for reading saccades, why are saccade landing positions not distributed accordingly? For example, when an initial interword progressive saccade is made, the peak of the landing site distribution, referred to as the "preferred viewing position" (Rayner, 1979), lies about halfway between word beginning and word centre. McConkie et al. (1988) have proposed that this apparent deviation between optimal and actual landing positions is the result of a small set of basic visuomotor principles. The most important of these principles, termed the saccadic range effect, consists of a tendency for saccades directed at near targets to overshoot and for saccades directed at far targets to undershoot the intended position, with the overall mean being diverted substantially to the left of the centre. For a range of launch sites relative to a target word, there is a linear relation between launch distance and mean landing position. This "landing site function" (Radach & McConkie, 1998) is now generally accepted as a fundamental property of eye movement control and has been implemented in several recent computational models (Rayner, Reichle, & Pollatsek, 2000; Reichle, Rayner, & Pollatsek, 1999; Reilly & O'Regan, 1998; Reilly & Radach, 2003). A second important low-level determinant for landing positions of initial interword saccades is word length. For each 1 letter increment in word length the mean landing position moves about 0.2 letters to the right, whereas for each letter variation in launch site the shift is in the order of about 0.5 letters. The word length effect can be attributed to a centre of gravity phenomenon (O'Regan, 1990; Vitu, 1991). Interestingly, parafoveal word length information appears not to be used to constrain orthographic or lexical word processing (Inhoff, Radach, Eiter, & Juhasz, 2003).

If there is consensus about the fact that visuomotor factors are the main force determining saccade amplitudes in reading, the question arises as to whether and how the programming of a saccade to the next word in the text is influenced by linguistic processing. Historically, the notion of linguistic control can be traced back to Hochberg (1970, 1976) who proposed that a process labelled "peripheral search guidance", together with "cognitive search guidance", controlled saccade amplitude programming. In the current literature, possible modulations of saccade amplitudes or landing positions as result of linguistic processing are often referred to as "cognitive landing site effects" (Underwood & Radach, 1998).

The hypothesis that linguistic properties of a parafoveal word may modulate the amplitude of interword saccades was first directly investigated in a series of studies by Everatt and Underwood (1992), Hyönä, Niemi, and Underwood (1989) and Underwood, Clews, and Everatt (1990). For these experiments, stimuli were developed either by dictionary counts or by asking participants to guess words on the basis of their first or last five letters. Stimulus words that were beginning-informative (e.g., "seamanship") versus end-informative (e.g., "microwave") were then presented in reading experiments. The main result obtained in these studies was that the eyes landed further into words when the informative part was at the end of the word as opposed to its beginning (see our discussion below for dissenting evidence).

At the time the effect of informativeness was interpreted as a consequence of lexical-semantic processing, claiming that fixations were "attracted" further into words with an informative ending. This position has been criticised on several levels. For example, Hyönä (1993) has questioned the methodology of the experiments, arguing that a number of linguistic variables were not well controlled. For example, in one study by Hyönä et al. (1989) stimulus words with informative beginnings had a productive derivational ending, whereas all words with informative endings were compounds consisting of two noun lexemes (generally referred to as NN-compounds). The proposal that lexical semantic information may be acquired from the second half of a parafoveal word and used for saccade amplitude programming is also in conflict with a large body of data on parafoveal semantic preprocessing in both single word recognition and reading (Balota & Rayner, 1991; Inhoff, 1982; Rayner, Balota, & Pollatsek, 1986; Rayner, White, Kambe, Miller, & Liversedge, 2003).

Hyönä (1993) has proposed "orthographic saliency",[1] denoting the orthographic regularity[2] of a word beginning as the factor most likely to be able to "pull" the eyes towards a critical region within a target word. He tested this hypothesis in an experiment using three types of stimuli: words with a derivational ending, compound words containing an identical first half but an orthographically less frequent ending, and words with an orthographically infrequent beginning (Hyönä, 1995). The main result was that saccades into words with more irregular beginnings landed about 0.3 letter spaces closer to the left word boundary. Much of this effect was due to fixations made on the space before the target word with a less frequent beginning. The interpretation suggested by Hyönä (1995) is that an extremely rare word beginning leads to a difficulty in parafoveal processing, triggering the occurrence of atypical saccades. The specific version of this hypothesis that he favoured was that instead of attempting to go to the centre of the next word, the system selects the space before the word as a saccade target that is both easy to attain und useful for further processing of the difficult letter cluster. A shift of saccade landing positions towards less regular word beginnings was also reported by Beauvillain, Doré, and Baudouin (1996; see also Beauvillain & Doré, 1998). They used an orthographic decision task involving the recognition of single words displayed to the right of a central fixation mark. In this task saccade landing positions were again further to the left for target words with orthographically irregular beginning characters. A similar effect was also obtained in a recent word recognition study by Kennedy, Pynte, and Ducrot (2002). Another way of looking at effects of orthography on initial landing positions is to introduce misspellings at word beginnings. In this case it is not regularity but the level of irregularity in terms of the difference to correct spelling that is varied. While White and Liversedge (2004) report a substantial leftward shift of landing positions effects for misspelled and thus orthographically illegal word beginnings, Pynte, Kennedy, and Ducrot (2004) did not obtain such an effect.

Reviewing the literature on the issue, Vonk, Radach, and van Rijn (2000) noted that cognitive landing site effects had so far only been obtained in studies where single sentences were presented, whereas attempts to find such effects in studies involving larger amounts of text had failed (Kerr, 1992; Radach & Kempe, 1993). This led them to consider the hypothesis that reading strategies used in particular experimental situations may play a role in the use of linguistic information for saccade programming. To test this idea, they embedded target words with a low versus high frequency initial trigram into story frames with either a neutral or a high degree of contextual constraint. The results were straightforward: There was a clear effect of orthographic regularity, amounting to

---

[1] It is useful to distinguish between the notions of orthographic saliency and informativeness of a word beginning. Saliency is usually computed as the frequency of the occurrence of a word initial *N*gram among all entries (tokens) in a word form count. In contrast, the number of different word forms (types) in which a word initial N-gram occurs can be seen as a measure of informativeness with respect to the lexical identity of the word. As could be expected, token and type n-gram frequencies are correlated in natural text (r=.62 for the initial quadrigrams of the target words used in this study). However, at least for languages like German containing a multitude of derivational endings it may be more appropriate to determine “informativeness” in terms of the number of word stems or lemmata a word beginning corresponds to. Despite these problems of defining an appropriate statistical measure, it appears reasonable from a theoretical point of view to assume that informativeness plays a relatively late role during the time course of orthographic word processing. Thinking in the framework of a multilevel interactive processing model, activation within and feedback from the level of lexical representations will be needed for informativeness to have an effect on further word processing and/or eye movement control.

[2]The terms saliency and regularity both carry a certain bias of interpretation. “Saliency” may suggest that a word with a rare *N*-gram, by definition, should be more noticeable. “Regularity” usually refers to the fact that a word conforms to the rules of orthography, in contrast to an irregular string, hence, the expression “high regularity” is problematic. Since no clear preference for either concept seems to have emerged so far, both terms are used in this paper, with “low regularity” being equivalent to “high saliency” and vice versa.

a shifting of 0.3 letters in the initial saccade landing position to the word beginning for words with less regular trigrams. In contrast, there was no effect of contextual constraint, a finding that has been confirmed by Rayner, Binder, Ashby, and Pollatsek (2001).[3]

Another type of linguistic landing site effect appears to have been found by Inhoff, Briihl, and Schwartz (1996) in a study designed to investigate the processing of word morphology in reading and naming tasks. They asked participants to read sentences containing three classes of target words: monomorphemic words (cathedral, arthritis), suffixed words (sainthood, heartless) and compound words with two 3- to 5-letter first constituents (timetable, gunpowder). An unexpected result emerged with respect to the initial interword saccades toward these target words. They went significantly further into compound words (mean landing position 4.4) as compared to suffixed (mean landing position 3.9) and monomorphemic words (mean landing position 3.7). We have recomputed the initial token and the type trigram and the first two bigram frequencies of the target words used by Inhoff et al. with the result that the observed effect is unlikely to be due to differences in orthographic saliency.

An alternative hypothesis that may account for this type of effect was first discussed by Rayner and Morris (1992), who considered the possibility that the initial part of a parafoveal word may be successfully processed and then skipped in some cases, creating the observed small shift in mean landing position. Further results suggesting the possibility of landing site effects on the level of morphemes or lexical constituents of complex words have been provided by Hyönä and Pollatsek (1998). In a series of experiments on the reading of Finnish noun-noun compounds, they found in one experiment that the initial saccade into 12–14 letter target words landed slightly closer to the word beginning when the lexical frequency of the initial constituent was low. To account for this result, they proposed a processing difficulty hypothesis, according to which a parafoveal low frequency word (or lexeme) will narrow down the span of effective preprocessing which in turn will lead to a shorter forward saccade (Hyönä & Pollatsek, 2000).

Looking at the literature, it is obvious that the evidence for cognitive landing site effects on saccades directed to the next word in the text is far from unequivocal. Even in the initial series of studies by Underwood et al. there was one experiment which showed no effect (Underwood, Bloomfield, & Clews, 1988). Using the same set of stimuli, Rayner and Morris (1992) failed to replicate the results of Underwood et al. (1990). In a more recent study,

Liversedge and Underwood (1998) attempted to test the hypothesis that a cognitive landing site effect may be increased when syntactic processing load on the foveal word is low and more resources can be allocated to parafoveal processing. However, in one experiment there was no main effect of word beginning saliency (initial trigram frequency) on the landing position of the initial saccade into target words. In a second experiment a small effect emerged only in a subanalysis including items with the shortest versus longest gaze durations in the pretarget region. With respect to the occurrence of higher order linguistic landing site effects, the situation is also contradictory. In contrast to the results of Hyönä and Pollatsek (1998) mentioned above, Andrews, Miller, and Rayner (2004) found no differences in landing positions for NN (nounnoun) compounds with varying lexical frequency.

Given this situation, it is obvious that more empirical evidence on the issue is needed. In addition to contributing to the empirical base for the discussion about cognitive landing site effects, the current study has two specific purposes. The first of these objectives originates from an observation made by Vonk et al. (2000). They varied the regularity of word beginnings in a far less extreme way than Hyönä (1995), who had used a number of loan words whose beginnings are virtually nonexistent in the rest of the Finnish vocabulary. Yet, in the Vonk et al. study a reliable cognitive landing site effect emerged that manifested itself in a shift of the whole distribution of landing sites rather than an increase of fixations at the very beginning of the less regular words. This appears to be in contradiction to Hyönä's proposal that cognitive landing site effects are due to a *discrete* deviation from normal oculomotor behaviour in cases where very unusual letter clusters are encountered. However, it does not address the related, more general issue of whether these effects are an exclusion from or, alternatively, part and parcel of normal reading behaviour. Typical for the prevailing view on this issue is the statement by Reichle, Rayner, and Pollatsek (in press) that “there appears to be general agreement that an orthographically irregular letter cluster at the beginning of a word results in the eyes' initial landing position *deviating* toward the beginning of the word” [italics added].

The alternative to this position is that orthographic landing site effects are of a graded nature: If there is a medium level of orthographic regularity for a certain class of words, it is possible that differences in either direction would lead to corresponding alterations in mean incoming saccade amplitude. It is quite obvious that such a result would have substantial theoretical implications. To allow for the investigation of this possibility, the present research used a variation of orthographic saliency on three rather than two levels.

The second purpose of this experiment was to reinvestigate the possibility that there are landing site effects at the lexico-morphological level. To this end, two classes of target words were used. One type of targets were morphologically regular single root nouns. The second class of target words consisted of

[3] There is one paper by Lavigne, Vitu, and d'Ydewalle (2000) suggesting that information on the semantic level may influence interword saccade landing positions. They found that when high frequency target words are in a prime—target relation with preceding words, saccades launched from near distances go further into these words. However, Rayner et al. (2001) did not get such an effect. Even if the Lavigne et al. results could be replicated, they would represent a rather infrequent exclusion in which several circumstances must act together to create an effect that would only very rarely occur in the reading of normal text.

nounnoun (NN) compounds that included four-letter initial lexemes that were matched on a number of linguistic properties to the first four letters of the single root nouns. In the case of words with medium orthographic regularity, it was even possible to create pairs of regular and compound words with identical initial quadrigrams (e.g., the medium-regular single root noun “Autopsie” and the medium-regular compound noun “Autokino"). Since the orthography of the first half of these two types of words was identical, differences in saccade landing positions between these two types of words could thus be attributed to their morphological structure.

METHOD

Participants

Twenty-four participants, all students at the University of Aachen, volunteered to take part in the experiment or received credit toward the fulfilment of a course requirement. All had normal or corrected-to-normal vision and were native speakers of German. Since the experiment was advertised as a study on semantic sentence processing, all participants were naïve about its purpose.

Materials

A total of 144 German target words served as stimuli, including 72 morphologically simple single root nouns and 72 NN compounds (see Appendix for sample sentences). Word length ranged from 8 to 10 letters with the distribution of lengths being exactly matched in all cells of the 3 (Regularity)x2 (Word type) design (cell means: 8.7 letters). In the case of the NN composita, the first stem was always four letters long. Word statistics on the orthographic, morphological, and lexical level were computed on the basis of the German CELEX word form corpus (Baayen, Piepenbrock, & van Rijn, 1993). The lexical frequency of the target words was relatively low, with means between 1.04 and 9. 42 per million, and did not differ between the cells of the 3 (Regularity)x2 (Word type) design. For both classes of target words the first four and three letter positions were used to compute the token quadrigram frequency and the token trigram frequency of word beginnings. Within each type of target words there were 24 stimuli with low, medium, and high saliency. As an example, for the regular nouns quadrigram frequency per million tokens was lower than 10, between 11 and 288, and larger than 288 for the low, medium, and high orthographic saliency stimuli respectively (see Table 1 for means and standard deviations). A group of 20 participants who did not take part in the reading experiment was asked to rate potential target words on a 7-point scale of subjective word familiarity from 1 (ubiquitous) to 7 (completely unknown). Only targets that were rated as relatively familiar (cell means between 2.28 and 2.69,

TABLE 1

Token frequency of word beginning quadrigrams and trigrams, mean ratings of word familiarity and mean number of lexical components for all cells of the 3 (Orthographic regularity)x2 (Word type) design. Word statistics are based on the German CELEX corpus. Standard deviations are given in parentheses

Condition	*Example*	*Initial quadrigram frequency*	*Initial trigram frequency*	*Word familiarity*	*Number of morphological components*
Nouns					
Low regularity	Rhabarber	5	41	2.3	1.3
		(5)	(70)	(0.6)	(0.5)
Medium regularity	Autopsie	129	369	2.5	1.5
		(93)	(379)	(0.7)	(0.7)
High regularity	Geschirr	900	2564	2.2	1.5
		(747)	(2742)	(0.6)	(0.5)
Compounds					
Low regularity	Filzstift	5	56	2.7	2.0
		(3)	(75)	(0.6)	(0.2)
Medium regularity	Autokino	129	369	2.6	2.0
		(93)	(379)	(0.7)	(0.2)
High regularity	Haustier	790	2049	2.4	2.0
		(655)	(2746)	(0.6)	(0.2)

no significant differences) were included into the final sample. Since one of the two factors varied in the present design is directly related to morphological complexity, the number of morphological components

(according to the CELEX definition) should differ substantially for the two types of target words. However, within the levels of orthographic saliency, the mean values in this measure were kept nearly identical. Table 1 gives an overview of the cell values for the most important linguistic parameters together with examples for the target words in each condition.

All targets were embedded in one-line declarative sentences with a line width between 71 and 84 characters. The critical words occupied sentence positions between four and six relative to the line beginning. The word preceding the target was an adjective of 5–9 characters length (mean length 7.46 letters, mean lexical frequency 25.4 per million, no significant differences between cells) and the word preceding this adjective was an article. The word length range for the adjectives was chosen on the basis of analysing a large corpus of reading data to maximise the proportion of cases with one fixation on the word before the target. The next word after the target was always a three-letter function word, followed by a noun. Sentences were presented in fixed random order with the constraint that two items belonging to the same condition would not be at adjacent positions within the list. Participants were asked to read at their normal pace such that they would be able to understand the main content of the text. To make sure that readers would follow this instruction, catch trials including sentences with a semantic inconsistency were added and participants asked to indicate the occurrence of gross irregularities in sentence meaning. These semantic irregularities only became apparent at the end of the respective sentences, whereas the initial part of the sentences that contained the target words was always correct.

Apparatus

Sentences were displayed on a Nokia 19-inch crt monitor running at a pixel resolution of 1024x768. Stimuli were shown in dark grey on a white background near the vertical centre of the screen. Text was displayed in nonproportional Courier font such that each letter maximally occupied 15 pixels horizontally and 25 pixels vertically. At a viewing distance of 71cm, each character subtended approximately 0.33° of visual angle. Eye movements were recorded via an Eyelink I video-based pupil tracking system. Viewing was binocular but eye movements were recorded from the right eye only. The recording system included a high speed video camera positioned below the monitored eye and held in place by head-mounted gear. The system has a relative spatial resolution in the order of a few minutes of arc and its absolute accuracy is better than 0.33°. Its output is linear over the vertical and horizontal range of the display. Fixation locations were sampled every 4 ms and used to determine basic measures of oculomotor activity during reading. The on-line saccade detector of the eye tracker was set to detect saccades with an amplitude of 0.15° or greater, using an acceleration threshold of $8000°/s^2$ and a velocity threshold of 30°/s.

Procedure

Participants were tested individually. To avoid gross head movements, the reader's head was stabilised by a chin rest. Head position was recorded by an additional head mounted camera and small movements were compensated online. At the beginning of the experiment a one-dimensional horizontal calibration routine was initiated when the participant pressed the space bar of a keyboard. During calibration, the reader was asked to fixate a sequence of three fixation markers as they appeared in fixed order at the horizontal midline of the screen. One of these markers was located near the left side of the screen, one at the horizontal screen centre and one near the right side of the screen. Calibration was immediately followed by a validation routine that determined the stability and accuracy of the initial measurement. Successful calibration was followed by the presentation of a fixation marker, consisting of a plus sign, that was shown at the left side of the screen. A second pressing of the space bar replaced the fixation marker with a line of text which remained visible until the sentence was read, which was signalled by the reader with a second space bar pressing. This self-paced sentence reading procedure was used throughout the experiment. The experimental session consisted of a training block with 28 practice trials including 4 semantically inconsistent sentences. The experimental sentences were divided into three blocks of 48 items plus 8 inconsistent sentences that were presented in a counterbalanced order. The experiment lasted between 45 and 60 min.

Data selection and data analyses

A target word was considered fixated when a fixation fell on one of its constituent letters or the blank space preceding it. Target fixation durations of less than 70ms and of more than 1500ms were removed from analyses. Also excluded were trials in which the first fixation on the target word was not preceded by a progressive saccade. Together with cases of blinks or track losses, these restrictions resulted in the rejection of 3.6% of all observations. Saccade landing positions, denoting the letter at which the eye landed, saccade amplitude, consisting of the number of letter spaces the eyes traversed to reach the target, and launch site, consisting of the distance (in letter spaces) between the last fixation location prior to the target's fixation and the blank space preceding that target, served as dependent variables to test for the occurrence of cognitive landing site effects. Landing positions of incoming saccades were rounded to a tenth of a character before averaging, with the space preceding the target word coded as position 0.1 to 0.9. Two commonly reported measures of target viewing durations were computed: First fixation durations consisted of the duration of the initial fixation on the word, irrespective of whether the target was subsequently refixated. Gaze durations included the time spent viewing the target, including the time spent refixating it during first pass reading, but excluding the duration of saccadic movements (see Inhoff & Radach, 1998, and Rayner, 1998, for discussions of oculomotor measures). Spatial and temporal eye movement parameters were subjected to 3 (Orthographic regularity)x2(Word type) analyses of variance (ANOVA) using subject *(F1)* and item *(F2)* variability in the computation of error terms.

RESULTS

The spatial measures of the saccade that placed the eyes on a target word, launch distance, saccade amplitude, and the resulting initial landing position, together with the duration of the first fixation on the target and the target's gaze duration are shown in Table 2. As can be seen, orthographic regularity had a profound effect on saccade amplitude specification. This was expressed in a highly significant main effect for landing positions of initial saccades into target words, $F1(2, 46)=11.47$, $p<.01$; $F2(2, 138)=5.81$, $p<.01$, and in a corresponding highly reliable main effect for saccade size, $F1(2, 46)=7.10$, $p<.01$; $F2(2, 138)=3.67$, $p<.05$. Moreover, the orthographic regularity effect could not be attributed to

TABLE 2

Spatial parameters of incoming progressive saccades and viewing duration measures for initial fixations on target words. Standard errors of means are given in parentheses

Condition	*Landing position*	*Saccade amplitude*	*Launch distance*	*Fixation duration*	*Gaze duration*
Nouns					
Low regularity	3.4	8.3	5.4	248	319
	(0.9)	(1.6)	(1.5)	(40)	(73)
Medium regularity	3.5	8.7	5.8	247	325
	(0.9)	(1.6)	(1.7)	(35)	(73)
High regularity	3.7	8.8	5.7	237	288
	(0.8)	(1.6)	(1.5)	(35)	(59)
Compounds					
Low regularity	3.3	8.5	5.7	270	372
	(0.9)	(1.6)	(1.5)	(42)	(91)
Medium regularity	3.6	8.6	5.6	270	357
	(0.9)	(1.7)	(1.7)	(40)	(88)
High regularity	3.8	8.8	5.6	249	314
	(0.7)	(1.6)	(1.6)	(34)	(76)

spatial properties of the pretarget fixation, as orthographic regularity had no effect on launch distance, $F1(2, 46)=0.71$, $p=.50$; $F2(2, 138)= 0.77$, $p=.46$.

In contrast to this, word type did not influence spatial saccade parameters in any of the analyses, including landing position, saccade amplitude, and launch site, all $F1$ and $F2<1$. The interaction between orthographic regularity and word type also did not approach statistical significance in the landing position data, $F1$ $(2, 46)=1.81$, $p=.18$; $F2<1$, or saccade size data, $F1$ and $F2<1$. There was, however, a marginally significant interaction in the launch site data, $F1(2, 46)= 2.80$, $p=.07$; $F2(2, 138)=1.55$, $p=.22$. The reason for this marginal effect becomes apparent when looking at Table 2. In the case of low regularity single root nouns, the mean launch position was relatively near, such that the modulation of saccade amplitude had to work against an opposing tendency. Consequently, the difference in landing position was somewhat smaller for simple nouns as for compounds while the difference in saccade amplitude showed the opposite (nonsignificant) tendency.

Initial fixation durations and gaze durations for the target indicated that orthographic regularity and the morphological status of the target influenced the time spent viewing it. Specifically, first fixation durations and gaze durations increased as the orthographic regularity of the beginning letter sequence decreased, $F1(2, 46)=12.04$, $p<.01$; $F2(2, 138)=6.37$, $p<.01$, and $F1(2, 46)= 44.22$, $p<.01$; $F2(2, 138)=8.56$, $p<.01$, respectively. Furthermore, first fixation durations and gaze durations were shorter when the fixated target noun contained one root morpheme than when it was a compound, $F1(1, 23)=46.66$, $p<.01$; $F2(1,$

138)=20.89, p<.01, and $F1(1, 23)=47.79$, p<.01; $F2(1, 138)= 17.41$, p<.01, respectively. Effects of orthographic regularity tended to be smaller for nouns with a single root morpheme than for compounds but the corresponding interactions did not approach significance in the items analyses, $F1(2, 46)=2.35$, p=.11; $F2$<1, for first fixation durations, and $F1(2, 46)= 7.10$, p<.01; $F2$< 1, for gaze durations.

Supplementary analyses

Landing site distributions are strongly affected by the fixation pattern on the previous word (Radach & Kempe, 1993; Vonk et al., 2000). Even when launch distance is controlled, landing sites are slightly shifted to the right when the prior word has been refixated and are substantially shifted to the left when the prior word has not been fixated (skipped). In the present data in 61.1% cases the prior word received one fixation, in 28.5% cases it was refixated, and in 10.4% it was skipped. To control for possible effects of this factor, the analyses of landing sites and saccade amplitudes were repeated for the subsample of cases with at least one fixation on the preceding word. For these data, the effect of orthographic regularity on landing positions was significant, $F1(2, 46)=11.85$, p< .01; $F2(2, 138)=5.57$, p<.01. The effect of word type was negligible, $F1$ and $F2$<1, as was the interaction of orthographic regularity and word type, $F1(2, 46)$ =2.10, p=.13; $F2$<1. Similarly, the regularity of word beginnings significantly influenced saccade amplitude, $F1(2, 46)=9.22$, p<.01; $F2(2, 138)$ =2.77, p=.06, but the effect of word type and the interaction of regularity and word type were negligible, all Fs<1. The distance from which the incoming initial saccade was launched was again neither related to regularity, $F1$ and $F2$< 1, nor to word type, $F1$ and $F2$<1, and the corresponding interaction once more did not reach statistical significance, $F1(2, 46)=2.32$, p=.109; $F2(2, 138)= 1.05$, p=.35. Overall, this subanalysis thus revealed an even more clear-cut effect pattern than the full set of data. Also, the suggestion of a slightly greater launch distance for saccades toward low regularity nouns with a single root morpheme had disappeared. In a further set of supplementary analyses including only cases where exactly one fixation was made on the preceding word essentially the same pattern of results was obtained. Means and standard errors for observation with at least one and exactly one fixation on the previous word are listed in Table 3.

One further key question about cognitive landing site effects is whether they are restricted to relatively near launch distances or generalise to the full range of launch sites. To address this issue, launch site distributions were computed individually and partitioned such that for each subject approximately half of the observations would fall into a “near” versus “far” launch category. This procedure resulted in mean launch distances of 3.7 letters (SD=1.7) for near launches and 7.8 (SD=2.9) letters for far launches, representing 53% and 47% of the total data set. Table 4 reports mean landing positions, saccade amplitudes and launch sites for far versus near launch distances.

TABLE 3

Spatial parameters of incoming progressive saccades for cases with at least one fixation on the prior word (89.6% of the data) and exactly one fixation on the prior word (61.1% of the data). Standard errors of means are given in parentheses

	At least one fixation on prior word			*Exactly one fixation on prior word*		
Condition	*Landing position*	*Saccade amplitude*	*Launch distance*	*Landing position*	*Saccade amplitude*	*Launch distance*
Nouns						
Low regularity	3.6	8.0	5.0	3.3	8.2	5.5
	(0.9)	(1.5)	(1.3)	(1.0)	(1.5)	(1.2)
Medium regularity	3.6	8.2	5.1	3.4	8.4	5.5
	(0.9)	(1.5)	(1.5)	(1.0)	(1.5)	(1.4)
High regularity	3.8	8.4	5.1	3.6	8.6	5.6
	(1.0)	(1.7)	(1.5)	(0.9)	(1.6)	(1.4)
Compounds						
Low regularity	3.4	8.0	5.1	3.2	8.2	5.6
	(0.9)	(1.6)	(1.3)	(1.0)	(1.6)	(1.3)
Medium regularity	3.7	8.3	5.1	3.5	8.4	5.5
	(0.97)	(1.6)	(1.5)	(1.1)	(1.5)	(1.4)
High regularity	3.9	8.3	4.9	3.6	8.4	5.3
	(0.9)	(1.6)	(1.4)	(0.9)	(1.5)	(1.2)

This supplementary analysis also confirmed the results of the main analyses with effects of orthographic regularity irrespective of launch distance. For near launches, the main effect of orthographic regularity on both landing position, $F1(2, 46)=7.86$, $p<.01$; $F2(2, 138)=3.08$, $p<.05$, and saccade amplitude, $F1(2, 46)=4.98$, $p<.05$; $F2(2, 138)=2.61$, $p=.07$, was significant, while there was no effect of word type: landing position: $F1$ and $F2<1$; saccade amplitude: $F1(1, 23)=2.40$, $p=.13$; $F2(1, 138)=2.68$, $p=.10$. The interaction of orthographic and word type properties was also not significant for landing position, $F1$ and $F2<1$, and saccade amplitude, $F1$ and $F2<1$. In the case of far launches, again, a significant effect of regularity on landing position was present, $F1(2, 46)=9.64$, $p<.01$; $F2(2, 138)=3.69$, $p<.05$, but there was neither an effect of word type, $F1(1, 23)=1.24$, $p=.28$; $F2(1, 138)=3.46$, $p=.07$, nor an interaction of regularity and word type, $F1$ and F2< 1. Saccade amplitudes showed a robust regularity effect in the $F1$ analysis, $F1(2, 46)=4.42$, $p<.05$, and a marginal effect in the $F2$ analysis, $F2(2, 138)=2.15$, $p=.12$. Word type again did not influence saccade amplitude, $F1$ and $F2 < 1$, and the interaction of the two factors was once more negligible, $F1$ and $F2 < 1$.

Spatial distributions of saccade landing positions are depicted in Figures 1 and 2 as a function of orthographic regularity. In addition to exhibiting the typical truncated normal distributions (Rayner, 1979; McConkie et al., 1988) the figures show the expected large effect of launch distance on landing position. For near

TABLE 4

Spatial parameters of incoming progressive saccades for initial fixations on target words computed separately for near versus far launch distances relative to the word beginning. Standard errors of means are given in parentheses

	Near launches			*Far launches*		
Condition	*Landing position*	*Saccade amplitude*	*Launch distance*	*Landing position*	*Saccade amplitude*	*Launch distance*
Nouns						
Low regularity	4.1	7.4	3.8	2.6	9.7	7.7
	(1.0)	(1.3)	(1.2)	(0.8)	(2.1)	(2.1)
Medium regularity	4.2	7.5	3.9	2.8	10.2	7.9
	(1.1)	(1.5)	(1.5)	(0.7)	(1.9)	(1.8)
High regularity	4.4	7.7	3.8	2.9	10.2	7.8
	(1.1)	(1.5)	(1.4)	(0.6)	(2.0)	(1.9)
Compounds						
Low regularity	4.1	7.3	3.7	2.6	9.9	7.9
	(1.1)	(1.4)	(1.0)	(0.7)	(1.9)	(1.9)
Medium regularity	4.1	7.4	3.8	3.0	9.9	7.5
	(1.1)	(1.7)	(1.5)	(0.8)	(2.0)	(2.0)
High regularity	4.3	7.5	3.1	3.0	10.5	8.0
	(1.0)	(1.3)	(1.3)	(0.8)	(2.2)	(2.2)

launches, the mean landing position was 4.2, whereas for far launches incoming saccades on the average landed at position 2.8. Thus, a difference of 4.1 in launch distance (see above) led to a difference in landing positions of 1.4. The proportion of about 0.35 between these two values is in the order of what would be predicted on the basis of results from large corpora of reading data (Radach & McConkie, 1998).[4] More importantly, it is evident from the figures that in both launch distance ranges the effect of orthographic regularity is a graded one; the distributions appear to be shifted as a whole. Only for far launches there is a suggestion of some extra fixations on the first letter and on the space preceding the target word as reported by Hyönä (1995).

[4] In the present study mean landing positions are computed empirically on the basis of saccades that actually landed on the target word. When taking into account cases of under- and overshooting the target word boundary by estimating the central tendencies of the respective normal distributions, a ratio of about 0.5 between variations in launch distance and landing position would be predicted (McConkie et al., 1988).

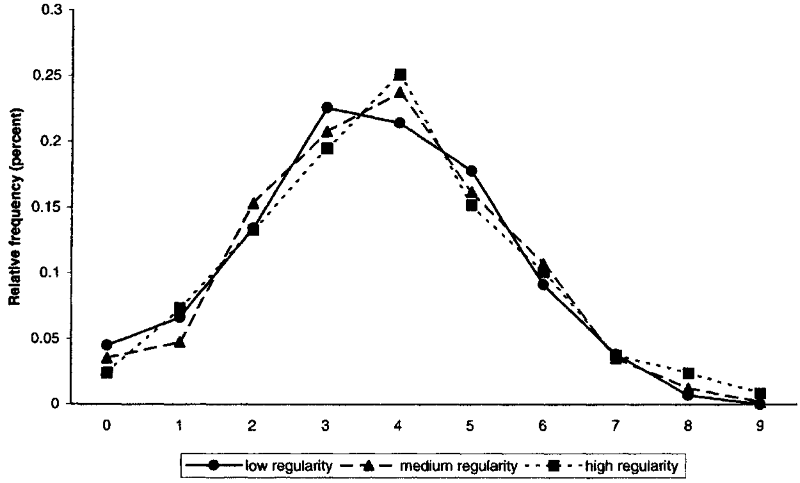

Figure 1. Spatial distributions of landing positions for saccades into target words with low, medium, or high orthographic regularity. Data for saccades that were launched from positions relatively near to the left boundary of the target word (mean launch distance: 3.7, standard deviation: 1.7).

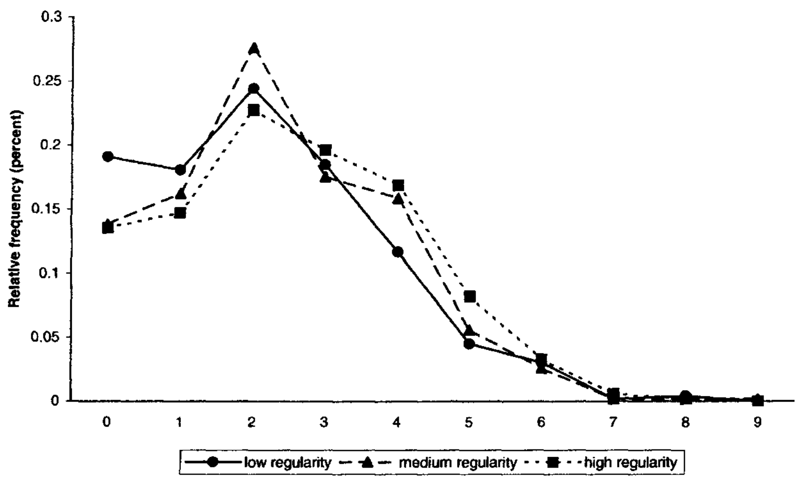

Figure 2. Spatial distributions of landing positions for saccades into target words with low, medium, or high orthographic regularity. Data for saccades that were launched from positions relatively far to the left boundary of the target word (mean launch distance: 7.8, standard deviation: 2.9).

DISCUSSION

The main purpose of the current study was to determine the influence of two types of linguistic information, orthographic regularity and morphological word structure, on saccade amplitudes and landing positions in normal reading. Toward this goal, a three-step variation of orthographic regularity was implemented by varying the token frequency of word initial quadrigrams and trigrams while the lexical frequency and familiarity was controlled. An analysis of word viewing time measures showed that this manipulation strongly influenced target processing. When it was fixated, readers spend more time on the target as the degree of orthographic regularity of the target's beginning letter sequence decreased. This confirms earlier studies that also found that words with a rare initial letter cluster require longer viewing durations (Lima & Inhoff, 1985; Vonk et al., 2000). In addition, readers spent more time viewing compound words than words with a single root morpheme, as reported in Inhoff et al. (1996).

Spatial properties of the saccade that reached the target revealed a different pattern of effects. Here, it was of no substantial consequence whether the critical word was a simple noun or a compound. Therefore, the current findings are not in harmony with another aspect of Inhoff et al.'s (1996) results, that showed larger incoming saccades for compounds than for monomorphemic controls. Our data are also difficult to reconcile with a particular interpretation of Hyönä and Pollatsek's (1998, 2000) findings, according to which interword saccades land further into words with a more frequent first constituent (see Andrews et al., 2004, for a compound reading study that does not find this effect). In the present experiment, the beginning quadrigram frequency is equivalent to the lexical frequency of the first constituent. Since landing positions were shifted to the right as orthographic regularity increased, it may appear as if landing position was modulated by word frequency. Critically, however, the shift in landing position was present to an equal degree when the beginning quadrigram was not a word. This is particularly compelling in the case of the medium regularity items where word beginning quadrigrams were identical. A similar result was obtained by Hyönä (1995) who compared Finnish compounds versus derived words with identical beginnings and also found no difference in landing positions. Taken together, these findings provide solid evidence against the idea that parafoveal lexical or morphological processing may have a modulating effect on saccade amplitude specification.[5] From a more methodological point of view they also suggest that in studies using comparisons between NN compounds, it may be quite difficult to exclude the possibility that a seeming morpho-lexical effect of initial constituent frequency on spatial saccade parameters is in fact due to processing on the orthographic level.

The data presented in this paper add to a growing body of evidence in favour of orthographic landing site effects. We have shown that saccades into orthographically more regular (less salient) words have longer amplitudes and the landing positions of these saccades are located further to the right. The size of this main effect (about 0.3 letters) is in the order of what has been reported before in studies that also found orthographic landing site effects in continuous reading (Hyönä, 1995; Vonk et al., 2000). Our results indicate that the effect is of a graded nature in two related but different respects: First, as in Vonk et al., the distributions are shifted as a whole rather than showing a discrete deviation from the norm in the case of irregular word beginnings as found by Hyönä (1995). Second, since our study used a three-level variation of orthographic regularity, we were able to demonstrate that the differences between low versus medium regularity words, and medium versus high regularity words, were in the same order of magnitude. Hence, there is good reason to assume that the effect is continuous over a wide range of orthographic processing load. In our view this excludes all accounts on the basis of visual or preattentive processing (see below). Our supplementary analyses of observations with saccades coming from relatively near versus relatively far launch positions further indicated that this influence of parafoveal word-beginning orthography on saccade specification is not restricted to targets that are very close to the current fixation (see Vonk et al. for a converging result).

Focusing on the fact that parafoveal orthographic processing influences the programming of saccades, the question arises how this effect can be accounted for. Current proposals on this issue have taken several routes. First, Hyönä (1993) suggested that irregular letter clusters may "pop out" and "attract" saccades towards them. This idea has also been promoted by Beauvillain and Doré (1998). This proposal is certainly appealing, as the notion of "pop out" has high face validity. However, going back to its original definition in the area of visual search, the concept was developed to account for the *automatic and parallel detection* of critical targets in an array containing distractors that differ on a physical feature dimension or visual feature changes in two successively shown displays (e.g., Treisman & Gelade, 1980). To our knowledge there is no direct evidence suggesting that this type of pop-out effect can also apply to visually similar letter strings with variations in particular linguistic properties.

Findlay and Walker (1999), in the framework of their saliency map theory of saccade generation, proposed that modifications of saccade amplitudes due to parafoveal processing may be accounted for in terms of an *intrinsic saliency* that

---

[5] One might argue that, if the beginning letter sequence corresponded to a lexeme, readers may have assumed that the parafoveal string is a "word", irrespective of the actual whole-word type (compound or single root noun). However, in this case landing positions should have been further to the left in the low or high regularity simple noun conditions where initial quadrigrams never corresponded to a potential lexeme. Also, if readers incorrectly assumed that the parafoveally visible beginning quadrigram of a single root noun was a morpheme, then subsequent viewing durations should have been relatively long—as there was a need to recover from garden pathing (see also Inhoff, 1989, and Lima, 1987, for evidence against morphological processing in the parafovea).

a stimulus may carry. The current results are also difficult to reconcile with this view, unless one would want to claim that we keep in memory a specific repertoire of rare letter clusters that are unlikely to form word beginnings. Another interpretation of intrinsic saliency could be to claim that all possible letter clusters are salient for a beginning reader and that the saliency is reduced with repeated encounters of a particular visual pattern (see Pollatsek & Rayner, 1989, for a critical discussion of visual template matching accounts of visual word recognition). However, any means *of gradually* differentiating a "salient" word beginning from a background of less salient letter strings must necessarily require a minimum of orthographic processing. However, as soon as the orthographic processor is activated, processing is no longer confined to the level of visual perception and concepts like "pop out", preattentive processing, and "intrinsic" saliency lose much of their explanatory appeal.

An alternative hypothesis was put forward by Hyönä and Pollatsek (1998, 2000) to account for their effect on landing position of initial compound constituent frequency. They suggested that high parafoveal processing load may reduce the extent of the perceptual span, which in turn would lead to a shortening of saccades. However, applying such a proposal to parafoveal processing may carry the danger of logical circularity: In the present context the perceptual span would be defined as the region around the current fixation position within which foveal and parafoveal letter discrimination takes place. Only *as a result* of this parafoveal orthographic letter processing can the system know that the beginning of the next word "does not compute". Also, there is direct empirical evidence according to which orthographic properties of the beginning letters of a parafoveally visible word do not influence the acquisition of useful letter-level information from the parafovea (Lima & Inhoff, 1985). In the experiment readers could or could not obtain useful parafoveal letter or word information from a parafoveally visible target word, and eye-movement-contingent display changes were used to show the intact target when it was fixated. In addition, the study controlled orthographic properties of the target's onset trigram so that it was consistent with a large number of words (e.g., roo) or with a relatively small number of words (e.g., dwa). The target words themselves (rooster and dwarf in the example) were matched on a wide range of lexical properties. Analyses of target viewing durations as the function of the type of previously visible target preview revealed virtually identical benefits from common (roo) and rare (dwa) word onset trigrams, suggesting that the processing of rare parafoveally visible word onset letters did not narrow the perceptual span. Consistent with this view, van de Weijgert (1993) found a uniform decrease in visual discrimination performance over a range of eccentricities rather than evidence for "tunnel vision" (see also Williams, 1988) in a memory comparison task that sought to manipulate nonvisual processing load.

Looking at basic oculomotor research there is a body of evidence dealing with graded effects on saccade amplitudes. In so-called double step experiments, two targets are shown in rapid succession and participants are asked to follow this seeming two-step movement as fast as possible. If the second target displacement goes in the same direction (e.g., to the right on a horizontal plane) and occurs within an interval approximately 70–180ms before the onset of the first saccade, this saccade consistently lands *in between* the first and the final target location (Becker & Jürgens, 1979; Ottes, van Gisbergen, & Eggermont, 1984; see Becker, 1989, for a review). As shown by Findlay and Harris (1984), for step eccentricities similar to the amplitude of reading saccades there is a continuous transition of the saccade amplitude toward the second target location, as the available reprocessing time increases.

This line of evidence was ingeniously applied to reading by Morrison (1984). Considering lexical access of a fixated word as the trigger for an attention shift and subsequent saccade initiation, he described a scenario where lexical access of an adjacent word N+l has taken place but the information about this arrives relatively late. Hence, a saccade that is being programmed to go to word N+1 cannot be cancelled and replaced with a saccade to word N+2. However, the critical information may arrive early enough to allow for the kind of spatiotemporal averaging described above, leading to a lengthening of the next saccade amplitude. Also, since in this case the subsequent saccade to word N+2 is already programmed, the fixation duration after the initial saccade can be reduced to a minimum latency, accounting for the very short fixation durations often observed in continuous reading (Radach, Heller, & Inhoff, 1999).

It is clear from the results obtained in research using the double step paradigm that the mechanism responsible for amplitude specification is prepared to take into account visual information before and *after* the decision to move has been made (see Deubel, O'Regan, & Radach, 2000, for a more detailed discussion). Cognitive landing site effects appear to suggest that this is true not only for visual but also for specific linguistic information. One way to implement such a mechanism would be to apply the Morrison scenario to data like those obtained in the present study. Assuming that the initial letter cluster is parafoveally processed, the system may sometimes make an attempt to "skip" over this word segment, initiating a saccade to go beyond these letters (Rayner & Morris, 1992). In the present experiment, mean initial fixation durations were indeed shorter for target words with more regular beginnings, as predicted by Morrison's reprogramming hypothesis. On the other hand, an inspection of the distributions underlying these means indicated that there was no increased frequency of very short fixation durations for highly regular beginning words. Nonetheless, this line of reasoning may in principle open a way to accommodate saccade amplitude modifications as a result of sublexical parafoveal processing within attention-based sequential processing models of eye movement control (Reichle et al., in press).

A parsimonious way to explain the observed orthographic landing site effects may be in terms of continuous feedback from ongoing orthographic processing to saccade amplitude specification. A theoretical framework allowing for an integration of visual and cognitive processing in the process of saccade generation has been proposed by Findlay and Walker (1999). They suggest that the specification of saccade targets is based on parallel processing and competitive inhibition within a spatial saliency map. In the recent Glenmore model by Reilly and Radach (2003), this takes the form of a saliency vector initially representing the visual configuration of letters within the perceptual span on a line of text. The model also implements a linguistic processing module including an interactive activation network for orthographic letter processing. Most importantly, results of processing within this network are fed back continuously to the saliency vector, leading to dynamic changes in local saliency values. In the current version of the model, these saliency values are integrated for each word, forming a preference list of potential target objects. When a saccade is triggered, it is always directed to the centre of the word with the highest saliency. Thus, the saccade generator is insensitive for properties of specific letter clusters. In a future version of the model saccade amplitude adjustments as a result of parafoveal orthographic processing could be made possible by taking local saliency fluctuations into account when specifying the saccade goal. The saliency maximum within the target word is likely to be further to the right for high regularity word beginnings at the point in time when the impending saccade to word N+1 is committed to action. This could effectively account for a graded effect of word initial regularity on the resulting saccade amplitude.

In sum, the present study has reported evidence in favour of a graded modulation of interword saccade amplitude in response to variations in the orthographic regularity of parafoveal word beginnings. The modification of current models of oculomotor control in reading to account for the relatively small orthographic landing site effects appears to require changes in model architecture and will certainly make these models more complex. More research into graded orthographic landing site modulations is needed to establish the phenomenon as an undisputed empirical fact. If this research leads to unequivocal results, the investment to incorporate this effect into current theories of eye movement control will begin to appear rewarding.

PrEview proof published online September 2003

REFERENCES

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*(1/2), 285–311.

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

Balota, D.A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influence of lexical variables. In D.Besner & G.W.Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 198–232). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Beauvillain, C., & Doré, K. (1998). Orthographic codes are used in integrating information from the parafovea by the saccadic computation system. *Vision Research, 38,* 115–123.

Beauvillain, C., Doré, K., & Baudouin, V. (1996). The "centre of gravity" of words: Evidence for an effect of word-initial letters. *Vision Research, 36,* 589–603.

Becker, W. (1989). Metrics. In R.H.Wurtz & M.E.Goldberg (Eds.), *The neurobiology of saccadic eye movements* (pp. 13–61). Amsterdam: Elsevier.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19,* 967–983.

Briihl, D., & Inhoff, A.W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and of exterior letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 55–67.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–148). Oxford, UK: Elsevier.

Brysbaert, M,Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychologia, 10,* 385–395.

Deubel, H., O'Regan, J.K., & Radach, R. (2000). Attention, information processing and eye movement control. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 355–376). Oxford, UK: Elsevier.

Everatt, J., & Underwood, G. (1992). Parafoveal guidance and priming effects during reading: A special case of the mind being ahead of the eyes. *Consciousness and Cognition, 1,* 186–197.

Findlay, J.M., & Harris, L.R. (1984). Small saccades to double-stepped target moving in two dimensions. In A.G.Gale & F.Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 71–78). Amsterdam: Elsevier.

Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioural and Brain Sciences, 22,* 661–721.

Hochberg, J. (1970). Components of literacy: Speculations and exploratory research. In H.Levin & J. P.Williams (Eds.), *Basic studies in reading*. New York: Basic Books.

Hochberg, J. (1976). Toward a speech-plan eye movement model of reading. In R.A.Monty & J.W. Senders (Eds.), *Eye movements and psychological processes* (pp. 397–416). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hyönä, J. (1993). *Eye movements during reading and discourse processing* [Psychological Research Rep. No. 65]. Turku, Finland: University of Turku.

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

Hyönä J., Niemi, P., & Underwood, G. (1989). Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 142–152.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Hyönä, J., & Pollatsek, A. (2000). Processing of Finnish compound words in reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 65–87). Oxford, UK: Elsevier.

Inhoff, A.W. (1982). Parafoveal word perception: A further case against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 137–145.

Inhoff, A.W. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 544–555.

Inhoff, A.W., Briihl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming and delayed naming tasks. *Memory and Cognition, 24,* 466–476.

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

Inhoff, A.W., & Radach, R. (2002). The biology of reading: Using of spatial segmentation in the reading of complex words. *Comments on Theoretical Biology, 7,* 121–138.

Inhoff, A.W., Radach, R., Eiter, B., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *Quarterly Journal of Experimental Psychology, 56A,* 803–827.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Kerr, P.W. (1992). *Eye movement control during reading: The selection of where to send the eyes.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, USA.

Lavigne, F., Vitu, F., & d'Ydewalle, G. (2000). The influence of semantic context on initial eye landing sites in words. *Acta Psychologica, 104,* 191–214.

Lima, S.D. (1987). Morphological analysis in sentence reading. *Memory and Language, 26,* 84–99.

Lima, S.D., & Inhoff, A.W. (1985). Lexical access during eye fixations in reading: Effects of word initial letter sequence. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 272–285.

Liversedge, S.P., & Underwood, G. (1998). Foveal processing load and landing position effects in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 201–221). Oxford, UK: Elsevier.

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

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

McConkie, G.W., & Zola D. (1987). Visual attention during eye fixations while reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 385–401). Hove, UK: Lawrence Erlbaum Associates, Ltd.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Nazir, T.A., Heller, D., & Suβmann, C. (1992). Letter visibility and word recognition: The optimal viewing position in printed words. *Perception and Psychophysics, 52,* 315–328.

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

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

O'Regan, J.K., Lévy-Schoen, A., Pynte, J., & Brugaillère, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 250–257.

Ottes, F.P., van Gisbergen, J.A.M., & Eggermont, J.J. (1984). Metrics of saccade responses to visual double stimuli: Two different modes. *Vision Research, 24,* 1169–1179.

Pollatsek, A., & Rayner, K. (1989). Reading. In M.I.Posner (Ed.), *Foundations of cognitive science* (pp. 401–436). Cambridge, MA: MIT Press.

Pynte, J.,Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1/2), 178–202.

Radach, R., Heller, D., & Inhoff, A.W. (1999). Occurrence and function of very short fixation durations in reading. In W.Becker, H.Deubel, & T.Mergner (Eds.), *Current oculomotor research: Physiological and psychological aspects.* New York: Plenum Publishers.

Radach, R., & Kempe, V. (1993). An individual analysis of initial fixation positions in reading. In G. d'Ydewalle & J.Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 213–226). Amsterdam: Elsevier.

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

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

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

Rayner, K., Balota, D.A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40(4),* 473–483.

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41,* 943–954.

Rayner, K., & Morris, R.K. (1992). Eye movement control in reading: Evidence against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 163–172.

Rayner, K., Reichle, E.D., & Pollatsek, A. (2000). Eye movement control in reading: Updating the E-Z Reader model to account for initial fixation locations and refixations. In A.Kennedy, R.Radach, D. Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 701–719). Oxford, UK: Elsevier.

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

Rayner, K., Well, A.D., Pollatsek, A., & Bertera, J.H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31,* 537–550.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z reader model. *Vision Research, 39,* 4403–4411.

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

Reilly, R.G., & O'Regan, J.K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research, 38,* 303–317.

Reilly, R., & Radach, R. (2003). Glenmore: An interactive activation model of eye movement control in reading. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429–456). Amsterdam: Elsevier.

Stevens, M, & Grainger, J. (2003). Letter visibility and the viewing position effect in visual word recognition. *Perception and Psychophysics, 65,* 133–151.

Treisman A.M., & Gelade G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Underwood, G., Bloomfield, R., & Clews, S. (1988). Information influences the pattern of eye fixations during sentence comprehension. *Perception, 17,* 267–278.

Underwood, G., Clews, S., & Everatt, J. (1990). How do readers know where to look next? Local information distributions influence eye fixations. *Quarterly Journal of Experimental Psychology, 42A,* 39–65.

Underwood, G., & Radach, R. (1998). Eye guidance and visual information processing: Reading, visual search, picture perception and driving. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 1–27). Oxford, UK: Elsevier.

Vitu, F. (1991). The existence of a center of gravity effect during reading. *Vision Research, 31,* 1289–1313.

Vitu, F., McConkie, G.W., Kerr, P., & O'Regan, J.K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41,* 3513–3533.

Vonk, W., Radach, R., & van Rijn, H. (2000). Eye guidance and the saliency of word beginnings in reading text. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 269–299). Oxford, UK: Elsevier.

van de Weijgert, E. (1993). Foveal load and peripheral task performance: Tunnel vision or general interference? In D.Brogan, A.Gale, & C.Carr (Eds.), *Visual search 2.* London: Taylor & Francis.

White, S., & Liversedge, S. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*(1/2), 52–78.

Williams, L.J. (1988). Tunnel vision or general interference? Cognitive load and attentional bias are both important. *American Journal of Psychology, 101,* 171–191.

APPENDIX

Sample sentences for each condition in the 3 (orthographic regularity)x2 (word type) design together with (literal) translations with target words written in italics. Note that in German all nouns have capital initial letters.

Simple noun target, low orthographic regularity Er stellte das gläserne *Aquarium* ins Wohnzimmer seiner neuen Wohnung. (He put the glass *aquarium* in the living room of his new apartment.)

Simple noun target, medium orthographic regularity Nächsten Monat soll das komplette *Fundament* des Hochhauses fertiggestellt werden. (By next month the complete *foundation* of the high-rise building should be finished.)

Simple noun target, high orthographic regularity Überraschend gewann der junge *Interpret* den Musikpreis der Stadt Hamburg. (Surprisingly the young *interpreter* [singer] won the music award of the city of Hamburg.)

Compound noun target, low orthographic regularity Für Säuglinge ist weiches *Lammfell* als Schlafunterlage von Geburt an zu empfehlen. (For babies soft *lambskin* as a sleeping pad from birth onwards is to be recommended.)

Compound noun target, medium orthographic regularity Der Beamte schleppte den schweren *Postsack* vom Erdgeschoss bis in die siebte Etage. (The clerk carried the heavy *mailbag* from the ground floor to the seventh level.)

Compound noun target, high orthographic regularity Der Tourist genoss den schönen *Rundblick* vom Leuchtturm der Nordseeinsel Helgoland. (The tourist enjoyed the nice *roundview* [panorama] from the lighthouse of the North Sea island of Helgoland.)

Example for a semantically inconsistant catch trial Zum Glück verfehlte der rote Dachziegel den Termin zur Anmeldung. (Fortunately the red roof tile missed the deadline for the registration.)

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 52–78

Orthographic familiarity influences initial eye fixation positions in reading

Sarah J.Wwhite and Simon P.Liversedge

University of Durham, UK

An important issue in the understanding of eye movements in reading is what kind of nonfoveal information can influence where we move our eyes. In Experiment 1, first fixation landing positions were nearer the beginning of misspelled words. Experiment 2 showed that the informativeness of word beginnings does not influence where words are first fixated. In both experiments, refixations were more likely to be to the left of the initial fixation position if the words were misspelled. Also, there was no influence of spelling on prior fixation durations or refixation probabilities, that is, there was no evidence for parafoveal-on-foveal effects. The results show that the orthographic familiarity, but not informativeness, of word initial letter sequences influences where words are first fixated.

There is considerable evidence to suggest that the visual characteristics of text largely determine where the eyes land during reading. The position of the first fixation on words is mainly dependent on word length, which is visually represented by the spaces between words (McConkie & Rayner, 1975; Morris, Rayner, & Pollatsek, 1990; O'Regan, 1979, 1980; Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998a). First fixations are most likely to land on the preferred viewing position (Rayner, 1979), which is between the beginning and the middle of words (Deutsch & Rayner, 1999; Dunn-Rankin, 1978; McConkie, Kerr, Reddix, & Zola, 1988). Nevertheless, there is substantial variability in the distribution of landing positions on words. McConkie et al. (1988) suggested that variability in landing positions is induced by random error as well as systematic range error related to launch site. In line with this, models of eye movements in reading state that the position of first fixations on words is largely determined by word length and systematic range error (O'Regan, 1990;Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, in press; Reilly & O'Regan, 1998).

Although it is clear that a substantial amount of the variability in first fixation landing positions on words can be explained by the visual characteristics of text (Vitu, O'Regan, Inhoff, & Topolski, 1995), there is also evidence indicating that linguistic processing of fixated and nonfixated text influences saccade computation. For fixated text, (1) lexical processing can influence the probability of making a refixation (Balota, Pollatsek, & Rayner, 1985; Inhoff & Rayner, 1986; Rayner, Sereno, & Raney, 1996), (2) syntactic processing (Frazier & Rayner, 1982) and clause wrap-up can influence the length of subsequent saccades (Hill & Murray, 2000; Rayner, 1975; Rayner, Kambe, & Duffy, 2000), and (3) syntactic and discourse processing can influence the probability of regressions (e.g., Garrod, Freudenthal, & Boyle, 1994; Liversedge, Paterson, & Clayes, 2002; Paterson, Liversedge, & Underwood, 1999; Rayner, Carlson, & Frazier, 1983).

Preprocessing of nonfixated text can also influence landing positions in reading. Studies have shown that landing positions are nearer the beginning of words or saccade lengths are shorter (1) when space but not letter (orthographic) information is presented (Inhoff, 1989; Morris et al., 1990; Rayner, Well, Pollatsek, & Bertera, 1982), (2) for words presented in uppercase letters (Inhoff, Starr, & Shindler, 2000b), (3) for words in which partially correct previews were presented before they were fixated (Inhoff, 1990), and (4) for nonsense letter strings presented in sentences (McConkie, Underwood, Zola, & Wolverton, 1985; Zola, 1984). The results of these studies indicate that nonfoveal processing beyond the level of visual word length information can influence where the eyes move. A number of studies have shown that lexical processing can influence where the eyes move in terms of the probability of fixating words (for a review of this area, see Brysbaert & Vitu, 1998). However, the issue to be investigated in this paper is what kinds of nonfoveal processing can influence where words are initially fixated. Some studies have suggested that nonfoveal processing of predictability within the context of the sentence (Lavigne, Vitu, & d'Ydewalle,

---

Correspondence should be addressed to S.J.White, Centre for Vision and Visual Cognition, Department of Psychology, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK. Email: S.J.White@durham.ac.ukThe order of authors is arbitrary. Sarah J. White acknowledges the support of a University of Durham Studentship. The experiments reported here form part of the first author's PhD thesis. The authors would like to thank Ralph Radach and Joel Pynte for helpful comments on this manuscript.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000204

2000; but see Rayner, Binder, Ashby, & Pollatsek, 2001; Vonk, Radach, & van Rijn, 2000) and morphology (Inhoff, Briihl, & Schwartz, 1996; Hyönä & Pollatsek, 1998; but see Andrews, Miller, & Rayner, 2004) can influence where words are first fixated. Although these studies suggest that nonfoveal lexical processing can influence fixation positions, word frequency has not been found to influence where words are first fixated (Rayner et al., 1996). Furthermore, the issue of whether nonfoveal sublexical processing can influence where words are first fixated remains controversial. The current paper focuses on how sublexical and lexical processing of nonfixated text can influence first fixation positions on words.

The first studies to investigate such landing position effects suggested that the distribution of informativeness (the importance of letter sequences for word recognition) within words could influence where words are initially fixated (Everatt & Underwood, 1992; Hyönä, Niemi, & Underwood, 1989; Underwood, Clews, & Everatt, 1990; Underwood, Hyönä, & Niemi, 1987). However, Hyönä (1993, 1995) argued that these studies confounded the variable of informativeness with orthographic, morphological, and semantic variables. Furthermore, Rayner and Morris (1992) failed to replicate Underwood et al.'s (1990) findings. However, an increasing number of studies have shown that orthographic nonfixated information can influence where words are first fixated. Hyönä (1995) presented words in Finnish sentences with either orthographically familiar or very unfamiliar word beginnings. The unfamiliar words included letters that rarely occur in Finnish. Hyönä found that fixations landed nearer to the beginning of the orthographically unfamiliar words, especially on the space before the word. Radach, Inhoff, and Heller (2004) and Vonk et al. (2000) also found that orthographic familiarity influences where words are first fixated in German and Dutch respectively. In support of these findings, artificial task experiments[1] have shown that orthographically unfamiliar initial letter sequences produce first fixation landing positions nearer the beginning of French words and nonwords than letter strings with familiar initial letter sequences (Beauvillain & Doré, 1998; Beauvillain, Doré, & Baudouin, 1996; Doré & Beauvillain, 1997). However, a number of studies have failed to demonstrate orthographic landing position effects in artificial tasks (Kennedy, 1998, 2000; Radach, Krummenacher, Heller, & Hofmeister, 1995), short passage reading tasks (Liversedge & Underwood, 1998), and corpus reading studies (Radach & Kempe, 1993; Radach & McConkie, 1998).

A number of possible explanations have been proposed to explain the modulation of landing positions on words by orthography. Hyönä (1993) suggested a pull assumption in which salient features (such as irregular orthography) "pop out" of nonfixated text and pull the eye towards them. Beauvillain et al. (1996) extended this idea by suggesting that irregular letter sequences "pop out" of nonfixated text and influence landing positions by adjusting the word length based saccade computation. More recently, Findlay and Walker (1999) suggested that medium—and long-term learning modifies the intrinsic salience of visual stimuli such as orthographic letter sequences. Intrinsic salience then contributes to a salience map in which the distribution of salience across the visual field determines the saccade target. Alternatively, the modulation of first fixation positions on words might be explained by an influence of processing difficulty. McConkie (1979) suggested that saccades are directed to regions of processing difficulty. Hyönä and Pollatsek (1998, 2000) suggested that processing difficulty reduces the perceptual span or the extent of preprocessing, and this shortens saccades (see also Hyönä, 1995).

Overall, there is mixed evidence for whether sublexical and lexical processing can influence where words are first fixated. A similar controversy exists over the issue of whether nonfixated information can influence fixation durations, sometimes referred to as "parafoveal-on-foveal" effects. A number of studies using both artificial and sentence reading tasks have reported that fixation durations are influenced by the characteristics of the following word (Inhoff, Radach, Starr, & Greenberg, 2000a; Inhoff et al., 2000b; Kennedy, 1998, 2000; Kennedy, Pynte, & Ducrot, 2002; Murray, 1998; Murray & Rowan, 1998; Pynte, Kennedy, & Ducrot, 2004; Underwood, Binns, & Walker, 2000). However, the direction of the effects in these studies is inconsistent and there are concerns about the generalisability of some of the findings (Rayner, White, Kambe, Miller, & Liversedge, 2003).

The aim of this study was to investigate whether sublexical and lexical processing can influence where words are first fixated in the reading of English sentences. Experiment 1 was designed to test whether orthographic familiarity influences where the eyes move. Experiment 2 was designed to test whether processing of word beginnings to generate possible lexical candidates can influence where words are first fixated. Both experiments also provide an opportunity to test whether orthographic and lexical characteristics of nonfixated words influence fixation durations on the previous word. Throughout this paper we will use the term "orthographically unfamiliar" to indicate that a letter sequence does not exist in the language or is infrequent. We will use the term "orthographically familiar" to indicate that a letter sequence frequently occurs in the language. Orthographic familiarity is measured using token frequency (the sum of the word frequencies of words including a particular letter sequence). We will use the term "informative" to indicate that a letter sequence is found at the beginning of a small number of words and "uninformative" to indicate that a letter sequence is found at the beginning of many words. Orthographic informativeness is measured using type frequency (the number of words including a particular letter sequence). In the present experiment, unless

[1] Artificial tasks do not involve the measurement of normal reading. However, advocates of such tasks usually argue that they are intended to generalise to reading. Such tasks often provide a greater degree of control over the location from which saccades are launched. Typically, two or three adjacent words are presented and participants then make a categorical decision.

otherwise specified, type and token frequency counts were position specific and case insensitive.

EXPERIMENT 1

A correctly spelled condition with frequent word initial trigrams (e.g., *agricultural*) was compared to four misspelled conditions. Three of the misspelled conditions had different degrees of orthographic familiarity in order that we might identify those characteristics of an orthographically unfamiliar string that influence saccade computation. The most unfamiliar misspelling condition formed an illegal unpronounceable word initial trigram (e.g., *ngricultural*). The word "illegal" indicates that no word in the English language begins with such an initial trigram. A second condition formed an illegal but pronounceable word initial trigram (e.g., *akricultural*). The third misspelled condition had word initial trigrams that occurred in the lexicon but which were unfamiliar (e.g., *aoricultural*). Following the evidence described above, it was predicted that these three misspelling conditions with unfamiliar initial letter sequences would produce first fixation landing positions nearer the word beginning compared to the correctly spelled condition.

The three unfamiliar misspelling conditions necessarily confounded the orthographic familiarity of the initial trigram with the presence of a misspelling. As a result, a fourth misspelling condition was included in which the word initial trigrams were frequent (e.g., *acricultural*). There were no differences in the initial bigram and trigram frequencies between the correctly spelled and high frequency misspelled conditions and so it was predicted that landing positions in these two conditions would be the same.

Method

Participants. Forty-five members of the University of Durham community participated in the experiment. All of the participants were native English speakers with normal or corrected to normal vision. The participants were paid to participate and all were naïve in relation to the purpose of the experiment.

Materials. Word frequencies and *n*-gram frequencies were based on the CELEX English word form corpus (Baayen, Piepenbrock, & Gulikers, 1995). There were 35 critical strings with a mean word length of 9.7 characters (range 8–13) and a mean word frequency of 41 counts per million (SD=53). Stimuli were chosen on the basis of position specific token frequency, but type frequencies followed similar patterns. *N*-gram frequencies were based on counts per 17.9 million because this is a more sensitive measure.

There were five spelling conditions that were manipulated within participants and items. In the baseline condition the critical string was spelled correctly with a high frequency initial trigram. In the high frequency, low frequency, and the illegal pronounceable misspelled conditions the second letter of the word was misspelled. In the illegal unpronounceable misspelled condition the first letter was misspelled if the original first letter was a vowel (21 items) and the second letter was misspelled if the original first letter was a consonant (14 items). There were 35 items in all of the conditions except for the illegal pronounceable misspelled condition in which there were 30 items.[2]

The token frequencies for the initial trigrams were high for both the correctly spelled (M=14,514, SD=17,465) and the high frequency misspelled (M= 11,345, SD=12,848) conditions. The stimuli were chosen primarily on the basis of token frequency but type frequencies for the initial trigrams also tended to be high for both the correctly spelled (M=75, SD=88) and the high frequency misspelled (M= 68, SD= 115) conditions. The position in the word at which the high frequency misspelled words became illegal was uncontrolled. The position ranged from four to seven characters and the mean position was 4.6 characters (SD=0.88). In the low frequency misspelled condition the initial trigram mean type (M=3, SD=3) and token frequencies (M=68, SD=63) were low. To summarise, the position specific token frequencies of the initial trigrams were within the seventh to the ninth deciles for the correctly spelled condition, the sixth to the ninth deciles for the high frequency misspelling condition, and the first to the fifth deciles for the low frequency misspelling condition. For the illegal pronounceable and the illegal unpronounceable misspelled conditions the initial trigram never occurred in the lower case corpus. For the correctly spelled and the high frequency misspelled conditions there were no significant differences between the initial trigram type or token frequencies (ts<1). There were significant differences in the initial trigram type and token frequencies between the correctly spelled and the high frequency misspelled conditions compared to the low frequency and illegal initial trigram conditions (ts #DXGT#3.3, ps<.01). The type and token initial bigram frequencies were high for the correctly spelled and high frequency misspelled conditions but low for the low frequency and illegal misspelled conditions. There were no significant differences in the initial or second (second and third letter) bigram type and token frequencies between the correctly spelled and the high frequency misspelled conditions (ts<1.5). There were significant differences in initial bigram type and token frequencies between the correctly spelled and the high frequency misspelled conditions compared to the low frequency misspelled conditions and the illegal misspelled conditions (*ts#DXGT#*3.8, ps<.01). There were no significant differences in nonposition specific type or token monogram frequency between the correctly spelled and the high frequency misspelled condition (ts<0.6) and the low frequency misspelled condition (ts<1.9).

[2] In Experiment 1, five critical words did not have illegal pronounceable misspellings because there were no suitable letters that could meet the necessary constraints. Therefore the items analyses were based on 30 items and the participants analyses were based on 30 items in the illegal pronounceable condition and 35 items in each of the other four spelling conditions.

The critical strings in each condition were placed in identical sentence frames. Each of the sentences was no longer than one line of text (78 characters) and the critical strings appeared approximately in the middle of the sentence. The words before (word *n*−1) and after (word *n*+1) the critical string were either five or six letters long and had medium to high frequencies.

Five lists of 171 sentences were constructed and nine participants were randomly allocated to each list. Each list included 34 experimental sentences of which 6 items were from the illegal pronounceable condition and 7 items were from each of the other four conditions. The conditions were rotated following a Latin square design. There were 30 misspelled filler sentences with the misspellings in a variety of word lengths and positions. There were also 107 filler sentences that were spelled correctly. Therefore of the 171 sentences, 57 contained a misspelling. Fifty-eight of the sentences were followed by a comprehension question to ensure that participants concentrated on understanding the sentences. The sentences were presented in a fixed random order with eight filler sentences at the beginning. See the Appendix for examples of experimental sentences and critical words.

Procedure. Eye movements were monitored using a Dual Purkinje Generation 5.5 eye tracker. Viewing was binocular but only the movements of the right eye were monitored. The sentences were displayed on a screen at a viewing distance of 70 cm. Three and a half characters subtended one degree of the visual angle. The resolution of the eye tracker is 10 min of arc and the sampling rate was every millisecond.

Participants were instructed to ignore the misspellings and to concentrate on understanding the sentence to the best of their ability. A bite bar and head restraint were used to minimise head movements. The initial calibration procedure lasted approximately 5 min and the calibration accuracy was checked after every few trials during the experiment. After reading each sentence the participants pressed a button to continue and used a button box to respond “yes” or “no” to comprehension questions. The entire experiment lasted approximately 45 min and participants were given two breaks.

Analyses. Fixations shorter than 80 ms that were within one character of the next or previous fixation were incorporated into that fixation. Any remaining fixations shorter than 80ms and longer than 1200ms were discarded. Analyses of word *n*−1, the critical string and word *n*+1included the space before the respective word in each case.

Five percent of trials were excluded due to either no first pass fixations on the sentence prior to word *n*−1 or tracker loss or blinks on first pass reading of word *n*–1 or the critical string. Five participants were replaced due to more than 15% of trials being excluded in this manner and one participant was replaced due to an error rate greater than 15% on the comprehension questions.

Results

Repeated measures analyses of variance (ANOVAs) were undertaken across the five spelling conditions, with participants *(F_1)* and items *(F_2)* as random variables. If there were significant main effects across both participants and items then simple effects were computed comparing the correctly spelled condition with each of the misspelled conditions. Paired samples t-tests were undertaken across both participants *(t_1)* and items *(t_2)* for comparisons between pairs of variables. The duration of the first fixation, gaze duration (the sum of fixations on a word before leaving it), and total time (the sum of all fixations within a word) were calculated for word *n*−1, the critical string and word *n*+1. For the critical string, landing positions, incoming saccade extent, launch site and frequency of refixations were analysed. The mean error rate on the comprehension questions was 5%.

Parafoveal-on-foveal effects. Table 1 shows the mean reading time measures on word *n*–1 and mean fixation durations prior to fixating the critical string. For word *n*–1 there were no significant effects of spelling on first fixation or gaze duration (F_s<1.2). There was no effect of the spelling of the critical string on the probability of refixating word *n*−1 (F_s<1). There was no significant effect of spelling on the duration of the fixation prior to the first fixation on the critical string for all of the data (F_s<1), for only those trials when the prior fixation was on word *n*−1, F_1(4, 176)=1.42, p=.231, *MSE*=1470; F_2<1, and for only those trials when the prior fixation was three or less characters away,[3] F_1(4,144)=1.27, *p*=.284, *MSE*=3556; F_2<1. These data provide no evidence of parafoveal-on-foveal effects.

Reading time measures. Table 1 shows the mean reading time measures on the critical string and word *n*+1. There was a significant effect of spelling on the reading time measures on the critical string for first fixation, F_1(4, 176)=5.28, *p* <.01, *MSE*=1803; F_2(4, 116)=4.43, *p*<.01, *MSE*=1727, gaze duration,[4] F_1(3.5, 152.2)=18.06, *p*<.01, *MSE*=12744; F_2(2.7, 79.6)=13.83, *p*<.01, *MSE*=17,027, and total time, F_1(3.2, 140.6)=33.21, *p*<.01, *MSE*=23,999; F_2(2.9, 84.3)=18.14, *p*<.01, *MSE*=37,272. For all of these measures the correctly spelled words were fixated for a significantly shorter time than the misspelled conditions (F_s#DXGT#4.6, p_s<.05). There were also significant differences between the high frequency and the illegal unpronounceable misspelled conditions on the critical string for first fixation duration, *t_1(44)*=2.02, *p*=.05; t_2(34)=2.23, p=.03, and gaze duration, t_1(44)=3.23, *p*<.01; t_2(34)=2.55, *p*= .02, which suggests that the words with more irregular misspellings were more difficult to recognise than the words with less irregular spellings.

There was no effect of spelling on the duration of the fixation after leaving the critical string (F_s<1). There was no significant effect of spelling on the first fixation on word *n*+1 (F_s<1) and although there was a significant effect of spelling on gaze duration on word *n*+1 across participants this was not significant across items,[5] F_1(4, 168)=3.44, *p*=.01, *MSE*=2309; F_2(4, 116)= 1.

93,p=.11, MSE=2571. However, spelling did significantly influence total reading times on word n+1, F_1(3.4, 151.1)=5.31, p<.01, MSE=5855; F_2(4, 116)=4.24, p<.01, MSE=5646. The correctly spelled condition produced shorter total reading times on word n+1 than each of the misspelled conditions (Fs#DXGT#8.7, ps<.01).

S

To summarise, there were no effects of the critical string on prior fixations. However, fixation durations were increasingly longer on the critical string, and for later measures on word n+1, the more irregular the misspelling. The misspelled words, especially the very irregular misspellings, were more difficult to process than the correctly spelled words. Previous studies have also shown that misspelled words (Inhoff & Topolski, 1994; Rayner, Pollatsek, & Binder, 1998b; Underwood, Bloomfield, & Clews, 1988), and words that are incorrect in the context of the sentence (Daneman, Reingold, & Davidson, 1995; Ehrlich & Rayner, 1981) produce longer fixation durations.

Landing positions. For the first fixation landing position analyses, the space before the critical string was classified as zero and the first letter of the string as one, etc. Table 2 shows the mean landing positions on the critical string. The mean first fixation landing positions on the critical string were 0.5 to 0.3 characters nearer the word beginning for the misspelled strings compared to the correctly spelled words. There was a significant effect of spelling on the mean first fixation landing position on the critical string, F_1(4, 176)=3.08, p=.02, MSE=0.75; F_2(4, 116)=2.91, p=.02, MSE=0.57. Compared to the correctly spelled condition, mean landing positions were significantly nearer the beginning of the word in the high frequency, F_1(1, 44)=7.46, p<.01, MSE=1.6; F_2(1, 29)=15.14, p<.01, MSE=0.69, low frequency, F_1(1, 44)=8.21, p<.01, MSE=1.01; F_2(1, 29)=4.97, p=.03, MSE=1.08, illegal pronounceable, F_1(1, 44)=14, p<.01, MSE=1.07; F_2(1, 29)=5.7, p=.02, MSE=1.54, and the illegal unpronounceable, F_1(1,44)=4.07, p=.05, MSE=1.3; F_2(1, 29)=6.1, p =.02, MSE=1.17 misspelled conditions. No other paired contrasts between the misspelled conditions were significant (ts<1.4). Figure 1 shows the distribution of landing positions for each condition. For all of the conditions, most fixations landed on the preferred viewing position (between the middle and the beginning of the word). The correctly spelled condition landing position distribution curve is shifted to the right of the misspelled condition curves. Clearly, readers processed the critical

[3] For Experiment 1, the analysis of saccades launched from three or less characters away was based on data from 37 participants due to failure to fixate these characters or excluded data.

[4] If a Mauchly test of sphericity was significant, the Greenhouse-Geisser Epsilon adjustment was used. Unless otherwise indicated, if the degrees of freedom do not correspond to the number of conditions and participants or items then the results were corrected for sphericity.

TABLE 1

Reading time measures for each condition in Experiments (Exp) 1 and 2. Mean first fixation duration (FF) and gaze duration (GD) on word n -1, word *n,* and word n+1. Total time (TT) on word *n* and word n+1. Fixation duration prior to fixating the critical string (prior fixation) for all the data (All), for saccades launched from word n-1 (Word n-1) and saccades launched from three or less characters from the beginnings of the critical string (=3). Fixation duration after leaving the critical string (Fixation n+1)

Exp	*Condition*	*Word n–1*		*Fixation n–1*			*Word n*			*Fixation n+1*	*Word n+1*		
		FF	*GD*	*All*	*Word n–1*	*≤ 3*	*FF*	*GD*	*TT*		*FF*	*GD*	*TT*
1	Correct	267	291	257	264	267	276	339	381	263	271	279	312
	High frequency misspelling	267	291	256	267	261	297	428	610	264	278	302	380
	Low frequency misspelling	260	281	253	255	245	301	478	632	269	285	314	366
	Illegal pronounceable misspelling	272	298	257	268	257	301	478	637	266	282	310	367
	Illegal unpronounceable misspelling	269	298	262	270	255	318	507	675	276	280	302	363
2	Correct	271	305	260	260	252	302	360	434	277	286	306	353
	Uninformative misspelling	261	292	256	261	257	340	518	785	280	292	321	405
	Informative misspelling	263	294	264	259	259	358	525	771	291	298	333	429

TABLE 2

Mean landing positions, incoming saccade extents, and launch sites for Experiments (Exp) 1 and 2. Standard deviations in parentheses

Exp	*Condition*	*Landing position*	*Saccade extent*	*Launch site*
1	Correct	4.4 (2.2)	9.1 (3.2)	4.7 (3.4)
	High frequency misspelling	4.0 (2.3)	8.9 (2.6)	4.9 (3.4)
	Low frequency misspelling	4.0 (2.1)	8.7 (2.6)	4.7 (3.1)
	Illegal pronounceable misspelling	3.9 (2.2)	8.8 (3.1)	4.9 (3.5)
	Illegal unpronounceable misspelling	4.1 (2.3)	8.8 (2.9)	4.7 (3.2)
2	Correct	3.5 (2.1)	8.4 (2.3)	4.9 (2.7)
	Uninformative misspelling	3.5 (2.1)	8.4 (2.5)	4.9 (2.9)
	Informative misspelling	3.6 (2.2)	8.3 (2.8)	4.8 (2.8)

string prior to direct fixation and misspellings produced landing positions nearer to the beginning of the critical string compared to when it was spelled correctly.

Incoming saccade extent and launch site. Table 2 shows the mean lengths of saccades into the critical string and mean launch sites. There was no significant effect of spelling on the length of the saccade into the critical string, $F_1(4, 176)=1.35$, $p=.253$, $MSE=1.06$; $F_2<1$. However, Table 2 shows that the mean saccade lengths into the critical string were longer for the correctly spelled condition than in any of the misspelled conditions. There was no effect of spelling on the position of the fixation prior to first fixating the critical string (Fs<1) and, in contrast to the means for saccade lengths, Table 2 shows no consistent pattern of differences in launch site between the correctly spelled condition and the misspelled conditions. The probability of skipping word n–1 before fixating the critical string (trials in which regressions were made from word n–1 were considered separately) was numerically greater for the misspelled conditions (high frequency: 0.17; low frequency: 0.11; illegal pronounceable: 0.14; illegal unpronounceable: 0.16) than for the correctly spelled condition (0.09). There was a significant main effect of spelling on the probability of skipping word n–1 $F_1(3.2, 141.8)=3.55$, $p=.01$, $MSE=202$; $F_2(4, 116)=3.26$, $p=.01$, $MSE=154$, and word n–1 was less likely to be skipped in the correctly spelled condition than the high frequency, $F_1(1, 44)= 10.5$, $p<.01$, $MSE=328$; $F_2(1,29)=11.93$, $p<.01$, $MSE=225$, and illegal unpronounceable, $F_1(1, 44)=9.07$, $p<.01$, $MSE=261$; F_2 $(1, 29)=4.83$, $p= .04$, $MSE=447$, misspelled conditions. There was no significant difference for the low frequency misspelling condition, $F_s<1.2$, and a significant difference across participants, $F_1(1, 44)=5.69$, $p=.02$, $MSE=172$, but not items, F_2 $(1,29)=2.33$, $p=.138$, $MSE=312$, for the illegal pronounceable misspelling condition. The

[5] In Experiment 1, the F1 analysis of first pass reading time on word n+1 was based on data from 43 participants due to skipping and excluded data.

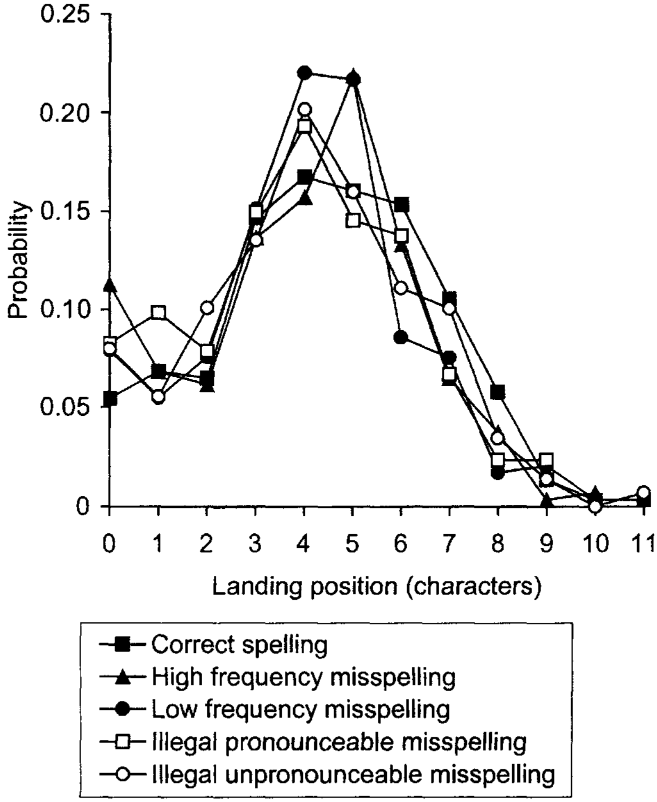

Figure 1. First fixation landing position distributions on the critical string for Experiment 1. Landing position zero is the space before the word and landing position one is the first letter of the word.

finding that spelling influenced the probability of skipping the previous word suggests that the misspellings may have attracted saccades from distant launch sites. Figure 2 shows the distribution of launch sites for each condition. Note that there tend to be slightly more saccades launched from seven or more characters from the critical string for the high frequency and illegal unpronounceable misspelled conditions compared to the correctly spelled condition.

Therefore, although there were no significant effects of incoming saccade extent or launch site, the results suggest that differences in both these factors may have contributed to the landing position effect. That is, nonsignificant differences in the mean incoming saccade extent and differences in the probability of skipping word *n*–1 may both have influenced first fixation positions on the critical string.

Frequency of refixations. Table 3 shows the probability of skipping or making one or more than one fixation on the critical string on first pass. There was a

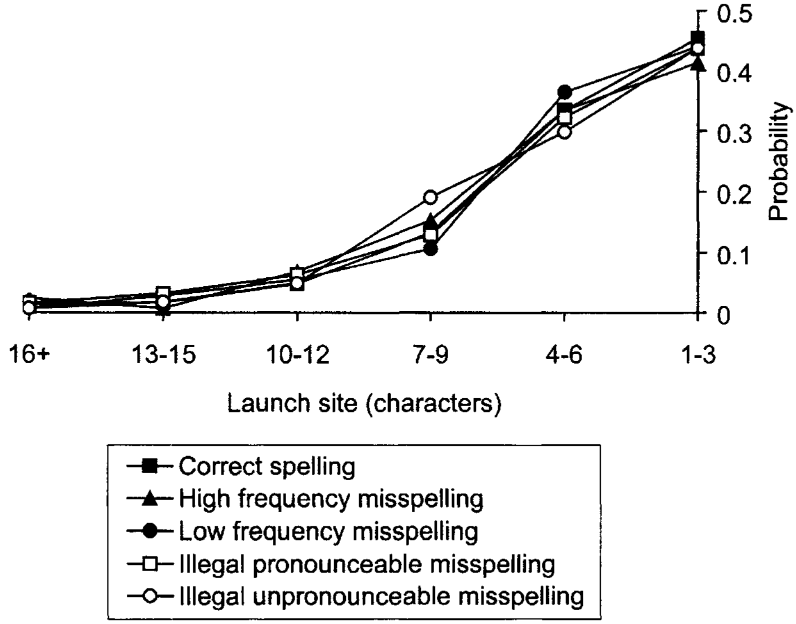

Figure 2. Launch site distributions for each condition for Experiment 1.

significant effect of spelling on the probability of refixating the critical string on first pass, $F_1(4,176)=8.21, p<.01, MSE=307$; $F_2(4, 116)=6.73, p< .01, MSE=287$. The correctly spelled condition was significantly less likely to be refixated than any of the four misspelled conditions (F_s#DXGT#4.6, $ps<.05$). Similar to the reading time measures, these results suggest that the misspelled words were more difficult to process than the correctly spelled words because they produced more first pass refixations on the critical string.

Table 3 shows the probability of making a refixation to the left of the initial fixation on the critical string for those trials in which multiple first pass fixations occurred on the critical word. There was a significant effect of spelling on the probability of making a refixation to the left of the initial fixation,[6] $F_1(4,104)$– 2. 77, $p=.03$, $MSE=1003$; $F_2(4, 88)=5.91$, $p<.01$, $MSE=666$. Compared to the correctly spelled condition, refixations were more likely to be to the left of the initial fixation in the low frequency, $F_1(1,26)=5.81$, $p=.02$, $MSE=2308$; $F_2(1, 22)=7.6$, $p=.01$, $MSE=1434$, and the illegal unpronounceable, F1(1, 26) =8.03, $p<.01$, $MSE=2134$; $F_2(1, 22)= 18.02$, $p<.01$, $MSE=1255$, misspelled conditions. For the illegal pronounceable misspelled condition the effect was significant across participants but not items, $F_1(1, 26)=4.27$, $p=.05$, $MSE= 1412$; $F_2(1, 22)=2.82$, $p=.107$, $MSE=1490$. There was no significant difference in the probability of refixating to the left of the initial fixation for the correct and high frequency misspelled words, $F_1(1, 26)=1.27$, $p=.27$, $MSE= 1848$; $F_2<1$. The results suggest

TABLE 3

Frequency of number of first pass fixations on the critical string for Experiments (Exp) 1 and 2. Frequency of first refixating to the left of the initial fixation on the critical string for multiple first pass fixation cases

		Fixation frequency			
Exp	*Condition*	*Skip*	*One*	*≥ Two*	*Leftward refixation*
1	Correct	.03	.72	.25	.34
	High frequency misspelling	.02	.64	.34	.34
	Low frequency misspelling	.02	.56	.42	.49
	Illegal pronounceable misspelling	.01	.58	.41	.46
	Illegal unpronounceable misspelling	.02	.55	.43	.56
2	Correct	.03	.74	.23	.16
	Uninformative misspelling	.02	.60	.38	.29
	Informative misspelling	.01	.59	.40	.39

that the more irregular misspelled items were more likely to produce regressive refixations than the correctly spelled condition.

Discussion

Both the position of the first fixation and the direction of refixations were influenced by the presence of misspellings. Therefore the results show that nonfoveal processing at least at the level of orthographic familiarity can influence where words are fixated. It is also interesting that the word before the critical word was more likely to be skipped if the critical word was misspelled. These results suggest that misspellings can attract saccades from distant launch sites. However, note that in contrast to these results, Pynte et al. (2004) found that irregular misspellings reduce the probability of skipping the previous word.

The effects of the misspellings on landing positions clearly show that they must have been processed during the previous fixation. However, importantly, there was no effect of the misspellings on prior fixation durations or refixation probabilities. That is, there were no parafoveal-on-foveal effects.

First fixation landing positions were also significantly nearer the word beginning for the high frequency misspelled condition compared to the correctly spelled condition. There were no differences between initial trigram frequencies

[6] In Experiment 1, for the analyses of the probability of refixating to the left, only participants and items that produced refixations in all of the conditions were included. Consequently the F_1 analysis was based on 27 participants and the F_2 analysis was based on 23 items.

in the correct and high frequency misspelling conditions and, as mentioned above, the position in the word at which the high frequency misspelled words became illegal was uncontrolled. Therefore letter sequences up to at least the first four letters of the high frequency misspelled strings had to be processed in order for the word initial letter sequence to be detected as irregular. Importantly, although the string would have to be processed up to at least the fourth letter, shorter infrequent or illegal letter sequences may have been detected (for example *quc* in *equcation*). Nevertheless, this result is surprising because Beauvillain and Doré (1998) showed that the letter sequence frequency of the second and third letters of a letter string did not influence fixation positions. The high frequency misspelling result suggests that letter sequences that are positioned further into the word than those tested by Beauvillain and Doré can influence fixation positions. The high frequency misspelling result is even more surprising because it appears to be at least partly determined by saccades being attracted from distant launch sites. That is, word $n-1$ is more likely to be skipped when there is a high frequency misspelling than when the critical string is spelled correctly. Note that saccades launched from distant launch sites tend to land nearer the beginning of words (McConkie et al., 1988), hence the numerically larger number of fixations on the space before the word for the high frequency misspellings (see Figure 1). These results imply that processing beyond the level of the orthographic familiarity of the word initial trigram can be undertaken on text presented more than one word from fixation, which is beyond or certainly towards the far edge of the region of text from which letter information can be processed (McConkie & Rayner, 1975; Rayner, 1975, 1998).

There are three possible explanations for the high frequency misspelling landing position result. First, the high frequency misspelled strings, or letter sequences within these strings, might have been identified as illegal. Second, the stimuli were chosen primarily on the basis of token, rather than type, frequency. Although there were no significant differences in type frequency, the high frequency misspelling condition had initial trigrams with a greater range and variation (range: 1–679, SD =115) than the correctly spelled condition (range: 8–421, SD=88) and therefore differences in type frequency (i.e., informativeness) might have influenced fixation positions. Third, given that the high frequency misspelling result was a surprising finding in relation to previous research, it may be spurious.

EXPERIMENT 2

Experiment 2 was designed to distinguish between the three alternative explanations for the high frequency misspelling result in Experiment 1. In order to do this, we constructed a set of materials in which a critical string was spelled

correctly with an uninformative initial trigram, misspelled with an uninformative initial trigram, or misspelled with an informative initial trigram. Each of the three explanations for the results of Experiment 1 described above generate different predictions for Experiment 2. First, if fixation positions are influenced by any kind of misspelling then landing positions should be nearer the beginning of the critical string in both of the misspelled conditions, compared to the correctly spelled condition. Second, if lexical preprocessing of a nonfixated word influences where words are first fixated then the informative ness of word beginnings might affect landing positions. Consequently, we might expect fixation positions to be nearer the beginning of the informative misspelled strings compared to the uninformative misspelled and correctly spelled strings. Third, if processing beyond the level of the orthographic familiarity of word initial trigrams does not influence where words are first fixated then there should be no difference in landing positions for letter strings with equally familiar initial letter sequences. That is to say, there should be no effect of spelling on landing positions.

Previous studies of the effects of orthographic regularity on landing positions (Beauvillain & Doré, 1998; Beauvillain et al, 1996; Doré & Beauvillain, 1997; Everatt & Underwood, 1992; Hyönä, 1995; Hyönä et al., 1989; Radach et al., 2004; Underwood et al., 1987, 1990; Vonk et al., 2000) have confounded the variables of orthographic familiarity and informativeness. Kennedy (2000) did carefully manipulate these, and other, variables in artificial task experiments but found no effects of letter sequence frequency on landing positions. Experiment 2 provides a test of the hypothesis that informativeness of the word initial trigam, independent of the orthographic familiarity of the word initial trigram, influences fixation positions.

Method

Participants. Twenty-four native English speakers at the University of Durham were paid to participate in the experiment. The participants all had normal or corrected to normal vision and were naïve in relation to the purpose of the experiment.

Materials. Word frequencies and *n*-gram frequencies were calculated using the CELEX English word form corpus (Baayen et al., 1995). All of the critical words were eight or nine characters long (M=8.5, SD=0.5) and the mean word frequency in counts per million was 17 (SD=29.5). There were three conditions that were manipulated within participants and items. The critical words were spelled correctly or the second letter was misspelled to create either an uninformative or informative initial trigram.

Position specific *n*-gram frequencies were calculated in counts per 17.9 million. The initial trigram token frequencies tended to be higher in the uninformative (M=8696, SD=10,545) and informative (M=6879, SD= 10,974) misspelled conditions compared to the correctly spelled condition (M= 4325, SD=5262), although these differences were not significant (ts<−1.8, *ps #DXGT#*. 9). The uninformative misspelled condition also tended to have higher type frequency initial trigrams (M=47, SD=26) compared to the correctly spelled condition (*M*= 32, *SD*=29) although this was not significant, $t(23)=-1.72$, $p= .1$. Importantly, the informative misspelled condition had significantly lower type frequency initial trigrams (*M*=6, *SD*=3) than both the correctly spelled condition, $t(23)=4.45$, $p<.01$ and the uninformative misspelled condition, $t(23) =7.51$, $p<.01$. The type and token frequencies of the initial bigrams followed a similar pattern. The type and token monogram frequencies of the misspelled second letter were significantly higher in the misspelled condition compared to the correctly spelled condition (t_s#DXGT#2). Importantly, the initial trigrams of the misspelled conditions were not more orthographically unfamiliar than the correctly spelled condition.

The 24 critical words were embedded in identical sentence frames for each condition. Each of the sentences was no longer than one line of text (78 characters) and the critical word appeared approximately in the middle of the sentence. The words before and after the critical word were either five or six letters long and had medium to high frequencies. Most of the sentences included context relevant to the critical word at the beginning of the sentence. See the Appendix for example experimental sentences and critical words.

Three lists of 96 items were constructed and eight participants were randomly allocated to each list. Each list included 24 experimental items of which 8 items were from each of the three misspelling conditions. The conditions were rotated following a Latin square design. There were 16 misspelled filler items with misspellings in a variety of word lengths and in a variety of positions within the word and the sentence. There were also 56 filler items that were spelled correctly. Therefore, of the 96 items, 32 contained a misspelling. Thirty-two of the sentences were followed by a comprehension question to ensure that participants concentrated on understanding the sentences. The sentences were presented in a fixed random order with six filler sentences at the beginning.

Procedure. The experimental procedure was the same as in Experiment 1. The entire experiment lasted approximately 30 min and participants were given one break.

Analyses. The analyses were the same as in Experiment 1. Of the trials, 1.4% were excluded due to either no first pass fixations on the sentence prior to word n–1 or tracker loss or blinks on first pass reading of word n–1 or the critical string.

Results

The results were analysed in the same manner as for Experiment 1. The mean error rate on the comprehension questions was 2.3%.

Parafoveal-on-foveal effects. Table 1 shows the mean fixation durations prior to fixating the critical string. There were no significant effects of spelling on first fixation, gaze duration, or total time for word n–1 (Fs< 1.1). There was also no difference in the probability of refixating word n–1 (Fs<1). There were no significant effects of spelling on the duration of the fixation prior to first fixating the critical word for all of the data, for saccades launched from word n–1 and for saccades launched from three or less characters from the critical word (F_s<1). Once again, the results show no evidence of parafoveal-on-foveal effects.

Reading time measures. Table 1 shows the mean reading time measures on the critical string and word n+l. For the critical string there were significant effects of spelling on first fixation, $F_1(2,46)=8.75$, $p<.01$, $MSE=19{,}235$; $F_2(2, 46)=6.59$, $p<.01$, $MSE=19{,}378$, gaze duration, $F_1(2, 46)=19.56$, $p< .01$, $MSE=207, 245$; F_2 $(2, 46)=18.96$, $p<.01$, $MSE=205, 442$, and total time, $F_1(2, 46)=54.42$, $p<.01$, $MSE=17{,}476$; $F_2(2, 46)=24.98$, $p<.01$, $MSE= 38{,}297$. For all measures reading times were longer on the two misspelled conditions compared to the correctly spelled condition (*F*s#DXGT#6.5, *ps*<.05). There were no significant differences in reading time between the uninformative and informative misspelling conditions (*ts*<1.6, *ps*#DXGT#.1).

Table 1 shows the mean reading times after leaving the critical string. There was no effect of spelling on the duration of the fixation after leaving the critical string or on first fixation durations on word n+1 (*F*s<1.2). There were also no effects of spelling on gaze duration on word n+1, $F_1(2, 46)=2.31$, $p=.11$, MSE =7719; F_2<1, or across items for total time (F_2<1) although there was a significant effect of spelling across participants for total time on word n+1, $F_1(2, 46)=3.8$, $p<.05$, $MSE=28{,}584$.

Reading times were longer on the misspelled words than the correctly spelled words and there were no reliable spillover effects. These data indicate that misspelled words were more difficult to process than the correctly spelled words.

Landing position. Table 2 shows the mean first fixation positions on the critical string in each of the conditions. There were no significant effects of spelling on landing position (*F*s<1). Figure 3 shows the distribution of landing positions for each of the conditions; note that most fixations landed on the preferred viewing position.

Incoming saccade extent and launch site. Table 2 shows the mean saccade extents and launch sites for each condition. There were no effects of spelling on the launch site or incoming saccade extent prior to fixating the critical string (*F*s <1). There was no effect of spelling on the probability of skipping word n–1 before fixating the critical string (trials in which regressions were made from word n−1 were considered separately) (Fs<1).

Frequency of refixations. Table 3 shows the probability of skipping or making one or more than one fixation on the critical string on first pass. There was a significant effect of spelling on the probability of refixating the critical string on first pass, $F_1(2, 46)=9.7$, $p<.01$, $MSE=211$; $F_2(2, 46)=6.69$, $p<.01$, $MSE=316$. The correctly spelled condition was significantly less likely to be refixated on first pass than the uninformative misspelling condition, $F_1(1, 23)= 9.88$, $p<.01$, $MSE=508$; $F_2(1, 23)=11.51$, $p<.01$, $MSE=471$, or the informative misspelling condition, $F_1(l, 23)=16$, $p<.01$, $MSE=443$; $F_2(l, 23) =10.65$, $p<.01$, $MSE=672$. Similar to the reading time measures, these results suggest that the misspelled

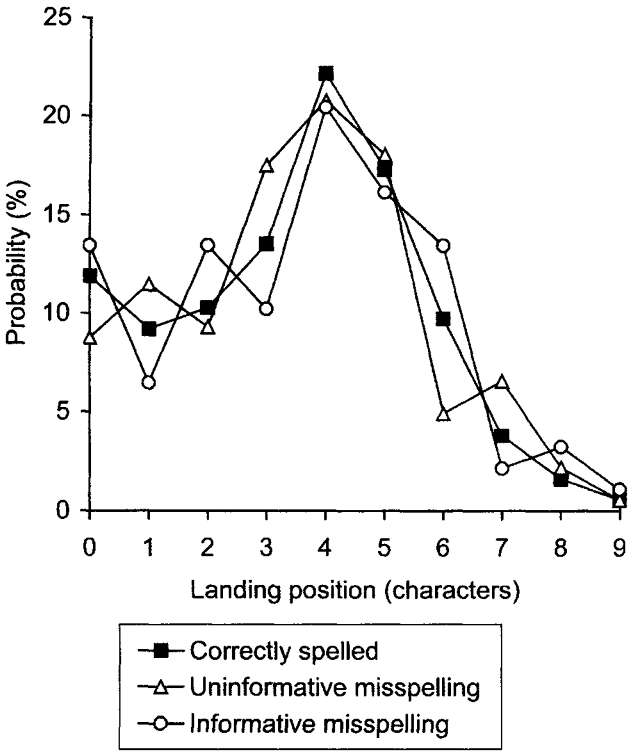

Figure 3. First fixation landing position distributions on the critical string for Experiment 2. Landing position zero is the space before the word and landing position one is the first letter of the word.

words were more difficult to process because they produced more first pass refixations.

There was also a significant effect of spelling on the probability of first refixating to the left of the initial fixation position,[7] $F_1(2, 38)=3.33$, $p<.05$, $MSE=2835$; $F_2(2, 36)=4.95$, $p<.05$, $MSE=3202$. Refixations to the left of the initial fixation position were significantly more likely in the informative misspelled condition compared to the correctly spelled condition, $F_1(1, 19)=8.71$, $p<.05$, $MSE=10{,}945$; $F_2(1, 18)=7.85$, $p<.05$, $MSE=11{,}299$. Leftward refixations also tended to be more likely in the uninformative misspelled condition compared to the correctly spelled condition but this effect was significant across items, $F_2(1, 18)=9.1$, $p<.05$, $MSE=7533$, but not participants, $F_1(1, 19)=2.53$, $p=.129$, $MSE=4835$. There was no difference in the probability of refixating to the left between the uninformative and informative conditions (ts< 1.2). The results again show that misspellings increase the probability of making regressive refixations.

Discussion

Experiment 2 clearly showed that neither the presence of a misspelling nor the informativeness of word initial letters reliably influenced where words were first fixated. The results do not support the suggestion that lexical factors influence first fixation positions on words. There is a clear inconsistency between the high frequency misspelling effect in Experiment 1 and the results of Experiment 2. It must be concluded that influences on landing positions beyond the level of the orthographic familiarity of the initial trigrams are, at best, unreliable. However, the results do show that misspellings influence the directions of refixations.

Similar to Experiment 1, the results also provide no support for the notion that nonfixated text influences fixation durations. There was also no difference in reading times on words that were misspelled with an informative or uninformative initial trigram. These results suggest that the time to process a misspelled word is not influenced by the number of possible candidates that are generated by the initial trigram.

GENERAL DISCUSSION

The aim of this study was to investigate whether orthographic familiarity and informativeness influence landing positions in the reading of English sentences. In accordance with previous evidence, our results clearly show that most fixations land left of the word centre on the preferred viewing position (DunnRankin, 1978; McConkie et al., 1988; Rayner, 1979) and that more distant launch sites are associated with landing positions nearer the beginning of words (Hyönä, 1995; McConkie, Kerr, & Dyre, 1994; McConkie et al., 1988; Radach & Kempe, 1993; Radach & McConkie, 1998; Rayner et al., 1996). The results of the present study provide no reliable support for the influence of lexical processing on fixation positions. The most important finding is that first fixation landing positions were nearer the beginning of words with unfamiliar orthography. To our knowledge, this is the first study to find an effect of orthographic familiarity on landing positions for English language sentences. The results are supported by other recent experiments that have shown that orthographic processing influences where words are first fixated in the reading of correctly spelled English sentences (White & Liversedge, 2003). These results are particularly striking since current models of eye movements in reading make no attempt to account for the influence of orthography on first fixation landing positions (O'Regan, 1990; Reichle et al, 1998, 1999, in press; Reilly & O'Regan,

[7] In Experiment 2, for the analyses of the probability of refixating to the left, only participants and items that produced refixations in all of the conditions were included. Consequently the F_1 analysis was based on 20 participants and the F_2 analysis was based on 19 items.

1998). Our experiments are a clear demonstration that these effects do occur and any complete model of oculomotor control in reading must account for them.

The results of Experiment 1 support a number of studies in languages other than English that have shown that word initial orthographic regularity influences where words are first fixated (Beauvillain & Doré, 1998; Beauvillain et al., 1996; Doré & Beauvillain, 1997; Hyönä, 1995; Radach et al., 2004; Vonk et al., 2000). However, in support of Rayner and Morris (1992), and against the claims of Underwood and colleagues (Everatt & Underwood, 1992; Hyönä et al., 1989; Underwood et al., 1987, 1990), the results provide no reliable support for the suggestion that informativeness, or the number of potential lexical candidates, can influence where words are first fixated.

Experiment 1 provided a strong test of the possibility that orthographic familiarity produces parafoveal-on-foveal effects and Experiment 2 provided a strong test of the possibility that informativeness produces parafoveal-on-foveal effects. However, both experiments showed no evidence of such effects. Previous studies using artificial tasks (Kennedy, 1998, 2000; Murray, 1998; Murray & Rowan, 1998), and sentence reading studies (Inhoff et al., 2000a, 2000b; Kennedy et al., 2002; Pynte et al., 2004; Underwood et al., 2000) have suggested that the characteristics of a word can influence fixation times on the previous word. However, as also reported in Rayner et al. (in press), in both the experiments reported here, there were no effects of the spelling of the critical string on the fixation duration prior to fixating the critical string, even when the prior fixation was three or less characters from the beginning of the critical string. Furthermore, the landing position effects in Experiment 1 show that the initial letter sequences were processed before the critical strings were fixated and yet there was no effect of the initial letter sequence on the prior fixation duration. Importantly, the absence of such effects in the present study at least suggests that there is no strong consistent parafoveal-on-foveal processing in normal sentence reading. That is, the results provide no support for the hypothesis that words are processed in parallel such that the sublexical and lexical characteristics of a nonfixated word can influence fixation durations or refixation probabilities on the previous word. Clearly, these data indicate that nonfoveal processing influences where the eyes move but does not influence when the eyes move to fixate a nonfoveal string.

Three types of possible explanations have been proposed to account for the effects of orthographic familiarity on fixation positions found in Experiment 1. First, Hyönä (1993) and Beauvillain and Doré (1998) suggested that irregular letter sequences "pop out" and attract saccades towards them. Second, Findlay and Walker (1999) proposed that visually unfamiliar (intrinsically salient) letter sequences could influence saccade computation within a salience map. The third possible explanation is that nonfoveal processing difficulty reduces the perceptual span and consequently shortens saccades (Hyönä, 1995; Hyönä & Pollatsek, 1998, 2000). In order for the processing difficulty hypothesis to explain the results of Experiment 2, it must be assumed that the uninformative and informative misspellings did not induce sufficient nonfoveal processing difficulty to influence saccades.

A further important question is whether the effects of orthography on landing positions form part of a strategy that is efficient for word recognition, or whether they are simply due to mandatory processes that function in all types of visual processing such as scene viewing or visual search. Legge, Klitz, and Tjan (1997) suggested that an ideal reader would use all available information to determine the length of a saccade in order to minimise uncertainty for word recognition. However, if this was the case then it is not clear why a number of studies have failed to find effects of within word characteristics on first fixation positions. The landing position differences in Experiment 1 were very small and one might reasonably assume that such differences would be much greater if landing positions were central to visual word recognition. Also, landing position effects such as increasing the probability of landing on the space before the word are not necessarily optimal for word recognition (McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; Vitu, O'Regan, & Mittau, 1990). If the influence of orthography on landing positions is not a strategy for efficient word recognition then an alternative possibility is that the effects might simply result from standard visual processes, as suggested by Findlay and Walker (1999).

The current experiments also yielded interesting results regarding the nature of refixations. O'Regan (1990) argued that the locations of refixations are determined by the position of the first fixation on a word in relation to the word length. However, in both Experiments 1 and 2 there were more refixations to the left for misspelled critical strings. In support of previous studies (Hyönä, 1995; Hyönä & Pollatsek, 1998; Pynte, 1996, 2000; Pynte, Kennedy, & Murray, 1991) the results suggest that the characteristics of a word influence the location of refixations. Once again, current models of eye movements in reading make no attempt to explain these effects.

In conclusion the results show that the orthographic familiarity, but not informativeness, of nonfixated words influences where words are first fixated and refixated. However, importantly the results provide no evidence for parafoveal-on-foveal effects.

PrEview proof published online September 2003

REFERENCES

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*(1/2), 285–311.

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390.

Beauvillain, C., & Doré, K. (1998). Orthographic codes are used in integrating information from the parafovea by the saccadic computation system. *Vision Research, 38,* 115–123.

Beauvillain, C., Doré, K., & Baudouin, V. (1996). The "centre of gravity" of words: Evidence for an effect of word-initial letters. *Vision Research, 36,* 589–603.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–148). Oxford, UK: Elsevier.

Daneman, M., Reingold, E.M., & Davidson, M. (1995). Time course of phonological activation during reading: Evidence form eye fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 884–898.

Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes, 14,* 393–21.

Doré, K., & Beauvillain, C. (1997). Latency dependence of word-initial letter integration by the saccadic system. *Perception and Psychophysics, 59,* 523–533.

Dunn-Rankin, P. (1978). The visual characteristics of words. *Scientific American, 238,* 122–130.

Ehrlich, S.F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behaviour, 20,* 641–655.

Everatt, J., & Underwood, G. (1992). Parafoveal guidance and priming effects during reading: A special case of the mind being ahead of the eyes. *Consciousness and Cognition, 1,* 186–197.

Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioural and Brain Sciences, 22,* 661–721.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language, 33,* 39–68.

Hill, R.L., & Murray, W.S. (2000). Commas and spaces: Effects of punctuation on eye movements and sentence parsing. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 565–589). Oxford, UK: Elsevier.

Hyönä, J. (1993). *Eye movements during reading and discourse processing* [Psychological Research Rep. No. 65]. Turku, Finland: University of Turku.

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

Hyönä J.,Niemi, P., & Underwood, G. (1989). Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 142–152.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Hyönä, J., & Pollatsek, A. (2000). Processing of Finnish compound words in reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 65–87). Oxford, UK: Elsevier.

Inhoff, A.W. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 15* 544–555.

Inhoff, A.W. (1990). Integrating information across eye fixations in reading: The role of letter and word units. *Acta Psychologica, 73,* 281–297.

Inhoff, A.W., Briihl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming and delayed naming tasks. *Memory and Cognition, 24,* 466–476.

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000a). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 221–246). Oxford, UK: Elsevier.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40,* 431–439.

Inhoff, A.W., Starr, M., & Shindler, K.*L.* (2000b). Is the processing of words during eye fixations in reading strictly serial? *Perception and Psychophysics, 62,* 1474–1484.

Inhoff, A, W., & Topolski, R. (1994). Use of phonological codes during eye fixations in reading and in on-line and delayed naming tasks. *Journal of Memory and Language, 33,* 689–713.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–179). Oxford, UK: Elsevier.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Lavigne, F., Vitu, F., & d'Ydewalle, G. (2000). The influence of semantic context on initial eye landing sites in words. *Acta Psychologica, 104,* 191–214.

Legge, G.E., Klitz, T.S., & Tjan, B.S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review, 104,* 524–553.

Liversedge, S.P., Paterson, K.B., & Clayes, E.L. (2002). The influence of *only* on syntactic processing of "long" relative clause sentences. *Quarterly Journal of Experimental Psychology, 55A,* 225–240.

Liversedge, S.P., & Underwood, G. (1998). Foveal processing load and landing position effects in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 201–221). Oxford, UK: Elsevier.

McConkie, G.W. (1979). On the role and control of eye movements in reading. In P.A.Kolers, M. E.Wrolstad, & H.Bouma (Eds.), *Processing of visible language* (pp. 37–48). New York: Plenum.

McConkie, G.W., Kerr, P.W., & Dyre, B.P. (1994). What are "normal" eye movements during reading: Toward a mathematical description. In J.Ygge & G.Lennerstrand (Eds.), *Eye movements in reading* (pp. 315–327). Oxford, UK: Elsevier.

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

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

McConkie, G.W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.

McConkie, G.W., Underwood, N.R., Zola, D., & Wolverton, G.S. (1985). Some temporal characteristics of processing during reading. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 168–186.

Morris, R.K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 268–281.

Murray, W.S. (1998). Parafoveal pragmatics. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 181–200). Oxford, UK: Elsevier.

Murray, W.S., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research, 27,* 1–22.

O'Regan, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics, 25,* 501–509.

O'Regan, K. (1980). The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception and Psychophysics, 28,* 112–117.

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

Paterson, K.B., Liversedge, S.P., & Underwood, G. (1999). The influence of focus operators on syntactic processing of *short* relative clause sentences. *Quarterly Journal of Experimental Psychology, 52A,* 717–131.

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 817–833.

Pynte, J. (1996). Lexical control of within-word eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 958–969.

Pynte, J. (2000). Words likely to generate many lexical candidates are granted an advantage: Evidence from within-word eye movements. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 47–64). Oxford, UK: Elsevier.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1/2), 178–202.

Pynte, J., Kennedy, A., & Murray, W.S. (1991). Within-word inspection strategies in continuous reading: Time course of perceptual, lexical and contextual processes. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 458–470.

Radach, R., Inhoff, A., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology, 16*(1/2), 27–51.

Radach, R., & Kempe, V. (1993). An individual analysis of initial fixation positions in reading. In G. d'Ydewalle & J.Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 213–226). Amsterdam: Elsevier.

Radach, R., Krummenacher, J., Heller, D., & Hofmeister, J. (1995). Individual eye movement patterns in word recognition: Perceptual and linguistic factors. In J.M.Findlay, R.Walker, & R. W.Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 421–432). Amsterdam: Elsevier.

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

Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica, 39,* 271–282.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–22.

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41,* 943–954.

Rayner, K., Carlson, A., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behaviour, 22,* 358–374.

Rayner, K., Fischer, M.H., & Pollatsek, A. (1998a). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.

Rayner, K., Kambe, G., & Duffy, S.A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology, 53A,* 1061–1080.

Rayner, K., & Morris, R.K. (1992). Eye movement control in reading: Evidence against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 163–172.

Rayner, K., Pollatsek, A., & Binder, K.S. (1998b). Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning Memory and Cognition, 24,* 476–497.

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

Rayner, K., Well, A.D., Pollatsek, A., & Bertera, J.H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31,* 537–550.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J.Hyönä, J.R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z reader model. *Vision Research, 39,* 4403–4411.

Reichle, E.D., Rayner, K., & Pollatsek, A. (in press). The E-Z reader model of eye movement control in reading: Comparisons to other models. *Behavioural and Brain Sciences.*

Reilly, R.G., & O'Regan, J.K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research, 38,* 303–317.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 247–268). Oxford, UK: Elsevier.

Underwood, G., Bloomfield, R., & Clews, S. (1988). Information influences the pattern of eye fixations during sentence comprehension. *Perception, 17,* 267–278.

Underwood, G., Clews, S., & Everatt, J. (1990). How do readers know where to look next? Local information distributions influence eye fixations. *Quarterly Journal of Experimental Psychology, 42A,* 39–65.

Underwood, G., Hyönä, J., & Niemi, P. (1987). Scanning patterns on individual words during the comprehension of sentences. In J.K.O'Regan & A.Lévy-Schoen (Eds.), *Eye movements from physiology to cognition* (pp. 467–477). Amsterdam: Elsevier.

Vitu, F., O'Regan, J.K., Inhoff, A.W., & Topolski, R. (1995). Mindless reading: Eye movement characteristics are similar in scanning letter strings and reading texts. *Perception and Psychophysics, 57,* 352–364.

Vitu, F., O'Regan, J.K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous texts. *Perception and Psychophysics, 47,* 583–600.

Vonk, W., Radach, R., & van Rijn, H. (2000). Eye guidance and the saliency of word beginnings in reading text. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 269–299). Oxford, UK: Elsevier.

White, S.J., & Liversedge, S.P. (2003). *Orthographic regularity influences the eyes' landing positions in reading.* Manuscript submitted for publication.

Zola, D. (1984). Redundancy and word perception during reading. *Perception & Psychophysics, 36,* 277–284.

APPENDIX

Examples of experimental sentence frames and critical strings for Experiments 1 and 2. A full set of materials is available from the first author on request.

Experiment 1

The slashes denote the correctly spelled, high frequency misspelled, low frequency misspelled, illegal pronounceable misspelled, and illegal unpronounceable misspelled conditions respectively.

> Farmers complained when local agricultural/acricultural/aoricultural/akricultural/ngricultural ground was contaminated. The gallery presented great exhibitions/ethibitions/ephibitions/ebhibitions/dxhibitions during the school holidays. The scientist worked in the large laboratory/liboratory/luboratory/lyboratory/lwboratory every day of the week.

Experiment 2

The slashes denote the correctly spelled, uninformative misspelled, and the informative misspelled conditions respectively.

Pip hated stairs and he even used the short escalator/encalator/eacalator down to the first floor. The builder gave a cheap estimate/entimate/eitimate before he realised the tiles were expensive. The musicians enjoyed playing in the small orchestra/occhestra/onchestra every Monday night.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 79–103

Word skipping in reading: On the interplay of linguistic and visual factors

Denis Drieghe

Ghent University, Belgium

Marc Brysbaert

Royal Holloway, University of London, UK

Timothy Desmet and Constantijn De Baecke

Ghent University, Belgium

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 predict ability were substantial. These findings are interpreted as evidence for the hypothesis that visual and language-related factors independently affect word skipping.

One of the current controversies in research on eye movements in readingconcerns the kind of information extracted from parafoveal vision and the waysin which this information influences subsequent eye movements (Starr &Rayner, 2001). A large body of research indicates that in reading information isextracted from the word next to the currently fixated word. This parafovealpreview 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 eyesland 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, isthe extent to which the parafoveal information determines the length of thesubsequent forward saccade out of the fixated word. In the present paper, wewill deal with the probability of fixating the next parafoveal word only, althougha very similar discussion exists about the landing position within the parafovealword.

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 socalled 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 look 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

---

Correspondence should be addressed to D.Drieghe, Dept. of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. Email: denis.drieghe@Ugent.beDenis Drieghe and Timothy Desmet are research assistants of the Fund for Scientific Research (Flanders, Belgium).Due to a software error which resulted in an incorrect calculation of the skipping data, this research has previously been presented with slightly different data (Drieghe & Brysbaert, 2001).We thank Keith Rayner, Françoise Vitu, Sarah White, and an anonymous reviewer for the many helpful comments on earlier drafts of this paper.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000141

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 (Gautier, O'Regan, & Le Gargasson, 2000; Vitu, 1991). 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 invisible boundary at the beginning of the word (e.g., Blanchard et al., 1989; Rayner et al., 1982).

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 (Altaribba, 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 et al., 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 paper, they are illustrated in Figure 1 (taken from Brysbaert & Vitu, 1998).

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 words, 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. 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. called this effect, by analogy to the previously defined optimal viewing position (e.g., O'Regan & Jacobs, 1992; O'Regan, Lévy-Schoen, Pynte, & Brugaillère, 1984), the Extended Optimal Viewing Position (EOVP) effect.

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 (1) the length of the word blobs[1] and the distance of the word

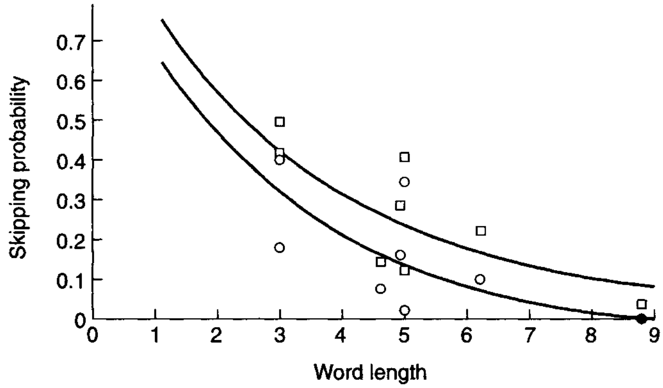

Figure 1. Skipping rate as a function of word length and contextual constraint (circle=predictable conditions; square=neutral conditions). Fitted curve based on nonlinear regression with exp(word length) and contextual constraint as predictors. The upper curve represents the best fit for the predictable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

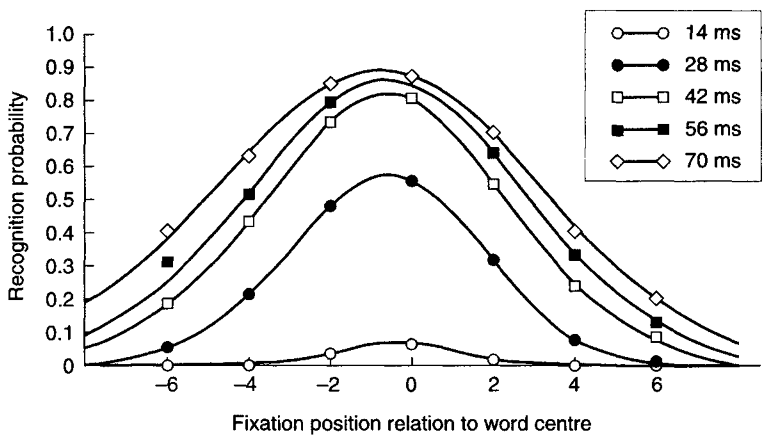

Figure 2. Probability of word recognition as a function of presentation duration and word position relative to fixation location. Empirical data and best fitting Gaussian distributions (reprinted from Brysbaert et al., 1996).

blobs from the fixation location, and (2) 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 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 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/Interaction 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

[1] Brysbaert and Vitu (1998) use this term for the parafoveal word to stress the limited visibility of the words in parafoveal vision.

programming time.[2] 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. 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.[3] 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, 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 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.

[2]This idea was originally formulated by Morrison (1984) and also features in the ideas of Henderson and Ferreira (1990).

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 two-letter word was strongly expected on the basis of the preceding sentence context, and we presented either this two-letter word, or an unexpected but still acceptable two-letter word, or an unexpected but acceptable four-letter word. We also constructed a second set of sentences in which a four-letter word was expected from the context and we presented this word, an intruder of the same length, or an intruder of two letters.

[3] 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).

The combination of these six 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 two- or four-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 two-letter words when a two-letter word was expected vs. when a four-letter word was expected, and by comparing the skipping rates for unexpected four-letter words when a four-letter word was expected vs. when a two-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.[4] 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 nonseparable verb *(vallen),* which in itself has another meaning (i.e., *to fall*). Because the meaning of separable verbs can differ from that of the nonseparable 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*[5] *de punker op."* (literally: *Because of his orange hair fell the punk up* [in reality: *Because of his orange hair the punk attracted a lot of attention*]).

The particles of separable Dutch verbs were used in our experiment as the target words[6] and the nonseparable 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 maximising 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 two-letter predictable word is shown in Table 1, one with a four-letter predictable word is shown in Table 2.

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 movementswere recorded from the right eye only. A high-speed video camera was used forrecording. It was positioned underneath the monitored eye and held in place byhead-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.[7] Half of the target sentences were built from a strongly expected two-letter word; half from a strongly expected four-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

---

[4]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.

[5] *viel* [fell] is the simple past singular tense of the verb *vallen* [to fall].

[6]Keep in mind that the term "target word" refers to the two- or four-letter particle parts of the separable verbs.

TABLE 1

Example sentence for a two-letter predictable word

Preceding context:	
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	
Continuations:	
2-letter word predictable	*op* volgens de instructies van de vroedvrouw.
2-letter word neutral	*na* volgens de instructies van de vroedvrouw.
4-letter word neutral	*vast* volgens de instructies van de vroedvrouw.
Following context:	
Daarna werd de vroedvrouw weggeroepen en stond ze er alleen voor.	
Translation:	
Hanneke started her internship as a nurse. After a long and hard training she was happy to finally experience how things were done in the real world. A midwife helped her on her first task. Hanneke	
Continuations:	
2-letter word predictable	*made* the bed according to the midwife's instructions.
2-letter word neutral	*imitated* the bed according to the midwife's instructions.
4-letter word neutral	*fastened* the bed according to the midwife's instructions.
Following context:	
Then, the midwife was called away and she was on her own.	

complementary length. To validate our stimulus materials, a group of 40 firstyear 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 two-letter word, the unexpected two-letter word, and the unexpected four-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 four-letter word, the unexpected four-letter word, and the unexpected two-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 two-letter and the one four-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 four-letter target words and one two-letter word the mean 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

[7] All materials are available from the first author upon request, denis.drieghe@Ugent.be

TABLE 2

Example sentence for a four-letter predictable word

Preceding 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:

4-letter word predictable	over en dan is de pudding niet lekker meer.
4-letter word neutral	*door* 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 put the milk on the heat, I heard my mother calling: Hurry or the milk will boil

Continuations:

4-letter word predictable	*over* and the pudding won't taste good anymore.
4-letter word neutral	*through* 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 had to start over again.

plausible. O'Regan (1990, 1992) hypothesised 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 quarter 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 9-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 pseudorandom 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 1 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., two letters and four 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 refixation. 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 two-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 two stimuli were excluded from the analyses; 7.7% more data were eliminated due to four possible reasons: (1) track loss, (2) the participant's first fixation in the fourth line was directly on the target or past the target, (3) 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 (4) 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 (*F*1-analyses) and stimulus materials (*F*2-analyses).

TABLE 3

Skipping probability and fixation time measures (in ms) on the target word as function of contextual constraint and target word length

Constraint	*Word length*	*Skipping probability*	*Single fixation duration*	*Total fixation duration*
Sentences with expected two-letter words				
Predicted	2	.79	219	236
Neutral	2	.74	232	385
Neutral	4	.56	235	413
Sentences with expected four-letter words				
Predicted	4	.55	230	283
Neutral	4	.46	239	329
Neutral	2	.71	230	289

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 two-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 four-letter words on the one hand and the two-letter words on the other hand. Contrasts showed that the difference between the predicted and the unpredicted two-letter words did not reach significance in a two-tailed test (2 pr vs. 2 unpr: $t1=1.53$, p#DXGT#.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 four-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 four-letter words and the unexpected four-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#DXGT#.10;4 pr vs. 2 unpr: $t1=-4.09$, $p< .001$; $t2=-3.38$, $p<.01$;4 unprvs. 2 unpr: $t1=7.72$, $p<.001$; $t2= -4.23$, $p< .001$).

The skipping rates for unexpected two-letter words were very similar when they replaced an expected two-letter word (74%) and when they replaced an expected four-letter word (71%). In a one-way between groups ANOVA run on the stimuli data they did not differ significantly from each other, $F2(1, 32)<1$, n.s. Although there seemed to be a bigger difference for the unexpected fourletter words when they substituted an expected four-letter word (46% skipping) than when they substituted an expected two-letter word (56%

skipping), the ANOVA showed that these two values did not differ significantly from each other either, $F2(1, 32)=1.16$, $MSE=0.04$, p#DXGT#.10.

To further examine the relative effects of word length and word predict ability, we ran a multiple linear regression analysis on the skipping rates for all the 34x3=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#DXGT#.10, nor was word frequency, $-1< t(91) <1$, n.s. Finally, average launch site per sentence failed to predict average word skipping rate too, $-1<t(91)<1$, 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 two-letter words, $F1(1, 129)=74.03$, $MSE=7.25$, $p<.001$, $F2\ (1, 47)= 27.39$, $MSE=5.42$, $p<.001$, and for the conditions with the expected four-letter words, $F1(1, 152)=108.18$, $MSE=6.95$, $p<.001$, $F2(1, 53)=87.86$, $MSE= 2.71$, $p<.001$.

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 two-letter words, the effect of target type was not significant for the single fixation durations, $F1(2, 38)=1.07$, $MSE=4125$, p#DXGT#.10, $F2(2, 20)=2.51$, $MSE=1563$, p#DXGT#.10. A different pattern was found for the total fixation durations: The main effect of target type was significant, $F1(2, 42)=9.19$, $MSE=21{,}857$, $p<.001$; $F2(2, 20)=21.32$, $MSE=6162$, $p<.001$, and the predictable two-letter words differed significantly from the neutral two-letter words and neutral four-letter words but the latter two were not significantly different from each other (2 pr vs. 2 unpr: $t1=-3.41$, $p<.01$; $t2=-2.80$, $p<.01$; 2 pr vs. 4 unpr: $t1=-7.03$, $p<.001$; $t2=-10.84$, $p<.001$; 2 unpr vs. 4 unpr: $t1<1$, n.s.; $t2<1$, n.s.).

For the single fixation durations of the expected four-letter words the effect of target type was not significant, $F1(2, 78)<1$, n.s.; $F2(2, 28)<1$, n.s. For total fixation durations, the main effect of target type was significant over participants, $F1(2, 80)=4.16$, $MSE=10{,}345$, $p<.05$, and over stimuli, $F2(2,28)= 4.85$, $MSE=4242$, $p<.05$. The neutral four-letter word target differed significantly from the predictable four-letter word but the other comparisons did not differ significantly from each other (4 pr vs. 4 unpr: $t1=-2.27$, $p<.01$; $t2= -3.28$, $p<.05$; 4 pr vs. 2 unpr: $t1<1$, n.s.; $t2<1$, n.s.; 4 unpr vs. 2 unpr: $t1= 2.04$, p#DXGT#.05; $t2=2.06$, p#DXGT#.05).

TABLE 4

Mean launch site in function of target type and skipping

	Launch site	
Word length constraint	*Skipping*	*No skipping*
2 predicted	5.82	9.35
2 neutral	5.29	9.20
4 neutral	5.63	7.20
4 predicted	5.07	8.29
4 neutral	5.19	7.77
2 neutral	5.19	8.44

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 spillover 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 two-letter words (12%) were less likely than regressions to four 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.

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 boundary.[8] It includes the first-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 5

Probability of regression to the target in function of target type and skipping

	Probability of regression	
Word length constraint	*Skipping*	*No skipping*
2 predicted	.125	.023
2 neutral	.228	.105
4 neutral	.275	.219
4 predicted	.153	.080
4 neutral	.208	.163
2 neutral	.166	.062

TABLE 6

Cumulative region reading time for the two-word region past the target in function of target type (in ms)

Word length constraint	*CRRT*
2 predicted	473
2 neutral	689
4 neutral	630
4 predicted	440
4 neutral	601
2 neutral	625

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 thatof unpredictable words to $100(e^{-0.34 \text{ wordlength}})-5$. Thus, for two-letter words, the expected skipping rates are 56% and 46% for predictable and unpredictable words respectively; for four-letter words, they are 31% and 21%. These can be compared with the data obtained in the present experiment (see Table 3). Expected two-letter words resulted in a skipping rate of 79%, unexpected two-letter words in a skipping rate of 72% (averaged across the two conditions with unexpected two-letter words). Expected four-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 two-letter words and four-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 two-letter word was more likely to be skipped (72%) than a predictable four 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 predictions about word length in a repeated measures design.

In the predictable four-letter word conditions a difference of 9% in skipping rate due to contextual constraint was observed. This effect of predict ability was reduced to a 5% difference in skipping rate in the two-letter word conditions. The fact that the effect was smaller in the two-letter word conditions may be due to ceiling effects: The 79% skipping in the predictable two letter word condition is among the largest thus far observed. As such, we believe that the 9% difference due to contextual constraint in the four-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).[9]

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 some 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 dis guise, certainly as far as word

[8] Measurements exist that are similar or the same as the CRRT, e.g., the Regression Path Reading Time. For a discussion on the benefits of such measurements see Liversedge, Paterson, and Pickering (1998).

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 et al. (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 et al., 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.

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 two-letter intruder, but it can be used to decide between this word and a four-letter intruder. Or formulated otherwise, an unexpected two-letter word should be more likely to be skipped in a sentence where a two-letter word is expected than in a sentence where a four-

[9] 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).

TABLE 7

Fixation durations prior to the target in function of target type and skipping (in ms)

	Prior fixation duration		
Word length constraint	*Skipping*	*No skipping*	*Difference*
2 predicted	206	203	3
2 neutral	202	193	9
4 neutral	198	224	−26
4 predicted	204	217	−13
4 neutral	205	216	−11
2 neutral	207	218	−11

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 two-letter words: 74% vs 71%; see column 1 of Table 3), the findings of the four-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 (2003) 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 preactivate 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 nonwords. No difference in skipping rate was found between the predictable words and the nonwords 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. (2003) 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. *subject*). 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. (2003), 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 programme 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 two-letter and four-letter words. This is in line with other evidence (e.g., McConkie & Dyre, 2000; Schroyens et al., 1999) that fixation durations are not always fully determined by the ongoing 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 spillover effect). As Schroyens et al. 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 summarise, 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, add further evidence to the view that verbal and visual factors on word skipping are controlled by two functionally different subsystems.

PrEview proof published online September 2003

REFERENCES

Altarriba, J., Kroll, J.F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory and Cognition, 24*(4), 477–492.

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

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*(3), 364–390.

Blanchard, H.E., Pollatsek, A., & Rayner, K. (1989). The acquisition of parafoveal word information in reading. *Perception and Psychophysics, 46*(1), 85–94.

Brysbaert, M., & Mitchell, D.C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *Quarterly Journal of Experimental Psychology, 49A,* 664–695.

Brysbaert, M, & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–147). Oxford, UK: Elsevier.

Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology, 10*(3), 385–395.

Calvo, M.G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical and contextual factors. *Spanish Journal of Psychology, 5*(1), 66–77.

Calvo, M.G., Meseguer, E., & Carreiras, M. (2001). Inferences about predictable events: Eye movements during reading. *Psychological Research, 65,* 158–169.

Clark, J.J., & O'Regan, K.J. (1999). Word ambiguity and the optimal viewing position in reading. *Vision Research, 39*(4), 843–857.

Drieghe, D., & Brysbaert, M. (2001, September). *Word skipping in reading: On the interplay of linguistic and visual factors.* Paper presented at the conference of the European Society for Cognitive Psychology, Edinburgh, UK.

Ehrlich, S.F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behaviour, 20,* 641–655.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636.

Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral & Brain Sciences, 22*(4), 661–721.

Gautier, V., O'Regan, J.K., & Le Gargasson, J.F. (2000). "The-skipping" revisited in French: Programming saccades to skip the article "les". *Vision Research, 40,* 2517–2531.

Henderson, J.M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for theories of reading eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*(2), 201–221.

Hochberg, J. (1975). On the control of saccades in reading. *Vision Research, 15*(5)*,* 620.

Hogaboam, T.W. (1983). Reading patterns in eye movement data. In K.Rayner (Ed.), *Eye movements in reading* (pp. 309–332). New York: Academic Press.

Hyönä, J. (1993). Effects of thematic and lexical priming on readers' eye movements. *Scandinavian Journal of Psychology, 34*(4), 293–304.

Inhoff, A.W., Radach, R., Eiter, B.M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *Quarterly Journal of Experimental Psychology, 56A*(5), 803–827.

Kempen, G. (1995). Processing separable complex verbs: Comments on Frazier, Flores d'Arcais, and Coolen (1993). *Cognition, 54,* 353–356.

Kerr, P.W. (1992). *Eye movement control during reading: The selection of where to send the eyes.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

Liversedge, S.P., Paterson, K.B., & Pickering, M.J. (1998). Eye movements and measures of reading time. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55–75). Oxford, UK: Elsevier.

Marslen-Wilson, W. (1989). Access and integration: Projecting sound onto meaning. In W.MarslenWilson (Ed.), *Lexical representation and process* (pp. 3–24). Cambridge, MA: MIT Press.

McConkie, G.W., & Dyre, B.P. (2000). Eye fixations during reading: Models of frequency distributions. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 683–700). Oxford, UK: Elsevier.

McConkie, G.W., Kerr, P.W., & Dyre, B.P. (1994). What are "normal" eye movements during reading: Toward a mathematical description. In J.Ygge & G.Lennerstrand (Eds.), *Eye movements in reading* (pp. 315–327). Oxford, UK: Elsevier.

McConkie, G.W., Kerr, P.W., Reddix, M.D., & Zola, D. (1988). Eye movement control during reading: 1 The location of initial eye fixations on words. *Vision Research, 28(*10), 1107–1118.

Morris, R.K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16*(2), 268–281.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 667–682.

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

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

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

O'Regan, J.K., Lévy-Schoen, A., Pynte, J., & Brugaillére, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10*(2), 250–257.

Pollatsek, A., Rayner, K., & Balota, D.A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics, 40*(2), 123–130.

Pollatsek, A., & Well, A.D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(3), 785–794.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 165–191). Oxford, UK: Elsevier.

Radach, R., & Kempe, V. (1993). An individual analysis of initial fixation positions in reading. In G. d'Ydewalle & J.Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp, 213–226). Amsterdam: North-Holland.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65–81.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*(1), 21–30.

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41*(1), 943–954.

Rayner, K., & Fischer, M.H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics, 58*(5), 734–747.

Rayner, K., & McConkie, G.W. (1976). What guides a reader's eye movements? *Vision Research, 16*(8), 829–837.

Rayner, K., Reichle, E.D., & Pollatsek, A. (1998). Eye movement control in reading: An overview and model. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 243–268). Oxford, UK: Elsevier.

Rayner, K.,Sereno, S.C., Morris, R.K., Schmauder, A.R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4,* 1–50.

Rayner, K., Sereno, S.C., & Raney, G.E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1188–1200.

Rayner, K., & Well, A.D. (1996). Effects of contextual constraint on eye movements during reading: A further examination. *Psychonomic Bulletin and Review, 3*(4), 504–509.

Rayner, K., Well, A.D., Pollatsek, A., & Bertera, J.H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31*(6), 537–550.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39*(26), 4403–4411.

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

Schreuder, R. (1990). Lexical processing of verbs with separable particles. In G.E.Booij & J.Van Marle (Eds.), *Yearbook of morphology 1990* (pp. 65–79). Dordrecht, The Netherlands: Foris.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movements during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology, 52A*(4), 1021–1046.

Schustack, M.W., Ehrlich, S.F., & Rayner, K. (1987). Local and global sources of contextual constraint in reading. *Journal of Memory and Language, 26,* 322–340.

Starr, M.S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences, 5*(4), 156–163.

Vitu, F. (1991). The existence of a center of gravity effect during reading. *Vision Research, 31,* 1289–1313.

Vitu, F., & McConkie, G.W. (2000). Regressive saccades and word perception in adult reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 301– 326). Oxford, UK: Elsevier.

Vitu, F., O'Regan, J.K., Inhoff, A.W., & Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception and Psychophysics, 57*(3), 352–364.

Vonk, W. (1984). Eye movements during comprehension of pronouns. In A.G.Gale & F.Johnson (Eds.), *Theoretical and applied aspects of eye movement research.* Amsterdam: North-Holland.

Yang, S.-N., & McConkie, G.W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research, 41,* 3567–3585.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 104–127

Do frequency characteristics of nonfixated words influence the processing of fixated words during reading?

Jukka Hyönä and Raymond Bertram

University of Turku, Finland

Are readers capable of lexically processing more than one word at a time? In five eye movement experiments, we examined to what extent lexical characteristics of the nonfixated word to the right of fixation influenced readers' eye behaviour on the fixated word. In three experiments, we varied the frequency of the initial constituent of two-noun compounds, while in two experiments the whole-word frequency was manipulated. The results showed that frequency characteristics of the parafoveal word sometimes affected eye behaviour prior to fixating it, but the direction of effects was not consistent and the effects were not replicated across all experiments. Follow-up regression analyses suggested that foveal and parafoveal word length as well as the frequency of the word-initial trigram of the parafoveal word may modulate the parafoveal-on-foveal effects. It is concluded that low frequency words or lexemes may under certain circumstances serve as a magnet to attract an early eye movement to them. However, further corroborative evidence is clearly needed.

One key issue in eye movement research on reading has been to determine the span of effective vision among competent adult readers. In other words, how much textual information can a reader process during an eye fixation? Adherents of speed reading contend that with adequate practice, readers could acquire semantic information during a single fixation from a whole line of text (or at least from several words). Recent eye movement research has convincingly shown that this estimate is clearly too optimistic. The span of effective vision does not extend more than up to about 20 characters around the fixation and is strongly asymmetric to the right, at least for Indo-European languages like English or Finno-Ugric languages like Finnish (see Rayner, 1998, for a review). The area from which semantic or lexical information may be extracted is clearly smaller than this.

According to the highly influential E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, in press), words are typically attended one at a time during reading, and lexical-semantic properties of the following word are not assumed to affect the processing of the fixated word. There is one exception to this. The word to the right of current fixation may sometimes be identified. This most likely happens when that word is short, frequent and highly predictable. As a consequence, the word to the right of fixation is skipped, and the duration of fixation prior to skipping is inflated (Pollatsek, Rayner, & Balota, 1986; but see Radach & Heller, 2000). Recently, Kennedy (1998, 2000; Kennedy, Pynte, & Ducrot, 2002) reported evidence that contradicts the view that words are attended sequentially one at a time. More precisely, Kennedy has provided evidence for what he calls a parafoveal-onfoveal effect: Lexical and orthographic features of the parafoveal (i.e., nonfixated) word influence the eye behaviour on the currently fixated word. This effect suggests that attention may not only incorporate the fixated word but may also spread to the adjacent word. It should be noted, however, that none of the Kennedy studies involved normal reading, but laboratory tasks were used (e.g., readers were asked to decide whether or not two words were synonymous).

Kennedy (1998) demonstrated that fixation time on word N was affected by features of word N+l. He reported an inverted frequency effect: Gaze duration on N was shortened when word N+l was a low-frequency word (no such effect was observed in Exp. 2 of Kennedy, 2000). He also observed an inverted effect of the length of the parafoveal word. When the parafoveal word was long, the gaze duration on the foveal word was shortened. This effect was replicated in Kennedy’s (2000) Experiment 2, but not in Experiment 1. Kennedy interpreted these rather surprising effects to reflect difficulties of parafoveal processing. When a parafoveal word is long or infrequent, its parafoveal processing is made more difficult, which results in the reader spending less time on the foveal word. It is as if a long or infrequent word attracts an early saccade to it. Kennedy et al. (2002) observed a similar result: Less fixation time was spent on short five-letter words, when the parafoveal word was low-frequency and possessed highly

---

Correspondence should be addressed to J.Hyönä, Dept. of Psychology, University of Turku, FIN-20014 Turku, Finland. Email: hyona@utu.fiThe financial support of Suomen Akatemia (the Academy of Finland) is gratefully acknowledged. We also thank Annu Haapakangas and Tuomo Häikiö for their help in the data analyses. We thank Wayne Murray and two anonymous reviewers for their insightftul comments on an earlier version of this article.

http://www.tandf.co.uk/jouraals/pp/09541446.html
DOI: 10.1080/09541440340000132

constraining (i.e., infrequent) initial letters. When the foveal word was long (i.e., nine letters), the nature of the parafoveal-on-foveal effect was qualitatively different. Now the low-frequency parafoveal words possessing highly constraining initial letters increased the time spent on the foveal word (this was true only, when the parafoveal word was short).

Inhoff, Radach, Stark, and Greenberg (2000a) reported a semantic parafoveal-on-foveal effect in a reading study. They found that when two adjacent words were semantically related, gaze duration on the previous word was shorter than when the two adjacent words were not related. A similar parafoveal-on foveal effect was observed by Murray (1998). It should be noted that in these two reading studies parafoveal processing difficulty slowed down foveal processing. In other words, the nature of the effect is opposite to what Kennedy typically has observed (with one exception; see also Kennedy, Murray, & Boissiere, 2004). The inconsistency and nonreplicability cast doubt on the generalisability of the parafoveal-on-foveal effects. Recently, Rayner, White, Kambe, Miller, and Liversedge (2003) reviewed the evidence in favour and against the effect and concluded that the wealth of the available evidence speaks against parafoveal-on-foveal lexical-semantic effects, and thus it would be premature "to accept that there is any strong evidence obtained from a natural reading task to suggest that semantic information is obtained from the parafovea, except in those cases when the parafoveal word is subsequently skipped" (pp. 229–230). Specifically, among other things, they pointed out four reading studies (Carpenter & Just, 1983; Henderson & Ferreira, 1993; Inhoff, Starr, & Shindler, 2000b; Rayner, Fischer, & Pollatsek, 1998), none of which found evidence in support of a parafoveal-on-foveal effect of frequency.

In the present study, we set out to examine if parafoveal-on-foveal effects may indeed be observed in normal reading. We were particularly interested in whether lexical characteristics of the word to the right of the currently fixated word would influence the processing of the fixated word. Such an observation may be taken as evidence for the view that lexical processing is not only confined to the fixated word but also includes the parafoveal word. Word and constituent frequency were used as a tool to study parafoveal lexical processing. It should be noted that the present study departs from previous studies in that it deals with normal reading, while employing a relatively large database and items that are perfectly matched (the foveal word is identical for each matched pair of parafoveal words).

It is well established that word frequency significantly contributes to the ease with which words are recognised during reading (see Rayner, 1998, for a review). For long, two-noun Finnish compounds (e.g., *joukkuehenki*=team spirit) it has been shown that not only word frequency, but also the frequency of the constituents as separate words (i.e., *joukkue* and *henki*) influence the time it takes to read these words when they are embedded in sentences (Hyönä & Pollatsek, 1998; Pollatsek, Hyönä, & Bertram, 2000). For the present study, we analysed the data from five experiments (previously published or submitted for publication), in which whole-word frequency or constituent frequency of compounds were manipulated. The experimental sentences, in which the target compounds appeared, were constructed so that the sentence frame up to target word was identical for each high- and low-frequency word pair. Thus, any differences in the eye behaviour on word N, the word before the target may be readily ascribed to the frequency difference in the target word pairs.

In the first three experiments reported below, the first-constituent frequency of compounds was manipulated while controlling for whole-word frequency, and in the last two experiments we varied the whole-word frequency while controlling for constituent frequency. Two opposite predictions can be made concerning the direction of possible parafoveal-on-foveal effects of frequency. If we were to replicate Kennedy (1998, 2000), we should find a paradoxical inverted frequency effect: Low-frequency compounds should attract an early saccade to them resulting in shorter processing time for the preceding word. On the other hand, if the effect manifests in a more orthodox manner, as was the case in Inhoff et al. (2000a), in Murray (1998), and in a subset of data of Kennedy et al. (2002), low-frequency target words should produce longer gazes on word N than high-frequency words.

As regards the possibility of finding a parafoveal-on-foveal effect of constituent vs. whole-word frequency, two alternative predictions seem plausible. First, given the acuity limits of foveal vision, it may be hypothesised that it would be more probable to find a parafoveal-on-foveal effect for the constituent frequency than for the whole-word frequency manipulation. On the other hand, exactly an opposite prediction may be made by assuming that it is easier to attend parafoveally to words separated by spaces than to morphological constituents that are visually less salient.

METHOD

Participants

The number of participants was 24, 30, 26, 30, and 24 for Experiments 1 through 5, respectively. All participants were university students recruited from the introductory course of psychology; participants spoke Finnish as their native language.

Apparatus

Eye movements were collected by the EYELINK eyetracker manufactured by SR Research Ltd (Canada). The eyetracker is an infrared video-based tracking system combined with hyperacuity image processing. There are two cameras mounted on a headband (one for each eye) including two infrared LEDs for illuminating each eye. The headband weighs 450 g in total. The cameras sample pupil location and pupil size at the rate of 250 Hz. Registration is monocular and is performed for the selected eye by placing the camera and the two infrared light sources 4–6 cm away from the eye. The spatial accuracy is better than 0.5°. Head position with respect to the computer screen is tracked with the help of a head-tracking camera mounted on the centre of the headband at the level of the forehead. Four LEDs are attached to the corners of the computer screen, which are viewed by the head-tracking camera, once the participant sits directly facing the screen. Possible head motion is detected as movements of the four LEDs and is compensated for online from the eye position records.

The target sentences were presented in Courier one at a time starting from the centre-left on the computer screen. The sentences extended a maximum of three lines of text; the critical word never appeared as the initial and final word of a text line. With a viewing distance of about 65 cm, one character space subtended approximately 0.5° of visual angle. A line of text comprised a maximum of 50 characters.

Materials

In all experiments matched sentence pairs were created for the low- and high frequency condition. The matched sentence frames were identical up to the word after the manipulated word. Thus, word N was identical between the low- and high-frequency conditions. All experiments were carried out in Finnish. An example sentence pair (from Bertram & Hyönä, 2003) is given below (the manipulated word appears in bold and word N in italics).

> *Low-frequency condition:* "Muistin samassa, *että* **sivuovi** oli jäänyt lukitsematta." "I suddenly remembered that the sidedoor was left unlocked."

> *High-frequency condition:* "Muistin samassa, *että* **pesukone** oli vielä tyhjentämättä." "I suddenly remembered that the washing machine was still not emptied."

There were a few sentence pairs, in which word N appeared as the initial word of the sentence; these sentence pairs were excluded from the analyses. Below we present further details about the materials of each experiment. The frequencies were computed on the basis of an unpublished newspaper corpus of 22.7 million words that was accessed by the WordMill software (Laine & Virtanen, 1999).

Experiment 1: First-constituent frequency manipulation. The manipulated targets consisted of two sets of two-noun compounds (12–14 characters long): One set had a frequent initial constituent (average lemma frequency of 551 per million) and the other set had an infrequent initial constituent (average lemma frequency of 9 per million). The constituent frequency refers to the frequency the constituent has as a separate word in Finnish. There were 22 words of each kind.

The two sets were equated for length and rated familiarity (for more information, see Exp. 2 of Hyönä & Pollatsek, 1998). The average initial trigram frequency (per thousand occurrences) of low-frequency compounds was 0.70 and that of high-frequency compounds was 1.24. These are absolute values that were computed with the help of the WordMill software (Laine & Virtanen, 1999). The mean length of word N was 6.3 characters for the 22 words included in the analysis; their mean lemma frequency was 3452 per million.

Experiment 2: First-constituent frequency manipulation for short and long compounds. A set of short (7–9 characters) and long (12–15 characters) two noun compounds were chosen that either had a relatively low- or a highfrequency initial constituent. The average frequency values (per million) were 23 and 25 for low-frequency long and short compounds, respectively; the respective values were 472 and 468 for the high-frequency conditions. There were 19 words in each of the four conditions. The conditions were equated for word frequency, second constituent frequency, average bigram frequency, and length of the first constituent (for further details, see Exp. 1 of Bertram & Hyönä, 2003). The average initial trigram frequencies (per thousand) were 0.95 and 0.96 for the low- and high-frequency long compounds; the respective means for the short compounds were 0.43 and 0.86. The N words in the long and short compound conditions were matched on length (7.0 characters); the average lemma frequency of N was 5780 per million for the long and 4727 per million for the short compound condition.

Experiment 3: First-constituent frequency manipulation for semantically transparent and opaque compounds. A set of 40 semantically transparent and 40 opaque compounds was selected. The words were two-noun compound words (12–15 characters long). For both sets, the frequency of the first constituent was manipulated: One set had a frequent first constituent and the other set had an infrequent first constituent. There were 19 words of each kind both for transparent and opaque words (for more information, see Exp. 1 of Hyönä & Pollatsek, 2003). By definition, a semantically transparent compound is one whose meaning can be directly derived from the constituent meanings by simply “gluing” them together. On the other hand, a compound word is semantically opaque, when its meaning cannot be computed by simply gluing together constituent meanings. In our set of opaque compounds either the meaning of the whole word was opaque (e.g., *kompastuskivi*=stumbling block), or the meaning of the first constituent was opaque (e.g., *verivihollinen*=blood enemy). The average initial trigram frequencies (per thousand) were 0.36 and 0.90 for the low- and high-frequency opaque compounds and 0.85 and 0.67 for the low- and high-frequency transparent compounds, respectively. The mean length of N was 5.5 and 6.8 for the transparent and opaque condition, respectively; the respective mean lemma frequencies were 4434 and 3020.

Experiment 4: Whole-word frequency manipulation for short and long compounds. A set of high- and low-frequency two-noun compounds were selected that were either short (7–9 characters) or long (12–15 characters). The frequency values (per million) for the frequent compounds were 23 and 22 for the long and short compounds, respectively; the respective values for the two low-frequency conditions were 2.4 and 2.3. There were 20 words of each kind in the short compound and 18 in the long compound condition. The words were equated across conditions for first-constituent frequency, second-constituent frequency, average bigram frequency, word length, and length of the first constituent (for further details, see Exp. 2 of Bertram & Hyönä, 2003). The average initial trigram frequencies (per thousand) were 0.69 and 0.74 for the long low- and high-frequency compounds and 0.66 and 0.63 for the short low-and high-frequency compounds, respectively. The mean length of N was 6.0 for the long compound and 5.2 for the short compound condition, respectively; the respective mean lemma frequencies were 6708 and 4833 per million.

Experiment 5: Whole-word frequency manipulation for compound and monomorphemic words. Whole-word frequency was manipulated separately for two-noun long compounds (11–15 characters) and length-matched mono morphemic words. The frequency values (per million) of the two low-frequency word sets were 0.4 and 0.7 for the compounds and monomorphemic words, respectively; the respective means for the high-frequency word sets were 41 and 45. There were 12 words in the compound word set and 13 words in the monomorphemic word set. Most of the monomorphemic words were loan words (e.g., *identiteetti*=identity). The four word sets were matched for word length, average bigram frequency, and beginning and final trigram frequency. More over, the two compound word sets were matched on both first constituent length and frequency as well as on second constituent frequency and length (for further details, see Exp. 2 of Pollatsek et al., 2000). The average initial trigra frequencies (per thousand) were 0.63 and 0.70 for the low- and high frequency compounds and 0.84 and 0.81 for the low- and high-frequency monomorphemic words, respectively. The average length of N was 6.6 characters and that of monomorphemic words was 7.2. The mean lemma frequencies of N were 2335 and 1079 for the compound and monomorphemic word condition, respectively.

Procedure

Prior to the experiment, the eyetracker was calibrated using a 9-point calibration grid that extended over the entire computer screen. Prior to each sentence, the calibration was checked by presenting a fixation point in a centre-left position of the screen; if needed, calibration was automatically corrected, after which a sentence was presented to the right of the fixation point.

Subjects were instructed to read the sentences for comprehension at their own pace. They were further told that periodically they would be asked to paraphrase the last sentence they had read to make sure that they attended to what they read. It was emphasised that the task was to comprehend, not to memorise, the sentences.

RESULTS

We analysed readers' eye fixation patterns of word N as a function of the frequency characteristics of words that appeared immediately to the right of N. The following eye fixation measures were employed: gaze duration, duration of final fixation, probability of refixation, and probability of skipping. Gaze duration is the summed duration of fixations on N prior to fixating the frequency-manipulated word to the right. The final fixation is the fixation made on N immediately before fixating the word to the right. When there is only one fixation on N, gaze duration and the duration of final fixation are equal. In addition to these two fixation time measures we also employed a measure of the probability of refixating N. Kennedy (2000) has shown that refixation rate may be a sensitive measure of parafoveal-on-foveal effects. It may be noted that refixation rate typically correlates quite strongly with gaze duration; when the probability of refixation increases, so does gaze duration. As the final measure we used the probability of skipping N, which is the rate of leaving the N word unfixated. Thus, this measure indexes possible processing done of the manipulated word N+l when fixating on the N word (provided, of course, that properties of N+l exert an effect).

We first report the results of three experiments, where the frequency of first constituent (or lexeme) was manipulated, followed by the results of two experiments, where whole-word frequency effects were examined. These analyses are followed by regression analyses computed on the pooled data set of all five experiments. It should be noted that in all experiments the N word was identical between the matched pairs of low- and high-frequency words.

Experiment 1: Effects of first-constituent frequency

The experiment included a set of long compounds with a low-frequency initial constituent and another set of compounds with a high-frequency initial constituent. Pairwise *t*-tests (both by participants and by items) were computed for the N words that appeared adjacent to the target compounds.[1] The data for N are presented in Table 1.

In order to be sure that a frequency manipulation can in principle exert an effect, a necessary requirement is that the parafoveal word exerts an effect when it is foveally inspected. In the present experiment this was clearly the case; gaze duration on low-frequency first constituent compounds was 87 ms longer than that on high-frequency first constituent compounds—a highly significant difference both by participants and by items (for more information, see Exp. 2 of Hyönä & Pollatsek, 1998).

Gaze duration. Gaze duration was practically identical for the N word preceding a compound with a low- or high-frequency first-constituent ($t_{1,2} < 1$).

TABLE 1

Gaze duration (in ms), duration of final fixation (in ms), probability of skipping, and probability of refixations for the N word, as a function of type of parafoveal compound word in Experiment 1 (standard deviations in parentheses)

Eye movement measure	*High-frequency first constituent*		*Low-frequency first constituent*	
Gaze duration	222	(54)	220	(58)
Duration of final fixation	180	(30)	181	(29)
Probability of refixation	.15	(.08)	.14	(.11)
Probability of skipping	.19	(.13)	.22	(.14)

Duration of final fixation. The duration of final fixation prior to fixating the compound word did not differ between the two compound word types ($t_{1,2}$<1).

Probability of refixation. The tested difference remained clearly nonsignificant ($t_{1,2}$< 1).

Probability of skipping. The probability of skipping N was greater in the low-frequency first-constituent compound condition, but the difference proved significant only in the participant analysis, t_1(23)=3.45, p=.002; t_2(2l)=1.46, p#DXGT#.1.

Experiment 2: Effects of first-constituent frequency for short and long compounds

In the next experiment, we examined first-constituent frequency effects separately for short (7–9 characters) and long (12–15 characters) compounds. It should be noted that there was a significant frequency effect in gaze duration for the compound words themselves, when they were foveally inspected (a 47 ms effect for long and a 20ms effect for short compounds; the latter effect was significant only in the participant analysis; see Exp. 1 of Bertram & Hyönä, 2003). The data for N appear in Table 2. ANOVAs were computed using frequency of N +l as a within-participant and a within-item variable and length of N+l as a within-participant but a between-item variable.

Gaze duration. The main effect of frequency was significant in the participant analysis, F_1(1, 29)=5.02, p<.05; F_2(l, 33)=1.19, p=.28, and the effect of word length was close to significant in the participant analysis, F_1(l, 29)=3.41, p<.08; F_2<1. Target words preceding compounds with a high-frequency first constituent

[1] Items that were fixated less than 50% of the time and their matched pair word were excluded from the item analyses of gaze duration across all experiments, because their average would be based on too few observations.

TABLE 2

Gaze duration (in ms), duration of final fixation (in ms), probability of skipping, and probability of refixations for the N word, as a function of type of parafoveal compound word in Experiment 2 (standard deviations in parentheses)

	Long compounds		*Short compounds*	
Eye movement measure	*High*	*Low*	*High*	*Low*
Gaze duration	241 (47)	245 (73)	244 (55)	252 (69)
Duration of final fixation	196 (15)	195 (17)	198 (20)	201 (24)
Probability of refixation	.16 (.17)	.17 (.21)	.20 (.20)	.20 (.21)
Probability of skipping	.22 (.24)	.24 (.26)	.18 (.24)	.18 (.25)

High=high-frequency first constituent; Low=low-frequency first constituent.

were gazed upon for less time than those preceding compounds with a low-frequency first constituent. Moreover, targets preceding long compounds were gazed upon for somewhat less time than those preceding short compounds. The Frequency x Length interaction was clearly nonsignificant ($F_{1,2}$<1).

Duration of final fixation. There was no main effect of frequency ($F_{1,2}$< 1), but the main effect of length proved significant in the participant analysis, $F_1(1, 29)=5.81$, $p=.02$; $F_2<1$. Targets preceding long compounds elicited somewhat shorter final fixations than those preceding short compounds. The Frequency x Length interaction remained nonsignificant ($F_{1,2}$< 1).

Probability of refixation. The main effect of length proved significant in the participant analysis, $F_1(1, 29)=6.51$, $p<.02$; $F_2<1$, indicating that words prior to short compounds elicited somewhat more second fixations than words prior to long compounds. The main effect of frequency and the Frequency x Length interaction remained nonsignificant (for both, $F_{1,2}$<1).

Probability of skipping. The main effect of frequency was not significant ($F_{1,2}$ < 1), but the main effect of word length turned out to be significant in the participant analysis, $F_1(1, 29)=16.66$, $p<.001$; $F_2<1$, indicating that words prior to long compounds elicited slightly more skips than words prior to short compounds. The Frequency x Length interaction was far from significant ($F_{1,2}$ < 1).

Experiment 3: Effects of first-constituent frequency for transparent and opaque compounds

In this experiment, first-constituent frequency was manipulated separately for semantically transparent and opaque long compounds. It should be noted that the frequency effect was comparable in size for the transparent and opaque compounds, when they were foveally inspected (a 46 ms effect for transparent

TABLE 3

Gaze duration (in ms), duration of final fixation (in ms), probability of skipping, and probability of refixations for the N word, as a function of type of parafoveal compound word in Experiment 3 (standard deviations in parentheses)

	Transparent compounds				*Opaque compounds*			
Eye movement measure	*High*		*Low*		*High*		*Low*	
Gaze duration	214	(53)	208	(47)	234	(47)	221	(45)
Duration of final fixation	183	(33)	183	(32)	188	(28)	185	(28)
Probability of refixation	.12	(.11)	.09	(.10)	.18	(.10)	.16	(.12)
Probability of skipping	.27	(.15)	.32	(.16)	.21	(.14)	.23	(.12)

High=high-frequency first constituent; Low=low-frequency first constituent.

and a 47 ms effect for opaque compounds; see Exp. 1 of Hyönä & Pollatsek, 2003). ANOVAs were computed for N treating frequency and transparency of N +1 as within-participant variables, while in the item analyses transparency was a between-item and frequency a within-item variable. The data are presented in Table 3.

Gaze duration. The main effect of frequency proved significant, $F_1(1, 24)= 5.36$, $p=.03$; $F_2(1, 21)=4.25$, $p=.05$. Gazes on N were shorter when the adjacent compound was low-frequency.

Duration of final fixation. No effect approached significance ($F_{1,2}< 1.1$).

Probability of refixation. The main effect of frequency approached significance in the participant analysis, $F_1(1,24)=3.18$, $p=.09$; $F_2(1, 36)= 1.92$, $p=.17$. There was a lower tendency for refixating N when the adjacent compound was low frequency.

Probability of skipping. A reliable frequency effect emerged, $F_1(1, 24)= 9.35$, $p=.005$; $F_2(1, 36)=7.04$, $p=.01$. The probability of skipping word N was higher when the compound word to the right was low-frequency.

Summary of results for the first-constituentfrequency experiments

Across the three experiments there was some evidence for a parafoveal frequency effect due to first-constituent frequency. There were five significant or marginally significant (typically the item analysis did not reach significance) effects; all except one was a negative frequency effect. In two experiments, the data on skipping probability demonstrated an inverted frequency effect; N was skipped more often when the initial constituent of the following compound word was infrequent. An analogous effect was observed once in gaze duration (i.e., gaze was shorter for the low-frequency condition). The reversed effect in gaze duration was accompanied by an analogous, but a statistically marginal effect in the probability of refixating N. Only once a marginal tendency for an orthodox frequency effect (i.e., gaze duration was somewhat longer for the low-frequency condition) was observed. The inverted frequency effects are compatible with the idea that low-frequency constituents attract the eyes toward them (see the Discussion).

In Experiment 2, in which the length of the parafoveal word was also manipulated, the participant analyses showed some indication for a parafovealon-foveal effect of length in all four measures. Long parafoveal words increased the probability of skipping N and decreased the probability of refixating N. In addition, they elicited shorter final fixations and gaze durations on N than short parafoveal words. These effects converge on the view that long parafoveal words attract an early saccade towards them. It should be born in mind, how-ever, that only a subset of items was responsible for these effects, as indexed by nonsignificant item analyses.

Experiment 4: Effects of whole-word frequency for short and long compounds

We now move to the examination of whole-word frequency effects. In the experiment presented below, word frequency was separately manipulated for short and long compounds. ANOVAs for N were computed treating frequency and length of N+l as within-participant variables, while in the item analyses length was a between-item and frequency a within-item variable. It should be noted that there was a reliable frequency effect in gaze duration both for short (a 48ms effect) and long (a 64ms effect) compounds, when they were foveally inspected (see Exp. 2 of Bertram & Hyönä, 2003). The eye fixation data for N are presented in Table 4.

Gaze duration. There was no main effect of frequency ($F_{1,2}$<1), but the effect of length was significant in the participant analysis, F_1(1, 29)=23.85, p< .001; F_2<1, indicating that words preceding the short compounds were gazed upon for somewhat shorter time than words preceding the long compounds. The Frequency x Length interaction was not significant, F_1(1,29)=2.97, p#DXGT#. 09; F_2(1, 30)=2.76, p#DXGT#.1.

Duration of final fixation. There was a main effect of frequency in the item analysis, F_1(1, 29)=1.60, p=.22; F_2(1, 30)=4.52, p<.05, and the effect of word length was significant in the participant analysis, F_1(1, 29)=12.46, p< .001; F_2(1, 30)=2.13, p=.16, indicating that the duration of final fixation on N prior to fixating a short compound was somewhat shorter than that prior to fixating a long compound. The Frequency x Length interaction was significant, F_1(1, 29) =9.89, p< .005; F_2(1, 30)=7.71, p<.01. Subsequent t-tests showed a significant effect of frequency for the long compounds, t_1(29)=2.86, p< .01; t_2(26)=2.62, p<. 02, but not for the short ones, t_1(29)=1.54, p=.14; t_2(34)<1. The frequency effect

TABLE 4

Gaze duration (in ms), duration of final fixation (in ms), probability of skipping, and probability of refixations for the N word, as a function of type of parafoveal compound word in Experiment 4 (standard deviations in parentheses)

	Long compounds				*Short compounds*			
Eye movement measure	*High*		*Low*		*High*		*Low*	
Gaze duration	200	(40)	206	(44)	193	(30)	187	(35)
Duration final fixation	181	(22)	169	(12)	167	(12)	169	(14)
Single fixation duration	188	(27)	177	(21)	169	(13)	171	(17)
Probability of refixation	.07	(.11)	.13	(.17)	.12	(.13)	.10	(.15)
Probability of skipping	.32	(.26)	.30	(.25)	.30	(.22)	.26	(.22)

High=high-frequency compound; Low=low-frequency compound.

for long compounds was an inverted effect with longer final fixation durations for the high-frequency than for the low-frequency condition.

The effects observed in final fixation duration may be confounded by the number of fixations made; when two fixations are made on the word, the second of the two is typically shorter than a single fixation (e.g., Hyönä, 1995; Kliegl, Olson, & Davidson, 1983; Underwood, Clews, & Everatt, 1990). Therefore, a separate analysis was conducted for the single-fixation trials.

Single fixation duration. There was a tendency for a frequency effect in the item analysis, $F_1(1, 29)=1.19$, $p=.28$; $F_2(1, 30)=3.51$, $p=.07$, and the effect of word length was significant in the participant analysis, $F_1(1, 29)=17.71$, $p< .001$; $F_2(1, 30)=3.64$, $p<.07$, indicating that the duration of final fixation on N prior to fixating a short compound was somewhat shorter than that prior to fixating a long compound. The Frequency x Length interaction was significant, $F_1(1, 29)=9.98$, $p<.005$; $F_2(1, 30)=6.12$, $p<.02$. Subsequent t-tests showed a significant effect of frequency for the long compounds, $t_1(29)=3.52$, $p=.001$; $t_2(26)=2.23$, $p=.03$, but not for the short ones, $t_1(29)=1.40$, $p=.17$; $t_2(34)< 1$. The frequency effect for long compounds was an inverted effect with shorter single fixation durations for the low-frequency than for the high-frequency condition.

Probability of refixation. Both the main effect of frequency, $F_1(1, 29)= 1.78$, $p=.19$; $F_2(1, 36)<1$, and the main effect of length, $F_1(1, 29)=1.04$, $p= .32$; $F_2 < 1$, remained nonsignificant. On the other hand, the Frequency x Length interaction proved significant, $F_1(1, 29)=12.13$, $p<.005$; $F_2(1, 36)= 6.69$, $p<.02$. Subsequent t-tests revealed a significant frequency effect for long compounds, $t_1(29)=3.49$, $p<.005$; $t_2(34)=2.53$, $p=.02$, indicating that the word to the left of a high-frequency compound elicited less refixations than the word to the left of a low-frequency compound. Such a frequency effect was not observed for the short compounds, $t_1(29)=1.47$, $p=.15$; $t_2<1$.

TABLE 5

Gaze duration (in ms), duration of final fixation (in ms), probability of skipping, and probability of refixations for the N word, as a function of type of parafoveal word in Experiment 5 (standard deviations in parentheses)

	Compound word		*Monomorphemic word*	
Eye movement measure	*High*	*Low*	*High*	*Low*
Gaze duration	208 (36)	208 (37)	214 (43)	216 (48)
Duration final fixation	182 (25)	172 (23)	188 (30)	189 (28)
Single fixation duration	190 (27)	180 (26)	193 (30)	196 (35)
Probability of refixation	.12 (.12)	.14 (.11)	.12 (.13)	.12 (.14)
Probability of skipping	.25 (.14)	.23 (.11)	.09 (.09)	.08 (.07)

High=high-frequency word; Low=low-frequency word.

Probability of skipping. Probability of skipping N was not reliably affected by frequency, $F_1(1, 29)=2.82$, *p*#DXGT#.10; $F_2(1, 36)=3.74$, *p*#DXGT#.06, or by length ($F_{1,2} < 1$) of the adjacent compound, nor did the Frequency x Length interaction prove significant ($F_{1,2} < 1$).

Experiment 5: Effects of whole-word frequency for compound and monomorphemic words

In the final experiment, word frequency was manipulated separately for long compounds and for length- and frequency-matched monomorphemic words. Both word types exerted a reliable frequency effect in gaze duration (a 82 ms effect for compounds and a 34 ms effect for monomorphemic words) when the words were foveally inspected (see Exp. 2 of Pollatsek et al, 2000). ANOVAs for N were computed treating frequency and word type of N+1 as within participant variables, while in the item analyses word type was a between-item and frequency a within-item variable. The data for N are presented in Table 5.

Gaze duration. No effect approached significance ($F_{1,2}<1.1$).

Duration of final fixation. The main effect of frequency approached significance $F_1(1, 24)=3.03$, $p=.09$; $F_2(1, 18)=3.07$, $p<.10$. The main effect was qualified by a Frequency x Word type interaction that remained marginal in the item analysis, $F_1(1,24)=5.35$, $p=.03$; $F_2(l, 18)=3.11$, $p<.10$. The frequency effect was apparent only for compound words, $t_1(24)=2.75$, $p<.02$; $t_2(14)= 1.94$, $p=.07$, but not for monomorphemic words ($t_{1,2}<1$); the final fixation on N was shorter when the adjacent compound word was low-frequency.

To unconfound final fixation duration from fixation frequency, a separate analysis was conducted for the single fixation duration.

Single fixation duration. The main effect of frequency was significant in the item analysis, $F_1(1, 24)=1.49$, $p=.23$; $F_2(1, 18)=5.70$, $p<.03$. The main effect was qualified by a Frequency x Word type interaction analysis, $F_1(1, 24)= 6.74$, $p<.02$; $F_2(1, 18)=5.56$, $p=.03$. The frequency effect is apparent only for compound words, $t_1(24)=2.58$, $p<.02$; t2(14)=3.29, $p<.01$, but not for monomorphemic words ($t_{1,2}<1$); the final fixation on N was shorter when the adjacent compound word is low-frequency.

Probability of reflxation. No effect approached significance ($F_{1,2}< 1.3$).

Probability of skipping. Only the main effect of word type proved significant, $F_1(1, 24)=96.76$, $p<.001$; $F_2(1, 23)=4.14$, $p=.054$, but the effect is not readily interpretable, because the N words were not matched across the two word types.

Summary of results for the whole-word frequency experiments

In both whole-word frequency experiments, the duration of final fixation demonstrated an inverted parafoveal frequency effect (low-frequency parafoveal words were associated with shorter final fixations on N than high-frequency words). This inverted effect emerged only for long compound words. It was also observed for the single fixation duration, thus confirming that it was not confounded by refixation probability. The inverted frequency effect is compatible with the view that low-frequency words are capable of attracting an early saccade to them. We also observed the refixation probability to be increased

when the following word was a long low-frequency compound, compared to the case when the following word was a long high-frequency compound (i.e., the effect is an orthodox frequency effect). This effect may be explained by the magnet view outlined in the Discussion. It should be noted, however, that this effect was not replicated in the other experiment.

Experiment 4, in which length of parafoveal word was also varied, showed some indication for a parafoveal-on-foveal length effect; gaze duration and single fixation duration on N was longer when the parafoveal word was long (it was only significant in the participant analysis). It may be noted that the effect is in the opposite direction to what was observed in Experiment 2, and it is compatible with a view that long words cause a parafoveal processing difficulty.

Follow-up regression analyses

To further examine the admittedly inconsistent pattern of effects, we conducted a set of regression analyses to see what factors might predict the size of the parafoveal-on-foveal effects of frequency. We tried to predict the difference in each dependent measure between the low—and high-frequency condition using the following predictor variables: the length of N, the difference in the initial trigram frequency of N+1 (i.e., high[N+1]—low[N+1]), and the frequency difference in N+1 (i.e., high[N+1]—low[N+1]); for the two last predictors, logarithmic values were used.[2] The Kennedy studies (1998, 2000; Kennedy et al., 2002) suggest that length of N and initial trigram frequency of N+l are variables that may potentially modulate the parafoveal-on-foveal effects. The frequency difference in N+1 was included to see whether the size of the parafoveal effect may be predicted by the relative magnitude in the manipulated frequency. Separate regression analyses were computed for the constituent frequency (Experiments 1–3) and word frequency experiments (Experiments 4 and 5), and they were computed using the word items as cases.[3]

Constituent frequency experiments (Experiments 1–3)

Separate regression analyses were computed for short and long compounds.

Long compounds. The three-predictor model did not reach significance for any of the dependent variables. For the difference in skipping rate, the model proved statistically marginal, $F(3, 75)=2.29$, $p=.085$. The regression coeffi cient was only significant ($p=.02$) for length of N; its semipartial correlation was −.26. The shorter the word, the more probable it was that there was a positive frequency effect in skipping rate.

Short compounds. Although the number of short compounds was considerably smaller than the number of long compounds, the three-predictor model proved much more successful for short compounds. First, the difference in gaze duration between the low and high frequency condition was reliably predicted by the model, $F(3, 15)=10.90$, $p<.001$. The regression coefficients proved significant for initial trigram difference ($p=.002$) and constituent frequency difference ($p=.026$) and was marginal for length of N ($p=.063$). The respective semipartial correlations were .53, −.36, and .29. Thus, the likelihood of finding a positive frequency effect in gaze duration was increased, the higher the initial trigram frequency was for high-frequency words in comparison to low-frequency words. In other words, a constituent-frequency effect is boosted by initial trigram frequency. Second, there was a greater tendency for observing a negative frequency effect when constituent frequency difference in N+l increased. Finally, it was more likely to find a positive frequency difference in N+l for long than short words. The analysis conducted on the difference in refixation probability demonstrated a similar pattern, $F(3, 16)=4.77$, $p=.015$, except that the regression coefficient for length of N did not reach significance. The semipartial correlations were .49, −.34, and .13, for initial trigram difference, frequency

[2] We also considered using frequency of N as a predictor, but as it correlated highly ($r=-.77$) with the length of N it was left out (length of N was a better predictor than frequency of N).

[3] We are aware that in these analyses subject variability is not taken into consideration. Lorch and Myers (1990) point out that, in the analysis based on item means, the estimates of the percentage of variance accounted for by the predictor variables may be inflated, but estimates of the population regression coefficients are unbiased.

difference, and length, respectively. With the other measures, skipping rate and final fixation duration, the model remained nonsignificant.

Word frequency experiments (Experiments 4 and 5)

As with the constituent frequency experiments, the regression analyses were conducted separately for long and short parafoveal words. In all analyses, the regression model remained clearly nonsignificant; this was true even when pooled analyses were computed for short and long words.

Summary of the regression analyses

Our three-predictor model did a considerably better job in predicting the relative size of the parafoveal frequency effect for the experiment, in which constituent frequency of relatively short compounds was manipulated, compared to the experiments with long initial constituents and to experiments where whole-word frequency was manipulated. This may be ascribed to visual acuity: Whole-word and constituent frequency of long compounds are more difficult to pick up from the parafovea than short (typically four letters) constituent morphemes, because words and long constituents extend further to the periphery.

The specific results of the regression analyses for the compounds with short initial constituents may be summarised as follows. First, a negative correlation was established between the relative constituent frequency difference in N+1 and the size of the parafoveal frequency effect, which was seen in gaze duration and refixation probability. Thus, the greater the frequency difference, the more likely it was to find a negative frequency effect. This suggests that a clearly low frequency parafoveal constituent attracts an early saccade to it thus reducing the fixation time on N. In the Discussion we offer an explanation for this finding based on the idea that properties of the parafoveal word may serve as "magnets" to draw the eyes to them.

Second, a positive first-constituent frequency effect was boosted by the low frequency constituent having an initial trigram that was also of low frequency; this was observed both in gaze duration and refixation probability. The finding suggests that in order to observe a positive parafoveal frequency effect two different properties need to co-occur: The parafoveal lexeme (or word) needs to be short, and it needs to have an infrequent initial trigram.

Third, length of N correlated with the size of the parafoveal frequency effect in gaze duration. It was more probable to find a positive parafoveal frequency effect when the foveal word was long and a negative effect when the foveal word was short. This finding should be coupled with a length effect on skipping probability that was found for long compounds of the constituent frequency experiments. Here the parafoveal effect was reversed: It was more likely to find a positive frequency effect (i.e., word N is skipped more often in the low frequency condition) with short foveal words. Taken together, these results suggest that for long foveal words the parafoveal effect may be observed while fixating on the word itself, whereas for short words the effect manifests itself while fixating on N. Visual acuity of parafoveal information most likely plays a relevant role here.

DISCUSSION

Table 6 summarises the results of the present study. As is apparent from the table, we did obtain parafoveal-on-foveal effects of frequency in reading, but the nature of effects is not consistent, and they do not replicate across experiments. First, in *skipping rate* we obtained an inverted frequency effect in two constituent frequency experiments, but the effect was not replicated in the experiments, in which word frequency was varied. The inverted effect suggests that N is skipped more often, when N+1 has an infrequent initial constituent. This is in line with a view that a low-frequency lexeme operates as a "magnet" that draws the eyes to it. By adopting this view, one may further assume that the magnet would operate more strongly, the lower the frequency is. In the follow-up regression analysis we did find evidence to support this: The relative constituent frequency difference in N+1 reliably predicted the effect size in gaze duration of N (but only when the parafoveal compound had a short initial constituent).

In *gaze duration,* we once observed a positive (nonsignificant in the item analysis), once a negative, and three times no parafoveal-on-foveal effect of frequency. A positive effect may be taken to suggest that readers allocate attention to lexical features of N+1 while fixating on N by spending longer time fixating N when N+1comprises an infrequent beginning lexeme. Thus, this result is consistent with the idea that parafoveal processing difficulty slows down foveal processing. In contrast, a negative effect may be interpreted to support a magnet view, according to which an infrequent beginning lexeme would pull the eyes toward it resulting in a shorter gaze duration and a lower probability of refixating N (the latter finding was also borne out in Experiment 3).

Follow-up regression analyses provided some evidence that both length of the parafoveal lexeme as well as length of the foveal word are factors capable of modulating the nature and existence of parafoveal-on-foveal effects of frequency. First, the evidence supporting the magnet view came from trials in which either the parafoveal lexeme or the foveal word was short. When the foveal word was short, it was more probable that low-frequency parafoveal beginning lexemes attracted the eyes to them, thus shortening the gaze duration. On the other hand, when the foveal word was long, it was somewhat more probable to observe a positive frequency effect (i.e., N is gazed upon for longer time and/or refixated more often when there is a low-frequency word or lexeme to the right; see Kennedy et al., 2002, for a similar pattern of results). Finally, length of N also modulated the effect observed in skipping rate for long initial constituents. It was more probable to find an inverted parafoveal frequency effect in skipping (i.e.,

TABLE 6

Summary of results of the five experiments

Experiment	*Word type*	*Frequency manipulation*	*Gaze*	*Final fixation*	*Skipping rate*	*Probability of refixation*
1	Long compounds	First constituent	No	No	Slightly negative	No
2	Long and short compounds	First constituent	Slightly positive	No	No	No
3	Long transparent and opaque compounds	First constituent	Negative	No	Negative	Slightly negative
4	Long and short compounds	Whole word	No	Negative (long compounds)	No	Positive (long compounds)
5	Long compounds and monomorphemic words	Whole word	No	Negative (compounds)	No	No

Positive=low-frequency condition greater than high-frequency condition; No=no frequency effect; Negative=high-frequency condition greater than lowfrequency condition (in skipping rate this inverted effect takes the shape of the low-frequency condition producing more skips).

more skips in the low-frequency condition) when N was short. All these effects may be accounted for by the magnet view by further assuming that the magnetic force varies as a function of visual eccentricity. When the eyes are closer to the magnet (as is the case when N or initial lexeme of N+1 is short), the magnet is capable of pulling the eyes to it, whereas when the eyes are farther away from the magnet (as is the case when N is long), they are only attracted closer to the magnet (i.e., the eyes remain on N thus lengthening the gaze duration on N). Finally, the analyses also suggested that the magnetic force may be stronger for short parafoveal lexemes— a suggestion in line with the idea that magnetic force may vary as a function of eccentricity. It should be noted that the magnetic view bears a lot of similarity to the process monitoring account recently proposed by Kennedy et al. (2002).

In the constituent frequency experiment where also parafoveal word length was manipulated we found some evidence suggesting that long parafoveal words would serve as magnets in attracting the eyes to them. However, this finding was due to only a few items as indexed by nonsignificant item analyses. Moreover, in the corresponding whole-word frequency experiment we obtained an opposite effect; long parafoveal words were associated with longer gaze durations on the foveal word—a finding consistent with the notion of parafoveal processing difficulty. Thus, these data lent no consistent support for the magnet account when it comes to pure word-length effects.

The finding observed in the regression analyses that infrequent initial trigrams (coupled with low-frequency initial lexeme) appearing parafoveally are associated with increased foveal processing also runs counter to the magnet view according to which infrequent trigrams would attract an early saccade towards them. The idea that orthographic saliency is picked up from the parafovea and is subsequently used in guiding the eyes not only emerges in the Kennedy studies (e.g., Kennedy et al., 2002), but it may also be found in Hyönä (1995), Inhoff et al. (2000b), Underwood, Binns, and Walker (2000), and White and Liversedge (2004). Hyönä and White and Liversedge reported evidence demonstrating that word N may be skipped when N+1 hosts an irregular or illegal word-initial letter cluster, while Underwood et al. observed increased fixation times for the foveal word, when the parafoveal word had an infrequent as opposed to a frequent word-initial trigram. A similar finding was reported by Inhoff et al., who found increased fixation times for the foveal word when the parafoveal item was a nonword consisting of quasirandom letters (i.e., orthographically highly infrequent letter sequences). The findings of Underwood et al. and Inhoff et al. are in line with what we obtained in the regression analyses but at odds with a number of other studies and with the prediction that may be derived from the magnet account.

When it comes to the magnet account, we readily admit that the evidence in support of it is at best suggestive, and it should thus only be regarded as a tentative hypothesis clearly in need of more empirical support. It makes a counterintuitive prediction in claiming that parafoveal processing difficulty should lead to shorter processing times for the foveal word. However, this does not necessarily imply that processing would be left uncompleted. What may be left uncompleted during one fixation, will most likely be completed during the next fixation. Moreover, Rayner, Inhoff, Morrison, Slowiaczek, and Bertera (1981) demonstrated in a foveal masking study that as little as 50 ms is needed at the beginning of each fixation in order for reading to proceed at a normal pace. If it were true, the magnet account would challenge all existing models of eye guidance in reading (see, e.g., Engbert, Longtin, & Kliegl, 2002; Inhoff et al., 2000; Reichle et al, in press; Reilly & Radach, 2003). Although most models would be capable of accommodating some parafoveal-on-foveal effects, as far as we can judge, no model would be able to account for the "magnetic" effects discussed above. We readily admit that the challenge is not a particularly serious one, and more converging evidence is needed to make the hypothesis more tenable.

The present study was motivated by the hope that we would be able to clarify the inconsistent pattern of results regarding possible parafoveal-on-foveal effects. Our hopes were not fulfilled. Instead, we fear our results may rather create further confusion than solve previous inconsistencies. Notwithstanding the inconsistencies, it is curious that parafoveal-on-foveal effects do pop up as often as they do in the present and other recent studies. Thus, we want to close by pointing to an apparent need for more empirical studies where parafoveal-onfoveal effects in reading are systematically studied. The present study and those of Kennedy (1998, 2000; Kennedy et al., 2002) suggest that when testing the possibility of two words being lexically attended at once factors such as length of the foveal word, length of the parafoveal word, and initial trigram frequency of the parafoveal word may have a significant impact on the probability of observing an effect. Thus, studies are needed in which one variable is manipulated at a time while controlling for the effects of other variables. It is left for future research to seek additional evidence bearing on our following tentative conclusions: (1) Low-frequency lexical items serve as magnets to attract the eyes either (a) directly to them when the previous word is short and easy to recognise, or (b) closer to them when the fixated word is long; (2) Parafoveal word (or lexeme) frequency may interact with initial trigram frequency in producing parafoveal-on-foveal effects in reading.

PrEview proof published online September 2003

REFERENCES

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from reading short and long Finnish compounds. *Journal of Memory and Language, 48,* 615–634.

Carpenter, P.A., & Just, M.A. (1983). What your eyes do when your mind is reading. In K.Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York: Academic Press.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed processing. *Vision Research, 42,* 621–636.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47,* 201–221.

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 142–152.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Hyönä, J., & Pollatsek, A. (2003). The role of semantic transparency in the processing of Finnish compound words. *Manuscript submitted for publication.*

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000a). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 221–246). Oxford, UK: Elsevier Science.

Inhoff, A.W., Starr, M., & Shindler, K.L. (2000b). Is the processing of words during eye fixations in reading strictly serial? *Perception and Psychophysics, 62,* 1474–1484.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–179). Oxford, UK: Elsevier Science.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kennedy, A., Murray, W.S., & Boissiere, C. (2004). Parafoveal pragmatics revisited. *European Journal of Cognitive Psychology, 16*(1/2), 128–153.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Kliegl, R., Olson, R.K., & Davidson, B.J. (1983). On problems of unconfounding perceptual and language processes. In K.Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 333–343). New York: Academic Press.

Laine, M., & Virtanen, P. (1999). *WordMill lexical search program.* Center for Cognitive Neuroscience, University of Turku, Finland.

Lorch, R.F., Jr., & Myers, J.L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 149–157.

Murray, W.S. (1998). Parafoveal pragmatics. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 181–199). Oxford, UK: Elsevier Science.

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

Pollatsek, A., Rayner, K., & Balota, D.A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics, 40,* 123–130.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 165–191). Oxford, UK: Elsevier Science.

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

Rayner, K., Fischer, M., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.

Rayner, K., Inhoff, A., Morrison, R., Slowiaczek, M, & Bertera, J. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167–179.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier Science.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

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

Reilly, R.G., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429–455). Amsterdam: Elsevier Science.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 247–268). Oxford, UK: Elsevier Science.

Underwood, G., Clews, S., & Everatt, J. (1990). How do readers know where to look next? Local information distributions influence eye fixations. *Quarterly Journal of Experimental Psychology, 42A,* 39–65.

White, S.J., & Liversedge, S.P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*(1/2), 52–78.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 128–153

Parafoveal pragmatics revisited

Alan Kennedy, Wayne S.Murray, and Claire Boissiere

University of Dundee, UK

The results of two experiments are reported, examining eye movements as participants read the initial sentence in a sentence-matching task. The sentences employed had a NP1-verb-NP2 construction and the pragmatic plausibility of the relationship between the verb and the two nouns was independently manipulated. The aim of the first experiment was to investigate the claim that the plausibility of a NP1-verb relationship influences reading time on NP1 even before the verb is directly inspected. The data confirm the existence of such "parafoveal pragmatic" effects, but suggest that sublexical properties of the particular nouns employed may also exert a parafoveal effect on foveal processing. Experiment 2 was carried out as a control. A contingent presentation procedure ensured that the critical verb remained masked until it was directly inspected. Parafoveal-on-foveal effects exerted by the verb were removed by this procedure, although effects relating to properties of the nouns remained. The results confirm the presence of processing interactions involving sublexical properties of the two nouns, even though these were quite widely separated. Overall, the results of the two experiments suggest that, for this task, there is a genuine parafoveal-on-foveal effect attributable to purely pragmatic relationships involving the initial noun and verb in the sentences employed. In addition, there is evidence of longer range parafoveal-on-foveal effects of orthographic properties of the words employed.

This paper deals with the sources of influence over foveal inspection strategies and, in particular, the way properties of parafoveal stimuli may affect foveal processing time. "Parafoveal-on-foveal" effects of this kind have been the subject of a lot of recent research, but the topic remains rather controversial. Under the assumptions of the "leave-on-completion" hypothesis, first proposed by Morrison (1984), parafoveal information only becomes available when attention has shifted from the locus of foveal processing. It cannot have an immediate effect, although it will determine any later parafoveal preview benefit. In more sophisticated variants of Morrison's model (e.g., Henderson & Ferriera, 1990; Reichle, Pollatsek, Fisher, & Rayner, 1998), it is accepted that some parafoveal information may influence foveal processing, but this is restricted to low-level properties such as the length of the next word, which determines saccade extent. Such information must by definition be available before foveal inspection ends. However, the notion of an attentional switch, operating serially from word to word, continues to play a significant role in these later models and higher level properties of parafoveal words, such as their syntactic category or meaning, cannot affect foveal processing time. Since evidence to the contrary would undermine the whole rationale of "leave-oncompletion" (Kennedy, Pynte, & Ducrot, 2002; Pynte, Kennedy, & Ducrot, 2004), the question as to whether such evidence is available, and the degree to which it is convincing, is of considerable theoretical significance.

There is a reasonable body of evidence in support of parafoveal-on-foveal effects at both a lexical and sublexical level (Inhoff, Radach, Starr, & Greenberg, 2000; Kennedy, 1995, 1998, 2000; Kennedy et al., 2002; Underwood, Binns, & Walker, 2000). But there are also reports of inconsistencies and of null effects (Carpenter & Just, 1983; Henderson & Ferreira, 1993; Hyönä & Bertram, 2004; Rayner, White, Kambe, Miller, & Liversedge, 2003; Vitu, Brysbaert, & Lancelin, 2004; White & Liversedge, 2004). Kennedy et al. (2002) argued that many of the inconsistencies relate to a failure to control for the powerful effect exerted by foveal word length on parafoveal processing. They argued that it is possible to account for concurrent foveal and parafoveal processing by postulating a mechanism sensitive to the rate at which information is acquired in parallel across a span which typically contains significant parts of at least two words. If a foveal word is relatively short, its parafoveal neighbour may be identified, but Kennedy et al. argued that this is not simply the source of a preview advantage; it is also reflected in concurrent foveal processing time. Certain parafoveal-on-foveal effects, therefore, relate to prior identification of a nonfixated word and the manner in which foveal inspection time is modulated in this case mirrors the

Correspondence should be addressed to A.Kennedy, Psychology Dept, The University, Dundee, DD1 4HN, UK. Email: a.kennedy@dundee.ac.ukThis paper was prepared with the assistance of Grant no. R000223650 from the UK Economic and Social Research Council to Alan Kennedy.

http://www.tandf.co.uk/joumals/pp/09541446.html
DOI: 10.1080/09541440340000187

relevant lexical property of the parafoveal item. Thus, for example, the presence of a low frequency parafoveal item will increase foveal inspection time. On the other hand, there are levels of processing possible that fall short of full lexical identification of a parafoveal word, and these lead to a more complex pattern of interaction. If the foveal stimulus is relatively short (and likely to be identified in a single fixation) the initial letters of a (long) parafoveal word may be targeted if they are particularly informative. In this case an early interword exit saccade (or GO response) may be launched. At first sight, this pattern of response appears paradoxical, because a more "difficult" parafoveal stimulus will be associated with shorter, not longer, foveal inspection times. If the foveal word itself is long, and likely to require more than a single fixation for identification, a different pattern of behaviour may be found. In this case, identical letter-level parafoveal information may act to increase the probability of a within-word foveal refixation. Such SHIFT responses have the effect of making the parafoveal stimulus more visible, but at the cost of increasing foveal inspection time.

Looked at this way, the pattern of parafoveal-on-foveal effects, obtained over a number of experiments, may be less inconsistent than at first appears: Many of the apparent inconsistencies in the direction of effects may simply relate to the fact the GO responses act to shorten foveal inspection time, whereas SHIFTS act to increase it. However, it is still the case that the evidence *for* such effects derives in the main from laboratory studies employing analogues of the reading process, whereas the evidence *against* (i.e., null effects) derives mainly from studies of normal reading. Obviously, the great advantage of laboratory tasks is the level of control they can achieve over relevant properties of text materials. But the degree to which such tasks can capture the essential cognitive processes employed in reading has become an important component in the debate over the existence and/or reliability of parafoveal-on-foveal interactions. It is known (but little understood) that several quite robust effects found under certain laboratory conditions vanish, or become extremely hard to demonstrate, in normal reading. For example, the optimal viewing position function relating gaze to initial landing position (O'Regan, 1990), which is extremely pronounced in measurements taken on single words, virtually disappears in normal reading (Vitu, 1991). In contrast, word frequency effects on first fixation duration are readily demonstrated in normal reading, but may disappear if the same experimental materials are processed as search stimuli (Rayner & Fischer, 1996). Researchers may have to reconcile themselves to the fact that significant aspects of eye movement behaviour may be dependent on the particular task used or the reader's approach to that task (see Kennedy, 2000, for further discussion).

In this context one experimental study has posed a particularly dramatic challenge to the leave-on-completion hypothesis. Murray and Rowan (1998; see also Murray, 1998) examined the effect of sentence plausibility on first-pass reading time. Participants carried out a sentence-matching task, which involved reading pairs of sentences and deciding whether they were physically identical. The second sentence was only presented after the first had been read. We shall discuss some of the peculiarities of this task in a later section, but it should be noted that since the decision cannot be made until the second sentence is visible, participants are free to deploy relatively normal reading habits while processing the first sentence. It was on this initial sentence that the critical eye movement measures were made.

(1) The savages smacked the child.
(2) The uranium smacked the child.

The experiment used materials like those in (1) and (2), each having a NP1-verb-NP2 construction. The plausibility of the relationship between NP1 and the verb and between the verb and NP2 was varied, and the extent of this was assessed by independent ratings. For example, the segment "savages smacked" is more plausible than "uranium smacked". Plausible and implausible verb-NP2 relationships were established in the same way (e.g., "smacked the child" vs. "smacked the money").[1] Across the experiment as a whole, the design crossed the set of plausible and implausible NP1-verb combinations with plausible and implausible NP2-verb combinations.

The results showed that the plausibility of particular combinations of noun and verb were reflected in measures of first-pass reading times in the relevant regions, and even in first fixation durations in these regions. Since plausibility draws on essentially pragmatic information, it is surprising that it influences such an early stage of processing. The outcome was not, however, inconsistent with the operation of a serial attentional switch. In contrast, one other aspect of the data certainly was. Murray and Rowan (1998) found that inspection time in NP1 itself (i.e., before the verb had been directly fixated) also reflected the plausibility of the NP1-verb combination. This parafoveal-on-foveal effect, apparently triggered by the plausibility of the NP1-verb relationship, implies very early access to high-level information. In terms of the discussion above, the direction of the effect, with increased times associated with relationships defined as "implausible", suggested that participants had identified the verb. As noted above, other studies have shown that the length, initial letter constraint, and even the frequency of a parafoveal word might influence current inspection time, but these are also properties of text likely to influence preview benefit (see Henderson, 1992, for a review). Assuming that a degree of parallel processing of adjacent words actually occurs in normal reading, it seems plausible that they may play a role in parafoveal "cross-talk". If computations involving real world knowledge were to be added to this list it would be difficult to rescue significant portions of the leave-on-completion hypothesis. It was this implication which motivated Rayner and colleagues to re-examine the effect, employing the materials devised by Murray and Rowan, but embedding them in normal text instead of using a sentence matching task (see Rayner et al., 2003). They found no evidence that inspection time on NP1 was influenced by its relationship with the (as yet uninspected) verb. The question arises, therefore, as to whether this outcome was because the effect itself is unreliable, whether it reflects some task-dependent process, or whether it is related to some other possibly uncontrolled factor in the original task or materials.

Regardless of these wider issues, there are puzzling aspects to the data of Murray and Rowan. First, plausibility had reliable effects on the duration of the very first fixation falling on the word (i.e., not simply on aggregated measures like gaze or first-pass reading time). It was demonstrably a *very* early effect; an outcome conflicting with observations suggesting that plausibility tends only to be reflected in "late" eye movement measures, such as second-pass reading time (see, for example, Pickering & Traxler, 1998). Direct modulation of fixation duration also conflicts with the suggestion made by Kennedy et al. (2002) that parafoveal-on-foveal effects on inspection time arise primarily through changes in refixation rate (and hence maximising the visibility of successive words). This is an equivalently "late" process in word recognition, which may be captured by measures such as gaze or first-pass reading time, but not reliably by fixation duration itself.

In this paper we explore these issues in the context of a replication of the study by Murray and Rowan (1998) and a critical extension to it. Our primary objective was to establish the reliability and direction of any obtained "parafoveal pragmatic" effects; a secondary objective was to determine the role, if any, played by sublexical properties of the materials employed.

EXPERIMENT 1

Method

Participants. Twenty-four volunteer undergraduate students of the University of Dundee took part. Participants were paid £3.00 for completing the experiment. All were native English speakers. For technical reasons, participants wearing spectacles were not recruited. Some of the participants had completed other eye movement experiments, but all were naïve with regard to the experimental hypotheses in the present study.

Materials and design. The experiment was a replication of the study carried out by Murray and Rowan (1998) employing identical materials and design. The same four "files" of experimental sentences were used, each file presented to different groups of participants in the same random sequences. Each file contained a different set of six examples of plausible and implausible relationships involving each particular noun and verb. Alternative nouns in each case were matched exactly for length and for frequency, using the Kuçera and

[1] In the example chosen here the contrast in plausibility is particularly vivid, but over the materials as a whole (see Appendix), the difference in plausibility was relatively subtle and sentences rated as "implausible" fell far short of being "anomalous".

Francis (1967) count. Mean log frequencies for plausible and implausible NP1 nouns were 7.67 and 7.66 respectively; and 23.4 and 20.4 for plausible and implausible NP2 nouns, with no significant difference between the plausible and implausible conditions in either case. The average length of NP1 nouns was 7.5 characters and the average length of NP2 nouns was 6.3 characters. The full set of experimental items is listed in the Appendix. One version of each of the experimental items was assigned to each of the four files ensuring that participants only saw one version of each base sentence. The design was thus a two-by-two factorial with NP1-verb and verb-NP2 plausibility as factors:

Plausible NP1-verb/Plausible verb-NP2	The savages smacked the child.
Plausible NP1-verb/Implausible verb-NP2	The savages smacked the money.
Implausible NP1-verb/Plausible verb-NP2	The uranium smacked the child.
Implausible NP1-verb/Implausible verb-NP2	The uranium smacked the money.

Each file also contained 12 practice items and an additional 18 fillers (of varying plausibility). Since the task involved sentence matching, a version of half of the sentences was prepared containing a changed word. This word was always exactly the same length as the word it replaced and frequently, but not invariably, contained a degree of orthographic overlap (e.g., thief/chief).

Procedure. The materials were presented using a high-resolution (8x16) monopitch font in white-on-black polarity on a monochrome monitor with a high-speed phosphor, running at 100Hz frame rate. Each trial began with a fixation marker. Once inspected, this was replaced by display of a sentence beginning three characters to its right. Participants were instructed to read the sentence and then press a button to reveal a second comparison sentence, aligned vertically below. The task was to signal, as accurately as possible, whether the two sentences were identical or not by pressing either a right- or left-hand button.

The display monitor was interfaced to a Control Systems Artist 1 graphics card mounted in an IBM-compatible computer. At the viewing distance of 500 mm, one character subtended approximately 0.3° of visual angle. Eye movements were recorded from the right eye using a Dr Bouis pupil-centre computation Oculometer interfaced to a 12-bit A-D device sampling X and Y position every 5 ms. This eye tracker has a resolution of better than 0.25 characters over the 60-character calibrated range (Beauvillain & Beauvillain, 1995). A dental wax bite bar and chin rest were used to minimise head movements. The eye movement recording system was calibrated prior to the presentation of each set of three items. Measures of fixation duration, fixation position, and intra- and interword saccade extent were computed off-line using statistical algorithms based on the effective resolution of the data for each individual subject with respect to the obtained noise in a given data set. On this basis, the effective resolution of the eye-tracking system was better than one character position.

TABLE 1

First-pass reading times in NP1. The first value in each cell is the first-pass reading time and the second value is the number of fixations on which the data are based. Figures in brackets give the first-pass reading times reported by Murray and Rowan (1998)

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	249 2.40 (244)	272 2.70 (238)	260 2.55 (241)
Implausible NP1-verb	284 2.59 (252)	275 2.63 (260)	280 2.61 (256)
Mean	266 2.49 (248)	274 2.66 (249)	

Results and discussion

Participants completed the task accurately, with less than 10% errors overall and no relationship between experimental condition and the number of errors made. But since we are considering here only the reading of the initial member of each pair and not the comparison reading or the nature of the comparison process, the following data are derived from all trials, whether the participant's final response was accurate or not.

Average first-pass reading time per word on the initial NP is shown in Table 1.[2] The focus of attention here is obviously the size of the NP1-verb plausibility effect. The outcome is very similar to that reported by Murray and Rowan (1998), which is shown in brackets. Overall reading time on the NP, *before the verb was fixated,* was 20 ms per word longer when the NP1-verb relationship was implausible. This main effect was reliable by participants and narrowly missed significance in the by-items analysis, $F_1(1, 20)=9.92$, $p<.01$; F_2 $(1, 20)=3.01$, $p=.09$. Interpretation of the main effect is, however, complicated by a significant interaction between NP1-verb and verb-NP2 plausibility, $F_1(1, 20)= 6.52$, $p<.05$; $F_2(1, 20)=6.23$, $p<.05$. The obtained NP1-verb plausibility effect was only present when the verb-NP2 relationship was defined as "plausible", $F_1(1, 20)=10.92$, $p<.01$; $F_2(1, 20)=5.70$, $p<.05$. The data from Murray and Rowan also showed some variability in the size of the NP1-verb effect with NP2 plausibility, but in their case the larger effect was with implausible verb-NP2 combinations and the interaction was not significant. They also suggested that modulation of the duration of the final fixation in this region effectively accounted for the entire first-pass reading time effect. In the present experiment the duration of the final fixation varied in line with the first-pass reading times, but the effects were smaller and were nonsignificant (all Fs< 2.1). (With plausible verb-NP2 combinations, the means for plausible NP1-verb and implausible NP1-verb were 229 and 256 ms respectively, and with implausible verb-NP2, the means were 235 and 237 ms.)

The data suggest that the pragmatic plausibility of the NP1-verb combination can be computed on the basis of parafoveal viewing and that these computations have an immediate effect on the current (foveal) fixation. But overlaying this effect there appears to be a powerful influence exerted from some property of the critical words assigned to NP2. An indication of the mechanism at work can be derived from an examination of the average number of fixations made on NP1, also shown in Table 1. There was no reliable effect of NP1-verb plausibility ($F1$ and $F2<1$), but fixation rate was significantly higher when the words assigned to the verb-NP2 relationship were defined as "implausible", $F1\ (1,20) = 5.48$, $p<.05$; $F2(1, 20)=10.23$, $p<.01$. The interaction was not significant, $F1=2.88$, $F2=1.6$. It seems undeniable that some property of the words assigned to NP2 acted to influence the outcome, but the outcome can hardly be interpreted in terms of sensitivity to the pragmatic plausibility of the verb-NP2 relationship. Over the materials as a whole, the distance from the centre of the first noun to the verb was 7.8 characters and this is within the range where the lexical status of a parafoveal word might be available. But the average centre-tocentre distance between the two nouns in NP1 and NP2 was 19.8 characters. There is no evidence that words can be identified at this distance, although letters may be (McConkie & Rayner, 1975; McConkie & Zola, 1984; Rayner & Bertera, 1979; Rayner, Well, Pollatsek, & Bertera, 1982). Consequently, it is not possible to argue that items several *words* (as distinct from characters) downstream were identified and pragmatic relationships between them computed. We will defer consideration of exactly what properties might be at work until the complete pattern of effects for the experiment has been presented.

Since changes in refixation rate tend to produce systematic changes in launch position, analyses were also carried out on the average saccade extent leaving NP1, the landing position on the verb and the effective launch position from NP1. The relevant data are shown in Table 2. In line with variation in the number of fixations, launch position in NP1 was significantly closer to the verb when the nouns in NP2 were defined as "implausible" (5.79 vs. 5.06 characters), $F1(1, 20) =11.40$, $p<.01$; $F2(1, 20)=7.38$, $p=.01$. There was a tendency for saccades to be a little shorter approaching the verb when the nouns in NP2 were defined as "implausible" (8.72 vs. 8.33 characters), but the effect was not reliable in the by-items analysis, $F1(1,20)=6.04$, $p<.05$; $F2(1, 20)= 1.72$. The two effects combined to produce a modest right-shift in initial landing position on the verb when NP2 was defined as "implausible" (3.27 vs. 2.93 characters), $F1(1, 20)=8.59$, $p<.01$; $F2(1, 20)=3.66$, $p=.07$. The nonsignificant effect of NP1-verb plausibility on fixation rate was also mirrored in a lack of effect of this variable on launch position, saccade extent, or landing position (all Fs<1.7).

[2] The measure of first-pass reading time used here is "total pass per word" as defined in Kennedy, Murray, Jennings, and Reid (1989) and discussed by Murray (2000). It is the sum of all fixations made from when the eyes first enter the defined zone until the eyes pass the right boundary of the zone, scaled by the number of words in the zone. The measure includes regressions out of the zone to the left. It has also been referred to as "regression path time" (Konieczny, 1996) and "gopast time" (Clifton, Bock, & Rado, 2000).

TABLE 2

Saccade extent leaving NP1, initial landing position on the verb, and launch distance (i.e., distance from the initial letter of the verb). The data are in units of character position and are shown as a function of NP1-verb and verb-NP2 plausibility

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	8.95, 2.97, 5.98	8.44, 3.38, 5.06	8.70, 3.17, 5.53
Implausible NP1-verb	8.48, 2.90, 5.58	8.22, 3.16, 5.06	8.35, 3.03, 5.32
Mean	8.72, 2.93, 5.79	8.33, 3.27, 5.06	

Turning to first-pass reading time on the verb itself, the data are shown in Table 3. Inspection time on the verb showed no continuing effect of NP1-verb plausibility ($F1$ and $F2<1$). In contrast, the overall effect of verb-NP2 plausibility was significant in the by-subjects analysis and very narrowly missed significance by-items, $F1(1, 20)=5.64$, $p<.05$; $F2(1, 20)=3.79$, $p=.06$. Since at this point the second noun has not been directly inspected, this difference must also be interpreted as a parafoveal-on-foveal effect. However, it is an inverted effect with respect to assigned plausibility: First-pass reading time was *faster* when the verb-NP2 relationship was defined as "implausible". There was also a significant interaction between NP1-verb and verb-NP2 plausibility, $F1(1, 20)=4.22$, $p<.05$; $F2(1, 20)=6.41$, $p<.05$. Subanalyses revealed that the (inverted) effect of verb-NP2 plausibility was only found following a plausible NP1-verb relationship, $F1(1, 20)=8.83$, $p<.01$, $F2(1, 20)=12.30$, $p<.01$.

A better understanding of the nature of the effect of NP2 plausibility can be gained by considering the data on the number of fixations on the verb shown in Table 3. As in the study of Murray and Rowan (1998), the verb was nearly always fixated at least once: Readers made at least one first pass fixation on the verb on more than 97% of the trials, and the number of times the verb was skipped did not vary in any systematic way with experimental condition (skipping rates were, PP: 1.4%, PI: 3.7%, IP: 4.4%, II: 2.2%). Nonetheless, analysis of the number of fixations revealed a highly significant interaction between NP1-verb and verb-NP2 plausibility, $F1(1, 20)=10.09$, $p<.01$, $F2(1, 20)=10.48$, $p<.01$, with fewer fixations on the verb when the following noun was defined as "implausible" (1.41 vs. 1.19), $F1(1, 20)=6.87$, $p<.05$; $F2(1, 20)=6.34$, $p<.05$.

Taken together, the pattern of refixations on NP1 and on the verb is similar to the pattern reported by Kennedy et al. (2002) for long and short foveal stimuli.

When the foveal stimulus was "long" (as with NP1 as a whole in the present materials) and the parafoveal item relatively remote, the number of within-word refixations (SHIFTS) increased, and this would have the effect of making the distant stimulus more visible. When the parafoveal stimulus was short (as with the verb in the present materials), an equivalent stimulus condition triggered an

TABLE 3

First-pass reading times on the verb. The first value in each cell is the first-pass reading time and the second value is the number of fixations on which the data are based.

Figures in brackets give the equivalent first-pass reading times reported by Murray and Rowan (1998)

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	355 1.41 (283)	294 1.19 (289)	325 1.30 (286)
Implausible NP1-verb	330 1.31 (314)	322 1.44 (299)	326 1.38 (307)
Mean	343 1.36 (299)	308 1.32 (294)	

immediate interword saccade (GO), decreasing the incidence of within-word SHIFTS.

First-pass reading times in the final NP itself are shown in Table 4. Murray and Rowan (1998) found no continuing effect of NP1-verb plausibility at this point, an outcome clearly replicated in the present data ($F1$ and $F2<1$). In contrast, they found a significant effect of verb-NP2 plausibility on the duration of the first fixation in this region and a trend towards an overall first-pass effect of 18 ms (in a conventional direction). While the 8 ms first-pass difference in the present study was in the same direction, it was not significant ($F1$ and $F2<1$).

We will now consider the outcome of the experiment in terms of the objectives set out in the introduction to this paper. We believe two claims are justified. First, NP1-verb plausibility does act to modulate inspection time on NP1 and this appears to be a genuine “parafoveal pragmatic” effect. The orthodox direction of the effect, with longer times associated with implausible relationships, is consistent with the claim that the verb was identified before execution of the exit saccade from NP1. It is clearly also consistent with the claim that NP1-verb plausibility acted to modulate NP1 inspection time. Second, some property or properties of the words assigned to NP2 also affected behaviour on NP1. This outcome is, of course, also a parafoveal-on-foveal effect, albeit one that was not predicted.

Since the source of influence is on foveal reinspection strategy, the considerations raised in the introduction to this paper suggest that the mediating variable involved some sublexical property (e.g., initial letter constraint or familiarity). Thus, with regard to the direction and reliability of the primary effect, the data from this experiment, while replicating in some respects the principal findings of Murray and Rowan (1998), do not match in every detail the pattern of effects found in that study. The critical question is whether differences between the two experiments are sufficiently large to call their conclusions into question. An obvious way of addressing this issue is to carry out combined analyses across the two experiments with the aim of identifying significant interactions.

TABLE 4

Mean first-pass reading time in NP2 as a function of both NP1-verb and verb-NP2 plausibility. Figures in brackets show the equivalent data in Murray and Rowan (1998)

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	211 (244)	218 (254)	214 (249)
Implausible NP1-verb	209 (240)	217 (265)	213 (253)
Mean	210 (242)	218 (260)	

Considering first-pass reading times in the initial NP, the main effect of NP1-verb plausibility in the combined analysis was highly significant, F1(1, 40)= 14. 38, p<.001; F2(1, 20) 20)=6.29, p<.05 and there was no hint of an interaction with experiment (F1 and F2<1). On the basis of this outcome, and in particular the lack of any two-way interaction, we wish to argue that the primary ' ‘pragmatic parafoveal effect" is real. Indeed, the size of the main effect of NP1-verb plausibility, before the verb was directly inspected, was actually larger overall in the replication than in the original study. However, the three-way interaction involving experiment, NP1-verb plausibility, and verb-NP2 plausibility was also significant, F1(1, 40)=6.80, p<.05; F2(1, 20)=6.19, p<.05, and this is an embarrassment. In Experiment 1 some property of the nouns allocated to NP2 influenced inspection time on both NP1 and on the verb and this stands in need of explanation. In fact two problems must be faced: First, how to account for the differences between the two experiments, and second, how to account for the observation that properties of a word many characters to the right (it will be recalled that the verb itself is identical across conditions) apparently influenced inspection time. We will deal with these two questions separately.

With respect to global differences between the two experiments, detailed analysis of the results leads to the conclusion that the two groups of participants approached the task in different ways, employing slightly different reading strategies.[3] There are several sources of evidence for this proposition: Overall, participants in the replication read the initial sentence more rapidly (1572 ms vs. 1623 ms) and although this difference was not itself significant, F1<1; F2(l, 20) =9.37, p<.01, it reflected a number of significant differences between the groups in the pattern of inspection deployed. In particular, the point in time at which various effects became evident in the two groups of participants differed. This can be demonstrated by comparing the performance of the two groups of participants on three measures in each of the three critical regions (i.e., NP1, verb and NP2).

Considering first the measure of initial fixation duration in a region, participants in the replication made significantly shorter fixations overall (233 vs. 253 ms), F1 (1, 40)=8.45, p<.01; F2(1, 60)=46.29, p<.01, but there was also a significant interaction between experiment and region, F1(2, 80)=5.51, p< .01; F2(2, 60)

=20.40, p<.01. Subanalyses showed that the behaviour of the groups of participants in the two experiments only differed significantly in NP1 and NP2 (185 vs. 206 ms, and 249 vs. 289 ms respectively). There was no difference in first fixation duration on the verb (265 ms vs. 263 ms). Turning to the measure of first-pass reading time, there was no overall difference between the two groups (270 vs. 265 ms; $F1$ and $F2$<1), but a highly significant Experiment x Region interaction $F1$ (2,80)=6.09, p<.01; $F2$(2, 60)=34.06, p <.01. First-pass reading time in NP1 and on the verb was longer in the replication than in the original study (270 vs. 249 ms, and 325 vs. 296 ms, respectively), but participants in the replication spent significantly *less* time in NP2 (214 vs. 250 ms). Finally, looking at the number of fixations contributing to first-pass reading time, participants in the replication made significantly more overall (1.88 vs. 1. 71), $F1$(1, 40)=5.61, p<.05; $F2$(1, 60)=46.82, p< .01, but again there was a significant Experiment x Region interaction, $F1$(2, 80)= 48.2, p=.01; $F2$(2, 60) =32.19, p<.01. In this case, subanalyses showed that the two groups only differed significantly in NP1 and the verb (2.58 vs. 2.19, and 1.34 vs. 1.18, respectively). The number of fixations determining first-pass time in NP2 did not differ (1.72 vs. 1.76).

In summary, the sample of participants in the replication study, for unknown reasons, (perhaps because they were derived from a different cohort of students, but also see Footnote 2) exhibited a more "mobile" reading strategy. They made more, but slightly shorter, fixations and distributed them in a different way across the materials. Can these apparent differences in reading strategy account for the differences between the experiments? More importantly, can it be argued that, notwithstanding these between-group differences, both experiments provide evidence that the plausibility of the NP1-verb relationship influenced eye movement behaviour before the verb was inspected? To answer these questions involves examining the way variation in reading strategy on and around the verb affected the unfolding pattern of effects.

Starting with the point just before the verb was directly inspected, analysis of the duration of the last fixation in NP1 showed a pattern of results similar to those found for first-pass reading time. The duration of these fixations was shorter in the replication study (264 vs. 239 ms), $F1$(1, 40)=6.83, p<.05; $F2$(l, 20) =20.43, p<.01, but despite this difference, a combined analysis over the two experiments showed a highly significant main effect of NP1 plausibility, $F1$(1, 40)=10.11, p<.01; $F2$(1, 20)=9.63, p<.001, and no interaction of this with experiment, $F1$=1.4, $F2$(1, 20)=3.33, p#DXGT#.05. This outcome is clearly consistent with Murray and Rowan's (1998) findings, but, as with the measure of first-pass reading time, there was also a significant three-way interaction between

[3] In fact, the replication experiment was run using updated control software employing a slightly different contingent initial fixation procedure. Participants in the replication were "released" from the fixation marker slightly faster.

experiment, NP1, and NP2 plausibility, $F1(1, 40)=5.48$, $p<.05$; $F2(1, 20)=5.85$, $p<.05$. First fixation duration on the verb itself did not differ between the two groups (262 vs. 265 ms; $F1$ and $F2<1$), but the extent to which the measure reflected NP1-verb plausibility effects did differ. The effect was strongly present in the original study (254 vs. 271 ms), but absent at this point in the replication (266 vs. 264 ms). The interaction with experiment was significant, $F1(1, 40)=5.96$, $p<.05$; $F2(1, 20)=7.09$, $p<.05$, and no other effects or interactions were significant at this point. At the end of the inspection period on the verb (i.e., in the measure of the duration of the final fixation falling on the verb) the data show a more consistent pattern of results, with a main effect of NP1-verb plausibility (259 vs. 269ms) significant by participants, $F1(1, 40)= 4.61$, $p<.05$, and nearly by items, $F2(1, 20)=3.78$, $p=.06$, and, crucially, no hint of an interaction with experiment ($F1$ and $F2<1$). In this measure there was also a significant parafoveal effect of verb-NP2 plausibility overall, $F1(1, 40)= 5.04$, $p<.05$; $F2(1, 20)=4.79$, $p<.05$, with slower times associated with nouns assigned to "plausible" relationships (270 vs. 259 ms) and again no interaction of this with experiment, $F1<1$; $F2=3.58$. The effects of NP1 and NP2 plausibility did not interact and there was no three-way interaction with experiment.

These more detailed combined analyses over the two experiments lead to the conclusion that the "conventional" effect of NP1-verb plausibility during inspection of the verb, reported by Murray and Rowan (1998), does replicate, but in a slightly different time frame across the two experiments. In addition, the data also reveal that during inspection of the verb *in both experiments,* there was a parafoveal NP2 effect on final fixation duration, although it was only the power of the combined analysis that enabled this to be clearly seen. Final fixations on NP2 were shorter in the replication than in the Murray and Rowan study (262 vs. 317 ms), but by this point too the effects were entirely consistent across experiments.[4] There was a significant effect of verb-NP2 plausibility in an orthodox direction (279 vs. 300 ms), $F1(1, 40)=8.51$, $p<.01$; $F2(1, 20)= 6.37$, $p<.05$, with no hint of an interaction with experiment, $F1=1.14$, $F2<1$, and no other main effects or interactions (all Fs<1.04).

In summary, both the "conventional" and the "unconventional" findings replicate across the two experiments, but they follow a slightly different time course. Adoption of a more mobile reading strategy in the replication study appears to have fractionally delayed the onset of "conventional" effects of plausibility, while other parafoveal effects (unrelated to plausibility) occurred either more robustly or earlier. Across the two experiments, the data appear overwhelmingly in favour of a "conventional" effect of NP1-verb plausibility, evident before the verb is directly inspected. It is present in measures of first pass reading time and in final fixation duration, and the data on verb skipping rate in both experiments demonstrate that it cannot be accounted for in terms of a serial attention shift model. Equally, there is also evidence that some property of NP2, in both experiments, influenced inspection time. This is shown consistently in the duration of the final fixation on the verb across both studies, but also in the duration of the later fixations falling on NP1 in the replication study.

We now turn to consider the second issue—the ubiquitous influence of verbNP2 "plausibility" on inspection time in NP1 and on the verb. Two considerations suggest that this is likely to relate to relatively low-level properties of the particular nouns assigned to the "plausible" and "implausible" conditions, rather than to plausibility *per se.* First, the distance between the final letter of NP1 and the initial letter of the noun defining NP2 (15.2 characters), while effectively ruling out access to lexical properties of the second noun, leaves open the possibility that low-level information, such as its initial letters, might be available during inspection of some parts of NP1. Second, sublexical properties of parafoveal words are known to influence the pattern of foveal refixation and hence inspection time (Kennedy, 2000; Kennedy et al., 2002). The most likely basis for the apparent effect of NP2 "plausibility" is, therefore, accidental variation in orthographic constraint or familiarity in the two classes of nouns used for plausible and implausible continuations.

In fact, Murray and Rowan (1998) considered in detail whether peculiarities of the particular lexical items they employed might account for the obtained "parafoveal pragmatic" effects. They concluded that this was unlikely on several counts: (1) The nouns were carefully matched for length and frequency using the Kuçera and Francis (1967) norms; (2) the verb was identical in each set of matched items; and (3) the design involved crossing combinations of nouns with the critical verb. Of course, in the nature of things, the set of plausible and implausible nouns must differ, and Murray (1998) considered one possible source of artefact, namely the fact that the set of implausible NPs tended to contain more nonhuman entities as agents. But there was no statistical support for the suggestion that this played a part in the obtained results. In any case, such a factor could hardly operate over the range implied for NP2 effects in the data from the present study.

Although Murray and Rowan (1998) did not set out to control for the initial letter constraint of the critical nouns, the materials were well matched on this measure. The number of words of the same length, plus or minus 2 characters, sharing the initial three letters was 57 and 35 for nouns allocated to "plausible" and "implausible" conditions in NP1, with values of 30 and 28 for NP2. These differences were far from significant (by Wilcoxon nonparametric tests).[5] In contrast, however, there were quite marked differences in the token familiarity of the initial trigrams of the critical nouns. This was assessed by computing a measure of trigram cumulative lexical frequency (CLF), adding the frequencies of all the words of the same length, plus or minus 2 characters, containing a given initial trigram (see Pynte et al., 2004). Overall, the initial letters of nouns

[4] First fixation duration on NP2 also showed a significant plausibility effect and no significant interaction with experiment.

allocated to “implausible” conditions were more “familiar” than those allocated to “plausible” conditions, with values of 29,563 and 20,715 respectively.

In the case of the nouns employed in NP1, the difference (38,047 and 26,644 for “implausible” and “plausible” respectively) approached significance, $z=1.84$, $p=.06$. Although nonsignificant, the difference for nouns employed in NP2 was equally large and in the same direction (21,223 and 14,930).

Experiment 2 set out to examine whether a factor such as cumulative lexical frequency (or perhaps some other “low-level” property) could account for either or both of the NP1-verb or verb-NP2 “plausibility effects”. There is a straightforward manipulation capable of arbitrating between possible effects of low-level information (e.g., properties of the lexical items themselves) and sources of syntactic/pragmatic information (involving noun-verb computations). If the verb (but no other part of the displayed sentence) is masked prior to its direct inspection, no effects of plausibility, involving either NP1-verb or verb-NP2 relationships, is possible. Any variation in inspection time or refixation rate must relate to properties of the individual words visible at that point. It is possible by this manipulation to control for *any* interpretation involving an appeal to very early (parafoveal) effects of pragmatic plausibility.

EXPERIMENT 2

Method

Participants. Thirty-six unpaid volunteer undergraduate students of the University of Dundee took part. All were native English speakers. As in Experiment 1, participants wearing spectacles were excluded. Some of the participants had completed other eye movement experiments, but none had taken part in Experiment 1. All were naïve with regard to the experimental hypotheses in the present study.

Materials, design, and procedure. The materials, design, and procedure were identical to those employed in Experiment 1 with two exceptions: The sampling rate was increased, with an X and Y value recorded every 2 ms, and a “boundary technique” (Rayner, 1975) was employed, such that during the initial display of the first sentence, while the eyes were to the left of an invisible boundary, defined by the space prior to the verb, this word was masked. When the eyes crossed this boundary, the verb was permanently unmasked. Masking was achieved by randomly relocating the pixels making up the constituent letters of a word, ensuring that the average brightness of the displayed stimulus in its clear and masked versions was exactly matched. Unmasking was achieved by writing directly to the video memory of the graphics control card and was not dependent

[5] A measure was also computed for words of any length, producing values of 87, 52, 57, and 56 respectively. There were no significant differences.

on the refresh cycle of the display. Thus, the average delay in executing the display change was under 7 ms (half a refresh cycle (5 ms) plus the average sampling and computation time) and the maximum delay less than 13ms. It was confirmed independently that unmasking invariably took place within a single screen refresh cycle.

Results and discussion

Participants completed the task accurately, with less than 10% comparison errors, but again these analyses, which relate to the reading of the initial member of each pair, are derived from all experimental trials. Discussion of these results will be taken along with the outcome of Experiment 1, reporting analyses of both experiments combined, where this is appropriate.

Average first-pass reading times on the first NP (i.e., while the verb was masked) are shown in Table 5. Obviously the 11 ms increase in reading time when the NP1-verb relationship was defined as implausible is greater than the predicted zero, although it is clearly nonsignificant, $F1(1, 32)=2.03$, $p=.2$; $F2(l, 20)= 1.34$, $p=.2$. There was also no effect of verb-NP2 plausibility on first-pass reading time in NP1, $F1$ and $F2$ both <1, and no interaction, $F1(1, 32)=1.38$; $F2(1, 20)=1.08$. Thus, ensuring the verb was no longer visible, removed (or, to put it more cautiously, reduced to nonsignificance) the main effect of NP1-verb plausibility found in Experiment 1. The manipulation also convincingly removed any interaction involving NP2 from the measure of first-pass reading time in NP1. These two conclusions were examined more formally by means of a combined analysis of variance over the two experiments. The second conclusion was clearly confirmed: The three-way interaction between experiment, NP1-verb plausibility, and verb-NP2 plausibility was significant, $F1(1, 52)=4.96$, $p<.05$; $F2(1, 20)= 10.26$, $p<.01$. Unfortunately, the first conclusion was less well supported: There was a main effect of NP1-verb plausibility, $F1(1, 52)=8.23$, $p<.01$; $F2(1, 20)= 4.14$, $p<.05$, but the interaction with experiment was not, in fact, significant ($F1$ and $F2$<1). It would appear, therefore, that the nonsignificant NP1-verb "plausibility effect" found in Experiment 2 was not reliably different from the significant effect found in Experiment 1.

Table 5 shows the average number of fixations made in NP1. In this case, although it cannot be attributed to an effect of manipulated plausibility, the NP1-verb "plausibility" effect (2.34 vs. 2.45 fixations) actually approached significance, $F1(1, 32)=3.27$, $p=.07$; $F2(1, 20)=4.83$, $p<.05$. The main effect of verb-NP2 "plausibility" and the interaction were not significant (all $F1$ and $F2$ <1). Table 6 also shows the average launch position, saccade extent and initial landing position on the verb (to be compared with the data in Table 2). The apparent tendency for saccade extent to be longer following inspection of a noun allocated to the implausible condition (7.96 vs. 7.63 characters) was not significant, $F1(1, 32)=3.47$, $p=.07$; $F2$<1, and the tendency towards a right-shift in initial landing position on the verb in the case of implausible NP1-verb

TABLE 5

Mean first-pass reading time in NP1. The first value in each cell is the first-pass reading time and the second value is the number of fixations on which the data are based. The data are shown as a function of NP1-verb and verb-NP2 plausibility although the verb was masked at this point

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	244 2.39	238 2.30	241 2.34
Implausible NP1-verb	245 2.46	259 2.44	252 2.45
Mean	245 2.42	249 2.37	

combinations (3.27 vs. 2.89) was also not significant, $F1(1, 32)=4.56$, $p<.05$; $F2(1, 20)=1.72$. In contrast to Experiment 1, the apparent variation in the number of fixations did not produce systematic changes in launch position from NP1 ($F1$ and $F2<1$).

"Plausibility effects" involving the verb are not possible while it is masked, but given the apparent changes in the number of fixations and the absence of an Experiment x NP1-verb plausibility interaction in the combined analysis, it is important to verify that the 11 ms difference in reading time as a function of "NP1-verb plausibility" was simply a random effect. To this end, a further analysis was carried out on the average duration of the final fixation on NP1 (i.e., just before the verb was unmasked). These data are shown in Table 7.

The main effect of "NP1-verb plausibility" was nonsignificant, $F1(1, 32)= 1.16$; $F2<1$, as was the main effect of "verb-NP2 plausibility", $F1(1, 32)=1.19$; $F2(1, 20)=3.35$, $p=.08$. However, the interaction involving NP2 was significant in the by-subjects analysis and narrowly missed significance by items, $F1(1, 32)=4.56$, $p<.05$; $F2(1, 20)=3.11$, $p=.09$. The interaction relates to an influence of verb-NP2 "plausibility" (213 vs. 236 ms) when the NP1-verb relationship was defined as "implausible", $F1(1, 32)=4.69$, $p<.05$; $F2(1, 20)=5.92$, $p<.05$.[6] Although statistically nonsignificant, the pattern of effects is too systematic to dismiss. On the other hand, the quotation marks are obviously necessary because, before the verb is unmasked, any obtained effect of plausibility must arise as a proxy effect of some uncontrolled property of the individual nouns. In this context, initial letter CLF suggests itself as a possible candidate. Its influence, assessed by means of analyses of covariance, is taken up in a later section.

The possibility that the effects on the final fixation in NP1 might be attributed to different factors in this experiment (where the verb is masked) from those which operated in Experiment 1, was investigated by means of a combined analysis over the two experiments. There was a significant main effect of experiment, $F1(1, 52)=3.99$, $p=.05$; $F2(1, 20)=9.00$, $p<.01$, with shorter final fixation durations preceding a masked verb. But, more critically, there was evidence that the interaction between NP1 and NP2 effects differed between the experiments, with a three-way interaction significant by participants, $F1(1, 52)=$

TABLE 6

Saccade extent leaving NP1, initial landing position on the verb, and launch distance prior to landing on the verb. The data are in units of character position and are shown as a function of NP1-verb and verb-NP2 plausibility

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	7.66, 2.85, 4.81	7.61, 2.93, 4.68	7.63, 2.89, 4.74
Implausible NP1-verb	8.03, 3.45, 4.58	7.89, 3.08, 4.81	7.96, 3.27, 4.69
Mean	7.85, 3.15, 4.70	7.75, 3.00, 4.75	

TABLE 7

Mean duration of the final fixation in NP1. The data are shown as a function of NP1-verb and verb-NP2 plausibility although the verb was masked at this point

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	7.66, 2.85, 4.81	7.61, 2.93, 4.68	7.63, 2.89, 4.74
Implausible NP1-verb	8.03, 3.45, 4.58	7.89, 3.08, 4.81	7.96, 3.27, 4.69
Mean	7.85, 3.15, 4.70	7.75, 3.00, 4.75	

TABLE 8

Mean first-pass reading time on the (unmasked) verb. The data are shown as a function of NP1-verb and verb-NP2 plausibility

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	7.66, 2.85, 4.81	7.61, 2.93, 4.68	7.63, 2.89, 4.74
Implausible NP1-verb	8.03, 3.45, 4.58	7.89, 3.08, 4.81	7.96, 3.27, 4.69
Mean	7.85, 3.15, 4.70	7.75, 3.00, 4.75	

6.12, $p<.05$, and approaching significance by items $F2(1, 20)=2.91$, $p=.1$. We will suspend considering the outcome of analyses of covariance investigating this possibility further until data on inspection time on the verb itself have been considered.

Average first-pass reading times on the unmasked verb itself are shown in Table 8. It would appear that there was about an 18ms "preview benefit" in Experiment 1 (with a mean first-pass time of 325 ms compared to the 343 ms in Experiment 2) but caution needs to be exercised concerning the magnitude of this effect, not simply because of the between-subjects comparison, but also because first-pass reading times might differently reflect the influences of pragmatic plausibility in the two experiments. Surprisingly, the 20ms main effect of NP1-verb plausibility on first-pass reading on the verb was not significant, $F1$ (1, 32)=2.40; $F2$=1.71. The 19 ms parafoveal-on-foveal effect of verb-NP2

TABLE 9

Mean first-pass reading time in NP2 (verb unmasked). The data are shown as a function of NP1-verb and verb-NP2 plausibility

	Plausible verb-NP2	*Implausible verb-NP2*	*Mean*
Plausible NP1-verb	205	220	213
Implausible NP1-verb	211	235	223
Mean	208	227	

plausibility, although in the same direction as that shown in Table 3, was also not reliable, $F1(1, 32)=1.40$; $F2<1$, and there was no significant interaction, $F1(1, 32)=1.39$; $F2(1, 20)=1.44$. These data suggest that masking may have acted to delay[7] or distribute the effects of NP1-verb plausibility. This was tested in an analysis of reading times on NP2.

Table 9 shows average first-pass reading times on the final NP. The 10 ms NP1-verb plausibility effect was not, in fact, significant, $F1(1, 32)=1.73$; $F2< 1$. It would appear that masking the verb had the effect of distributing the NP1-verb plausibility effect in the form of two (nonsignificant) differences: on the verb itself, and on NP2. This was confirmed in an analysis combining the average first-pass reading time on the verb and NP2. Both the main effect of NP1-verb plausibility, $F1(1, 32)=4.44$, $p<.05$; $F2(1, 20)=3.28$, p=.08, and the verb-NP2 plausibility effect approached significance, $F1(1, 32)=8.18$, $p< .01$; $F2(1, 20)=2.67$, $p=.1$, but there was no interaction ($F1$ and $F2<1$).

One possible conclusion to be drawn from the fact that the set of words in the "implausible" conditions began with orthographically more familiar sequences is that this may have been responsible for changes in the number of refixations on the initial NP (something which also would be reflected in changes in inspection time[8]). As raised in the introduction to this paper, it is possible that a "magnetic" attraction of sublexical properties of parafoveally presented nouns acts to modulate foveal reinspection strategy (and hence inspection time). The crucial question, however, is whether systematic, albeit accidental, variation in initial letter familiarity in the present materials could be responsible for some or all of the otherwise inexplicable effects of "plausibility" on reading time when the verb was masked. Are apparent parafoveal NP1-verb effects related to CLF differences in the sets of nouns used in NP1? Can the otherwise inexplicable long-range effects of verb-NP2 "plausibility" be attributed to the CLF of the nouns employed in NP2 (albeit located some 15 characters to the right of the

[6] The "NP1-verb plausibility effect" when the verb-NP2 relationship was implausible (214 vs. 235 ms) was significant in the by-subjects analysis, F1(1, 32)=6.01, $p< .05$; F2(l, 20) =1.67.

point of current fixation)? The role played by the initial letter familiarity of the critical nouns was assessed formally by means of a reanalysis of those significant effects found in the two experiments, entering CLF of the items as a covariate. This was achieved by carrying out $F2$ analyses between items.[9]

It will be recalled that in Experiment 1 the measure of first-pass reading time in NP1 show a near-significant main effect of NP1-verb plausibility, but an interaction involving properties of NP2. The between-items reanalysis (without a covariate) adequately reproduces this pattern, with $F2(1, 80)=3.57$, $p=.06$ for the main effect and $F2(1, 80)=3.05$, $p=.08$ for the interaction. Entering CLF values for the nouns in NP1 as a covariate actually increased the significance of the main effect ($F2=4.35$, $p<.05$) and entering CLF values for the nouns in NP1 and NP2 had no effect on the interaction ($F2=3.04$). We may conclude that the effect of NP1-verb plausibility on first-pass reading time on NP1 in Experiment 1 was not attributable to variation in CLF of the nouns. This variability actually acted to increase error variance. This appreciably strengthens the possibility that the obtained parafoveal NP1-verb effects are real.[10]

In contrast, reanalyses of the number of fixations in NP1 leads to the conclusion that the variation apparently attributable to properties of verb-NP2 "plausibility" in Experiment 1 was actually related to CLF of the nouns used in NP2. An overall by-items analysis reproduces the modest (but nonetheless counterintuitive) main effect of verb-NP2 plausibility, $F2(1, 80)=3.53$, $p=.06$. However, entering CLF of the nouns as a covariate reduced this to nonsignificance, $F2(1, 80)=2.26$, $p=.14$. The apparent variation in launch position attributable to verb-NP2 plausibility meets the same fate. This approaches significance in the by-items analysis, $F2(1, 80)=3.63$, $p=.06$, but is reduced to nonsignificance with CLF of the nouns entered as a covariate, $F2(1, 80)=1.96$, $p=.17$.

With regard to Experiment 2, the most puzzling effect relates to the apparent verb-NP2 "plausibility effect" on final fixation duration in NP1 ("puzzling", because the verb was not visible at this point). In the by-items analysis of Experiment 2, this main effect of verb-NP2 "plausibility" was reduced from a value of $F2=3.35$ to $F2(1, 80)=1.89$, $p=.18$. However, crucially in the present context, entering the CLF of nouns allocated to NP2 as a covariate eliminated the effect entirely ($F2<1$). We conclude that the apparent effects of verb-NP2 relationships arose artefactually from parafoveal analysis of the initial letters of the critical NP2 noun. It should be recalled that although these letters were

---

[7] The outcome is consistent with the finding of Inhoff and Rayner (1986) that when a prior word is masked, frequency effects are not evident in initial fixation duration.

[8] Pynte et al. (2004) argue for separate influences over foveal inspection time from parafoveal informativeness and foveal orthographic familiarity.

[9] It needs to be borne in mind that there is some loss of power treating the design in this way, because the variables matched across item sets cannot be treated as "within-items".

relatively remote (12–15 characters to the right), the intervening text was in part masked and probably offered less interference than might normally occur.

It appears, therefore, that an account for some of the "long-range" effects can be provided by an appeal to variation in at least one sublexical property of the relevant nouns. But it is equally clear that variation in this property *cannot* account for the parafoveal pragmatic effect found in NP1. Indeed, variation in CLF appeared to act to reduce the reliability of that effect. We cannot, of course, claim that CLF is exactly the property the system is sensitive to. Measured CLF may be something partially correlated with the underlying mechanism. It is also entirely possible that the system might be affected by different orthographic properties foveally and parafoveally (see Footnote 7). Nonetheless, we are left with a plausible explanation of the "unconventional" long-range effects: They relate to an orthographic property (CLF or something similar) of the distant noun. This effect may be more marked when the verb is masked, but is probably present even with intervening text visible.

Can we claim that the parafoveal effect of NP1-verb relationships on NP1 inspection time (as distinct from refixation rate) is genuinely a pragmatic influence? Clearly, the effect is unrelated to CLF. On the other hand, there appears to be some hint of an effect, albeit nonsignificant, in Experiment 2, where it cannot possibly relate to plausibility. The outcome sits disturbingly in "no mans land", neither different from, nor the same as, the effects previously shown. One possibility is that some proportion of the original effect reported by Murray and Rowan (1998) was driven by (unknown) orthographic differences between the two classes of nouns used in NP1, with perhaps an additional genuine parafoveal pragmatic effect. Another possibility is that the pattern of effects present in Experiment 2 was generated in a different way from that found in previous studies. While we cannot be sure, there is some evidence in favour of this latter possibility: The pattern in Experiment 1 was driven principally by faster reading times in the plausible-plausible combination, while in Experiment 2 it was driven by *slower* times in the implausible-implausible cell, with a significant three-way interaction between experiment, NP1, and NP2 plausibility. Also, of course, if it were a "standard" foveal orthographic effect, we would have expected to see it in the results of Rayner et al. (2003). We are therefore inclined to suggest—albeit tentatively—that it is something about the masking procedure itself that potentiates the influence of long-distance orthographic properties, in which case parafoveal-on-foveal interactions involving low-level properties of the word under current inspection become plausible.

[10] An identical conclusion can be drawn from an equivalent reanalysis of the data of Murray and Rowan (1998). The strongest parafoveal-on-foveal effect in that study related to the duration of the final fixation on NP1. Entering CLF of NP1 as a covariate in a (between-items) analysis *increased* the significance of the effect from $F2(1, 80)=14.53$, $p<.01$ to $F2(1,80)=16.15$, $p<.01$.

GENERAL DISCUSSION

As the introduction to this paper makes clear, the claim that properties of an uninspected parafoveal word influence foveal inspection time has proved controversial. In part this is because abandoning the commitment to serial allocation of attention in favour of some form of parallel processing model is theoretically costly. In part it is because the effects, if real, are relatively small and sometimes difficult to demonstrate; sensitive to a range of interacting variables; and possibly task-specific (see Hyönä & Bertram, 2004). Taken together, the results of the present two experiments demonstrate four specific parafoveal-on-foveal effects. (1) In Experiment 1, first-pass reading time in NP1 is influenced by the plausibility of the NP1-verb relationship. The effect is consistent across this study and the one conducted by Murray and Rowan (1998), and appears in both first-pass reading time and the duration of the last fixation in both cases. It appears not to be an effect attributable to variation in lexical or sublexical properties of the nouns employed in either NP1 or NP2. The effect is in an orthodox direction, with longer reading time in NP1 when the following verb, before it is directly inspected, combines with it to form an implausible relationship. (2) In Experiment 1, the number of fixations made on the words comprising NP1 was apparently modulated by the plausibility of the verb-NP2 relationship. We believe we have satisfactorily demonstrated that this effect is attributable to an accidental association between assigned "plausibility" and variation in the cumulative lexical frequency (or "familiarity") of the initial string of letters in the nouns employed in NP2. While the outcome rules out interpretation in terms of plausibility, it is, nonetheless, a parafoveal-on-foveal effect—and one that is remarkably strong, given that the initial letters in NP2 were physically quite remote. (3) Similarly, variation in the duration of the final fixation on NP1 in Experiment 2 appears to be attributable to sublexical properties of the nouns involved, and this modulation must be interpreted as a parafoveal-on-foveal influence, although perhaps one potentiated by the effects of the masking procedure itself. (4) Finally, in Experiment 1 there is what appears to be a parafoveal orthographic effect exerted from NP2 influencing reading time on the verb itself. Combined analyses across these data and those of Murray and Rowan, show a consistent effect of this variable on the duration of the final fixation on the verb. Clearly, this is from a location closer to the critical noun, but it is nonetheless another instance of a parafoveal-on-foveal effect. We interpret this as a sublexical effect because of its direction and the fact that "conventional" effects of genuine plausibility show up in the duration of later fixations falling in NP2.

In summary, concurrent processing interactions involving properties of both adjacent and nonadjacent words are ubiquitous in this task. At a sublexical level, effects can be found across a variety of regions in both studies. It is, nonetheless, necessary to be cautious when assigning the modulation of foveal inspection time in NP1 to rated plausibility of the NP1-verb combination. The effect clearly replicates across experiments, even though the participants appeared to deploy somewhat different reading strategies, and we are inclined to interpret it as a genuine effect of pragmatic plausibility. But it is one which is overlaid by other low-level orthographic effects. Certainly, some orthographic effects are to be found in one cell of the design in Experiment 2, but we attribute the form and magnitude of these to the effects of the masking procedure. The most cautious claim is that the effect replicates and we are strongly inclined to attribute it to high- rather than low-level processes.

We also raised in the introduction to this paper the general question as to whether such effects are *only* found in artificial laboratory tasks and obviously this cannot be settled with the present data. It is possible to argue that the first sentence displayed in this matching task was read in a normal fashion because measures of average fixation duration, reading time and saccade extent, in fact, fall well within the normal range. Equally, it must be accepted that the task may have permitted, or even encouraged, a rather shallow level of processing, with a focus on those physical properties necessary to make a decision on presentation of the second sentence. To the degree that the range of attention may be wider in search tasks than in normal reading, the obtained "long-range" effects of properties of NP2 imply that sentence matching may not engage strictly normal reading processes. Against this stands the clear-cut evidence of the effects of the pragmatic plausibility of the text—the task cannot be all that shallow. Even with the different reading strategies apparently deployed by participants in the study by Murray and Rowan (1998) and in the present replication (Experiment 1), many effects appear identically in both studies and differences, where these occur, seem to be attributable to slight differences in the onset time of the effects. In other words, the effects are robust, although what we have called "strategy" may influence the exact point at which they occur.

In the final analysis, the distinction between the properties of normal reading and other "reading-like" tasks may turn out to be unhelpful. Our working hypothesis is that normal reading must engage the component processes isolated in many laboratory tasks, including the one employed here (e.g., the identification of letters and letters sequences, lexical processing, the assignment of syntactic structure, etc). Indeed, it would be very strange indeed if the process of normal reading were seen as necessarily restricted to a "one word window", with resources available to broaden the window available, just so long as the reader is asked to complete some other task. In which case, the more interesting question would be why such resources are less often deployed in normal skilled reading.

Experiment 2 was carried out to restrict the range of possible interpretations of the data in Experiment 1. If the verb in a NP1-verb-NP2 sequence is masked before its direct inspection, it is impossible for its properties (apart from its physical length) to have any influence on NP1 inspection time. The manipulation was clearly successful and provided further evidence of the important role played by preview in sentence processing. The mask delayed computation of the NP1-verb relationship until after the verb was directly inspected. Less obviously, this mask-induced processing delay appeared to persist for several words, with the verb-NP2 plausibility effect only emerging gradually. Masking the verb also seemed to act to increase the long-range influence of NP2 on the processing of NP1. We conclude that properties of parafoveal words do influence foveal inspection time and that masking may increase, rather than reduce, the range over which such effects occur. Since masking is routinely used in experimental work to control for preview, it is important to note that the present data suggest it has other, less specific, effects. In fact it should be borne in mind that if readers routinely take in information from more than one word at a time, the presence of masking might not simply deny access to certain information. The process of denial itself may influence the *way* in which the text is inspected.

To what degree do the present data qualify the conclusions of Murray and Rowan (1998) regarding parafoveal pragmatic effects? Orthographic factors exert an influence on inspection time, but statistical control for one property, which appears to exert a long-range effect, served to strengthen, rather than weaken, the parafoveal effect of NP1-verb plausibility. We are led to suggest that the nonsignificant (but admittedly nonzero) effect found in Experiment 2, when the verb was masked, resulted from a complex interaction of low-level properties of the currently fixated noun and similar long-range effects, potentiated by the presence of masking. If this is the case, then properties of an as yet uninspected verb in combination with NP1 (i.e., a genuinely pragmatic effect) do modulate foveal inspection time, although the extent to which this is specific to sentence matching, or indeed to laboratory tasks in general, remains open. But, if the effect is real, it would be remarkable if such a process could not be involved in normal reading. Researchers involved with eye movements and reading have long known that effects do not always unfold at exactly the same point in time and show up in exactly the same situation and on exactly the same measure across different studies. Typically that has not led to the conclusion that the effects are not real and it seems unreasonable to conclude that parafoveal-onfoveal effects do not exist simply because they show up more readily in some measures than others. Quite apart from the evidence of long-range effects related to orthographic properties of the words employed, the results of the two experiments considered here are consistent with the operation of a genuine parafoveal-on-foveal effect attributable to purely pragmatic relationships involving the initial noun and verb in the sentences employed.

PrEview proof published online September 2003

REFERENCES

Beauvillain, C., & Beauvillain, P. (1995). Calibration of an eye movement system for use in reading. *Behavior Research Methods, Instruments and Computers, 55,* 1–17.

Carpenter, P.A., & Just, M.A. (1983). What your eyes do while your mind is reading. In K.Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–305). New York: Academic Press.

Clifton, C., Bock, J., & Rado, J. (2000). Effects of the focus particle only and intrinsic contrast on comprehension of reduced relative clauses. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 591–620). Oxford, UK: Elsevier Science.

Henderson, J.M. (1992). Visual attention and eye movement control during reading and picture viewing. In K.Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 260–283). New York: Springer-Verlag.

Henderson, J.M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47,* 201–221.

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology, 16*(1/2), 104–127.

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy. R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 221–246). Oxford, UK: Elsevier Science.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40,* 431–439.

Kennedy, A. (1995). *The influence of parafoveal words on foveal inspection time.* Paper presented at the AMLaP-95 conference, Edinburgh, UK.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–223). Oxford, UK: Elsevier Science.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kennedy, A., Murray, W.S., Jennings, F., & Reid, C. (1989). Comments on the generality of the principle of minimal attachment. *Language and Cognitive Processes, 4,* 51–76.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Konieczny, L. (1996). *Human sentence processing: A semantics oriented parsing approach.* Unpublished doctoral dissertation, University of Freiburg, Germany.

Kuçera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

McConkie, G.W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.

McConkie, G.W., & Zola, D. (1984). Eye movement control during reading: The effect of word units. In W.Prinz & A.F.Sanders (Eds.), *Cognition and motor processes* (pp. 63–74). Berlin, Germany: Springer-Verlag.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Murray, W.S. (1998). Parafoveal pragmatics. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 181–199). Oxford, UK: Elsevier Science.

Murray, W.S. (2000). Sentence processing: Issues and measures. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds). *Reading as a perceptual process* (pp. 649–664). Oxford, UK: Elsevier Science.

Murray, W.S., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycho-linguistic Research, 27,* 1–22.

O'Regan, J.K. (1990). Eye movements and reading. In E.Kowler (Ed.), *Eye movements and their role in visual and cognitive processses* (pp. 395–453). Amsterdam: Elsevier Science.

Pickering, M.J., & Traxler, M.J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 940–961.

Pynte, J., Kennedy, A., & Ducrot (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1/2), 178–202.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K., & Bertera, J.H. (1979). Reading without a fovea. *Science, 206,* 468–469.

Rayner, K., & Fischer, M.F. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics, 58,* 734–747.

Rayner, K., Well, A.D., Pollatsek, A., & Bertera, J.H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31,* 537–550.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier Science.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 247–268). Oxford, UK: Elsevier Science.

Vitu, F. (1991). The influence of parafoveal pre-processing and linguistic context on the optimal landing position effect. *Perception and Psychophysics, 50,* 58–75.

Vitu, F., Brysbaert, M., & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology, 16* (1/2), 154–177.

White, S.J., & Liversedge, S.P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*(1/2), 52–78.

APPENDIX

Experimental materials

The savages/uranium smacked the child/money

The carpenter/ambulance cleared his throat/finger

The fugitive/pedigree groomed the horses/forest
The constable/islanders answered the telephone/orchestra
The tutor/trout delivered his sermon/bubble
The robbers/newborn rehearsed the part/meal
The hostess/charity weighed the carcass/sleeves
The lecturer/princess delivered the packages/wardrobe
The burglars/doorbell amused the crowd/grass
The comedians/lyricists picked the lock/lift
The labourer/bacteria stalked the tiger/algae
The housewife/alligator loaded the rifle/chair
The scientist/guerrilla arrested the criminals/ambulance
The policeman/therapist tested the chemicals/saxophone
The tailor/Libyan took the tickets/scenery
The bishop/knight carried the luggage/buttons
The guard/saint cooked the meal/beer
The hunters/bishops stacked the bricks/tulips
The servant/lawyers locked the gates/drums
The porter/rebels heard the organ/dolls
The vicar/beast corrected his pupil/giant
The soldiers/treasury were tipped by the customers/musicians
The politician/expedition wore the costumes/cylinder
The butcher/witches welcomed the guests/horses

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 154–177

A test of parafoveal-on-foveal effects with pairs of orthographically related words

Françoise Vitu

CNRS, Université René Descartes, Boulogne-Billancourt, France

Marc Brysbaert

Royal Holloway, University of London, UK

Denis Lancelin

CNRS, Université René Descartes, Boulogne-Billancourt, France

One of the main controversies in the field of eye movements in reading concerns the question of whether the processing of two adjacent words in reading occurs in sequence, or in parallel. To distinguish between these views, the present experiment tested the presence of parafoveal-on-foveal effects with pairs of orthographically related words (or neighbours that differed by a single letter) in a controlled but reading-like situation. Results revealed that fixation times on a foveal target word were shorter when the target was accompanied by an orthographically similar parafoveal word than when the parafoveal word was dissimilar. Furthermore, the size of the effect tended to vary with both the relative frequency of target and parafoveal words, and the position of the critical letter. These results were interpreted in the framework of a pure parallel processing hypothesis, where the processing of adjacent words is only limited by visual acuity, and the respective lexical properties of the foveal and the parafoveal words.

Since the early 1990s, a large body of research has been devoted to the study of visual attention in reading. One of the main issues relates to the question of whether the visual information from two adjacent words is processed in parallel or in sequence. As suggested in several models of eye guidance in reading, it may well be that the processing of the fixated word comes first and only when this word has been identified does the processing of the next word in parafoveal vision begin (Henderson & Ferreira, 1990; Morrison, 1984; Rayner, Reichle, & Pollatsek, 1998; Reichle, Pollatsek, Fisher, & Rayner, 1998). Several different mechanisms have been proposed although the basic principle remains the same. An attentional process filters the visual information and only the information in the attention spotlight is processed up to the word level. At the beginning of a fixation, the spotlight (usually) is on the fixated word and, when the fixated word is identified, it shifts to the next word. Parafoveal processing is then initiated, and will last until a saccadic eye movement brings the eyes onto this word (or onto the next word if the attentional spotlight has shifted again in the mean time). According to Morrison (1984), the programming of a saccade immediately follows an attention shift. In contrast, the E-Z Reader model proposed by Reichle et al. (1998) assumes that saccadic programming starts as soon as a word familiarity check has been completed, which means that it always precedes an attention shift. Thus, in both models, parafoveal processing time cannot exceed the saccadic programming time, and according to the E-Z Reader model, it is even shorter, being then a function of the difficulty of the foveal word.

On the other hand, we could also assume that rather than being processed in sequence, all words within the perceptual span are processed in parallel. This view has been expressed in different variants depending on the role given to attention. At one extreme, and still very close to the sequential attention shift hypothesis, the gradient-shift assumption posits that attention is distributed over several words, with more resources being allocated to the word in the fovea and gradually diminishing towards the word(s) in the parafovea. In addition, the attention distribution is modulated by the processing difficulty, with more resources being absorbed by words that are difficult to process (Engbert, Longtin, & Kliegl, 2002; Inhoff, Radach, Starr, & Greenberg, 2000a; Inhoff, Starr, & Shindler, 2000b; see also, McConkie, 1979; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999). At the other extreme, a pure parallel hypothesis envisages

---

Correspondence should be addressed to F.Vitu-Thibault, CNRS, Université René Descartes, Laboratoire de Psychologie Expérimentale, 71 avenue Edouard Vaillant, 92 774 BoulogneBillancourt Cedex, France. Email: Francoise.vitu@psycho.univ-paris5.frThe present project was developed with the great help of E.Simonatti (Erasmus Program 1999– 2000, University of Padova, Italy) and A.Francisco (University René Descartes) who participated in the preparation of the list of stimuli, and the running of preliminary studies (not reported here). The stimuli were selected under the advice of L.Ferrand and J.Grainger, whom we would like to thank. We would also like to thank W.Ward (Fourward Optical Technologies, Inc.) for his great help in setting up and maintaining the eye tracker, and J.van Rensbergen, P.de Graef, and P. van Diepen (Catholic University of Leuven, Belgium) for their great help in the use of the eye movement software. Finally, we would like to thank A.Kennedy, S.Liversedge, K.Rayner, and an anonymous reviewer for their helpful comments on an earlier version of the manuscript.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000178

that attention does not intervene at all in the processing of the visual information in reading, and that visual acuity is the primary variable that determines which word in the perceptual span is identified first, and when the eyes should move (see Schroyens et al., 1999, for a similar view; see also Schiepers, 1980). Due to visual acuity constraints, parafoveal word processing lags behind the processing of the fixated word. This ensures that in most instances, the winner of the competition is the fixated word, and therefore questions the need for selective processing. In between the above two views, a process-monitoring hypothesis was recently proposed, which does not clearly state the role of visual attention in the processing of adjacent words, but makes the proposal that a signal to move the eyes may result from difficulty being encountered in the processing of either the foveal or the parafoveal word, rather than from the foveal word only (Kennedy, 1998; Kennedy, Pynte, & Ducrot, 2002). Thus, there would be some kind of trade-off in the time spent fixating each of two adjacent words. Furthermore, how much a word is processed in parafoveal vision would be self determined rather than being a function of the time required to process the fixated word, and/or how much resources are allocated to the foveal word.

In favour of a parallel processing view, several arguments were recently put forward which question the validity of the basic assumptions made in sequential processing models. First, Schroyens et al. (1999) noted that parafoveal preview effects get larger with increasing fixation time, suggesting that the processing of the parafoveal word is not limited to the saccadic programming time as was postulated by Morrison (1984). This result is also opposite to the E-Z Reader's assumption that parafoveal processing is initiated later as the difficulty of the foveal word increases since this would predict that parafoveal preview benefit decreases as fixation time increases (Reichle et al., 1998). In addition, the finding that parafoveal processing is less when the foveal load is high (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Schroyens et al., 1999) does not seem to originate from a sequential processing of the foveal, and parafoveal words. Rather, this would be the result of an interaction between processing associated with both the foveal and the parafoveal word that would operate while the eyes are still on the foveal word, or once they have landed on the parafoveal word. Indeed, as noted by Schroyens et al., foveal-on-parafoveal effects are associated with spillover effects (or instances where processing of the foveal word is not terminated by the time the eyes land on the parafoveal word), and they emerge only at rather short fixation times (less than 240ms). Furthermore, the likelihood of these effects depends on the frequency of the parafoveal word. These findings, which question but cannot completely rule out the hypothesis of sequential processing, call for clearer evidence for or against a sequential view.

To distinguish between a sequential vs. parallel view for the processing of both foveal, and parafoveal word information, several studies tested whether the processing associated with a fixated word is influenced by the characteristics of the word located in parafoveal vision, starting from the assumption that only a parallel view can predict the presence of parafoveal-on-foveal effects. The effects of different variables were tested including the visual presentation mode for the parafoveal word (i.e., upper vs. lower cases), the orthographic legality of the parafoveal letter sequence, the length and frequency of the parafoveal word or the frequency of the initial letter sequence in the parafoveal word, the semantic relatedness between the foveal and parafoveal words, and the plausibility of the parafoveal word in the sentence (Baccino, Lavigne, Gragnic, & Vitu, 2001; Henderson & Ferreira, 1993; Inhoff et al., 2000a, 2000b; Kennedy, 1998, 2000; Kennedy et al, 2002; Murray, 1998; White & Liversedge, 2004; see also Altarriba, Kambe, Pollatsek, & Rayner, 2001; Bradshaw, 1974). Across studies, parafoveal-on-foveal effects do not appear to be systematic. When the low-level visual and orthographic properties of the parafoveal stimulus are manipulated, parafoveal-on-foveal influences can be observed which result in longer fixation times with unexpected visual or orthographic patterns in the parafovea (Inhoff et al., 2000b). However, this was not confirmed in a recent study (see White & Liversedge, 2004). Furthermore, when it comes to higher level lexical or semantic variables, the picture gets even less clear. First, while some studies fail to report an effect of the parafoveal word frequency on foveal viewing inspection time (Carpenter & Just, 1983; Henderson & Ferreira, 1993), other studies indicate either an increase or a decrease in foveal processing time with low-frequency parafoveal words (Kennedy, 1998, 2000). Second, several studies indicate the presence of semantic and pragmatic parafoveal-on-foveal effects in both isolated word recognition and sentence reading, but the effects are often conditional upon the distance of the parafoveal word from the fixation location and/or the viewing conditions (Baccino et al., 2001; Inhoff et al., 2000a, 2000b; Murray, 1998). Finally, some results suggest to the contrary that semantic information cannot be extracted in the parafovea, and that semantic parafoveal-on-foveal effects are unlikely to occur (Altarriba et al., 2001; see Rayner, White, Kambe, Miller, & Liversedge, 2003, for a review).

The finding of lexical and semantic parafoveal-on-foveal effects may, how ever, be critical to distinguish between parallel and sequential processing hypotheses. Indeed, low-level visual and orthographic parafoveal influences could be accommodated in sequential attention shift models, assuming that attention is not always perfectly focused on the foveal word at the beginning of a fixation, and that low-level visual information can be extracted in the parafovea before it is completely filtered out, or selective foveal processing operates. As pointed out by Treisman and Souther (1986), the fact that letter migration effects may occur in the processing of two orthographically similar words presented simultaneously in parafoveal vision (see Mozer, 1983) is not strict evidence for a parallel processing of both words. Rather, this would indicate that the filtering of the irrelevant visual information occurs only after the letter level of word processing.

According to Kennedy et al. (2002), the discrepancy in the reported lexical and/or semantic effects comes from the fact that in most studies length and frequency of the foveal word were not controlled, and their data support to some extent this contention. Another possibility relates to the fact that parallel processing does not necessarily imply parafoveal-on-foveal influences. Depending on the respective time course of the processing associated with both foveal and parafoveal words, the characteristics of the parafoveal word may or may not affect the time needed to process the fixated word, and these may either facilitate or interfere with foveal word processing. A primary constraint relies on visual acuity, or the fact that more visual information accumulates from the foveal word, which results in it to be processed more rapidly than the parafoveal word (see Schiepers, 1980). Additional constraints come from the respective lexical properties of each word, and how many of these properties are common to both words. In past attempts to test the presence of parafoveal-on-foveal effects, the pairs of words selected for a presentation in fovea and parafovea differed on several dimensions (orthography, phonology, etc.), and several word properties that are critical to word recognition were not controlled. Both foveal and parafoveal words were therefore processed with different unpredictable speed, and predictions can hardly be made of whether the manipulated characteristics of the parafoveal word should influence (and how) processing of the fixated word.

To shed some light on the question of sequential vs. parallel processing, and to determine whether the occurrence of parafoveal-on-foveal effects is conditional upon the time course of processing associated with both the foveal and the parafoveal word, we investigated parafoveal-on-foveal effects with pairs of stimuli that differed by only one characteristic. As in Segui and Grainger's (1990) orthographic priming study, pairs of orthographic neighbours (or words that differ by a single letter, as defined by Coltheart, Davelaar, Jonasson, & Besner, 1977) were selected, with one of the words being a higher frequency neighbour of the second word. Instead of being presented in sequence as in the original experiment, both prime and target words were presented simultaneously in parafoveal and foveal vision respectively (in the present experiment, the previously so-called prime will be referred to as the parafoveal word). For comparison, the parafoveal word of each related pair was replaced in a control condition with an orthographically unrelated word, but of a similar frequency as the related parafoveal word. In addition, to manipulate a lexical variable, the frequency of both parafoveal word and target was varied, using as a parafoveal word the higher or lower frequency word of a pair, and vice versa for the target. For sake of clarity, the frequency factor will from now on, be referred to as target frequency. Care was taken that both words in a pair were never semantically related, and homophones were avoided as much as possible.

There is a debate going on in the literature whether orthographic neighbours help or hinder the processing of a target word. In general, it is thought that orthographic neighbours facilitate processing (Andrews, 1997), except when a not-fully processed high-frequency neighbour precedes a low-frequency target (see Ferrand, 2001; Grainger & Jacobs, 1996, for complete reviews). This is the case when the high-frequency neighbour is used as a masked prime (Segui &

Grainger, 1990) or when the high-frequency neighbour is used as the parafoveal preview in a boundary technique (i.e., the parafoveal view is present as long as the reader's eyes have not landed on the word; once the eyes cross the boundary in front of the word, the preview is changed to the target word; Pollatsek, Perea, & Binder, 1999).

Our prediction is quite straightforward. According to a sequential model of word processing, the qualities of the parafoveal word will not have an effect on the fixation time on the foveal word, although there might be two exceptions to that scheme. First, as mentioned above, early facilitation effects due to feature/letter similarity (or a shortening in fixation times in related compared to unrelated cases) could still be reconciled with a sequential hypothesis as long as the effects do not vary with the relative frequency of the target and parafoveal words. Another exception may also occur if the parafoveal word is more likely to be skipped in one condition than in the other, in which case the preceding fixation duration should increase according to the E-Z Reader model. In contrast, a pure parallel processing view does predict the presence of parafoveal-onfoveal influences, and it additionally assumes that the size and/or the direction of the effects should vary with the relative frequency of both words. In particular, given that low-frequency words are usually associated with longer fixation times, and that high-frequency words are more easily processed in parafoveal vision than low-frequency words, parafoveal preview would be more efficient when low-frequency targets are paired with high-frequency parafoveal words than in the opposite condition (see Inhoff & Rayner, 1986; Schroyens et al., 1999; Vitu, 1991). This could result in greater facilitation in the former case as the majority of findings on orthographic priming point in the direction of facilitation (see for a review, Ferrand, 2001). On the other hand, if a word in parafovea can be compared with a not-fully processed word, then an inhibition may occur at least when a low-frequency target is accompanied by a higher frequency neighbour in the parafovea (see Segui & Grainger, 1990). Finally, the gradient-shift hypothesis proposed by Inhoff et al. (2000b; Inhoff et al., 2000a) also predicts the presence of parafoveal-on-foveal influences, but these would result in an overall facilitation when both words are similar, that would be independent of the relative frequency of target and parafoveal words. Since more resources are allocated to low-frequency foveal words, parafoveal processing should be reduced in those instances, and the amount of facilitation due to feature/letter similarity should not differ from when a high-frequency target is presented with a low-frequency parafoveal word. It must be noted that unlike sequential attention shift models, the gradient-shift hypothesis does not imply that parafoveal-on-foveal effects would be limited to a short time window at the beginning of the fixation.

EXPERIMENT

In our experiment, we tested the presence of parafoveal-on-foveal effects with pairs of isolated words, but in a reading-like situation. On each trial, two four- or five-letter words were presented simultaneously. These two words were either orthographic neighbours or unrelated words. The unrelated trials were made by taking the "foveal" target word of the neighbour pair and replacing the "parafoveal" word by an unrelated word of the same length and frequency. In all related pairs, one of the words was a higher frequency neighbour of the other word (e.g., avec [with] vs. aveu [confession]). In half of the trials, the highfrequency neighbour was used as target and the low-frequency neighbour as parafoveal word; in the other half the assignment was reversed (counterbalanced across participants). So, we had four types of stimuli (first word=target; second word=parafoveal word): "avec-aveu", "avec-pipe" [pipe], "aveu-avec", and "aveu—mais" [but]. Pairs of neighbours were chosen that had different deviating letters. For the four-letter words, a distinction was made between: (1) different first letter, (2) different middle letter, or (3) different last letter. For the five-letter words, we were able to have an equal number of words that differed at each letter position.

To mimic a normal reading situation, the two words were inserted between two x-letter strings. Participants were asked to jump from the first string of xletters to the first word (or target), then to the second (or parafoveal) word, and finally to the last x-letter string. Their task was then to determine whether one of the two words referred to an animal (none of the test trials contained a word referring to an animal; these were filler trials). In the present situation, reading of the target word was therefore embedded in a series of left-to-right eye movements, and as suggested by previous studies, the eye movement pattern that characterises the reading of our isolated words should be similar to the pattern observed in forward text reading (Kennedy, 2000; Vitu, 1993). In normal text reading, word processing times range from an average gaze duration of 320 ms for low-frequency words to an average gaze duration of 260 ms for highfrequency words (Schilling, Rayner, & Chumbley, 1998). On the other hand, it must be noted that the target and parafoveal words, although being initially presented in parafoveal vision, were masked until an eye movement that crossed an invisible boundary in front of the target word was detected (see Rayner, 1975). This was done to ensure that the visual information related to either word started being extracted at the same time in all conditions.

Method

Participants. Forty-eight psychology students from the University Rene Descartes, and the Catholic University of Paris (Ecole des Psychologues Praticiens) participated in the experiment which was run at the Laboratory of Experimental Psychology, in Boulogne-Billancourt, France. Participants were between 20 and 30 years old, they were all native French speakers, and they had normal, or corrected-to-normal vision (in the latter case, only participants wearing glasses were accepted). All participants were naïve regarding the purpose of the experiment.

Stimulus materials. Pairs of target and parafoveal words were either four or five letters long. For the four-letter words, 120 pairs of orthographic neighbours were selected from the French corpus *Trésor de la langue française* (1971) using the following criteria. Each pair differed in only one letter, controlled as much as possible over the different positions of the word: 40 differed in the first letter, 40 in either the second or the third, and 40 in the last letter (see example stimuli in Table 1). Both words differed in terms of their frequency of occurrence in such a way that one of the words was a higher frequency neighbour of the other word. The mean and the median frequency for the high frequency words corresponded to 265, and 48 occurrences per million (with a minimum and a maximum of 4.6 and 9600 occurrences per million); these values were 7.1 and 2.2 occurrences per million for the lower frequency words (with a minimum and a maximum of 0.1 and 109 occurrences per million).

In the experiment, each word of the neighbour pair could serve both as target and as parafoveal word (counterbalanced over participants). Therefore, two sets of 120 related words were derived from the original list, one where the target words corresponded to the high-frequency neighbours, and one where the target words corresponded to the low-frequency neighbours. Two corresponding sets of 120 unrelated items were then constructed, using the same words as targets, but replacing each parafoveal word with an unrelated word that was matched in frequency to the orthographically related word it was replacing (see Table 1). So, the mean and the median frequency of the first set of unrelated words (replacing the high-frequency neighbours) were 250 and 41; those of the second set of unrelated words (replacing the low-frequency neighbours) were 5.9 and 2.1. Unrelated words had no more than two letters in common with the target word, and these two letters were never at the same position.

All in all, there were four conditions, two with a high-frequency target paired with a low-frequency related or unrelated parafoveal word, and two with a lowfrequency target paired with a high-frequency related or unrelated parafoveal word: “avec-aveu”, “avec-pipe”, “aveu-avec”, “aveu-mais” (where the first word corresponds to the target, and the second word to the parafoveal word). In both related and unrelated word pairs, words were never semantically related, and only a few pairs of related words were homophones (2%).

For the five-letter words, a list of 160 pairs of orthographic neighbours was constructed, with words differing in the first, second, third, fourth, or fifth letter. In each pair, both words differed in terms of their frequency of occurrence. The mean and median frequency corresponded to 97, and 40 occurrences per million for higher frequency words (with a minimum, and a maximum of 3.4, and 1624 occurrences per million), and they corresponded to 3.6, and 1.2 occurrences per

TABLE 1

Example pairs of four-letter orthographically similar words that differ by their first, third, and last letter, with the corresponding pair of unrelated words

	Target word	*Parafoveal word*
Letter 1		
Related	pour (for)	four (oven)
Unrelated	pour (for)	clan (group)
Letter 3		
Related	bord (border)	bond (step)
Unrelated	bord (border)	clos (ended)
Letter 4		
Related	mien (mine)	miel (honey)
Unrelated	mien (mine)	clou (nail)

The first word corresponds to the target word, and the second word to the parafoveal word. In the example, the target word is of a higher frequency than the corresponding parafoveal word. Translation is given in parentheses.

million for lower frequency words (with a minimum, and a maximum of 0.01, and 63 occurrences per million). Two sets of 160 related word pairs were derived from the original list with the target corresponding either to the lower or higher frequency word of the pair. These two sets were matched with two sets of 160 unrelated word pairs where higher, and lower frequency target words respectively were each paired with an orthographically unrelated word of a similar frequency as the related parafoveal word; the median for the corresponding lower, and higher frequency parafoveal words corresponded to 1. 3, and 37 occurrences per million, respectively (mean of 3.3, and 122 occurrences per million). Both words had no more than two letters in common, and these two letters were never at the same position in both words. In both related, and unrelated word pairs, the words were never semantically related.

Two sets of filler items were made, one consisting of 80 pairs of four-letter words, and the other consisting of 100 pairs of five-letter words. Of these filler items, respectively 28 and 36 contained an animal name; the other fillers were pairs of orthographically unrelated items in order to conceal the experimental manipulation. For each word length, the proportion of animal names was 13%, and the proportion of related words was 34%. It must be noted that the pairs that contained an animal name were in half the cases orthographically related, and in the other half unrelated, to discourage the participants from adopting specific strategies.

For both word lengths, two sets of 10 practice trials were prepared for presentation at the beginning of each block of trials (see below). These did not contain an animal name, and were all orthographically unrelated. In addition, two lists of 30 four- and five-letter practice trials were constructed for presentation at the beginning of the experiment. Here, both related and unrelated pairs were mixed with pairs of words that contained an animal name.

Design. For the four-letter words, a 2x2x3 within-subject design was used, with orthographic relatedness (related vs. unrelated), frequency of the target word (higher or lower frequency than the parafoveal word), and position of the critical letter (first, middle, or last letter of the word) as independent variables. For five-letter words, a 2x2x5 within-subject design was defined, with orthographic relatedness (related vs. unrelated), frequency of the target (higher or lower frequency than the parafoveal word), and position of the critical letter (first, second, third, fourth, or fifth letter of the word) as independent variables. To ensure that all words were seen in the different conditions across participants, and that each word or its associate was seen only once by each participant, a latin square design was used for both word lengths.

There was a total of 200 pairs of four-letter words, and 280 pairs of five-letter words. These were each presented in two separate blocks of trials, equated in terms of orthographic relatedness, target frequency, position of the critical letter, and number of filler, and of practice trials. Each block started with 10 practice trials, followed by the rest of the items. Words and conditions were presented in a random and different order for each participant in each block. For half the participants, the two blocks made of four-letter words were presented first, followed by the two blocks of five-letter words, and for the other half, four-, and five-letter words were presented in the reverse order. Within each word length, the order of the two blocks of trials was counterbalanced across participants. The first block of trials for each word length was preceded by a separate block of 30 practice trials to familiarise participants with the task, but also with the length of the presented words.

Apparatus. Eye movements were recorded using a fifth-generation Dual Purkinje Image eye tracker (Fourward Optical Technologies, Inc.), sampling the right eye position every millisecond with a spatial accuracy of 1 min of arc (Cornsweet & Crane, 1973). The eye tracker was interfaced with two IBMcompatible microcomputers. The first computer recorded the eye movement parameters, and analysed them online, using the software developed at the University of Leuven by Van Rensbergen and de Troy (1993). The second computer controlled the visual presentation of the stimuli. Eye movement parameters were continuously sent to the second computer, so that the visual display could be changed contingent on the position of the eyes.[1] The first computer was interfaced with two response buttons. The decision to start the next trial was based on the signal of the button press, which indicated the decision made by the participant.

Stimuli were displayed in graphics mode on a 17-inch CRT monitor with 60 Hz refresh rate. Each character space and each letter subtended respectively 0.37°, and 0.30° of visual angle, at a distance of 1075 mm from the participants' eyes. Viewing was binocular.

Procedure. When a participant arrived in the laboratory, he or she was seated in an adjustable chair. A bite bar was prepared to minimise head movements. After setting up the eye tracker for the participant, a calibration phase began. Calibration was made using 15 points presented successively on the entire screen (5 points on both diagonal axes, and 5 points around the central horizontal axis of the screen). The first calibration point was presented in the left upper corner of the screen until the participant pressed a button, which made the point disappear, and appear at another location. Participants were asked to press the button only when they were fixating very precisely at the displayed dot location. If the calibration was not satisfactory (the correlation between the actual and the estimated eye location was less than .99 for both horizontal and vertical coordinates), another calibration phase was initiated. Otherwise, a block of trials began.

Each trial began with the presentation of two vertically aligned fixation bars that participants were asked to fixate. When the computer detected that the eyes were fixating within a region of about half a character on each side of the fixation bars, the fixation bars were removed, and an x-letter string centred on the fixation bar was presented. Simultaneously, three masks that corresponded respectively to the target, and parafoveal word and an additional x-letter string were presented to the right of the fixation bar. Each mask consisted in a series of four or five random dot patterns (depending on word length) that contained approximately the same number of pixels as the letters of the corresponding stimuli. The four simultaneously presented stimuli were equally spaced (by about one and a half characters). As soon as a saccade crossed an invisible boundary that corresponded to the last letter of the first letter string, the two words and the last letter string became visible (see for the boundary paradigm, Rayner, 1975). Participants were asked to read both words silently, and then to move their eyes to the last letter string. When the computer detected a fixation to the right of the beginning of the last letter string, and in addition that a delay of 112 ms had elapsed from the beginning of the fixation, all stimuli were removed, and a question mark was displayed at the bottom of the screen, which told participants they could indicate whether an animal name had been presented. After one or the other button press, a "C"or a "F" was displayed together with a number, which indicated respectively whether the participant's response was correct or false, and the number of the past trial. After a delay of 2 s, the next trial began.

Data analysis. Data analyses examined the first pass eye behaviour on the target word (or first word of a pair) using the following selection criteria: (1)

---

[1] The display changes that were contingent on the execution of a saccade in a particular zone of the screen occurred on average at about two-thirds of the saccade duration (or at about 16 ms from the beginning of the saccade); display changes failed to occur during a saccade in about 6% of all trials across participants and conditions.

the display change that made the two words and the last letter string visible occurred only during a saccade, (2) there was no button press before the stimuli were displayed, or during fixation on the target word, (3) there was no blink or other signal irregularity during the first eye pass on the target word, as well as during the fixation preceding, and following the first eye pass, and (4) the first fixation on the target word corresponded to the first fixation following the display change (or the appearance of both words). In addition, although there was no selection based on where the eyes initially landed in the target word, fixations on the first x-letter string in front of the target word were not included in the analyses.

After selection, the amount of data available for analysis was greatly reduced for a few participants in several conditions. Since the number of items for each critical letter position in each condition was originally low (5–10 for four-letter words, and 8 for five-letter words), we decided to combine four-, and five-letter words in the reported analyses. Furthermore, critical letter positions were grouped in two classes, referred to as outer and inner letter positions. These included respectively the first and last letter of four-, and five-letter words, and letters 2–3 in four-letter words, and 2–4 in five-letter words. The present grouping procedure was based on previous data indicating that the visibility of the first and last letter of a word in parafoveal vision is enhanced compared to the visibility of inner letters, these being more strongly affected by lateral masking (Bouma, 1973). Furthermore, as suggested by Perea (1998), orthographic priming effects may vary between inner, and outer letter positions.

In all analyses, means or proportions were calculated for each participant, and these were then averaged across participants, such that the weight of each individual participant's contribution to the global mean was similar. Analyses of variance were run on the means for each participant or each item in each condition, and when there were missing data points, these were replaced with the mean of the corresponding condition. F values will be reported by participant (F_1), and by item (F_2).

Results

In a first control analysis, the distribution of initial landing sites in the target words was plotted in the different conditions (not reported here). This revealed first that in most instances the eyes landed towards the centre of the words. An analysis of variance run on the mean initial landing sites in the different conditions revealed that none of the effects were significant (Fs 1.29), which ensured that in all conditions, the presented words were seen in similar visual conditions.

The *mean gaze duration* on the target words (or summed fixation duration before the eyes leave the word) was calculated as a function of orthographic relatedness, and target frequency, for both inner, and outer critical letter positions.

Results showed significant main effects of target frequency and orthographic relatedness, $F_1(1, 47)=16.15$, $p<.0005$, $F_2(1, 278)=51.25$, $p<.0005$, and $F_1(1, 47)=9.95$, $p<.005$, $F_2(1, 278)=12.75$, $p<.0005$, respectively, and a significant three-way interaction, $F_1(1, 47)=5.87$, $p<.05$, $F_2(1, 278)=4.20$, $p< .05$. Figure 1 illustrates these effects. In this figure, it is clear that the effect of target frequency was present for both outer and inner critical letter positions, with high-frequency targets being read faster than low-frequency targets, $F_1(1,47)= 17.71$, $p<.0005$, $F_2(1, 143)=31.69$, $p<.0005$, and $F_1(1, 47)=9.18$, $p<.005$, $F_2(1, 135)=20.09$, $p<.0005$, respectively. For the effect of orthographic relatedness which indicated an overall facilitation, it was almost completely due to the low-frequency target words when the critical letter corresponded to an external letter, while it was mostly due to the high-frequency target words when the critical letter was an inner letter. The effect of orthographic relatedness was significant for both outer, and inner letter cases, $F_1(1, 47)=7.60$, $p<.01$, $F_2(1, 143)=5.86$, $p<.05$, and $F_1(1, 47)=4.02$, $p<.05$, $F_2(1,135)=7.08$, $p<.01$, respectively. The interaction between orthographic relatedness and target frequency was marginally significant for outer letters in the participant analysis, $F_1(1, 47)=2.80$, $p<.10$, $F_2(1, 143)=2.17$, and it was significant for inner letters in the participant analysis, $F_1(1, 47)=3.98$, $p<.05$, $F_2(1, 135)=2.05$.

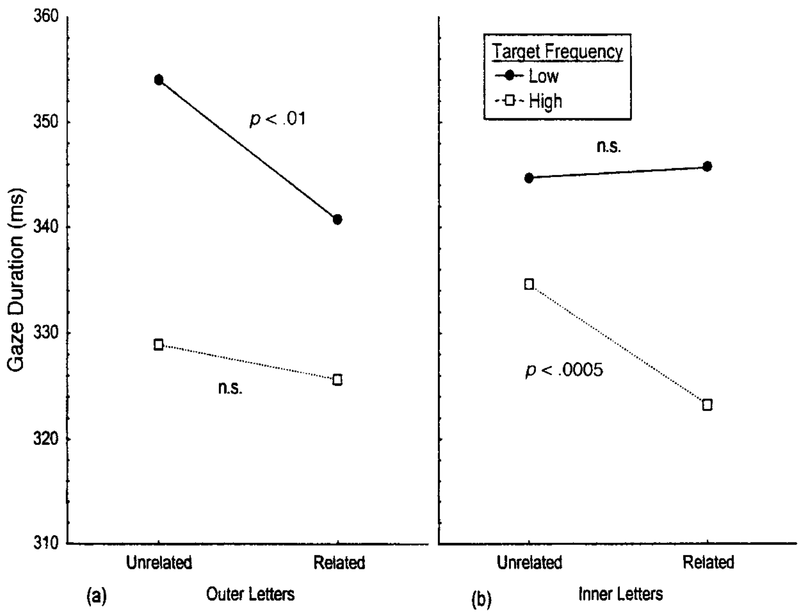

Figure 1. Mean gaze duration (in ms) on target words of high and low frequency as a function of orthographic relatedness, and separately for outer (a) and inner (b) critical letter positions. This analysis was run across four—and five-letter words.

TABLE 2

Likelihood of refixating the target word as a function of target frequency, and orthographic relatedness, separately for outer and inner critical letter positions. This analysis was run across four- and five-letter words

	Outer letters		*Inner letters*	
	Related	*Unrelated*	*Related*	*Unrelated*
Low-frequency targets	.27	.26	.27	.26
High-frequency targets	.26	.25	.26	.24

Since the gaze duration corresponds to the delay that elapses between the onset of the target word in foveal vision, and the moment the eyes move to the next word, the effects observed with this dependent variable may either result from differences in the refixation rate (or likelihood of making more than one fixation) or from differences in the duration of individual fixations. The data reported in Table 2 indicate that the target words were refixated in only 26% of the cases, and that the effects of orthographic similarity observed in the measure of gaze duration were not due to variations in the *refixation likelihood* since none of the effects reached significance in either participant or item analysis (*F*s 2. 65).

The results obtained with the *mean duration of single fixations* on four-, and five-letter words closely resembled those observed in the gaze duration measure (see Figure 2). The effect of target frequency was significant in both participant and item analyses, $F_1(1, 47)=8.53$, $p<.005$, $F_2(1, 278)=13.79$, $p<.005$, although being significant only for outer letter positions, $F_1(1, 47)=18.87$, $p<.0005$, $F_2(1, 143)=16.99$, $p<.0005$, and $F_1(1, 47)=1.60$, $F_2(1, 135)=1.22$, respectively for outer and inner letters. The effect of orthographic relatedness was marginally significant in the participant analysis, $F1(1, 47)=3.30$, $p<.10$, and significant in the item analysis, $F_2(1,278)=4.37$, $p<.05$. In addition, the three-way interaction was significant in the participant analysis, $F_1(1, 47)= 7.91$, $p<.01$, and marginally significant in the item analysis, $F_2(1, 278)=2.74$, $p<.10$. Looking at the data for both positions of the letter changed, we obtained a pattern that looked very similar to the one displayed in Figure 1. The only real difference between both sets of findings was for low-frequency targets with a parafoveal word that differed by one inner letter; this now revealed a tendency for an inhibition with orthographic relatedness, while the general tendency was again facilitating. The effect of orthographic relatedness was significant for outer letters in the participant analysis, $F_1(1, 47)=7.46$, $p<.01$, $F_2(1, 143) = 1.50$, and marginally significant in the item analysis of inner letters, $F_1<1$, $F_2(1, 135)=3.05$, $p<.10$. The interaction between orthographic relatedness and target frequency was significant only for inner letters in the participant analysis, $F_1(1, 47)=4.24$, $p<.05$; other *F*s 2.24.

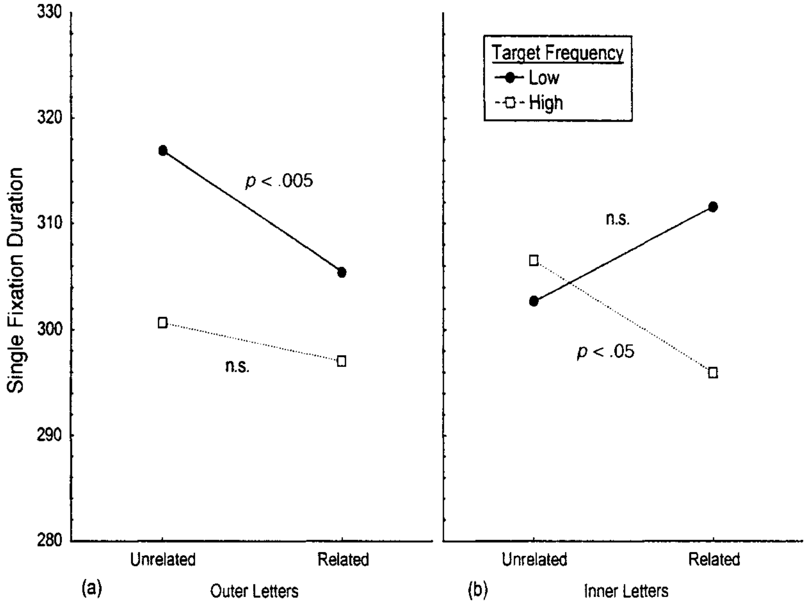

Figure 2. Mean duration of single fixation (in ms) on target words of high and low frequency as a function of orthographic relatedness, and separately for outer (a) and inner (b) critical letter positions. This analysis was run across four- and five-letter words.

As can be seen in Figure 3, the *mean duration of first fixations* (which includes instances where the first fixation on the target word was or was not followed by an additional fixation), revealed again the same trends with target frequency and orthographic relatedness, but in a much weaker manner. The effect of target frequency was significant in both participant and item analyses, $F_1(1, 47)=14.80$, $p<.0005$, $F_2(1, 278)=12.72$, $p<.0005$, although being significant only for outer letter positions, $F1(1, 47)=23.21$, $p <.0005$, $F_2(1, 143)=16.33$, $p<.0005$. The effect of orthographic relatedness was never significant (Fs<2.30), and the three-way interaction was only marginally significant in the participant analysis, $F_1(1,47)=3.20$, $p<.10$; $F_2(1, 278)= 1.30$.

Discussion

The aim of the present experiment was to examine the role of visual attention in the processing of two adjacent words, one in foveal vision and one in parafoveal vision, and to determine whether both words are processed in parallel or in sequence. For this purpose, a test of the presence of parafoveal-on-foveal effects due to orthographic relatedness was conducted. A foveal target word was presented in a reading-like situation, and it was accompanied in parafoveal vision by either an orthographically very similar word (a neighbour that differed by a single letter) or a completely different word. In addition, the relative

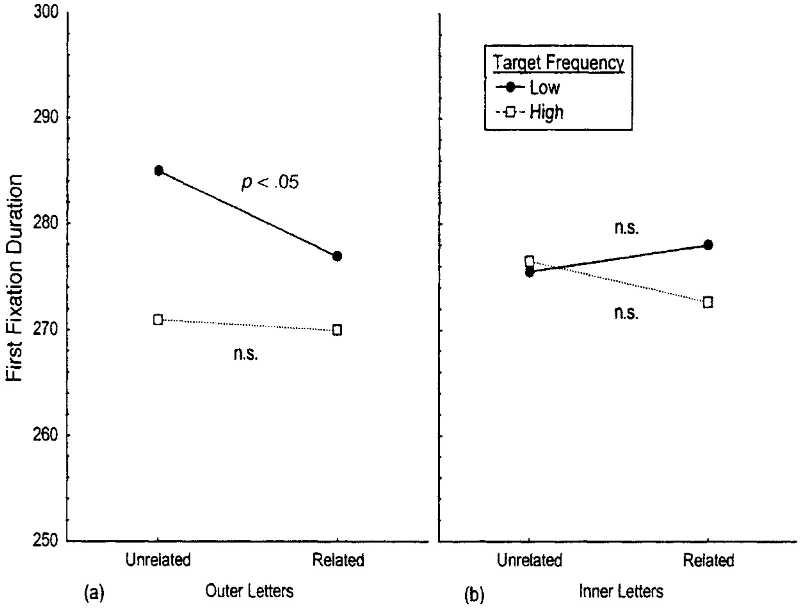

Figure 3. Mean duration of first fixations (in ms) on target words of high and low frequency as a function of orthographic relatedness, and separately for outer (a) and inner (b) critical letter positions. This analysis was run across four- and five-letter words.

frequency of the target and the parafoveal word was manipulated such that the target corresponded either to the higher or the lower frequency word of the pair. According to a sequential attentional model, the nature of the parafoveal word should not have an effect on the gaze duration of the target word, because processing of the foveal word is supposed to be *finished* before the attention spotlight switches to the next word (Henderson & Ferreira, 1990; Morrison, 1984; Reichle et al., 1998). However, as noted above, an early facilitation due to orthographic similarity may still be reconciled with a sequential view, but only if the effect is similar between high-, and low-frequency parafoveal words. In contrast, a pure parallel processing view predicts an interaction between the processing of parafoveal and foveal words (see Schroyens et al., 1999). This interaction should be reflected in a differential effect of orthographic similarity with the relative frequency of both words. In between both views, distributed parallel processing unambiguously predicts the presence of parafoveal-on-foveal influences of an orthographic type, but it cannot predict differential effects depending on the relative frequency of target and parafoveal words (Engbert et al., 2002; Inhoff et al., 2000a, 2000b).

The present study yielded two results. A first significant finding was that parafoveal-on-foveal influences of an orthographic type do occur when reading pairs of isolated words. In conditions where the foveal target word shared orthographic features with the parafoveal word, the gaze duration, and to a certain extent the duration of single fixations, were shorter than when both words were orthographically dissimilar. The second, but less systematic finding relied on the possibility that orthographic parafoveal-on-foveal effects may vary with the relative frequency of target and parafoveal words. A tendency for differential effects of orthographic relatedness with word frequency was indeed observed at least in the measures of gaze duration and single fixation duration, but this proved to be nonsignificant in most cases, and the picture was made more complex, as the effects also varied with the position of the critical letter in the words. A pattern of data still came out significantly in both gaze duration, and single fixation duration (i.e., the three-way interaction), which basically revealed a tendency for greater facilitation with either low- or high-frequency target words depending, respectively, on whether the critical letter corresponded to an external or an internal letter.

A priori, these results may not allow us to distinguish between alternative views for the processing of adjacent words. First, the observation of an overall facilitation due to orthographic relatedness, which is in line with previous reports of visual, and orthographic parafoveal influences (Inhoff et al., 2000b), only tells us that the processing of two adjacent words can occur in parallel at least up to the letter level. It is therefore compatible with either pure parallel or distributed processing views. It is also not completely opposed to a sequential view, as the possibility that attention may not be perfectly focused on the foveal word at the very beginning of a fixation can be envisaged in sequential processing models. On the other hand, the unexpected combined influence of target frequency and critical letter position on the likelihood of parafoveal-on-foveal effects does not unambiguously indicate that parallel processing extends beyond the letter level. Indeed, the possibility that the pattern of data may be due to perceptual rather than lexical processes cannot be completely excluded. In particular, since both four-, and five-letter words were mixed in the reported analysis, data selection may have resulted in more data from a given word length in some conditions.

The observation that parafoveal-on-foveal effects were not present across all dependent variables, and particularly that these failed to come out significant in the measure of refixation likelihood, and first fixation duration may however, constitute one case against sequential processing models. If parafoveal-onfoveal effects were due to attention being not perfectly focused on the foveal word at the very beginning of an eye fixation, or before selective processing occurs, then these effects should emerge relatively early in the time course of an eye fixation. They should therefore be more likely to occur in the measure of refixation likelihood, and first fixation duration, as these may correspond to earlier indicators of ongoing processing than single fixation or gaze durations. In line with this assumption, previous studies indicate that the duration of first fixations in instances where the word receives two consecutive fixations is relatively short, and shorter on average than the duration of single fixations (see Vitu, McConkie, Kerr, & O'Regan, 2001; Vitu & O'Regan, 1995). This suggests first that the programming of a within-word refixation arises relatively early, and second, that the duration of first fixations when averaged over one and several fixation cases is biased towards shorter time ranges than single fixation or gaze duration.

To examine in more detail the time course of parafoveal-on-foveal effects, a quartile analysis of the gaze duration in the different conditions was conducted. For this analysis, only five-letter target words were considered, since controlling for the length of the words may help by reducing the range of possible interpretations for the observed phenomena. Results presented in Figure 4 reveal that a facilitation effect due to orthographic similarity emerges over time for low-frequency targets in the case of outer letters, and for high-frequency targets in the case of inner letters. It is only for gaze durations in the range of 350–400 ms (around the third quartile) that the outer-letter curves differentiate between high- and low-frequency targets, and at about 320 ms (median) that a difference starts emerging between high- and lowfrequency targets for the case of inner letters. An analysis of variance run separately on each quartile actually revealed that the three-way interaction involving orthographic relatedness, target frequency, and critical letter position (with two levels) was significant only for the third quartile, $F_1(1,47)= 4.76$, $p<.05$. These results confirm the notion that orthographic parafovealon-foveal influences may be only lately determined, which suggests that our failure to report significant effects with first fixation duration resulted from the time course of the effects in relation with the characteristics of the dependent variable. Given the time range of first fixation durations, which averaged at about 272 ms for five-letter words, only a hint of an effect could be captured with this dependent variable.

The present observation that orthographic influences from the parafovea take time to develop, and actually do not emerge earlier than 320–350 ms from the beginning of an eye fixation, leads us to reject the hypothesis that sequential processing was at the origin of the reported effects. In addition, the similarity in the pattern of data between instances where four- and five-letter words were mixed, and instances where only five-letter words were considered (see for comparison, Figure 1b and Figure 4c, e-f) suggests that the combined influence of target frequency, and critical letter position on the likelihood of parafoveal-on-foveal effects does not result from a confounding with word length, and it cannot therefore be solely attributed to perceptual factors. The question however remains as to whether pure parallel processing or distributed processing best accounts for the observed effects, and in a relative manner, which processes underlie the role of the critical letter position. We have at present no definitive explanation for the obtained pattern of data, although we could not find a way to account for it in the framework of distributed processing models (see Inhoff et al., 2000a, 2000b). The assumption that more resources would be allocated to the foveal word when it is difficult to process cannot account for the fact that parafoveal-on-foveal effects tend to be larger with low- than high-frequency targets when the words differ by their first or last letter.

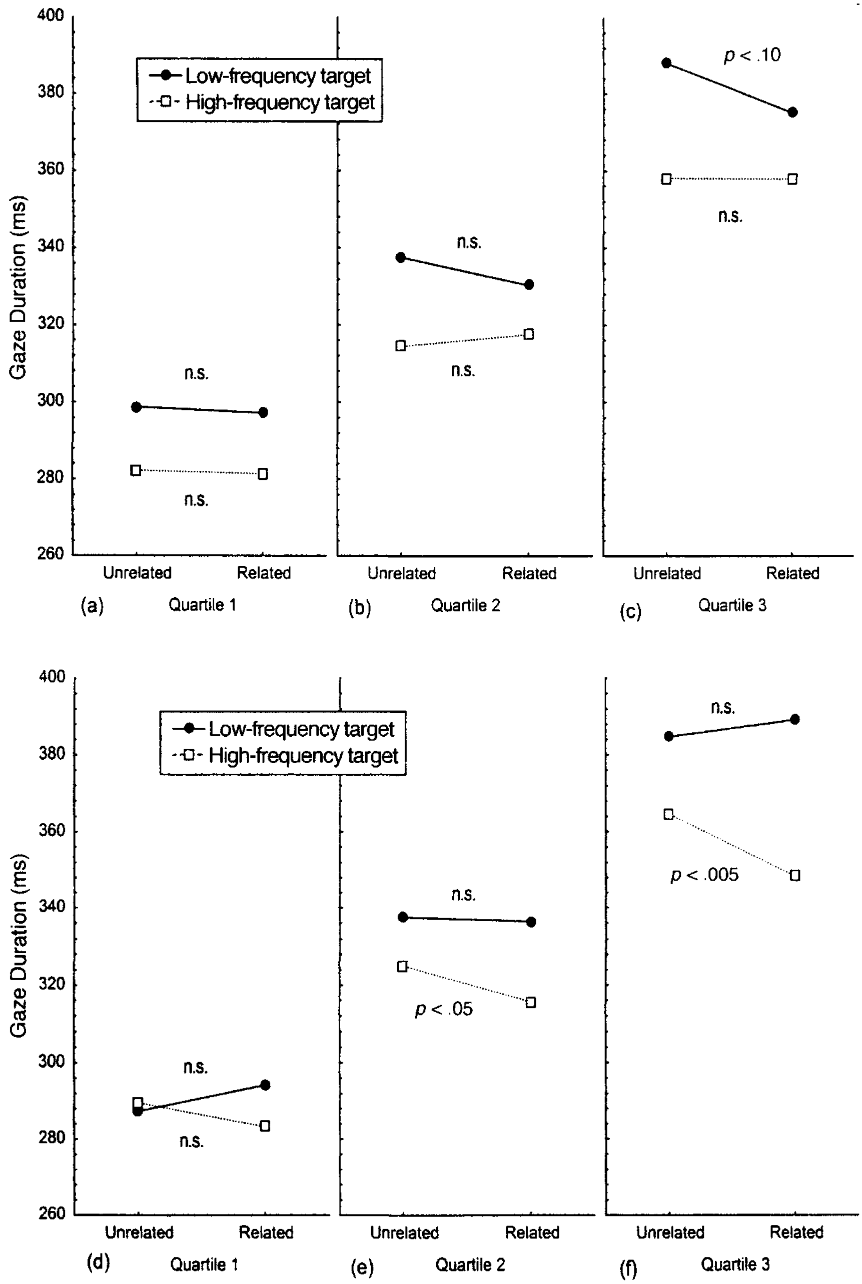

Figure 4. First, second, and third quartile of the gaze duration on five-letter target words of high and low frequency as a function of orthographic relatedness, and separately for outer (a-c) and inner (d-f) critical letter positions.

A pure parallel processing view may be a better candidate at accounting for the present set of findings. This makes no recourse to the notion of attention, and envisages the processing of two adjacent words in terms of a competition between the lexical candidates that become activated on the basis of both foveal and parafoveal letter information. In this view, all letters are extracted simultaneously, and the primary constraint is that visual information is sampled with a higher resolution, and therefore faster from the centre of the foveal region than from the parafoveal region (see Schiepers, 1980; Schroyens et al., 1999). An additional constraint relates to the respective properties of the foveal and parafoveal words, which affect the rate of ongoing processing associated with each word, and therefore determine the presence or absence of parafoveal-on-foveal effects. As suggested by previous studies, word frequency may be one of those critical factors. First, the frequency of the fixated word directly influences the gaze duration on the word, which in turn determines the amount of parafoveal processing (Schroyens et al., 1999). Second, the frequency of a parafoveal word largely influences the rate of parafoveal processing (Inhoff & Rayner, 1986; Vitu, 1991). Third, when both foveal and parafoveal words are orthographically similar, their respective frequency determines which word gets activated first, and therefore whether the parafoveal word can compete with the foveal word (see Segui & Grainger, 1990).

Starting from this, we may expect that in the present experiment, parafoveal preview was more efficient when low-frequency targets were paired with highfrequency parafoveal words than in the opposite condition. However, whether this resulted in the former case in greater facilitation or in inhibition (or a null effect) did not only depend on the respective frequency of foveal and parafoveal words, or their level of competitiveness, but also on the position of the critical letter that distinguished both words, or the ease with which this letter could be detected (Bouma, 1973). First, when the critical letter corresponded to the first or last letter of the word, its visibility was enhanced, and the possibility of noticing where both words differed was increased. This helped in the rejection as a candidate for the foveal location, one competing candidate (i.e., a higher frequency neighbour) that was quickly activated from the parafovea. As a result, facilitation was favoured when low-frequency targets were paired with a highfrequency parafoveal word. In a reverse manner, the condition where a high frequency target was associated with a lower frequency parafoveal word tended to produce no effect, probably because rapid identification of the foveal word coupled with the possibility of detecting the critical letter resulted in an inhibition of the parafoveal word unit, neutralising the facilitation that would normally arise from having two sets of similar features/letters in foveal and parafoveal vision respectively.

On the other hand, when the critical letter corresponded to an inner letter, its visibility was reduced due to lateral masking. This first allowed facilitation due to feature/letter similarity to arise when high-frequency targets were paired with low-frequency parafoveal words. In opposition, this produced no facilitation, or a tendency for an inhibition in the condition where a high-frequency parafoveal word accompanied a low-frequency target. Indeed, the parafoveal word was processed faster than the foveal word, and it acted as a strong competing candidate. As suggested by Segui and Grainger's (1990) masked priming study, when a candidate gets more likely, inhibition of its lower frequency neighbours including the target word may develop interfering with the identification of the target. The reason why our results failed to reveal a significant inhibition is not clear at present, although this may be related to timing constraints or the fact that inhibitory effects take more time to emerge than facilitation effects. Indeed, in another experiment (unpublished data using the same paradigm) we succeeded in obtaining inhibitory effects in conditions where the words were read at a lower pace than in the present experiment. In the same manner, Perea (1998) reported inhibitory priming effects for the case of inner letters, but in a perceptual identification study where participants' responses were not time constrained.

It is clear that the above set of assumptions will need to be tested further, and that the present work was only an initial approach in trying to distinguish between alternative views for the processing of adjacent words. Further research may benefit from trying to replicate the present study with more observations, and also from attempting to determine whether the effects suggested in the present paper generalise to normal reading situations. At the same time, we think that the present results may not be specific to the recognition of pairs of isolated words. Our experimental situation although corresponding to an over-simplified reading task that involves none of the syntactic or semantic high-level processes specific to natural reading, was relatively successful in reproducing the dynamics of oculomotor scanning that characterise reading. The gaze durations observed in our experiment were on average about 50 ms longer than the gaze durations observed in normal reading, indicating that the two RT distributions have considerable overlap. In addition, the effect still prevailed when the analysis was limited to single fixations, which had an average duration very similar to the gaze durations reported by Schilling et al. (1998).

PrEview proof published online September 2003

REFERENCES

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception and Psychophysics, 63*(5), 875–890.

Andrews, S. (1997). The effects of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review, 4,* 439–461.

Baccino, T., Lavigne, F., Gragnic, C., & Vitu, F. (2001). *Parafoveal-on-foveal priming during eye movements: Semantic effects.* Paper presented at the 10th European conference on Eye Movements, Turku, Finland.

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research, 13,* 767–782.

Bradshaw, J.L. (1974). Peripherally presented and unreported words may bias the perceived meaning of a centrally fixated homograph. *Journal of Experimental Psychology, 103*(6), 1200– 1202.

Carpenter, P.A., & Just, M.A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York: Academic Press.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornič (Ed.), *Attention and performance VI* (pp. 535–555). London: Academic Press.

Cornsweet, T.N., & Crane, H.D. (1973). Accurate two-dimensional eye tracker using first and fourth Purkinje images. *Journal of the Optical Society of America, 63,* 6–13.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636.

Ferrand, L. (2001) *Cognition et lecture: Processus de base de la reconnaissance des mots écrits chez l'adulte.* Bruxelles, Belgium: DeBoeck Université.

Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Henderson, J.M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psy-chology: Learning, Memory and Cognition, 16*(3), 417–429.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47,* 201–221.

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000a). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 301–326). Oxford, UK: Elsevier.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40,* 431–439.

Inhoff, A.W., Starr, M., & Shindler, K.L. (2000b). Is the processing of words during eye fixations in reading strictly serial? *Perception and Psychophysics, 62*(1), 1474–1484.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection times: Evidence for a processing trade-off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–223). Oxford, UK: Elsevier.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A*(2), 429–455.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Kennison, S.M., & Clifton, C.C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology; Learning, Memory, and Cognition, 21,* 68–81.

McConkie, G.W. (1979). On the role and control of eye movements in reading. In P.A.Kolers, M.E. Wrolstad, & H.Bouma (Eds.), *Processing of visible language (Vol. 1).* New York: Plenum.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Mozer, M.C. (1983). Letter migration in word perception. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 531–546.

Murray, W.S. (1998). Parafoveal pragmatics. In G.Underwood (Ed.), *Eye guidance in reading and scene perception.* Oxford, UK: Elsevier.

Perea, M. (1998). Orthographic neighbors are not all equal: Evidence using an identification technique. *Language and Cognitive Processes, 13*(1), 77–90.

Pollatsek, A., Perea, M., & Binder, K.S. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1142–1158.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65– 81.

Rayner, K., Reichle, E.D., & Pollatsek, A. (1998). Eye movement control in reading: An overview. In G.Underwood (Ed.), *Eye guidance in reading and scene perception.* Oxford, UK: Elsevier.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier Science.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157.

Schiepers, C.W. (1980). Response latency and accuracy in visual word recognition. *Perception and Psychophysics, 27*, 71–81.

Schilling, H.E. H., Rayner, K., & Chumbley, J.L. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26,* 1270–1281.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology, 52A*(4) 1021–1046.

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 65–76.

Treisman, A., & Souther, J. (1986). Illusory words: The roles of attention and of top-down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception and Performance, 12*(1), 3–17.

Trésor de la langue française. (1971). Nancy, France: CNRS.

Van Rensbergen, J., & de Troy, A. (1993). *A reference guide for the Leuven dual-PC controlled Purkinje eyetracking system (Tech. Rep.).* Laboratory of Experimental Psychology, Catholic University of Leuven.

Vitu, F. (1991). The influence of parafoveal preprocessing and linguistic context on the optimal landing position effect. *Perception and Psychophysics, 50,* 58–75.

Vitu, F. (1993). The influence of the reading rhythm on the optimal landing position effect. In G. d'Ydewalle & J.Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 181–192). Amsterdam: North Holland.

Vitu, F., McConkie, G.W., Kerr, P., & O'Regan, J.K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41*(25– 26), 3511–3531.

Vitu, F., & O'Regan, J.K. (1995). A challenge to current theories of eye movements in reading. In J. Findlay, R.W.Kentridge, & R.Walker (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 381–392). Amsterdam: Elsevier.

White, S.J., & Liversedge, S.P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*(1/2), 52–78.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 178–202

The influence of parafoveal typographical errors on eye movements in reading

Joel Pynte

CNRS and University of Provence, France

Alan Kennedy

University of Dundee, UK

Stephanie Ducrot

CNRS and University of Provence, France

Three experiments are reported, examining the effects of a typographical error in parafoveal vision on aspects of foveal inspection time and saccade targeting. All the experiments involved reading for comprehension. A contingent presentation procedure ensured that typographical errors were restored to their correct form before they were viewed in foveal vision: They were never available for foveal processing. In Experiment 1, the error was formed by replacing the first letter of the target word with a second occurrence of its second letter, producing an illegal nonword. This manipulation had no significant effect on foveal inspection time, but lowered the probability that a short word ("de" or "du") prior to the target would be skipped. In Experiment 2 the familiarity of the target's initial letters was maintained constant across conditions. This manipulation removed the target 1 skipping effect, suggesting that the outcome of Experiment 1 was due to orthographic rather than lexical illegality, but revealed shorter foveal inspection times as a function of the presence of the error. Experiment 3 manipulated lexical and sublexical properties of the parafoveal typing error. Properties of the parafoveal error again influenced prior foveal inspection times. The pattern of results suggested that the determining properties were sublexical rather than lexical. The results as a whole are incompatible with a view of information processing in reading in which foveal processing remains immune from concurrent parafoveal influences.

This paper deals with the controversial question of the nature of some of the cognitive events responsible for triggering a saccade during reading. In spite of a great deal of research, a number of questions remain unresolved. Are the eyes targeting specific locations within the word about to be fixated? For example, is there a specific location which the system targets as a point where processing needs to be carried out? Alternatively, should saccades be seen primarily as movements initiated when processing at a specific location has been completed, for example when lexical access has been achieved? We aim to contribute to this debate by examining how the presence of a typographical error in the parafovea influences eye movements. Typographical errors[1] are a useful source of evidence, because they are frequently associated with highly salient illegal sequences of letters, which may readily be identified parafoveally as potential targets for an eye movement. They are, nonetheless, difficult to manipulate experimentally because the presence of numerous typos in experimental text inevitably alerts participants and lowers the ecological validity of the task. The work reported here avoids this problem with a contingent presentation technique, which ensured that typos could never be fixated directly. That is, a typo would be present, but only in parafoveal vision: The critical word was restored to its correct spelling when directly fixated.

In attempting to assess the possible effects of typos, it is necessary to distinguish between the timing of saccades and their targeting (i.e., the "when?" and "where?" decision). There is a large body of evidence suggesting that these two decisions may be under independent sources of control, with low-level factors, such as word length, being the primary influence over saccade targeting, and higher level linguistic factors affecting saccade latency (Findlay & Walker, 1999; Rayner, 1998). Although this distinction has been important theoretically, it remains unclear whether the conclusion is forced that the *only* influences over "when?" decisions are derived from foveal processing operations (a conclusion that seems inevitable if higher level linguistic properties are only available under direct foveal inspection). The three experiments reported here examine this question by looking at the influence on foveal processing exerted by typos that were only ever present in parafoveal vision. Four cases can theoretically be distinguished regarding the cognitive consequences of "encountering" a parafoveal typo: (1) The typo may not be processed until the currently fixated word has been identified; (2) the typo may be processed prior to identification of a fixated word (the precise mechanism involved in achieving this is irrelevant in

---

Correspondence should be addressed to J.Pynte, Laboratoire Parole et Langage, Université de Provence, 29, Avenue Robert Schuman, 13621, Aix-en-Provence, France. Email: pynte@up.univ-aix.frThis paper was prepared with the assistance of Grant no. R000223650 from the UK Economic and Social Research Council to Alan Kennedy.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI:10.1080/09541440340000169

the present context), but the system waits until completion of foveal processing before initiating a saccade; (3) processing of the typo in parafoveal vision may act as the signal to initiate a saccade, with the typo as target (following Kennedy, 2000, we shall refer to this as a GO response); (4) processing of the typo takes place, but this leads to a competition between a tendency to remain at the current fixation point (a STAY response) in order to complete identification of the foveal word, a tendency to execute a GO response, and a tendency to move the eye towards the typo in order to increase its visibility (a SHIFT response; see Kennedy, Pynte, & Ducrot, 2002, for a discussion). The first two cases are probably indistinguishable empirically and encapsulate what we shall refer to as the "leave-on-completion" hypothesis. This is an essential ingredient of all serial attention-switching models of eye movement control (Henderson & Ferreira, 1990; Morrison, 1984; Rayner, Reichle, & Pollatsek, 2000; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999). With regard to the "when?" decision (determining the point in time when a saccade is launched to the next word) cases (1) and (2) both predict no effect of parafoveal typos on foveal processing. Predictions for these two cases with regard to the "where?" decision (where to go next) are much less clear cut, since it is always possible to maintain that target selection may occur after a currently fixated word has been identified. In contrast, cases (3) and (4) are more straightforward and predict, in principle, an effect of parafoveal typos on both the "where?" and "when?" decisions. These cases encapsulate what we shall refer to as the "targeting hypothesis". In both cases the presence of a (parafoveal) typo will have an influence on concurrent foveal processing, modulating the pattern of inspection and having (at least consequential) effects on foveal inspection time. The decision to STAY, GO, or SHIFT, as consequence of parallel parafoveal processing, underlies a rather complex pattern of "parafoveal-on-foveal" effects (Kennedy, 1995, 1998, 2000; Kennedy et al., 2002).

Although not the primary objective of this paper, the experiments to be reported bear on the nature of the processing operations carried out in the parafovea. This, too, has proved to be a highly controversial topic. Clearly, some processing must occur to the right of the currently fixated word, since the system needs to determine where to fixate next. However, the mechanisms involved need not to be the same as those carried out on the fixated word itself. Under the "leave-on-completion" hypothesis, the target of each succeeding fixation is determined by default (usually the next word). In this case only "low-level" information, such as the location of spaces on the line of text, needs to be extracted from the parafovea. Only "preattentive" processes (Hebb, 1949; Neisser, 1967) are involved, and visual attention is entirely devoted to the word under current inspection (see Morrison, 1984; Rayner et al., 2000; Reichle et al., 1998, 1999, for extensive recent developments of this notion). In contrast, under the

[1] In the interests of conciseness these will henceforth be referred to as "typos".

targeting hypothesis, a specific location must be determined to act as a target and this almost certainly involves processing the parafovea to a deeper level. Models in which attentive processing is not confined to the currently fixated words are more in line with this view. For example, in the model described by Engbert, Longtin, and Kliegl (2002), processing occurs across a region greater than a single word, with the rate of acquisition of information being a function of eccentricity (see also, Schiepers, 1980; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999).

Parallel attentive processing is not an essential prerequisite of target selection, however. Potential targets may emerge from preattentive processes taking place in parallel with foveally focused attention. From this point of view, the allocation of visual attention to the right of a fixated word would be the consequence, and not the cause, of target selection. For example, Hyönä (1995) and Beauvillain and Doré (1998) suggest that salient features such as irregular letter sequences may "pop out" in the parafovea and attract saccades towards them. Hyönä and Bertram (2004) put this in terms of a "magnetic attraction" exerted by specific stimulus configurations. The idea is consistent with the model developed by Findlay and Walker (1999), which proposes that intrinsically salient stimuli, such as unfamiliar letter sequences, determine saccade location within a "salience map" (i.e., the distribution of salience across the visual field determining the saccade target in a "winner-takes-all" computation). Whether potential targets emerge through attentive or preattentive processes is thus an open question and one which will not be further addressed in this paper. We wish to emphasise that in using the expression "targeting hypothesis" we remain neutral as to whether attentive (e.g., strategic) mechanisms are at work or not.

The data bearing on the "where?" decision in favour of either the targeting or the leave-on-completion views remain so far inconclusive. An important source of evidence is provided by experiments reporting orthographic effects on initial landing position. The fact that first fixations in reading are usually located near the beginning of words (the so-called preferred landing position or preferred viewing location; Rayner, 1979) has often been considered as a counterargument to the targeting hypothesis. Obviously, if the eye always lands at approximately the same location, there is little need for a targeting mechanism. But a growing body of evidence makes this position difficult to sustain. Data from a number of languages suggest that initial landing position is shifted about one third of a character to the left in words with unfamiliar beginning (Hyönä, 1995; Hyönä & Pollatsek, 1998, 2000; Kennedy, 2000; Kennedy et al., 2002; Radach, Inhoff, & Heller, 2003; Vonk, Radach, & van Rijn, 2000; White & Liversedge, 2004). Unfortunately, this shift in landing position does not clinch the argument in favour of a targeting mechanism. An alternative explanation, which does not involve a target selection mechanism, is to assume that the area where attention is allocated can shrink or expand as a function of global processing difficulty, involving both foveal and parafoveal properties—the so-called *processing difficulty hypothesis* (Hyönä, 1995; Hyönä & Pollatsek, 2000). This predicts a parafoveal influence over saccade targeting, as a result of changes in the "width" of the attentional spotlight acting to modulate the landing position of interword saccades (see Kennedy et al., 2002, for further discussion).

A second source of evidence can be looked for in skipping effects. It is well known that interword saccades are *not* necessarily directed towards the next "blob" and that the determinants of word-skipping are not restricted purely to low-level information (Rayner, 1998; see also Brysbaert & Vitu, 1998, for a review and meta-analysis). Words, especially short and predictable ones, are often skipped (O'Regan, 1979; Rayner & Duffy, 1986), and this points to the existence of a rather sophisticated target-selection mechanism. The question arises, however, as to whether skipping occurs because the word following the skipped word is actually targeted, in line with the targeting hypothesis, or simply because processing has been completed on the skipped word, in accordance with the leave-on-completion view. Identification in the absence of actual foveal inspection is indeed possible if it is admitted that visual attention can be decoupled from visual inspection. For example, in Morrison's model (Henderson & Ferreira, 1990) visual attention moves ahead of overt inspection and, in the case of short and high-frequency words, identification may be achieved before the eye reaches the corresponding region. In such circumstances, the motor program could be modified in order to reach the next word on the line, and this would readily account for skipping effects. Even if attentive processes are time-locked to the currently-fixated word, with no parallel parafoveal processing at all (either attentive or preattentive), there may be time left for parafoveal processing prior to saccade initiation, enough to allow the system to select which word to fixate next, or which word not to fixate, and which location to aim for. For example, in E-Z Reader model (Reichle et al., 1998) attention is usually allocated to the next word before saccade initiation and some (attentive) parafoveal processing can thus take place on that word, after foveal processing has been completed, but at a point in time when saccade extent can still be adjusted as a function of properties of the to-be-fixated word.

A more promising way of distinguishing between "leave-on-completion" and "targeting" derives from a more detailed analysis of foveal inspection strategies, that is, strategies developed on a currently fixated word, before an interword saccade has been initiated. Kennedy (1995, 1998) has claimed that the same factors that exert an influence on interword saccades (e.g., initial letter constraint in an as yet unfixated parafoveal word) also have an immediate effect on current foveal processing, and, in particular, on within-word refixation strategy. This has been confirmed by a number of recent studies, although with many inconsistencies in the direction of differences reported (Hyönä & Bertram, 2004; Inhoff, Radach, Starr, & Greenberg, 2000; Murray, 1998; Underwood, Binns, & Walker, 2000). Kennedy et al. (2002) argue that most of these inconsistencies can be explained by lack of effective control over the respective length of foveal and parafoveal words. They report data suggesting that inspection strategy is determined over a sequence of words (typically two), in order to optimise overall visibility and processing. This view clearly contradicts the leave-on-completion hypothesis, which states that refixations are primarily aimed at foveal processing completion (see Reichle et al., 1998, for a discussion concerning possible mechanisms).

The interaction between, for example, patterns of intraword reinspection and inspection time means that the investigation of targeting patterns becomes inevitably bound up with measures of processing time. However, there is little evidence so far suggesting that properties of a given potential target can *directly* influence the latency of the saccade directed towards it. Evidence for such parafoveal-on-foveal timing effects would provide a strong argument in favour of the targeting hypothesis which, unlike the leave-on-completion view, predicts parafoveal influence over both "where?" and "when?" decisions. A primary aim of the present study was thus to shed light on possible sources of parafoveal control over the "when?" component of a possible saccade-targeting mechanism. The operation of such a mechanism was examined in the context of a task calling for relatively normal reading, but which exposed readers to parafoveal typos. The three experiments reported successively refine the factors playing a (parafoveal) role on targeting, from effectively *anything,* to lexical and sublexical properties of the critical target words.

Participants read short French sentences like (1) for comprehension. Inspection strategies were analysed across three successive words, defined as target-2 ("chaîne"), target-1 ("de"), and target ("vélo") respectively. The target-1 word was invariably a short article or preposition and was highly predictable. In the "typo" condition, an error was introduced into the defined target word by replacing its first letter with another letter, resulting in a nonword. A contingent presentation procedure (the "boundary technique" introduced by Rayner, 1975) ensured that the typo was only present in parafoveal vision: As a saccade crossed the space before the target word, the critical letter was permanently restored. Thus, although the materials contained many typos, these were never directly inspected and participants were probably unconscious of their presence.

(1) II répare la chaîne de vélo avec un tournevis.

In line with the various theoretical positions outlined above, we predict that the presence of a parafoveal typo will influence foveal inspection strategy. Variation in target-2 inspection time should be observed, depending on whether the potential target for the next interword saccade in the parafovea is or is not a typo. Given that the relevant words were quite short, the targeting hypothesis predicts shorter (initial) inspection times on a foveal word in the presence of a parafoveal typo, either as a result of execution of a GO response or as a by product of competition between STAY, GO, and SHIFT responses. In addition to this main prediction, which concerns the "when?" component of eye movement control, two typo effects were expected in relation to the "where?" component. Typos should exert an influence on target landing position. In particular, typos comprising highly salient visual configurations (e.g., illegal letter sequences) should exert a kind of "magnetic attraction" for saccades. Moreover, the presence of a typo should influence skipping rate associated with the short target-1 word. In the context of the targeting hypothesis, the critical question is whether skipping the target-1 word results from a targeting mechanism directed towards the following word (i.e., the defined target) or relates to processing of the skipped word itself. In order to answer this question, properties of the target-1 word were maintained constant across conditions. Only properties of the target itself were manipulated. Any modulation in target-1 skipping rate will thus have to be attributed to parafoveal influence interfering with the "where?" decision.

EXPERIMENT 1

In this first experiment the typo manipulation was designed to produce as "salient" a visual configuration as possible, within the constraints of the contingent procedure. Thus, the typo, when present, comprised an illegal sequence of letters in the French language. Consequently, the resulting letter string defining the (parafoveal) target was a nonword with unfamiliar and irregular initial letters. We claim this manipulation represents an appropriate condition to determine whether *any* properties of the parafoveal stimulus influence concurrent foveal inspection and targeting strategies.

Method

Participants. Thirty-two Psychology students, from the University of Provence, Aix-en-Provence, took part in the experiment. All were native speakers of French. All had either normal or corrected-to-normal vision. Participants were not paid for taking part.

Task and apparatus. A trial consisted of the presentation of a sentence in normal upper- and lowercase French script on a display monitor interfaced to Pentium 2 PC computer. Participants were instructed to read the sentence carefully and decide whether it made sense or not. They responded by pressing one of two buttons located under their left and right hands. At the viewing distance of 500 mm, one character subtended approximately 0.3° of visual angle. Eye movements were recorded from the right eye using a Dr Bouis pupil-centre computation Oculometer interfaced to a 12-bit A-D device sampling X and Y position every 2 ms. This eye tracker has a resolution of better than 0.25 characters over the 60-character calibrated range (Beauvillain & Beauvillain, 1995). The contingent change from "typo" to "no-typo" was achieved by writing directly to the video memory of the graphics control card and was not dependent on the refresh cycle of the display. Thus, the average delay in executing the display change was under 7 ms (half a refresh cycle—5 ms—plus the average sampling and computation time) and the maximum delay less than 13 ms. It was confirmed independently that the change took place within a single screen refresh cycle. The display change was implemented in both the typo and no-typo condition (i.e., the critical letter of the target word was replaced by itself). At the beginning of a trial, a fixation point appeared at the left of the screen, three characters to the left of the position to be occupied by the initial letter of the sentence. Sentence presentation occurred 200 ms after a stable fixation of this point for longer than 150 ms. During the experiment, the participant's head was restrained with a dental-composition bite bar. A calibration procedure was carried out every four trials in which participants fixated in turn five points located across the region to be occupied by the stimulus material. If the calibration data fell within tolerable limits, the experiment continued; if not the calibration process was repeated.

Materials. Sixty-four experimental sentences, based on sixteen different lexical contents were prepared. All the sentences began with a pronoun, and then comprised a verb and an article followed by a sequence of three critical words (the sequence "chaîne de vélo" in example 1). Visual inspection of these words was analysed in detail. Hereafter, the second noun of this sequence ("vélo" in example 2) will be referred to as the target word, the preposition ("de" in example 2) as target-1, and the first noun ("chaîne" in example 2) as target-2.

(2) II répare la chaîne de vélo avec un tournevis.[2]
(2') II répare la chaîne de éélo avec un tournevis.
(3) II enlève la chaîne du vélo avec un tournevis.[3]
(3') II enlève la chaîne du éélo avec un tournevis.

Four different sentences were derived from each lexical content. In the two "typo" versions, a typographical error was formed by replacing the first letter of the target word with a second occurrence of the second letter (examples 2' and 3'). Two syntactic structures were compared involving the contrast between "de" and "du".[4]

Procedure. Four experimental lists were prepared to balance the association of items with conditions across participants. Each list comprised 16 experimental sentences (4 per condition), all derived from a different lexical content. Those lexical contents that were associated with a given condition in a given list were associated with another condition in another list. In each list, experimental sentences were preceded by eight training sentences and were mixed with 40 fillers of various syntactic structures. Sentences were presented using a different random order for each participant. The contingent presentation procedure ensured that the typo (if present) was displayed only while the eyes were to the left of an invisible boundary defined as the space between the target-1 word and the target. The properties of the target's initial letters thus varied over time as a

---

[2] He is mending the bicycle chain with a screwdriver.

[3] He is removing the chain from the bicycle with a screwdriver.

function of the position of the eye relative to the invisible boundary. To ensure strict comparability between display conditions, the contingent procedure was also employed in the "no-typo" condition (in which case the critical letter was replaced with itself).

Results and discussion

In terms of the hypotheses set out in the Introduction, assuming the existence of a typo effect, this could be expected to influence inspection strategy on the target-2 word, the probability of skipping the target-1 word, or landing position on the target itself. The relevant data are summarised in Table 1. With regard to inspection strategy on the target-2 word, three measures are presented: first fixation duration, refixation probability, and gaze duration, defined as the sum of all fixations recorded between the moment when the eye entered the word to the moment when it left it, either to left or right. The table also shows the first fixation duration and gaze duration on the target word, and the probability of a regression back from the target to either target-1 or target-2. An analysis of variance was carried out for each of these measures, using a 2 (typo vs. no-typo) x2 ("de" vs. "du") factorial design. No effect involving syntactic structure, either main effect or interaction, reached significance in item and participant analyses and this factor will not be considered further.[5] No significant effect of typo was obtained for any of the measures recorded on target-2, with $F1<1$; $F2(1, 15)=1.35$ for first fixation duration; all Fs<1 for refixation probability; and all Fs<1 for gaze duration.

There was a clear typo effect on the probability of skipping the target-1 word, although the direction of the effect was surprising. The word "de" or "du" was significantly *less* likely to be skipped when the typo was present (0.49 vs. 0.62), $F1(1, 28)=6.44$, $p<.05$; $F2(1, 15)=8.91$, $p<.01$. This variation in skipping probability is obviously equivalent to the claim that more saccades landed on the target-1 word in the typo condition. Notwithstanding the lack of significant differences in inspection strategy on the target-2 word, to get a better understanding of the reasons for the modulation in the target-1 word skipping rate, it is helpful to examine differences in launch site in the target-2 word in the typo and no-typo conditions. Launch position was slightly closer in the no-typo condition (−4.4 vs.−5.0 characters), but this modest difference was not reliable, $F1(1, 28)=3.11$, *p*#DXGT#.05; $F2(1, 15)=2.80$, *p*#DXGT#.05, an outcome consistent with the lack of significant effect on refixation rate. We conclude that the variation in skipping rate was driven directly by the acquisition of

[4]The phrase "de vélo" is necessarily a noun complement, whereas "du vélo" is preferentially a verb complement. This linguistic contrast and the difference between the monotransitive and ditransitive verbs employed (see Frenck-Mestre & Pynte, 1997, and Pynte & Prieur, 1996, for a discussion) is the subject of another investigation and will not be directly considered here.

TABLE 1

Target, target-1, and target-2 inspection strategies in the four experimental conditions (Experiment 1: Illegal initial bigram in the typo condition)

	Target-2			*Target-1*	*Target*			
	ff	*rf*	*gaze*	*skip*	*flp*	*ff*	*reg*	*gaze*
De								
Typo	255	0.24	324	0.45	3.9	282	0.34	280
No-typo	275	0.28	346	0.67	3.8	263	0.26	283
Du								
Typo	260	0.17	298	0.52	4.1	292	0.35	303
No-typo	250	0.17	286	0.56	4.1	253	0.22	283

ff=target-2 and target first fixation duration (ms); rf=target-2 refixation probability; gaze = target-2 and target gaze duration (ms); skip=target-1 skipping probability; flp=target initial landing position; reg=target regression probability.

information signalling the presence of a typo two words "downstream" from the current point of fixation.

Before examining the target word itself, it should be kept in mind that the typo itself could never be directly inspected. Any variation in behaviour attributable to the typo manipulation must reflect computations made before the target was fixated. First fixation duration on the target following the typo condition was longer (287 vs. 258 ms), an effect that can be interpreted in terms of a preview advantage for the no-typo condition, but the difference was not reliable in the by-items analysis, $F1(1, 28)=14.51$; $F2<1$, and disappeared in the analysis of gaze duration, $F1(1, 28)=1.08$; $F2<1$. There was also no shift in target landing position for saccades directed towards it in its "typo" form, all *F*s <1. However, no very strong conclusion can be drawn from this null result. Landing position is known to be dependent on launch site (Hyönä, 1995; McConkie, Kerr, & Dyre, 1994; McConkie et al., 1988; Radach & Kempe, 1993; Radach & McConkie, 1998; Rayner et al., 1996; White & Liversedge, 2004) and the significant variation in the target-1 skipping rate means that the launch site of the saccade entering the target word was subject to systematic variation. The prior presence of a typo did exert a systematic influence on target inspection strategy; not on initial landing position, but on the probability of a regression from the (restored) target back to either target-1 or target-2. This probability was significantly greater following a typo (0.35 vs. 0.24), $F1(1, 28)= 9.34, p<.01$; $F2(1, 15)=11.79, p<.01$.

[5] In fact, the slightly longer gaze durations on target-2 words with "de" constructions approached significance, F1(1, 28)=*7.33, p*<.05; *F2*(l, 15)=4.23, p*#DXGT#*.05, but no important consequences flow from this in the context of the present paper.

There were more regressions in the typo condition than in the no-typo condition. This effect can probably be explained in terms of the inevitable lack of preview benefit in the typo condition and is consistent with the (albeit nonsignificant) elevation in first fixation duration. Long first fixations were apparently followed by a regression rather than by a refixation (or a right-directed exit saccade) in the typo condition. In a sense, the system behaved as if "surprised" on landing on the target by the absence of the typo, but this outcome does not comment directly on the possible existence of a targeting mechanism.

The only outcome relevant to the operation of a targeting mechanism relates to target-1 skipping probability. The outcome is significant because it was obtained by manipulating properties of a word following the skipped word, not properties of the skipped word itself. The lack of any significant variation in refixation probability and the nonsignificant changes in launch position certainly suggest a targeting mechanism (i.e., an intention to land on the target rather than simply a decision not to process the target-1 word). However the changes in skipping rate can hardly be described in terms of target "attraction". Had the eyes been attracted to the typo, as predicted by the targeting hypothesis, *fewer* target-1 landings, not more, should have been observed in the typo condition. The fact remains, however, that the presence of a typo in the parafovea exerted an influence on the probability of skipping. However, our main hypothesis concerning parafoveal-on-foveal timing effects is not confirmed. Prediction concerning landing position stands in "no man's land", because there is at least a hint of a late effect of the (restored) typo on fixation duration. But before examining the wider theoretical significance of the outcome, it is worthwhile looking in more detail at the properties of the altered target in the present experiment which may have caused the changes in skipping rate. As noted above, our primary aim in employing this particular typo manipulation was to produce a highly salient sequence of illegal letters; that is, a "magnetic" stimulus configuration (see Hyönä & Bertram, 2004) comprising a powerful potential target for our putative targeting mechanism. But the direction of the effect exerted by the manipulation clearly suggests that some other mechanism must have been at work. The typo manipulation actually resulted in two forms of illegality: Since the sequence "éé" is illegal in French it follows that "éélo", as any other sequence containing the sequence "éé", is illegal as well. Thus, the manipulation produced both orthographic and lexical illegality ("éélo" is a nonword). To determine how these two properties might differentially have influenced eye movement control in the task, some control is necessary for both the lexical status of the target item and for the the statistical properties of its initial letters across the typo and no-typo conditions. Experiments 2 and 3 set out to do this.

EXPERIMENT 2

Experiment 2 was similar in Experiment 1 in most respects, including the use of the contingent display procedure. The only difference between the two experiments related to the target initial bigram in the typo condition which was orthographically legal (and relatively familiar) in Experiment 2.

Method

Participants. Thirty-two Psychology students at University of Provence participated in the experiment. None had taken part in Experiment 1. All were native speakers of French. All had either normal or corrected-to-normal vision. Participants were not paid for taking part.

Materials. The typo was no longer formed by replacing the first letter of the target word with a second occurrence of the second letter. Another letter was used instead. The changed letter string was invariably a nonword (the string "célo" in example 4 is a legal nonword in French). Familiarity was assessed by adding the frequencies of all the words containing a given bigram (Content, Mousty, & Radeau, 1990) at their initial positions (cumulative lexical frequency: CLF). The familiarity of the initial bigram of the target word was matched, with mean CLF values for the initial bigrams in the typo and no-typo conditions 5360 and 6483 (per million) respectively.

(4) II répare la chaîne de célo avec un tournevis.

Results and discussion

The results are summarised in Table 2. An equivalent analysis strategy to that adopted in Experiment 1 was followed, using measures of first fixation duration, refixation probability, and gaze duration on the target-2 word; the probability of skipping the target-1 word; and landing position, first fixation duration, gaze duration, and regression probability on the target itself. There were no significant differences attributable to the manipulation of syntactic structure and this factor will not be discussed further. In contrast to the outcome of Experiment I, there was a significant influence of the typo manipulation on inspection time on the target-2 word. This parafoveal-on-foveal effect took the form of shorter time in the typo condition and the difference was significant in both first fixation duration (267 vs. 311 ms), $F1(1, 28)=6.20$, $p<.05$, $F2(1, 15)=9.29$, $p<.01$, and gaze duration (317 vs. 356 ms), $F1(1, 28)=5.98$, $p<.05$, $F2(1, 15)= 5.14$, $p<.05$. The difference in refixation probability was not significant, all Fs <1. In contrast to Experiment 1, no typo effect was present in the measure of skipping probability on the target-1 word, $F1<1$; $F2(1, 15)=1.26$; or on regression probability from target, Fs<1. As for target landing position, first fixation duration, and gaze duration, all Fs were either <1 or close to 1.

The modulation of inspection time on the target-2 word in this experiment goes along with a failure to reproduce the effect of typos on skipping rate found in Experiment 1. It is obviously necessary to reconcile these two patterns of data.

TABLE 2

Target, target-1, and target-2 inspection strategies in the four experimental conditions (Experiment 2: Legal initial digram in the typo condition)

	Target-2			*Target-1*	*Target*			
	ff	*rf*	*gaze*	*skip*	*flp*	*ff*	*reg*	*gaze*
De								
Typo	268	0.22	327	0.69	3.7	293	0.28	300
No-typo	327	0.17	376	0.62	3.6	278	0.33	276
Du								
Typo	266	0.17	306	0.60	3.2	304	0.30	283
No-typo	295	0.15	336	0.56	3.7	297	0.33	286

ff=target-2 and target first fixation duration (ms); rf=target-2 refixation probability; gaze = target-2 and target gaze duration (ms); skip=target-1 skipping probability; flp=target initial landing position; reg=target regression probability.

Since the target-2 word in the present experiment was inspected in a single fixation in more than 80% of cases, it was possible to carry out a separate analysis of variance for single-fixation trials (i.e., excluding cases where the word was subjected to refixation). The mean values corresponding to the single- and several fixation cases are presented in the first two columns of Table 3. Foveal single fixation durations were shorter in the presence of the parafoveal typo (278vs. 325ms), $F1(1, 28)=8.01$, $p<.01$; $F2(1, 15)= 4.81$, $p<.05$. Since single fixations are both first and last fixations, the outcome suggests that, although there was no significant overall effect of the typo on the probability of skipping the target-1 word, a typo effect may have been present in the latency of the saccade entering the target. To examine the latency of saccades aimed to the target it is obviously necessary to distinguish between final fixations preceding saccades that skipped the target-1 word and those that led to a landing on target-1. The corresponding mean values are presented in Table 3 (last two columns). In the case of saccades directed to the target (skipping the preceding word), latency was significantly shorter when the parafoveal typo was present than when it was absent (269 vs. 312 ms), $F1(1, 28)=5.74$, $p<.05$; $F2(1, 15)=8.26$, $p=.01$. In marked contrast, the latency of saccades which landed on the target-1 word showed no equivalent effect whatsoever (239 vs. 240ms), $F\text{s}<1$. It is worth noting in passing that fixation durations preceding a skip were longer than those prior to no skip (see Rayner, 1998, for further discussion).

Combined analyses

In both Experiments 1 and 2 there is clear evidence of parafoveal-on-foveal effects, but the locus of these apparently differed in the two experiments. In Experiment 1 the effect of a parafoveally presented typo was only evident in

TABLE 3

Target-2 first (ff) and last (lf) fixation duration as a function of inspection strategies (Experiment 2: Legal initial digram in the typo condition)

	ff		*lf*	
	sgl	*rf*	*skip*	*nsk*
De				
Typo	284	222	263	246
No-typo	337	238	321	237
Du				
Typo	272	237	274	232
No-typo	313	226	303	243

sgl=first fixation duration if only one fixation was made on target-2; rf=first fixation duration if at least one refixation was made on target-2; skip=last fixation duration if target-1 was to be skipped; nsk=last fixation duration if target-1 was not to be skipped.

target-1 skipping rate, although coupled with two late effects which indicated sensitivity to the corrected typo: a (nonsignificant) increase in target fixation duration, and an increased probability of making a regression from the target. In Experiment 2, parafoveal-on-foveal effects were more obviously present in various measures of target-2 inspection time. The question arises as to how the null effects in Experiment 1 on these measures are to be interpreted. Is the lack of an effect for a particular measure a reflection of different processing mechanisms or simply the result of a lack of statistical power? To address this question new analyses of variance were carried out, combining the data of both experiments. The relevant measures were those set out in Tables 1 and 2 (i.e., target-2 first fixation duration, refixation probability, and gaze duration; target-1 skipping probability; and target initial landing position, first fixation duration, gaze duration, and regression probability). Experiments were treated as a between-participant/within-item factor. If, for a given measure, the discrepancy observed between experiments reflects genuinely different processing mechanisms, an interaction involving the factors of typo and experiment should be observed in the combined analysis.

We will consider first measures which showed an effect, but no interaction. There was a significant main effect of typo on the measure of first fixation duration on the target-2 word, with shorter times associated with the presence of a typo (262 vs. 287 ms), $F1(1, 56)=6.30$, $p<.05$; $F2(1, 15)=7.66$, $p<.05$. The difference in the same direction (274 vs. 309ms) in gaze duration also approached significance, $F1(1, 56)=2.59$; $F2(1, 15)=4.09$. The interaction of typo with experiment, although significant in the by-participants analysis of first fixation duration, was far from signifiant by items, $F1(1, 56)=3.99$, $p<.05$; $F2<1$ and $F1(1, 56)=1.61$; $F2(1, 15)=1.49$ for first fixation and gaze duration, respectively. Thus, although these effects were significant in one experiment and not in the other, we need not conclude that the pattern differed significantly between the two. However, for two measures, behaviour in the two experiments did differ significantly. There was a significant Typo x Experiment interaction in the measure of probability of skipping the target-1 word, $F1(1, 56)=535$, $p< .05$; $F2(1, 15)$ 15)=7.11, $p<.05$, and a significant interaction in the probability of executing a regressive saccade back from the target, $F1(1, 56)=6.89$, $p<.05$; $F2(1, 15)=8.89$, $p<.01$. For all other measures, either $F1$ or $F2$ was close to 1.0. We conclude that the experiments differed as far as target-1 skipping probability and target regression probability are concerned. With regard to first fixation duration on the target-2 word (the primary parafoveal-on-foveal effect), a typo effect was possibly present in both experiments, although we only have clear evidence for it in Experiment 2.

The result obtained for target-2 first fixation duration strongly suggests that some property of the target word was taken into account at the moment when the decision to leave the target-2 word was being made. Since the initial bigram familiarity of the target was controlled in Experiment 2, the most plausible determining factor is lexical status (i.e., the fact that the target was temporarily a nonword in the typo condition), although an alternative possibility is poor control for some other sublexical factor. For example, although the first bigram was controlled and the second one (corresponding to letter positions 2 and 3) was maintained constant in the typo and no-typo conditions, the initial trigram, as a whole, was not. Post-hoc measurements revealed a trigram-based CLS of 660 and 1219 per million for the typo and no-typo conditions, respectively. The difference is quite substantial and can possibly account for the obtained effects. Moreover, it could be the case that the target initial trigrams were more (or less) "constraining" (Lima & Inhoff, 1985) or "informative" (Kennedy et al., 2002; Pynte, Kennedy, & Murray, 1991) in the typo and no-typo conditions. These two notions both refer to the number of lexical candidates consistent with a given word's initial letters (i.e., the type frequency of this sequence of letters) and will be considered as equivalent hereafter. Post-hoc measurements confirmed that there was, indeed, a difference: The mean number of same-length lexical candidates consistent with the target initial trigram was less in the typo condition (2.00 vs. 4.14). That is, the altered initial stimulus configuration in the typo condition was consistent with fewer real words. Both initial letter familiarity (Kennedy, Murray, & Boissiere, 2004) and/or initial letter constraint (Kennedy, 2000; Kennedy et al., 2002; Underwood et al., 2000) may act to produce parafoveal-on-foveal timing effects, although, as outlined in the introduction, this remains a somewhat contentious claim. Experiment 3 set out to distinguish between possible lexical and sublexical sources of influence on saccade targeting and timing.

EXPERIMENT 3

Method

Participants. Twenty-four Psychology students at the University of Provence, Aix-en-Provence, volunteered for the experiment. None had participated in the previous experiments. All were native speakers of French. All had either normal or corrected-to-normal vision. Participants were not paid for taking part.

Task and procedure. The task and procedure were identical to those employed in Experiments 1 and 2.

Materials. As in the previous experiments, we were interested in inspection strategies developed on three critical words referred to as target-2, target-1, and target (e.g., the sequence “commencer a vrombir” in example 5). Only properties of the target were manipulated and, as in the previous experiments, the target-1 word was always very short and predictable. There were 96 six-, seven-, or eightletter word targets (no-typo condition) and 96 nonword targets (typo condition) used in the experiment. Half the word targets were “constrained” and the other half were “unconstrained”, defined in terms of their initial trigram. Constrained and unconstrained words had similar lexical frequency (mean=15 and 11 per million, respectively, $F<1$). A “constraining” trigram was compatible with only one word of a given length (plus or minus one character), namely the word target itself. In contrast, “unconstraining” trigrams were compatible with 25 words of similar length on average. Each constrained word was paired with an unconstrained one (of similar lexical frequency) sharing its second bigram (i.e., letters located at positions 2 and 3). For example, “vrombir” was paired with “croupir”. The sequence of letters v,r,o, form a constraining initial trigram (“vrombir” is the only word in French, among six-, seven-, and eight-letter words, beginning with “vro”), and the sequence c,r,o forms an unconstraining initial trigram. Moreover, “vrombir” and “croupir” share the letters r and o at positions 2 and 3, respectively. Other positions were not controlled (the fact that the final bigram is also identical in the examples given is a coincidence). Each target word was embedded in a sentence. Paired words were embedded in the same base sentence at the same location (examples 5 and 6).

(5) Les engins vont commencer a vrombir dans l’atelier de réparation.[6]
(5') Les engins vont commencer a crombir dans l’atelier de reparation.
(6) Les engins vont commencer a croupir dans l’atelier de réparation.[7]
(6') Les engins vont commencer a vroupir dans l’atelier de reparation.

[6] The engines are going to start humming in the workshop.

[7] The engines are going to start to go rotten in the workshop.

In half the cases a typo was introduced in the target word by replacing its initial letter with the first letter of its paired word (resulting in a nonword). For example, the typo version of “vrombir” was “crombir” and the typo version of “croupir” was “vroupir”. Given that paired words shared their second bigram, this meant that the initial trigram of any item in the typo version was identical to the initial trigram of the corresponding paired word (compare examples 6 to 5’ and 5 to 6’). Since the manipulation involves matching both changed and unchanged versions, it was only possible in practice to cross “lexical” (i.e., wordness) and “sub-lexical” (i.e., informativeness and/or familiarity) properties. No attempt was made to cross constraint with familiarity. Indeed, it would have been extremely difficult to achieve this. Thus, the “constrained” words were also “unfamiliar”.

Design, In summary, four conditions were compared, depending on both the lexical status of the target and the properties of its initial trigram before and after the eyes had reached it. In the two no-typo conditions, the target remained a legal word over sentence presentation. Moreover, the same initial trigram (either constraining or unconstraining) was available both foveally and in the parafovea. These two conditions will be referred to as the constrained (C) and the unconstrained (U) conditions, respectively.[8] In contrast, in the typo conditions, the properties of its initial trigram (and hence lexical status of the target) varied over time. A parafoveal constrained nonword (e.g., “vroupir”) was turned into a foveal unconstrained word (“croupir”), whereas a parafoveal unconstrained nonword (e.g., “crombir”) was turned into a foveal constrained word (“vrombir”). The two typo conditions will be referred to as the constrained/unconstrained (CU) and the unconstrained/constrained (UC) conditions, respectively.

Four experimental lists were prepared in order to balance the association of items with conditions across participants. Each list comprised 48 experimental items embedded in base sentences mixed with 48 filler sentences and preceded by 8 training sentences. In any given list, a given base sentence was presented with either a constrained target word or with the corresponding paired unconstrained word and either contained a typo or not. It should be borne in mind that, in any case, typos were only present while the eyes were to the left of the defined invisible boundary. Participants were randomly distributed in four groups, each group being associated with a given experimental list. Each participant was thus presented with all the base sentences and all the conditions (each sentence only once, 12 repeated measures for each condition).

[8]We shall refer to this manipulation as “constraint”, although, as noted, it represents a combination of initial letter constraint and familiarity.

Results and discussion

The results are summarised in Table 4. As in the previous experiments, three measures were recorded on the target-2 word (first fixation duration, refixation probability, and gaze duration); one measure on the target-1 word (skipping probability); and four measures recorded on the target itself (first landing position, first fixation duration, gaze duration, and the probability of making a regression). An analysis of variance was carried out for each of these measures, using a 2 (word vs. nonword target)x2 (constrained vs. unconstrained target) factorial design. There was no significant main effect of the typo manipulation for any of these measures, all $Fs<1$. This outcome effectively rules out the lexical status (word vs. nonword) of the target item in the parafovea as the property responsible for reliable foveal or parafoveal-on-foveal effects.

TABLE 4
Target, target-1, and target-2 inspection strategies in the four experimental conditions (Experiment 3)

	Target-2			*Target-1*	*Target*			
	ff	*rf*	*gaze*	*skip*	*flp*	*ff*	*reg*	*gaze*
No-typo								
C	249	0.35	345	0.66	3.2	310	0.13	335
U	276	0.36	359	0.63	3.0	297	0.15	345
Typo								
UC	272	0.36	363	0.61	3.5	323	0.12	348
CU	250	0.32	350	0.61	3.1	291	0.18	333

ff=target-2 and target first fixation duration; rf=target-2 refixation probability; gaze=target-2 and target gaze duration; skip=target-1 skipping probability; flp=target first landing position; reg = target regression probability; C=parafoveally and foveally constrained target; U=parafoveally and foveally unconstrained target; UC=parafoveally unconstrained and foveally constrained target; CU = parafoveally constrained and foveally unconstrained target.

Turning to the effect of properties of the target's initial letters, it should be recalled that, because of the contingent presentation procedure, these varied during inspection, depending on the position of the eyes. Assessing the influence of this factor thus involves different types of comparison for the different measures (depending on whether the target was available in the parafovea or in the fovea at each stage of the inspection process). At the point in time when participants were inspecting the target-2 word and/or determining whether to skip the target-1 word, no contingent change had occurred. Accordingly the relevant comparisons involve grouping together conditions C and CU on the one hand, and U and UC on the other hand (i.e., the two conditions in which the target's initial trigram was constraining vs. the two conditions in which it was unconstraining). The duration of target-2 initial fixation was significantly shorter when the (yet to be fixated) target initial trigram was constraining than when it was unconstraining (250 vs. 274ms), $F1(1, 20)=16.34$, $p<.001$; $F2(1, 44)= 9.84$, $p<.01$. The effect, although in the same direction, was not reliable in the measure of gaze duration, $F1<1$; $F2(1, 44)=1.73$, but reappeared significantly when postregression fixations were included in the measure, $F1(1, 20)=4.54$, $p = .05$; $F2(1, 44)=5.11$, $p<.05$. No significant effect was found for either the measure of target-1 skipping probability or target initial landing position, $Fs<1$.

To answer the question as to whether properties of the target's initial trigram influenced inspection of the target word itself, it is necessary to group together conditions C and UC on the one hand and conditions U and CU on the other hand. This is to take account of the fact that, in the typo condition, at the point when the target was eventually fixated, its initial trigram had been changed. Constrained items had been turned into unconstrained ones, whereas unconstrained items had been turned into constrained ones. The relevant comparison was significant for the measure of first fixation duration, with longer initial fixations in the constrained conditions than in the unconstrained ones (317 vs. 294 ms), $F1(1, 20)=8.53$, $p<.01$; $F2(l, 44)=6.31$, $p<.05$. Note that this significant effect is in a direction which is the reverse of the significant parafoveal-on-foveal effect found for initial fixations on the target-2 word, where the unconstrained conditions were associated with longer fixation durations. There was a near-perfect trade-off, with the combined effect of constraint over the two inspection sites effectively zero (567 vs. 568 ms). This aspect of our results will be further examined in the General Discussion section. No significant effect was found for gaze duration, Fs<1, and regression probability, $F1(1, 20)=2.16$; $F2(1, 44)=1.43$.

The shortening of inspection time on the target-2 word produced as a result of the parafoveal typo manipulation in Experiment 2 was replicated in the present experiment. The factor responsible for the effect was the informativeness and/or orthographic familiarity of the target's initial letters, not the lexical status of the target in the parafovea. Obviously, this does not mean that the target lexical status did not contribute in some way to the difference in target-2 inspection time in Experiment 2, since lexical status was to an uncontrolled degree confounded with informativeness in that experiment. But, when lexical and sublexical factors are crossed, as it was the case in Experiment 3, it is patently sublexical properties that exert the stronger influence. In fact, we did not obtain any evidence that lexical status ("wordness") mattered at all.

GENERAL DISCUSSION

The experiments reported in this paper show some of the ways in which foveal inspection strategies vary as a function of different properties of words and nonwords only present in parafoveal vision. In order to make the task as close as possible to normal reading, properties of a single target item were manipulated while participants were engaged in a sentence comprehension task.

Thus, from the participants' point of view, the experiment involved reading for comprehension, even though a typographical error was present from time to time (a situation readers are frequently confronted with). Since the contingent presentation procedure ensured that the correct text was restored as soon as the eyes reached the corresponding region, typos were never available for foveal processing. We believe this greatly increased the ecological validity of the task. The presentation of overt typos in experimental materials (as, for example, in the studies by White & Liversedge, 2004) is almost bound to trigger task-specific strategies (i.e., the participants are bound to form hypotheses as to why the materials contain errors). In general terms, therefore, the present data add further weight to the proposition that parafoveal-on-foveal effects are not restricted to highly artificial laboratory tasks. Our hypotheses met a complex fate, however, and must be discussed in more detail.

We raised two questions in the context of the targeting hypothesis: (1) Which parameter(s) of visual inspection (if any) did the typo manipulation exert an influence on; and (2) which aspect (if any) of the stimuli was the system sensitive to? In Experiment 1, the typo manipulation involved replacing the first letter of the target word with a second occurrence of its second letter, thus producing an illegal nonword (e.g., "chaîne de éélo"). A significant difference was found in the probability of skipping the target-1 word (i.e., the word "de" in the previous example), with a lower skipping probability in the typo condition than in the no-typo condition. In order to determine whether the source of the obtained effect was lexical or sublexical in nature ("éélo" is illegal at both levels), the familiarity of the target's initial letters was maintained constant across conditions in Experiment 2 (e.g., "chaîne de vélo" vs. "chaîne de célo"). This manipulation removed the target-1 skipping effect, suggesting that the outcome of Experiment 1 was due to orthographic illegality, rather than to lexical illegality. However, another interesting typo effect was revealed by Experiment 2: Inspection times on target-2 were shorter (not longer) in the typo condition, relative to the no-typo condition. Was this effect due to lexical illegality? That is, did participants detect that the letter string in the parafovea was a nonword in the typo condition and execute an earlier saccade (a GO response)? This seems unlikely. The decision as to whether a sequence of letters is not a word is known to take quite a long time in lexical decision experiments (and response times are usually longer for "nonword" than for "word" decisions). A more plausible explanation is that a sublexical property of the target led to an early saccade in the typo condition, and the post-hoc measurements of initial letter constraint appear to confirm this. We conclude that what determined the time at which a saccade was executed in Experiment 2 was not parafoveal "wordness", but some more easily computed property defining the "attractiveness" or "oddity" of the unfixated stimulus.

Experiment 3 was carried out to clarify this point. The number of lexical candidates likely to be generated by the initial letters of the target was systematically varied in Experiment 3. As in Experiment 2 a modulation of target-2 inspection time was found as a function of the manipulated (parafoveal) property. Moreover, this effect was present regardless of the lexical status of the target word in the parafovea, crucially suggesting that lexical illegality is not a necessary condition for sublexical properties to exert an influence. As noted above, in order to maximise the contrast, type frequency was confounded with token frequency in Experiment 3 and it is possible that the observed differences in first fixation duration on the target-2 word relate to a combination of "familiarity" and "informativeness". This is not to deny that the distinction is of some theoretical consequence,[9] but it is not critical in the context of the present series of experiments, since our aim was simply to oppose lexical (i.e., word vs. nonword) and sublexical (i.e., informativeness or familiarity) determinants.

We will now turn to consider the outcome in terms of the proposed targeting hypothesis, offering some comparisons with predictions from the leave-oncompletion hypothesis. First, in Experiment 1, the lack of any significant effect on landing position, as well as the effect on the probability of skipping the target-1 word (with a lower, not higher, skipping rate associated with the typo condition) seems inconsistent with the targeting hypothesis, or, at least, with its strongest version, which states that visually salient features exert some kind of "magnetic attraction" (see case 3 in the introduction). Nonetheless, the obtained skipping effect indicated that interword saccades may be affected by the presence of an orthographic illegality in the parafovea and it is extremely difficult to account for this in the framework of the leave-on-completion hypothesis. According to that hypothesis, skipping occurs when processing of the to-beskipped word has been completed in parafoveal vision and before the eyes have a chance to reach it. From this perspective it would need to be argued that target-1 processing was completed sooner in the no-typo condition than in the typo condition and there are no grounds at all for making this claim. Properties of the short target-1 word were constant across conditions: Only properties of the target were manipulated. It follows that target-1 processing difficulty cannot be responsible for the effect. Nonetheless, it might be possible to save the leaveon-completion account by arguing that the observed reduction in skipping probability in the typo condition was the result of a late inhibition process, occurring after the decision to skip target-1 had been made. For example, in the E-Z Reader model of eye movement control in reading (Reichle et al., 1998), attention is allowed to shift to the parafovea after the decision to move the eye has been made. Parafoveal information becomes available at this time and if a typo were to be detected early enough (for example during the "labile" stage of motor programming), the saccade could possibly be inhibited, thus accounting for the observed reduction in skipping probability.

[9] The locus of type and token frequency effects may be different, with token frequency exerting its influence at an early (possibly preattentive) stage and type frequency playing a role at a later (possibly attentive) stage involving candidate generation.

The results of Experiments 2 and 3 are even more damaging for the leave-oncompletion hypothesis. The difference found for target-2 single fixation duration in Experiment 2 suggests that target-directed saccades were initiated earlier in the typo condition than in the no-typo condition (the combined analyses involving Experiments 1 and 2 suggest this effect may have been present in Experiment 1 also). This outcome clearly contradicts the notion that completion of foveal processing is the sole determinant of interword saccade latencies. In line with the targeting hypothesis, it must be assumed that some particular property of the target word was taken into account before the decision to leave the target-2 word was made. Unlike the skipping effect in Experiment 1, it is not possible to attribute this difference to the operation of a late adjustment mechanism. First, it was quite an early effect. Second, shorter (not longer) inspection times were associated with the typo condition. All of which suggests that the relevant mechanism must operate before the saccade was initiated. Once the decision to move the eyes has been made it may be possible to modify the landing site of the saccade, or possibly delay its execution, but it is not possible to get the saccade started even earlier! The fact that saccade execution was apparently speeded up in the typo condition can only be interpreted as an indication that parafoveal information interfered with the decision to leave the target-2 word.

The observation that the presence of a typographical error in parafoveal vision acts to *speed* foveal processing time seems paradoxical, but the paradox is only apparent. First, it should be recalled that lexical illegality *per se* must be dismissed as a relevant factor explaining typo effects. Typos seem to exert their influence at a sublexical level. Target informativeness (and/or orthographic familiarity) shortened inspection time on the target-2 word in both the typo and no-typo conditions of Experiment 3. Second, and equally important, typos were not invariably associated with paradoxical processing advantages. It is possible to point to evidence of adverse effects. For example, in Experiment 1 the reduction in target-1 skipping probability observed in the typo condition was such an adverse effect, with a less effective inspection strategy (more fixations) in the typo condition relative to the no-typo condition.[10] Similarly, in Experiment 3, the same manipulation that apparently speeded up target-2 processing was traded off against an opposite effect when the target was eventually fixated (i.e., first fixation duration was longer, not shorter, for informative items). We believe a plausible account for this outcome can be derived from the fact that "constraining" target beginnings in Experiment 3 were both lexically more informative and orthographically less familiar than

[10] A referee has pointed out that this outcome is compatible with psychophysical data provided by Rao, Zelinsky, Hayhoe, and Ballard (2002). In some circumstances the eyes are not moved with one saccade to a specified target in the parafovea but take an inevitable intermediate step on the way there. The illegal letter sequences found in Experiment 1 may have been salient enough to trigger this kind of response.

"unconstraining" beginnings. These two factors may exert their influence at different times during visual inspection. When the eyes were still left of the target (the more constrained the target the less time spent on target-2), the relevant factor was possibly informativeness (e.g., the number of lexical entries compatible with the target word's initial letters). In contrast, the relevant factor was possibly orthographic familiarity when the target word was processed in foveal vision (shorter first fixations for less constrained target words).

In conclusion, visually salient features (Experiment 1) seem to influence fixation positions without saccades being directed towards them. Unfamiliar or informative regions in the parafovea (Experiments 2 and 3) also seem to attract the eyes, but this seems to translate into a shortening of foveal processing time, rather than a landing position effect. As a whole, these results are consistent with a version of the targeting hypothesis in which inspection strategies are developed over several words, while the system is continuously monitoring the sublexical properties of potential targets in the parafovea and computing the best moment-to-moment viewing position, in order to optimise overall visibility. This notion has proved useful for interpreting various parafoveal-on-foveal effects reported in the literature (see Kennedy et al, 2002, for further discussion) and provides a parsimonious account for the diversity of sublexical parafoveal factors that appear to exert an influence on foveal inspection strategies. It also accounts for the apparently paradoxical moment-to-moment changes in the direction of the effects observed in the present studies. From this perspective, the shortening of target-2 inspection time observed in Experiments 2 and 3 suggests that the eyes are affected by the presence of lexically informative and/or orthographically unfamiliar regions in the parafovea. Similarly, the fact that participants apparently preferred to land left of target in the typo condition in Experiment 1 may merely mean that this was possibly the most convenient location for processing an illegal typo in the conditions of the experiment.

PrEview proof published online September 2003

REFERENCES

Beauvillain, C., & Beauvillain, P. (1995). Calibration of an eye movement system for use in reading. *Behavior Research Methods, Instruments and Computers, 55,* 1–17.

Beauvillain, C., & Doré, K. (1998). Orthographic codes are used in integrating information from the parafovea by the saccadic computation system. *Vision Research, 38,* 115–123.

Brysbaert, M., & Vitu, F. (1998). Word-skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 135–147). Oxford, UK: Elsevier.

Content, A., Mousty, P., & Radeau, M. (1990). Brulex: Une base de données lexicales informatisée pour le francais écrit et parlé. *L'Année Psychologique, 90,* 551–566.

Engbert R., Longtin A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636.

Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22,* 661–721.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50A,* 119–148.

Hebb, D. (1949). *The organization of behaviour.*New York: Wiley.

Henderson, J. M, & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

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

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology, 16*(1/2), 104–127.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Hyönä, J., & Pollatsek, A. (2000). Processing of Finnish compound words in reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 65–87). Oxford, UK: Elsevier.

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 221–246). Oxford, UK: Elsevier.

Kennedy, A. (1995). *The influence of parafoveal words on foveal inspection time.* Paper presented at the AMLaP-95 conference, Edinburgh, UK.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–223). Oxford, UK: Elsevier.

Kennedy, A. (2000). Attention allocation in reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 193–220). Oxford, UK: Elsevier.

Kennedy, A., Murray, W.S., & Boissiere, C. (2004). Parafoveal pragmatics revisited. *European Journal of Cognitive Psychology, 16*(1/2), 128–153.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Lima, S.D., & Inhoff, A.W. (1985). Lexical access during eye fixations in reading: Effects of word initial letter sequence. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 272–285.

McConkie, G.W., Kerr, P.W., & Dyre, B.P. (1994). What are "normal" eye movements during reading: Towards a mathematical description. In J.Ygge & G.Lennerstrand (Eds.), *Eye movements in reading* (pp. 315–327). Oxford, UK: Elsevier.

McConkie, G.W., Kerr, P.W., Reddix, M.D., & Zola, D. (1988). Eye movement control during reading: 1 The location of initial eye fixations on words. *Vision Research, 27,* 227–240.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Murray, W.S. (1998). Parafoveal pragmatics. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 181–199). Oxford, UK: Elsevier.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton Century Crofts.

O'Regan, J.K. (1979). Eye guidance in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics, 25,* 501–509.

Pynte, J., Kennedy, A., & Murray, W. (1991). Within-word inspection strategies in continuous reading: Time course of perceptual, lexical and contextual processes. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 458–470.

Pynte, J., & Prieur, B. (1996). Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes, 11,* 165–191.

Radach, R., Inhoff, A., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology, 16*(1/2), 27–51.

Radach, R., & Kempe, B. (1993). An individual analysis of initial fixation positions in reading. In G. d'Ydewalle & J.Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 213–225). Amsterdam: North-Holland.

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

Rao, R.P. N., Zelinsky, G.J., Hayhoe, M.M., & Ballard, D.H. (2002). Eye movements in iconic visual search. *Vision Research, 42,* 1447–1463.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

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

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

Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity and lexical ambiguity. *Memory and Cognition, 14,* 191–201.

Rayner, K., Reichle, E.D., & Pollatsek, A. (2000). Eye movement control in reading: Updating the E-Z Reader model to account for initial fixation locations and refixations. In A.Kennedy, R. Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 701–719). Oxford, UK: Elsevier.

Rayner, K., Sereno, S.C., & Raney, G.E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 88–120.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39,* 4403–4411.

Schiepers, C. (1980). Response latency and accuracy in visual word recognition. *Perception and Psychophysics, 27,* 71–81.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Visual attention and eyemovement control during reading: The case of parafoveal processing. *Quarterly Journal of Experimental Psychology, 52A,* 1021–1046.

Vonk, W., Radach, R., & van Rijn, H. (2000). Eye guidance and the saliency of word beginnings in reading text. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 269–299). Oxford, UK: Elsevier.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 247–268). Oxford, UK: Elsevier Science.

White, A., & Liversedge, S. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*(1/2), 52–78.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 203–225

Attention allocation to the right and left of a fixated word: Use of orthographic information from multiple words during reading

Matthew S.Starr and Albrecht W.Inhoff

State University of New York at Binghamton, USA

Eye movement contingent display changes were used to manipulate the information to the right of a fixated target in one experiment and to its right and left in a second experiment. Experiment 1 showed large increases in target viewing duration when orthographically illegal information was visible to the right of the target. Experiment 2 yielded smaller effects of orthographically illegal information in the parafovea on target viewing. Experiment 2 also showed that readers obtained useful information to the right of a fixated target irrespective of the type of information that was visible to the target's left. These results indicate that orthographic information is obtained from more than one word prior to saccade programming, and that the processing of orthographic information to the right and left of a fixated target is functionally independent.

Since reading differs from spoken language comprehension in a number of fundamental aspects, learning to read requires the acquisition of task specific skills, among them the spatial selection of successive words of text for processing. Because the range of high-acuity vision is relatively small, skilled reading also requires the coordination of this spatial selection process with the programming of corresponding eye movements (saccades). Not all saccades move the eyes from a fixated word to the next word in the text, however. Some skip over it and land on a subsequent word, and some (regressions) are directed at a previously read word instead (see Rayner, 1998).

In between saccades, the eyes are relatively stationary (fixated), and visual and linguistic properties of a fixated word, e.g., its phonological properties, frequency of occurrence, or predictability within a sentence context, influence the time spent viewing it. Acquisition of useful information is not confined to the fixated word, however. Readers also obtain information from the next (parafoveal) word in the text and less time is spent viewing it when it was visible prior to its fixation than when it had been masked, a phenomenon referred to as *parafoveal preview benefit* in the literature (see Rayner, 1998, for a comprehensive review).

One class of eye movement control models, sequential attention shift models, provides an elegant account for effects of word properties on fixation durations, parafoveal preview benefit, word skipping, and regressions to the word to the left of fixation. The key assumptions of these models are that an attentional spotlight controls the spatial selection of a word for processing and that spatial selection and saccade programming are coordinated (Henderson & Ferreira, 1990; Inhoff, Pollatsek, Posner, & Rayner, 1989; Morrison, 1984; Pollatsek, Reichle, & Rayner, 2003; Rayner & Pollatsek, 1989; Rayner, Reichle, & Pollatsek, 2000; Reichle, Pollatsek, Fisher, & Rayner, 1998).

According to the original model (Morrison, 1984), attention confines linguistic processing to one word at a time and sustains its recognition, so that engagement and disengagement of attention defines the onset and offset of a selected word's linguistic processing. Once a word is recognised, the spotlight of attention is shifted to the next word in the text. Three other assumptions explain the linkage between covert shifts of attention and overt eye movements: (1) More than one spatial shift of attention can be initiated during a fixation, (2) each shift of attention initiates the programming of a corresponding saccade, and (3) saccade programming is relatively time consuming, involving a phase during which saccade programming can be cancelled and a phase during which saccade planning has reached a "point of no return".

The yoking of saccade programming to shifts of attention accounts for effects of word properties on fixation durations. Since saccade programming is more time consuming than the shifting of attention, attention will generally arrive at the parafoveal word before the eyes do, thus explaining parafoveal preview benefits

Correspondence should be addressed to A.Inhoff, Dept of Psychology, State University of New York at Binghamton, Binghamton, NY 13902, USA. Email: inhoff@binghamton.eduMatt Starr is currently working for IBM's human factors division. This research is based on parts of his doctoral dissertation which was submitted to the Graduate School of the State University of New York at Binghamton. The work was supported by Grant SBR-9810169 from the National Science Foundation and preparation of this paper was supported by a NIH Grant 525872 PostDoctoral Fellowship to Matt Starr. We thank Lars Placke, Alan Kennedy, Keith Rayner, and Erik Reichle for their helpful comments on an earlier version of the manuscript.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI:10.1080/09541440340000150

when a parafoveally visible word is subsequently fixated. Moreover, the parafoveal word can be skipped when the shifting of attention toward it leads to its recognition before the programming of the corresponding saccade has reached a point of no return. A second shift of attention during the fixation, directed at the subsequent word, will then cancel the originally planned saccade and direct the eye to the next word in the text instead.

A more recent version of the model (Rayner et al., 2000; Reichle et al., 1998) also accommodates regressions to the left of the target. In contrast to the original model, Reichle et al.'s family of E-Z Reader models no longer assumes that shifts of attention and saccade programming coincide. According to the model, allocation of attention to a visual word initiates a sequence of two functionally distinct processes, familiarity checking, which involves the computation of global linguistic word properties, and lexical completion, which culminates in word recognition. New to the model is that interword saccade programming begins after an attended word's familiarity checking has been completed. As in earlier models, word recognition leads to a shift of attention.

As in the original model, properties of a fixated word will influence its fixation duration and readers will obtain useful parafoveal information when the shifting of the spotlight of attention precedes the execution of a corresponding eye movement. The magnitude of the parafoveal preview benefit is now influenced by the ease with which the fixated target is recognised, however. A difficult lexical completion will delay the shifting of attention to the next word, thus decreasing preview benefits (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999). Further, when processing of the fixated word is particularly time consuming, the eyes may even arrive at the next word before a corresponding shift of attention has been completed. A regression is programmed under these conditions and the eyes are redirected to the attended word. Consistent with this, Binder, Rayner, and Pollatsek (1999) showed that readers obtain useful information from the word to the left of fixation when the eyes regressed back to it.

According to the model, a reader can thus obtain useful information to the right or left of a fixated target word during a fixation. Nevertheless, processes that demand attention, i.e., familiarity checking and lexical completion, are confined to one word at a time. A fixated word's familiarity cannot be checked, for instance, when the lexical processing of a word to the left fixation is still in progress. Analogously, the checking of the familiarity of the next word in the text cannot be initiated while the lexical completion process of a fixated word is still in progress. Since the programming of a saccade to the next word in the text is initiated before a corresponding shift of attention occurs, the model predicts that neither a parafoveal word's familiarity nor its meaning can influence saccade programming.

The results of a number of recent studies indicated, however, that familiarity- and identity-defining information from the next word in the text can influence the programming of a saccade to it. In one of our experiments (Inhoff, Radach, Starr, & Greenberg, 2000a), readers spent more time viewing a target word, e.g., *mother's,* when the next word in the text was semantically unrelated, e.g., *garden,* than when it was either repeated, *mother,* or a close semantic associate, *father*. The movement of the eyes leaving the target was thus a function of the parafoveal word's structure and meaning. Other results confirmed effects of the next word's orthography on target viewing (Inhoff, Starr, & Shindler, 2000b; Kennedy, 1998, 2000; Rayner, 1975; Underwood, Binns, & Walker, 2000), and effects of the next word's pragmatic sentence role (Murray, 1998).

In view of this, we (Inhoff et al., 2000a, 2000b) proposed that allocation of attention is not necessarily confined to one word at a time. Although an attentional gradient may give preferential processing to a fixated word, some attention may be allocated to spatially adjacent words as well. This will be at the expense of the fixated target, increasing its viewing duration. We refer to this view as the *attention-gradient hypothesis* in the following. Kennedy (1998, 2000) also proposed that readers acquire useful information from a parafoveal word before the saccade to it is programmed. Processing difficulties in the parafovea decreased target viewing durations in his experiments, however, and parafoveal processing difficulties were assumed to attract attention. This is referred to as the *attention attractor hypothesis*.

Effects of parafoveal processing difficulty on target viewing have not converged across studies and have led to conflicting hypotheses. Moreover, some studies have not shown any effects of a parafoveal word's frequency of occurrence on target viewing (Henderson & Ferreira, 1993; Rayner, Fischer, & Pollatsek, 1998), and Rayner, White, Kambe, Miller, and Liversedge's (2003) comprehensive review of relevant literature questions the validity of parafovea-on-target effects in general. Specifically, Rayner et al. (2003) concluded that it has not been conclusively demonstrated that lexical properties of a parafoveally visible word influence target viewing. They further suggest that any such effect is due to "oculomotor error", i.e., instances in which the eyes were directed at a word adjacent to the target but landed on the target instead. Although Rayner et al. (2003) did not rule out that orthographic properties of a parafoveally visible letter string can influence target viewing, they also proposed that oculomotor error is likely to cause or inflate such an effect as well.

Experiment 1 replicated a key condition of our earlier study (Inhoff et al., 2000b) in which an orthographically illegal string of letters was shown to the right of a fixated target (ROF). Based on our earlier findings, we predicted that such an orthographically illegal letter string would again increase target viewing durations. To determine whether this anticipated effect is due to the random letter string's orthographic illegality or its lack of lexical status, Experiment 1 also included a pseudoword condition in which an orthographically legal nonword occupied the ROF location. In a third condition, a contextually inconsistent legal word occupied the ROF location, to determine whether the meaning of the ROF word influenced saccade programming. The ROF word was fully visible throughout sentence reading in a control condition. Similar data selection criteria were applied as in earlier studies. Additional supplementary analyses sought to exclude potential effects of oculomotor error. This was accomplished by removing all instances from the data in which the eyes may have been directed at a word flanking the target but landed on the target instead.

EXPERIMENT 1

Method

Participants. Twenty-four undergraduate students at the State University of New York at Binghamton received experimental course credit for their participation. All students had uncorrected vision and were native English speakers.

Materials. Forty-eight sentences were constructed, each containing three critical words: a pretarget word, also referred to as the left of fixation (LOF) word, a target word, and a posttarget word, also referred to as the right of fixation (ROF) word. The full set of materials is listed in the Appendix. To increase the likelihood that readers obtained useful information from the words to the right and left of fixation, all three critical words were closely associated (see Inhoff et al., 2000a). For instance, in the triplet *steaming bath water* the target, *bath,* is closely related to its LOF and ROF words. The relatedness of the two word pairs of each triplet was determined by asking 18 independent participants to rate (on a scale of 1 to 10) the degree of semantic relatedness of each pair (e.g., *steaming/bath* and *bath/water*). Stimuli were eliminated if the rating for the LOF/target pair differed by more than two points from the rating for the ROF/target pair, and triplets were removed if their overall mean relatedness rating was less than six. Target, LOF, and ROF words had mean word frequencies of 44 (SD=64), 53 (SD=68), and 65 (SD=95) words per million, respectively (Kučera & Francis, 1967). These means were not significantly different (p#DXGT#.11). LOF word lengths ranged from 3 to 11 letters, with a mean of 6.29 (SD=1.58), and ROF word lengths ranged from 3 to 10 letters, with a mean of 5.63 (SD=1.48). Target words ranged from 3 to 8 letters, with a mean of 5.32 (SD=1.11). The critical three-word sequence ranged from 14 to 21 letters and was embedded in a sentence that did not contain any lexical or syntactic ambiguities. Each sentence ranged in length from 50 to 78 characters and always occupied a single line of text. None of the three critical words occupied the first or last position in a sentence.

Four ROF conditions were created. In the three experimental conditions, the ROF word was either replaced with an orthographically dissimilar string of quasirandom letters (e.g., *bath bvkln*), an orthographically legal pseudoword (e.g., *bath wafar*), or a context-inconsistent word (e.g., *bath night*) until the target was read. A version of Rayner's (1975) boundary technique was used with the space *following* the target word constituting the display change boundary. Saccades that moved the eyes onto this blank space or to its right initiated a

Control condition:

I can't wait to relax in some *steaming bath water* after work.
*

I can't wait to relax in some *steaming bath water* after work.
*

I can't wait to relax in some *steaming bath water* after work.
*

Random letter condition:

I can't wait to relax in some *steaming bath bvkln* after work.
*

I can't wait to relax in some *steaming bath bvkln* after work.
*

I can't wait to relax in some *steaming bath water* after work.
*

Pseudoword condition:

I can't wait to relax in some *steaming bath wafar* after work.
*

I can't wait to relax in some *steaming bath wafar* after work.
*

I can't wait to relax in some *steaming bath water* after work.
*

Inconsistent word condition:

I can't wait to relax in some *steaming bath night* after work.
*

I can't wait to relax in some *steaming bath night* after work.
*

I can't wait to relax in some *steaming bath water* after work.
*

Figure 1. Text visibility during pretarget (LOF word), target, and posttarget (ROF word) fixations in the control, random letter, pseudoword, and inconsistent word conditions in Experiment 1. Each fixation is indicated by an asterisk below the line of text.

display change, so that the intact ROF word was shown upon the fixation of its location in all three experimental conditions. In a control condition, the movement of the eyes across the boundary replaced the previously visible posttarget word with the identical posttarget word so that intact target and ROF words were visible throughout sentence reading (e.g., *bath water*). A sample sentence and illustrations of the four conditions are shown in Figure 1.

Each subject was presented with a total of 96 sentences, of which 48 were experimental (12 sentences in each condition) and 48 were fillers. The 48 experimental sentences were counterbalanced across the four parafoveal conditions.

Apparatus. Eye movements were recorded using a fifth-generation dual-Purkinje SRI eye-tracking system (Crane & Steele, 1985). Viewing was binocular but eye movements were only recorded from the right eye. The system has a relative spatial resolution of <10 minutes of arc and its output is linear over the vertical and horizontal range of the visual display. Analog input from the eyetracker was digitised via a Data Translation 2801-A A-to-D converter housed in a personal computer. The computer controlled the visual display and measured horizontal and vertical fixation coordinates every 2ms. These data were used to determine the size of a lateral saccade to the nearest tenth of a character space and the duration of a fixation to the nearest 2 ms.

Sentences were displayed in VGA mode (640 x 480 pixels) on a NEC 5FGe monitor with 0.28mm dot pitch. All stimuli were shown in light green on a black background in Courier type font. The distance between the reader's eyes and the monitor was set at 65 cm, so that each character of text subtended approximately 0.3° of visual angle. Eye movement contingent display changes replaced each letter of the ROF preview with the intact ROF word when the eyes moved to the right of the target. The different types of ROF previews were shown again, however, when the eyes moved back onto the target or other pretarget words. A custom developed VGA driver routine was used to refresh 111 of the 480 vertical pixel lines of the monitor at a rate of 200 Hz, which resulted in a mean display change implementation lag of approximately 5ms.

Procedure. All participants were tested individually in a dimly lit room. When a participant arrived in the laboratory, a bite bar was prepared for the reader to reduce head movements during the experiment. A two-dimensional calibration of the eye-tracking system began the experiment, and the reader was asked to fixate on each of five monitor positions (left top, right top, left bottom, right bottom, screen centre). Each position consisted of a 12 (vertical) x9 (horizontal) pixel square box. To focus the eyes at each square, a sequence of random numbers, ranging from 1 to 9, with a maximal extension of 10x7 pixels, was shown inside the square with each digit being visible between 100ms and 400ms. The reader's fixation location was sampled for 150ms after each mouse key activation by the experimenter when a box had been visible. This was repeated for each of the five monitor positions. The X/Y A-to-D converter values for the five sampled screen fixations were then mapped onto the corresponding monitor locations. After calibration, five fully illuminated 12x8 pixel squares were shown, one near each of the four corners of the monitor and one in the centre of the screen. The reader's fixation location was indicated on the screen via a 12x8 pixel cross that moved in synchrony with the eyes. To check the accuracy of the calibration, the reader was asked to fixate each of the five illuminated locations and was reminded that the task was not to move the cursor onto the illuminated positions but to merely look at each position. The calibration was considered successful when no more than 12 (vertical) and 8 (horizontal) blank pixels intervened between each of the five calibration check positions and the location of the eye movement contingent fixation cursor.

After calibration, the reader was asked to fixate a 5x4 pixel marker at the left side of the screen and to depress a mouse button to display a sentence. Activation of the mouse button replaced the fixation marker with a sentence and started the recording of eye position. After a reader finished reading a sentence, she/he again depressed the mouse button that erased the line of text, terminating the recording of eye movements for that trial, and displaying three 5x4 pixel boxes which served to check calibration. Once the calibration was checked, an additional press then displayed the next sentence.

Data analyses. The target word was considered fixated when the line of gaze fell on one of its constituent letters or the blank space preceding it. Instances in which the eye tracker failed to follow the eyes, in which the saccade reaching the target was not launched from pretarget text, and in which the eyes regressed to the target after posttarget viewing were excluded from analyses. Outlying fixation durations of less than 50ms and of more than 2000ms were also excluded. Altogether 19% of the trials were removed. Excluded cases were treated as missing values (not as zero values).

To determine effects of a parafoveally visible item on target viewing, target viewing durations were examined at a function of the type of available ROF information. In addition, we examined the time spent viewing the subsequent ROF word to determine parafoveal preview benefits. Although eye movement recordings can be used to compute a large number of measures (Inhoff & Radach, 1998; Inhoff & Weger, 2003), we concentrated our analyses on measures that are likely to index the success of the initial stages of visual word recognition. Specifically, we computed first fixation duration and gaze duration for the target and the subsequently fixated ROF word. First fixation duration consisted of the duration of the first fixation on a critical word irrespective of the properties of the next saccade, and gaze durations consisted of the time spent viewing the word until another word was fixated, i.e., it included the duration of intraword refixations. Instances in which a critical word was not fixated were not included in these data. Skipping rates were analysed separately, as were regressions to target and posttarget words. Omnibus ANOVAs were applied to each set of data to determine whether effects of parafoveal previews on target and posttarget reading were reliable. This was followed by additional paired comparisons that sought to determine the nature of the ROF effect on target viewing and the use of the different types of ROF previews during subsequent ROF reading.

To determine whether ROF effects may be caused by—or inflated by—oculomotor error, supplementary analyses were applied to a subset of the target data. Specifically, we excluded all instances in which the pretarget fixation was within two character spaces of the target and all instances in which the first target fixation fell on one of its last two characters. A pretarget fixation that was relatively close to the target could have been directed at the target but, due to oculomotor error, may have landed on the pretarget instead. Similarly, a saccade that placed the eyes at one of the ending characters of the target may have been directed at the posttarget word but placed it on the target instead. This resulted in the exclusion of 45% of the trials.

TABLE 1

Viewing times, skipping rates, and regression rates for target words as a function of parafoveally visible information to the right of the target

	Parafoveal ROF preview during target viewing			
Oculomotor measure	*Control*	*Random letters*	*Pseudoword*	*Inconsistent word*
FFD	289 (14)	323 (22)	305 (14)	276 (9)
Gaze	331 (16)	386 (25)	332 (18)	309 (12)
Corrected FFD	266 (18)	326 (19)	314 (15)	277 (13)
Corrected gaze	305 (21)	398 (28)	345 (19)	317 (16)
Skipping rate	12 (3)	8 (3)	11 (3)	14 (3)
Regression rate	1 (0.3)	5 (0.2)	2 (0.1)	2 (0.3)

Standard errors of the mean are shown in parentheses. FFD refers to first fixation durations and gaze to target gaze durations. Skipping and regression rates are shown in percentage values.

Results

Target reading. First fixation durations, gaze durations, skipping rates, and regression rates are shown in Table 1 as a function of the type of available ROF preview.

As in Inhoff et al. (2000b), parafoveal preview influenced target viewing—first fixation durations and gaze durations were shorter in the control than in the three experimental conditions. This was expressed in a reliable ROF effect in the omnibus ANOVAs for first fixation durations, $F1(3, 66)=2.40$, $p<.06$, $F2(3, 138)$ [1]$=3.52$, $p<.025$, and for gaze durations, $F1(3, 66)=4.82$, $P<.01$, $F2(3, 138)=5.32$, both $p<.01$.

Paired comparisons revealed that relative to the control condition, target gaze durations increased substantially when a string of random letters occupied the ROF location, $t1(22)=2.67$, $t2(46)$[1]$=2.31$, both $p<.025$. Only one other comparison, that of the contextually inconsistent word condition with the control condition, approached significance in the items analysis but not in the subjects analysis, $t1(22)=1.22$, $p<.2$, $t2(46)=1.91$, $p<.07$. None of the other comparisons approached significance, all *p*#DXGT#.22.

First fixation and gaze durations for the truncated set of data, from which instances of potential oculomotor error were removed, are also shown in Table 1. Again, a robust ROF omnibus effect emerged for corrected first fixation durations, $F1(3, 66)=4.25$, $p<.01$, and corrected gaze durations, $F1(3, 66)= 5.21$, $p<.01$ (no $F2$ statistics were computed due to the relatively small number of trials). Paired

[1] Inadvertently, no data were collected for one of the items.

TABLE 2

Viewing times, skipping rates, and regression rates for posttarget (ROF) words as a function of parafoveally visible information to the right of the target

	Posttarget (ROF-word) viewing							
Oculomotor measure	*Control*		*Random letters*		*Pseudoword*		*Inconsistent word*	
FFD	308	(16)	330	(15)	305	(13)	334	(17)
Gaze	361	(18)	400	(24)	356	(18)	398	(22)
Skipping rate	16	(3)	4	(1)	9	(3)	8	(2)
Regression rate	16	(4)	23	(4)	20	(4)	23	(5)

Standard errors of the mean are shown in parentheses.

comparisons confirmed deleterious effects of a string of random letters to the right of the target, $t1(22)=2.69$, $p<.025$ and $t1(2)=3.66$, $p<.01$, for first fixation durations and gaze durations, respectively. Furthermore, this set of data showed a marginally significant increase in target first fixation durations and gaze durations when a pseudoword replaced the word to the right of the target, $t1(22)=2.50$, $p<.025$ and $t(22)=2.02$, $p<.06$, respectively. None of the other effects approached significance, all p#DXGT#.2.

Target skipping rates and regression rates are also shown in Table 1. Although skips occurred less often in the random letter ROF condition than the other conditions, no reliable effect emerged from the omnibus ANOVAs, $F1(3, 66)=1.93$, $p<.13$, $F2(3, 138)=2.14$, $p<.1$. Regressions back to the target were also slightly more common in the random letter condition than the other conditions but none of the analyses revealed a reliable difference, all $p$values #DXGT#.18.

Posttarget (ROF word) reading. First fixation durations, gaze durations, skipping rates, and regression rates for posttarget words are shown in Table 2.

The results revealed the familiar parafoveal preview benefit effect, with shorter posttarget first fixation durations and gaze durations when the preview provided useful orthographic information, as occurred in the control and pseudoword conditions. The omnibus effects of preview type were reliable only for gaze durations, however, $F1(3, 66)=2.39$, $p<.08$, $F2(3, 138)=3.00$, $p<.05$, in that gazes were shorter in the control condition than in either the random letter condition, $t1(22)=2.15$, $p<.05$, $t(46)=2.44$, $p<.025$, or the inconsistent word condition, $t1(22)=2.79$, $p<.01$, $t2(46)=2.35$, $p<.025$. The difference between the control and pseudoword conditions was negligible, both p-values #DXGT#.5. The parafoveal preview benefits for the control and pseudoword conditions were sizable, more than 20ms, in the first fixation durations but they were not reliable in the corresponding omnibus analysis, both p-values #DXGT#.2.

Preview also influenced the rate of ROF word skipping, $F1(3, 66)=7.77$, $F2(3, 138)=10.07$, both $p<.01$. Paired comparison showed higher skipping rates in the control condition than in the random letter condition, $t1(22)=4.65$, $t2(46)=4.91$, both $p<.01$, the inconsistent word condition, $t1(22)=3.00$, $t2(46)=3.68$, both $p<.01$, and the pseudoword condition, $t1(22)=2.94$, $t2(46)=2.78$, both $p<.01$. The type of ROF preview had no reliable effect on the rate of regression back to the ROF word.

Discussion

The results of Experiment 1 replicate a key finding of Inhoff et al.'s (2000b) study, again showing large and robust increases in target viewing durations when a sequence of random letters was visible to its right. Critically, this occurred even when instances of potential oculomotor error were excluded. The effect of a pseudoword at the ROF location on target viewing was less clear cut. No reliable effect emerged in the main analysis, suggesting that orthographic properties of information at the ROF location, rather than its lexical status, influenced target viewing durations. However, a subset of data did reveal a pseudoword effect, indicating that the lexical status of the letter string at the ROF location can influence target viewing under some conditions. Neither the main nor the supplementary analyses revealed an effect of ROF meaning on target viewing under the conditions of this experiment.

In spite of their limits, parafovea-on-target effects in Experiment 1 are theoretically informative. They favour a view in which readers obtain familiarity defining information from the ROF location before a saccade to that location is programmed. Moreover, the increase in the target viewing durations in the random letter condition of Experiment 1 indicates that an unfamiliar letter sequence at the ROF location did not attract the eyes during sentence reading. Instead, it delayed the execution of a saccade to it, as maintained by the attention-gradient hypothesis.

The time spent reading of the posttarget (ROF) word confirms well established parafoveal preview benefit effects, with shorter viewing durations when the preview revealed useful orthographic information than when it did not (Rayner, 1998). It also showed that the target skipping rates were higher in the control condition than in any of the experimental conditions, thus suggesting that the lexical status and meaning of a parafoveally visible letter sequence influenced saccade programming on some trials.

EXPERIMENT 2

The E-Z Reader model and the attentional-gradient hypothesis differ in how they account for the acquisition of information to the left of a fixated target. According to the E-Z Reader model, this can occur on occasions when the execution of a saccade precedes a corresponding shift of attention. According to our attention-gradient hypothesis, readers can allocate some attention to the left of a fixated target while the target is being processed. Two studies have thus far been conducted, one (Binder et al., 1999) that supports the E-Z Reader model and one (Inhoff et al., 2000a) that is partially compatible with the E-Z Reader and with the attention-gradient hypothesis.

In Binder et al.'s (1999) study, experimental conditions were created in which a base word to the left of fixation was changed when the eyes moved to its right. No change occurred in a control condition. The LOF change was confined to a single fixation and the base word reappeared during all subsequent fixations. The main result revealed longer viewing times on the base word when it was reread in the experimental condition, indicating that readers had obtained some information from the previously inserted LOF word. Consistent with the E-Z Reader model, the base word change did not influence the rate of regressions to the base word location, and it did not influence the duration of the fixation following base word reading, thus suggesting that it was the initial (first pass) base word processing that controlled the regression back to it.

Our experiment (Inhoff et al., 2000a) implemented a change to the left of fixation in which a short base word, e.g., *cat,* was replaced with a graphemically similar word, e.g., *rat,* when the eyes moved to the right of it. A decision concerning the identity of the base word was made after sentence reading. Although readers generally chose the base word, they chose the replacement word more often than a graphemically similar control word, *bat,* that never appeared in the sentence. Similar to Binder et al. (1999), our results thus indicated that readers can obtain effective information from a LOF location. Moreover, similar to Binder et al., and consistent with the E-Z Reader model, LOF processing was not accompanied by increased postchange fixation durations. However, in violation of the E-Z Reader model, effective information was obtained from the LOF location even when the eyes did not regress back to it.

One goal of Experiment 2 was to determine whether familiarity-defining, rather than identity-defining, information to the left of a fixated word can influence target viewing durations. This should not occur according to the E-Z Reader model, but should according to the attention-gradient hypothesis. Specifically, in view of the results of Experiment 1, allocation of attention to a string of random letters to the left of the target should increase target viewing durations.

The second goal of Experiment 2 was to determine whether LOF processing influences the acquisition of useful information at the ROF location. According to the E-Z Reader model, the processing of words at the LOF and ROF locations should be functionally independent. According to the attention-gradient hypothesis, difficulties processing a letter string at the LOF location should impede the acquisition of useful information at the ROF location.

The experiment involved the orthogonal manipulation of information to the right and left of a fixated target. Specifically, three viewing conditions were created, one in which the LOF word was replaced with a string of random letters upon the target's fixation, one in which the ROF word was replaced with a string of random letters, and one in which both ROF and LOF words were replaced. No change occurred in a control condition. Target viewing durations could thus be examined as a function of the type of available LOF (and ROF) information to determine whether orthographically illegal information to the left (and right) of the target influenced target viewing durations, and posttarget (ROF word) viewing durations could be examined to determine whether the presentation of orthographically illegal information to the left during target viewing impeded the acquisition of useful ROF information thus decreasing parafoveal preview benefits.

Method

Participants. Twenty-eight undergraduate students at the State University of New York at Binghamton received experimental course credit for their participation. All had uncorrected vision and were native English speakers. The results of one participant were lost due to equipment failure.

Materials and apparatus. The materials and apparatus of Experiment 2 were identical to the materials and apparatus of Experiment 1.

Procedure. The procedure of Experiments 1 and 2 was identical with one exception. The display change boundary now consisted of the blank space preceding the target word rather than the space following it. That is, an intact experimental sentence was visible until the target was reached in all experimental conditions. When the eyes reached or traversed the boundary, the LOF word, ROF word, or both words were changed. In all conditions, this involved a replacement of the original word(s) with a length-matched string of graphemically dissimilar and orthographically illegal letters. The original LOF word reappeared when the eyes moved to the right of the target, but the change was not latched, i.e., it was reimplemented when the target was refixated. Again, no display changes occurred in a control condition. Figure 2 shows a sample sentence and item visibility at the pretarget (LOF), target, and posttarget (ROF) location when these locations were fixated in the four experimental conditions.

Data analyses, The same selection criteria were applied as in Experiment 1. In addition, all trials were excluded in which the eyes regressed to the LOF position. Altogether 24% of the trials were removed. Since information at the LOF location was manipulated during target viewing, we also computed one additional measure, the rate of regressions to the LOF word. ANOVAs with the factors LOF word (visible or replaced) and ROF word (visible or replaced) were applied to the different sets of target and posttarget data.

Results

Target reading. First fixation durations and gaze durations for target words that received at least one fixation are shown as a function of the available parafoveal information in Table 3.

Control condition:
I can't wait to relax in some *steaming bath water* after work.
*
I can't wait to relax in some *steaming bath water* after work.
*
I can't wait to relax in some *steaming bath water* after work.
*

LOF condition:
I can't wait to relax in some *steaming bath water* after work.
*
I can't wait to relax in some *hdpwndgs bath water* after work.
*
I can't wait to relax in some *steaming bath water* after work.
*

ROF condition:
I can't wait to relax in some *steaming bath water* after work.
*
I can't wait to relax in some *steaming bath bvkln* after work.
*
I can't wait to relax in some *steaming bath water* after work.
*

Dual change condition:
I can't wait to relax in some *steaming bath water* after work.
*
I can't wait to relax in some *hdpwndgs bath bvkln* after work.
*
I can't wait to relax in some *steaming bath water* after work.
*

Figure 2. Text visibility during pretarget (LOF word), target, and posttarget (ROF word) fixations in the LOF, ROF, and dual change conditions in Experiment 2. Each fixation is indicated by an asterisk below the line of text.

The effects of LOF and ROF changes were not reliable in the first fixation durations, all *p*#DXGT#.2. They interacted, however, in that first fixations were shorter in the control condition than in the three experimental conditions, although the effect was reliable only across subjects, $F1(1, 26)=5.80$, $p<.025$, $F2$ $(1, 46)= 2.31$, $p<.14$. Paired comparisons with the control condition as baseline revealed longer first fixations (in the subjects analysis) in the LOF condition, $t1$ $(26)= 3.12$, $p<.01$, $t2(46)=1.58$, $p<.12$, and the ROF condition, $t1(26)=2.69$, $p<.025$, $t2=0.8$, *p*#DXGT#.4. The difference between the dual change condition and baseline did not reach significance, all p-values #DXGT#.2, and none of the differences between the three experimental conditions was reliable, all p-values #DXGT#.2.

Gaze durations showed a different pattern of condition means. There was a small LOF effect of 9ms that was not statistically reliable, $F1(1, 26)=1.99$, *p*#DXGT#.17, $F2<1$, and a slightly larger ROF effect of 15ms that also failed to reach significance, $F1(1, 26)=2.31$, *p*#DXGT#.14, $F2(1, 46)=1.48$, *p*#DXGT#.2.

TABLE 3

Viewing times, skipping rates, and regression rates for target words as a function of parafoveally visible information to the right and left of the target

	Parafoveal information during target viewing			
Oculomotor measure	*Control*	*LOF change*	*ROF change*	*Dual change*
FFD	263 (7)	281 (10)	282 (9)	273 (11)
Gaze	286 (10)	303 (11)	309 (14)	311 (15)
Corrected FFD	264 (7)	287 (11)	273 (10)	285 (12)
Corrected gaze	287 (11)	305 (11)	304 (17)	328 (18)
Skipping rate	20 (3)	17 (3)	20 (3)	17 (3)
Regression rate	8 (3)	9 (3)	11 (3)	18 (3)

Standard errors of the mean are shown in parentheses.

Effects of LOF and ROF changes no longer interacted, $F1$ and $F2<1$, and the added effects of LOF and ROF changes yielded a 25 ms difference between the control and dual change conditions.

To determine whether LOF and ROF effects on first fixation durations were potentially compromised by oculomotor error, we applied the same data corrections as in Experiment 1, which resulted in the exclusion of approximately 50% of the trials. Corrected first fixation duration now revealed a 18ms LOF effect, $F1(1, 26)=5.61$, $p<.025$, and a 4ms ROF difference, $F1<1$, and ROF and LOF effects no longer interacted, $F1<1$. Corrected gaze durations revealed a 21 ms LOF effect, $F1(1, 16)=4.08$, $p<.054$, and a slightly smaller 20ms ROF effect, $F1(1, 26)=2.23$, $p<.15$. Again, these effects combined additively, yielding a 41 ms difference between the control and dual change condition.

The target's skipping rate and regression rates are also shown in Table 3. Since the critical three-word sequence was fully visible prior to target viewing in all four conditions, target skipping rate should not be influenced by the different types of pre- and posttarget visibility during the target's fixation, which was confirmed by the data, all *p*-values #DXGT#.2. Regressions following target viewing were, however, a function of parafoveal information. Specifically, regression rate tended to be higher when a LOF change occurred than when no change occurred although the effect was not reliable across items, $F1(1, 26)=4.60$, $p< .05$, $F2(1, 46)=1.22$, *p*#DXGT#.2. Regression rate was distinctly higher, however, when a ROF change occurred, $F1(1, 26)=4.86$, $p<.05$, $F2(1, 46)=10.16$, $p< .01$. The effect pattern also showed a small interaction of LOF and ROF effects, with a particularly high regression rate in the dual change condition, although this effect was not statistically reliable, $F1(1, 26)=1.44$, *p*#DXGT#.2, $F2(1, 46)=2.44$, $p<.13$. Regressions to the target location thus occurred primarily in the ROF condition.

TABLE 4

Viewing times, skipping rates, and regression rates for posttarget (ROF) words as a function of the previously visible information to the right and left of the target

	Posttarget (ROF-word) viewing							
Oculomotor measure	*Control*		*LOF change*		*ROF change*		*Dual change*	
FFD	269	(10)	263	(11)	313	(13)	308	(13)
Gaze	294	(12)	288	(14)	364	(17)	362	(20)
Skipping rate	22	(3)	21	(3)	10	(2)	11	(3)
Regression rate	5	(1)	11	(3)	18	(4)	19	(4)

Standard errors of the mean are shown in parentheses.

While a sequence of random letters to the right of the target increased regressions back to the target, it inhibited regressions back to the pretarget location. This was expressed in a lower regression rate to pretarget words in the ROF condition (8%) than the control condition (14%), the LOF condition (12%) and the dual change condition (16%). The corresponding interaction was reliable, $F1(1, 26)=7.48$, $p<.01$, $F2(1, 46)=4.37$, $p<.05$, although paired comparisons, with the control condition as baseline, revealed significantly lower regression rates in the ROF conditions across subject only, $t1(26)=2.53$, $p< .025$, $t2(46)=1.56$, $p<.13$. None of the other paired comparisons approached significance, all *p*#DXGT#.3.

Posttarget (ROF) reading. First fixation durations, gaze durations, skipping rates, and regression rates are shown in Table 4 as a function of LOF and ROF visibility during prior target viewing. The 2x2 ANOVA revealed a robust parafoveal preview benefit for first fixation durations, $F1(1, 26)=23.29$, $F2(1, 46)=36.21$, both $p<.01$, and for gaze durations, $F1(3, 79)=33.89$, $F2(1, 46)=52.50$, both $p<.01$. Critically, the visibility of an orthographically illegal letter string to the left of the target did not influence the usefulness of a ROF preview. The main effect of the LOF change and its interaction with the ROF change were both negligible, all *F*s<1.

As can be seen in Table 4, the parafoveal visibility of the posttarget word during target reading influenced its skipping which was more common when the ROF word was visible than when it was replaced with an orthographically illegal letter string, $F1(1, 26)=82.91$, $F2(1, 46)=33.65$, both $p<.01$. Again, the main

effect of the LOF change and its interaction with ROF change effects were negligible, all *F*s<1. Readers were also more likely to regress to the posttarget word when its preview consisted of a string of dissimilar letters, $F1(1, 26)= 11.20$, $F2(1, 46)=22.43$, both $p<.01$. Effects of a LOF change during target viewing on regressions to the posttarget word were small and not statistically reliable, $F1(1, 26)=2.30$, *p*#DXGT#.14, $F2<1$, as was the interaction of LOF and ROF effects, $F1(1, 26)=1.75$, $p<.2$, $F2(1, 46)=1.98$, $p<.2$. Together, the four sets of posttarget data converge, showing that orthographic information to the left of a fixated target word had virtually no influence on the processing of orthographic information available at the ROF location during target viewing.

Discussion

The effects of an orthographically illegal letter sequence at the LOF location on target viewing were relatively small. Although they were statistically reliable in the subjects analyses of the full and corrected set of data, they failed to reach significance in the items analysis. The effects of an orthographically illegal letter sequence at the ROF location on target viewing tended to be even smaller and they were reliable only in the subjects analysis of the full set of first fixation duration data. Nevertheless, large and robust parafoveal preview benefits emerged during the subsequent ROF word reading. Moreover, the acquisition of useful parafoveal ROF information during prior target viewing was not influenced by the type of information at the LOF location, indicating that LOF and ROF processing were functionally independent during target viewing.

Together the target viewing data are again difficult to reconcile with the E-Z Reader model. All analyses showed small LOF and ROF effects, and effects of LOF changes were robust in the subjects analysis of the full and corrected set of data. Yet, the data are also difficult to reconcile with the attention gradient hypothesis, as neither the LOF nor ROF effects showed a robust pattern of effects across subjects and items.

Regressions to pretarget and target words were primarily influenced by the presence of orthographically illegal letters to the right of the target. This decreased regressions to the pretarget location and increased regressions to the target location. This suggests that a random letter sequence at the ROF location attracted a saccade before target processing was completed, thus initiating the programming of a saccade back to the target. A random letter string at the ROF location could have attracted other "premature" saccades to the right that were not followed by a regression. Since these cases were included in the computation of the target's first fixation duration and gaze duration, they could have diminished magnitude of ROF and dual change effects.

GENERAL DISCUSSION

Together, Experiments 1 and 2 show effects of parafoveally visible orthographic information on target viewing, but the effect pattern was not fully consistent across experiments. In Experiment 1, information to the right of a fixated target influenced target viewing when it consisted of a sequence of orthographically illegal letters. This effect was robust even when more restrictive data selection criteria were applied that excluded instances of potential oculomotor error. Moreover, these restrictive selection conditions also revealed an effect of lexical status that was not present in the full set of data.

The effect of orthographically illegal letters to the right of a fixated target was considerably smaller in Experiment 2, and it was no longer statistically reliable. Although orthographically illegal letters to the left of the target influenced target viewing durations, these effects were reliable only across subjects but not across items.

The discrepancy in the effects of an orthographically illegal letter sequence on target viewing durations in the two experiments is reminiscent of other parafoveaon-target discrepancies that can be found in the literature (Rayner et al., 2003). However, two seemingly minor procedural differences between Experiments 1 and 2 offer a principled account for the changes in the effect pattern.

First, orthographically illegal ROF letter strings were visible throughout sentence reading and no display change occurred until after the eyes moved to the right of the target in Experiment 1. Readers could thus have noted the orthographic irregularity at the ROF location even before the target was fixated. Consistent with this view, target skipping rates were slightly—though not signiflcantly—lower in the random letter condition. In Experiment 2, the intact ROF word was shown throughout much of sentence reading, and the string of random letters at the ROF location appeared only when the eyes moved at the target location. In contrast to Experiment 1, readers were thus not exposed to illegal information at the ROF location before the target was reached.

Second, the viewing of the target was not accompanied by any display change in the three experimental conditions of Experiment 1. In contrast to this, viewing of the target in Experiment 2 always initiated a display change in the three experimental conditions that could have popped out. Theeuwes, Kramer, Hahn, and Irwin (1998) showed that a visually distinct new object in the parafovea attracted attention. Even when task irrelevant, this new object attracted a subpopulation of "reflexive" saccades with relatively short latencies.

The lack of a robust ROF and dual change effect on target viewing durations in Experiment 2 could thus be explained if the sudden appearance of a string of orthographically illegal letters was particularly likely to pop out when it was presented to the right of the target. It then triggered saccades with relatively short latencies to that location on a subset of trials thus decreasing overall target gazes in the ROF and dual change conditions. As noted before, the target regression rates of Experiment 2 provide support for this contention. On a distinct subset of ROF and dual change trials the eyes regressed back to the target word, presumably because the eyes had prematurely left it before its processing was completed.

According to these considerations, Experiment 2 is likely to underestimate parafoveal effects on target viewing. Since numeric effects of parafoveally visible orthographic information on target viewing were present, however, and were highly reliable in Experiment 1, we are inclined to conclude that the current results favour the view according to which attention can be allocated to more than one word before a saccade shift to the next word in the text is programmed.

Consistent with the attention-gradient hypothesis, difficulties processing the next word in the text delayed saccade execution. Although, consistent with the attractor hypothesis, irregularities of the parafoveal stimulus may pop out under some conditions and attract a fixation.

In defence of the E-Z Reader, it can be argued that the current results do not apply to normal reading because the stimulus materials were highly unusual, the critical word sequence always consisting of a sequence of closely associated words. This objection fails on logical grounds, however. The semantic relatedness between successive words should be completely hidden to an oculomotor control system in which the spotlight of attention is confined to one word (or subword unit) at a time, and in which saccades are programmed to an area from which no orthographic information can be obtained while the target is processed.

Another defence of the EZ Reader model cannot be ruled out, however. Specifically, it could be argued that strings of parafoveally visible random letters in our experiments were not just orthographically illegal but deviated from a parafoveal word in other perceptually distinct ways (see Reichle, Rayner, & Pollatsek, in press). For instance, a random letter sequence could have letters at unusual positions, e.g., an "x" at the string's first letter position, and anomalous bigrams that form unusual visual configurations. Word recognition processes during reading could thus be strictly serial and progress from one word to the next but unusual visual cues in the parafovea may interfere with the success of familiarity checking and lexical completion.

A key finding of Experiment 2, the independence of LOF and ROF processing, also appears to favour the E-Z Reader model over the attention-gradient hypothesis. If, as argued, a string of random letters to the left of the target attracted attention, then less attention should have been allocated to the ROF location, which should have—but did not—diminish subsequent parafoveal preview benefits. The damage to the attention-gradient hypothesis is not irreparable, however. Temporal overlap in the processing of LOF (pretarget), target, and ROF (posttarget) words need not entail that the processing of these words is synchronised and that all three words compete for the same type of resource. At the onset of a target's fixation, processing of the LOF word should be much closer to completion than processing of the target, as little or no need may exist any longer for the acquisition of visual information from the ROF location— though a change in visual information will still be noted. Conversely, ROF processing may primarily involve the use of visual information for the activation of prelexical and lexical forms. Hence, words to the right and left of the target can influence target viewing under some conditions, because there is some overlap in the processing of immediately adjoining words, but processing of words at the LOF and ROF locations is sufficiently distinct to prevent any functional overlap. Recently, Engbert, Longtin, and Kliegl (2002) proposed a computational model of eye movement control during reading, SWIFT, that is predicated on this type of processing assumption.

We thus take the position that the current results favour a view according to which the recognition processes of spatially adjacent words can temporally overlap. The resulting cross-talk is relatively confined, however, encompassing no more than two words. Furthermore, expression of this cross-talk is not clear cut. It can be influenced by viewing conditions, by data selection criteria, item properties (Inhoff et al., 2000a, 2000b), and possibly task demands.

PrEview proof published online September 2003

REFERENCES

Binder, K., Rayner, K., & Pollatsek, A. (1999). Extractions of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performancve, 25*(4), 1162–1172.

Crane, H., & Steele, C.M. (1985). Generation V Purkinje-image eye-tracker. *Applied Optics, 24,* 527–537.

Engbert, R., Lontin, A., & Kliegl, R. (2002). A dynamic model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636.

Henderson, J.M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

Henderson, J., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*(2), 201–221.

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

Inhoff, A.W., Pollatsek, A., Posner, M.I., & Rayner, K. (1989). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology, 41A,* 63–89.

Inhoff, A.W., Radach, R., Starr, M.S., & Greenberg, S. (2000a). Allocation of visuo-spatial attention and saccade programming during reading. In A.Kennedy, R.Radach, D.Heller, & J. Pynte (Eds.), *Reading as a perceptual process.* Oxford, UK: Elsevier.

Inhoff, A.W., Starr, M.S., & Shindler, K. (2000b). Is the processing of words during a fixation of text strictly serial? *Perception and Psychophysics, 62,* 1474–1484.

Inhoff, A.W., & Weger, U. (2003). Advancing the methodological middle ground. In J.Hyönä, R. Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 335–343). Amsterdam: Elsevier Science.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–180). Oxford, UK: Elsevier Science.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kennison, S.M., & Clifton, C. (1995). Determinants of parafoveal preview in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 68–81.

Kučera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Murray, W.S. (1998). Parafoveal pragmatics. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 181–200). Amsterdam: North-Holland.

Pollatsek, A., Reichle, E.D., & Rayner, K. (2003). Extensions of the E-Z Reader model. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 361–390). Amsterdam: Elsevier Science.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 3,* 372–422.

Rayner, K., Fischer, M., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*(8), 1129–1144.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice-Hall. Rayner, K., Reichle, E.D., & Pollatsek, A. (2000). Eye movement control in reading: Updating the E-Z reader model to account for initial fixation locations and refixations. In A.Kennedy, R. Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 701–720). Oxford, UK: Elsevier Science.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier Science.

Reichle, E., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E., Rayner, K., & Pollatsek, A. (in press). The E-Z reader model of eye movement control in reading: Comparison to other models. *Behavioral and Brain Sciences.*

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology, 52A,* 1021–1046.

Theeuwes, J., Kramer, A.F., Hahn, S., & Irwin, D.E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science, 9,* 379–385.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighboring words. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 247–268). Oxford, UK: Elsevier Science.

APPENDIX

Sentences used in the experiment. Left-of-fixation (LOF), target, and right-of-fixation (ROF) words are italicised.

Every time I get groceries I get a *ripped paper sack* to carry home.
The building had a single *broken window pane* on the north side.

I wonder who will have *the fastest race car* in tonight's competition?
I can't wait to relax in some *steaming bath water* after work.
Bill constructed the *wooden door frame* quickly and efficiently.
My brother's *twisted ankle joint* pained him during the marathon.
We decided to buy the cheap *plastic curtain rods* for the dining room.
I'd hate to witness such a *terrible traffic accident* in person.
The weatherman predicts *icy road conditions* for tomorrow evening.
Where else can you buy such *delicious apple juice* for that price?
The infant was fascinated by the *delicate flower petals* in the garden.
A vacation in an *elegant casino resort* was just what we'd needed.
Why don't you get rid of that *outdated computer screen* tomorrow?
The rebels staged a *surprise ambush attack* on the Capitol.
The glade was full of *white birch trees* and hearty pines.
I can't stand to eat *the flaky bread crust* on my sandwiches.
Sadly, he began to display *chronic cancer symptoms* last week.
The government predicted a *rampant virus outbreak* in Rwanda.
You could almost read by the *bright moon light* last night.
I was relieved to finally see a *budding spring flower* in the garden.
It seemed like the *fragrant rose scent* covered the whole room.
I found a *rusty metal hammer* near the old construction site.
The rescue team pulled the *dusty coal miner* out of the collapsed shaft.
The job of *circus lion tamer* is filled with danger.
I couldn't hear over all of the *noisy cattle market* sounds.
Be careful when approaching the *busy traffic light* down there.
Take out that *dirty garbage pail* before it begins to smell.
The team couldn't lift the *heavy stone weight* any further.
I can't find *my favorite spice variety* any more.
The farmer was overjoyed at the *abundant wheat crop* this season.
We forgot to return the *rented video tape* to the store.
Francois recommended the *fashionable dress shirt* for spring.
Run and get my *medical doctor kit* before it's too late!
The spray is ineffectual against the *pesky insect swarm* out here.
We began a regimen of *physical fitness exercise* just last month.
In Ireland, we viewed some *beautiful castle artwork* in the country.
Becky cherished the *joyful wedding picture* from last year.
The snow at the *winter ski resort* begins to melt in March.
I cut myself on that *sharp knife edge* yesterday.
We encountered *swampy marsh lands* in the Everglades.
The squirrel fell off the *leafy tree branch* and onto the porch.
My mother made the *gold tassel ornament* when she was younger.
The child relished the *sweet candy lollipop* until it was gone.
Pricking your finger on a *sharp needle point* can be painful.
I wish I had a *soft feather pillow* instead of this hard one.
Your new *office desk chair* looks exceedingly comfortable.

There's nothing like a *burning wood fire* in the winter.
Drive slowly on the *rough road surface* down there.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 226–261

Saccade generation during reading: Are wordsnecessary?

S.-N.Yang and G.W.McConkie
Beckman Institute, University of Illinois at Urbana-Champaign, USA

Most current theories of eye movement control during reading are word based in multiple ways: They assume that saccade onset times result from word-based processes, and that words are involved in selecting a saccade target. In the current study the role of words was examined by occasionally replacing the text with one of five alternate stimulus patterns for a single fixation during reading, and observing the effects on the time, direction, and length of the saccade that ends that fixation. The onset times of many saccades are unaffected by replacing spaces with random letters, thus removing visible word-units; also, the effects of this removal on saccade length is not different than that of having space-delimited nonwords. It does not appear that words play a critical role in generating saccades. The results are compatible with the Competition/Interaction theory of eye movement control during reading (Yang & McConkie, 2001).

Reading is a highly skilled task that requires the close coordination among various perceptual, cognitive, and motor processes (Rayner, 1998; Rayner & Pollatsek, 1989). This coordination is sufficiently great that it has been possible to use eye movement recording to study the perceptual and cognitive processes that take place during reading. However, the exact nature of the coordination is currently in dispute (Kliegl & Engbert, 2003; O'Regan, 1989; Reilly & Radach, 2003; Rayner, 1998; Yang & McConkie, 2001). Three basic assumptions are made by most recent theories of eye movement control during reading: First, the text is visually attended in a sequential, word-by-word manner (*sequential attention assumption:* Henderson & Ferreira, 1993; Morrison, 1984; Reichle, Pollatsek, Fisher, & Rayner, 1998); second, when the eyes are moved they are sent to an attended unit, a word (*word-based targeting assumption:* McConkie, Ker r, Reddix, & Zola, 1988; O'Regan & Lévy-Schoen, 1987; Reichle, Rayner, & Pollatsek, 1999)[1]; third, each saccade is triggered by some cognitive event and thus its onset time reflects the time of that event (*cognitive saccade triggering assumption:* Just & Carpenter, 1980; Morrison, 1984; Reichle et al., 1998). The primary dispute has been in the nature of the saccade triggering event, whether it be the lack of needed visual information to resolve currently attended letters or words (McConkie, 1979), completion of the identification of a word (Henderson & Ferreira, 1990, 1993; Morrison, 1984), completion of an estimate of the familiarity of a word (Reichle et al., 1998, 1999), or completion of the processing that is enabled by a word (Just & Carpenter, 1980). A computational model based on the above three assumptions has been very successful in accounting for many phenomena observed in eye movement patterns of normal reading (Reichle et al., 1998, 1999).

The three assumptions, and the theories embodying them, all give a prominent place to the role of words in eye movement control, a position that we will refer to as *word-based control*. Words are prominent in most alphabetic writing systems, being separated by spaces and thus serving as visually distinct objects. It is generally assumed that these word objects are the units of attention, the units to which attention is shifted, the targets to which saccades are sent, and the units of language processing.

There is considerable evidence supporting a word-based control position. Words can be identified fastest when the eyes are near (somewhat left of) their centre, referred to as the optimal viewing position (OVP; O'Regan, 1994; O'Regan & Jacobs, 1992) and the eyes tend to be sent to positions near the centres of words more frequently than to their beginnings or ends, showing a preferred viewing position usually just to the left of the word centre (Rayner, 1979), though the actual landing position also depends on the location from which the saccade was launched (McConkie et al, 1988; Radach & McConkie, 1998). In addition, the frequency of refixating a word depends strongly on the location of the first fixation on the word, being lowest when the initial fixation is near its centre (McConkie et al., 1988; O'Regan, 1994; Radach & McConkie, 1998; Rayner, 1979). Replacing spaces with a distinctive character reduces

---

Correspondence should be addressed to S.N.Yang, Beckman Inst., University of Illinois, 405n. Matthews Ave., Urbana, IL 61801, USA. Email: syang3@brain.riken.go.jpThis research was supported by a grant from the Research Board of the University of Illinois at Urbana-Champaign, by a Beckman Graduate Fellowship to the first author and by the National Science Foundation under Grant No.9624396. Gary Wolverton did the computer programming that made the research possible.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000231

reading speed by 10– 20%, but this reduction is 44–75% when spaces are replaced by random letters that make the word boundaries more difficult to locate (Epelboim, Booth, Ashkenazy, Taleghani, & Steinman, 1997; Epelboim, Booth, & Steinman, 1994; Rayner, Fischer, & Pollatsek, 1998; Rayner & Pollatsek, 1996).

Chinese readers, for whom the lack of spatial markers for words is the normal condition, still read at approximately the same rate as people reading alphabetic languages when measured by the content read (Sun, Morita, & Stark, 1985).

Chinese readers show no preferred viewing position, though refixations are less frequent following an initial fixation at the word's centre, presumably because during that fixation the word boundaries are identified on some other basis (Yang & McConkie, 1999). Adding spaces between words to make them perceptually distinct does not facilitate skilled Chinese reading (Liu, Yeh, Wang, & Chang, 1974; but see Hsu & Huang, 2000), though it may help children who are learning to read (Yang, 1998). Apparently, Chinese readers develop oculomotor strategies that do not involve the use of words: Inserting spaces may actually interfere with these strategies, indicating that perceptually isolated word units may not be necessary for normal saccadic activity to occur in reading.

The above observations about reading in alphabetic languages are consistent with current theories that propose that early visual processes parse the text into word units as a basic perceptual representation, and that readers select a word to send the eyes to on each eye fixation, perhaps determined by the word being attended at the time at which a saccade is programmed, and with the eyes landing at a position in the word that is determined by the distance from which the eyes were launched (McConkie et al., 1988; Reichle et al., 1999) and by certain characteristics of the word stimulus itself (Beauvillain, Dore, & Baudouin, 1996). Whether Chinese readers have an alternate procedure for fast parsing of character strings into words, or for controlling saccades on some basis not directly involving word units, is a matter requiring further research (Tsai & McConkie, 2003).

A recent study (Yang & McConkie, 2001) appears to provide further evidence that word-like units are critical for normal eye movement control during reading. Using the single fixation replacement (SFR) method, in which the text is occasionally replaced by an alternate stimulus pattern for a single critical fixation as people are reading, the effects of certain stimulus characteristics on eye movement decisions were examined. Since the display changes occurred during saccades, they were not directly perceived because of saccadic suppression (Inhoff, Starr, Liu, & Wang, 1998; Matin, 1974). Yang and McConkie (2001) observed the effects of various alternate stimulus patterns on the time at which

---

[1] Some theorists suggest that the eyes are simply sent to the next word, without invoking the concept of attention (Just&Carpenter, 1980; O'Regan, 1994).

the saccade was made that ended the critical fixation (i.e., on the fixation duration), and on where the eyes were sent on that saccade.

In their study, the effects of various stimulus patterns on the fixation durations (referred to here as saccade onset times) were analysed by comparing hazard curves of these data under different experimental conditions. Hazard curves are derived from the same data as fixation duration frequency distributions and are simply an alternate way to represent the same data. Whereas a proportional frequency distribution (i.e., saccade onset time density function) is obtained by dividing the number of cases in each time bin by the total number of cases, the hazard curve is produced by dividing the number in each bin, in increasing temporal order, by the total number of cases minus the number of cases in which a saccade has occurred prior to the beginning of this bin. Thus, hazard curve values indicate, for each successive time bin, the proportion of *surviving fixations* for which a saccade occurs during that time period.

Consider a set of data in which 20 saccades occur during each of two time bins, but with 1000 surviving fixations at the beginning of the first bin (hazard level: 20/1000=0.02 so 2% of all remaining fixations end during this period) but with only 100 surviving fixations at the beginning of the later bin (hazard level: 20/100=0.20 so 20% of all remaining fixations end during this period). In this case, the hazard curve indicates a much higher level of saccadic activity during the later than during the earlier time period, even though the actual number of saccades produced is the same in each. Thus, the hazard curve can be taken as indicating the saccade generating activity level at different times during the fixation, while the frequency distribution shows the product of this saccade generating activity. The hazard curve is particularly useful in examining changes in the level of saccadic activity over time during fixations, thus indicating the dynamics of these saccade decisions. Hazard curves for saccade onset times in reading tend to show an initial period during which the curve is very low and rising slowly, a second period during which a rapid rise in saccadic activity occurs, and a third period in which the hazard level reaches asymptote, often then dropping somewhat (McConkie & Dyre, 2000).

Yang and McConkie (2001) found that replacing the text with unbroken lines of Xs (called the X-minus condition, labelled as X−, with the minus indicating the lack of between-word spaces) produced no effect on the saccade activity level, as compared to the control condition, for the first 125 ms. It then caused a near-complete cancellation of saccades for the next 75 ms, the period when the hazard curve normally rises quickly. The same results were found with a blank page and with each letter replaced by an underline. The saccades that did occur during this early period, in spite of the abnormal stimulus characteristics, are referred to as *early* saccades and are assumed to be those observed by Morrison (1984) and others, which have been attributed to saccade decisions made during the prior fixation. Yang and McConkie suggested that the massive cancellation of saccades, which will be referred to as the *X−saccade cancellation effect,* or just *X−effect* (X-minus effect), might be due to the lack of perceptible word units, as expected by word-based control theories, where the lack of spaces would eliminate the possibility of performing normal saccade-related activities such as attending to a word unit, judging the familiarity of a word, identifying a word, shifting attention to a word or selecting a word as the target of the next saccade. Without word-like units, the saccade would have to be produced on some other basis, which apparently required a considerable amount of time, since the mean fixation duration in that condition was 414ms as compared to 212ms in the control condition, and with many fixations over 800 ms.

The first purpose of the current study was to replicate the X–effect and to test the above explanation by using the same method with other space-less conditions: normal text with spaces replaced by random letters, and random letter strings containing no spaces. If the word-based control explanation of the X–effect is accurate, the X–effect should also be found in these other conditions, as well.

The second purpose of the study was to test between predictions derived from word-based control theories and those derived from an alternative theory proposed by Yang and McConkie (2001; McConkie & Yang, 2003), called the Competition/Interaction (C/I) theory, for responses to SFR manipulations of the text. While both of these types of theories assume that words are basic language processing units in reading, and that their disruption creates serious processing problems that are evident in eye movement data, word-based control theories give words a critical role in the actual generation of saccades while C/I theory does not.

The cognitive saccade triggering assumption, which is basic to most current word-based control theories, predicts that if the saccade triggering event is delayed or eliminated by a stimulus pattern that does not support normal, saccade-related processing activities, the saccade will be delayed, thus creating a longer fixation duration. This would be seen as a failure of the saccade onset time hazard curve to rise at its normal time, or to rise as steeply as normal. The word-based targeting assumption typically assumes that a saccade is sent to the currently attended word; thus, if there are no word units in the stimulus array, the normal saccade generating process will fail and a saccade must be produced on some other basis, presumably requiring additional time for recovery from the initial failure and again showing a delay in the rise of the hazard curve level. Removing spaces should also cause a substantial change in where the eyes are sent, observed as changes in directions and lengths of saccades.

The C/I theory, in contrast, postulates an active saccadic system that initiates saccades after random waiting times. The saccadic activity level varies over the course of the fixation as a result of activation and inhibition, in a manner that is revealed by the hazard curve, and that can be represented as being the result of changes in a parameter that controls the momentary mean waiting time. Cognitive activities can affect this saccadic activity level in three ways (McConkie & Yang, 2003). First, changing the reading strategy can influence the values of parameters that determine general properties of saccadic activity, including the asymptotic level of the waiting time parameter. Decreasing this parameter leads to generally longer fixations. Second, if problems are encountered in some processing centre, inhibition of the saccadic system occurs which can cancel or delay some or all of the saccades that have not yet been executed. Any saccade that occurs prior to the arrival of inhibition is not influenced by it. The onset time and degree of this inhibition can be directly observed as reductions in the hazard curves for experimental conditions, as compared to the control condition. Third, if a saccade is delayed for a sufficient time (at least about 325ms) then there is time for direct cognitive control to occur, which can directly determine the saccade properties and onset time. These changes can also be observed in the hazard curve, as well as in time-related plots of saccade length and direction (McConkie & Yang, 2003; Yang, 2002). The frequency distribution of fixation durations is assumed to be a mixture of distributions of three types of saccades (Feng, 2001), early saccades that cannot be affected by current stimulus conditions, normal saccades that occur at their normal times even though they were vulnerable to being inhibited (e.g., they either occurred before inhibition or in spite of it), and late saccades that occur following the inhibition of an earlier saccade, and which may, in some cases, be cognitively determined.

These two types of theories make different predictions about the effects on saccadic activity that will result from stimulus manipulations that interfere with basic reading processes. Word-based control theories predict that removing spaces or other information critical to saccade generation will delay all but early saccades and substantially change the properties of the saccades when they do occur. The hazard curve will represent this as a failure of the curve to rise in its normal manner. The C/I theory predicts that, in spite of processing problems that occur, saccades will still be made at their normal times with normal properties, up until the time that inhibition is produced as a response to those problems. From this perspective, the hazard curve should rise normally, but then drop, relative to the control condition, at the time inhibition begins. Thus, word-based control theories attribute longer fixations on words that produce processing difficulties to a delay in the time of the critical saccade-triggering processing event, thus delaying the saccade onset time. In contrast, the C/I theory gives a probabilistic account, accounting for a longer mean fixation duration as resulting from a greater proportion of the fixations being cancelled or delayed by inhibition, plus the possibility of a few unusually long fixations being directly delayed on a cognitive basis.

Yang and McConkie (2001) reported hazard curve changes that are more like those expected by the C/I theory than by the cognitive saccade triggering assumptions. They also presented evidence that many (and perhaps most) non-early saccades are not delayed by having text replaced by random letters or even by Xs, so long as words are delimited by spaces or perceptually distinct characters. The present study was designed to replicate these results, and to examine the effects of the single fixation elimination of spaces on saccade onset times during reading. If the removal of spaces by replacing them with random letters in normal text and in random letter stimuli produces results like the X –effect observed in the earlier study, this would provide evidence for the importance of word units in saccade generation. On the other hand, if this manipulation affects the onset times of only some saccades, those that occur after a particular time interval, this would support the C/I account in terms of inhibition-based effects. The basic question here is whether, in reading, saccades must be induced by the occurrence of some perceptual or cognitive event, or whether they are generated on some basis not directly related to momentary cognitive events, being inhibited if processing problems are encountered. This is, of course, a very basic issue for theories of saccade control during reading.

In presenting saccade onset time hazard curve data, Yang and McConkie (2001) aggregated all saccades, whether forward (rightward) or regressive (leftward). However, from a neurophysiological perspective it is appropriate to consider the saccade generating system as being divided into two interconnected parts related to the left and right visual fields, thus plotting hazard curves for forward and regressive saccades separately. For example, behavioural and neurophysiological evidence shows that in the antisaccade paradigm, in which people must move their eyes in a direction opposite the direction indicated by some signal, the movement-related activity of a saccade initially planned toward the signal must be suppressed in order to plan a new saccade in the opposite direction. The rise in activity opposite the signalled direction is temporally correlated with the drop in activity toward the signalled direction (Everling, Dorris, Klein, & Munoz, 1999). In addition, when a saccade is made to one location, competition from a visual stimulus at the corresponding location in the opposite hemifield increases saccade latency (Walker, Deubel, Schneider, & Findlay, 1997). Finally, pharmaco-physiological evidence shows that there are horizontally oriented inhibitory interneurons within and between the superior colliculus of the two hemispheres of the brain (Meredith & Ramoa, 1998), revealing a possible neuronal mechanism for the competition of saccade programming to opposite directions. By plotting separate hazard curves for forward and regressive saccades, it is possible to observe the activity of these two saccade-generating mechanisms. Since hazard curves that are based on the same total set of saccades are additive (Luce, 1986), forward and regressive saccade hazard curves can be summed to obtain data corresponding to that reported in earlier studies. Partitioning the data in this manner makes it possible to observe changes in saccade likelihood towards the two directions at different times during critical fixations. Of particular interest, from a neurophysiological perspective, is to see whether changes in saccadic activity are bilateral (occurring in both hemispheres similarly), unilateral (occurring in only one hemisphere), or competitive (increased activity in one hemisphere coupled with decreased activity in the other).

TABLE 1 Text patterns used to substitute the whole page of original text during the critical fixation

Display conditions	*Single-line examples*
Segmented	
Orig+	He is persuaded that a white people, of
Rand+	Rl ui mgbvikhsh lxbx l tmgyj ejzhpw, gl
X+	Xx xx xxxxxxxxx xxxx x xxxxx xxxxxx, xx
Unsegmented	
Orig−	Hexiscpersuadedlthatmazwhiteopeople,qof
Rand−	Rlxuicmgbvikhshllxbxmlztmgyjoejzhpw,qgl
X−	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

METHODS

Participants

Thirty-six college students and adults from the University of Illinois community (25 females, 2 left-handed; average age=22.9 years) participated in this experiment for pay or course research credit. All were native English speakers with normal or corrected-to-normal vision. The experiment lasted about 45–60 minutes.

Materials

Participants read 47 nine-line pages (triple-spaced, 80 character lines, Courier New font having 8 x 16 pixel space for each character on a 640 x 480 display) of an early twentieth-century novel, *The Mystery of Sasassa Valley,* written by Sir Arthur Conan Doyle. Nine alternative versions of each page were created as display conditions by substituting letters, symbols, or spaces for the original characters in the text. An example of each of these conditions is shown in Table 1. The first part of condition names indicates the nature of the letters, including Original text (Orig), Random letter strings (Rand, with each letter replaced by a randomly selected letter), and X strings (X, with each letter replaced by an X). A plus or minus at the end of the name indicates that the spaces between words remained in the text (*Orig*$^+$, *Rand*$^+$, *X*$^+$) or were replaced by other characters, Xs in the case of X strings (*X*−), and randomly selected letters in the Orig (*Orig*−) and Rand (*Rand*−) conditions.[2]

Apparatus

Eye movements were recorded using a head-mounted SR Inc. EyeLink system having a sampling rate of 250 Hz and high spatial resolution (Reingold &

Stampe, 1999). Software was written to examine the incoming data and identify the onsets of saccades in real time. Text was displayed on a Viewsonic 21-inch monitor, controlled by a Diamond S3 Trio64V2/DX display controller card having 2 MB of video memory, and being refreshed at 60 Hz. Text was displayed in black on a white background. The screen outside of the display area was also in black. At the beginning of each trial, the original text page and three alternate versions of the page were loaded into video memory. Changing from the display of one page to another was accomplished by simply panning to a different region of video memory, which occurs very quickly. This change could occur at any point during a refresh cycle, not just at its completion. Thus, a new image was completely written onto the monitor within 17ms of the time it was requested at the beginning of a saccade.

The participants were seated in a quiet room with dimmed light, about 80 cm from the CRT, with text displayed within a 40x30 cm central area of the screen. At this distance, there were 2.8 letters per degree of visual angle. No head restraint was used in this experiment since absolute spatial position of the eyes was not essential.

Procedures

Participants were told about the types of display changes that would occur, and were told that their task was simply to read the text for comprehension and in preparation for occasional test questions, ignoring the text changes insofar as possible. Following a 9-point calibration, each participant received a practice trial with three sample pages and two comprehension questions.

The participants then read the text with seven breaks, answering two multiple-choice questions about the story content during each break. They controlled the progress of the experiment by pressing buttons to advance to the next page and to answer questions. During each break, after answering the questions, a drift correction was conducted and the participant was recalibrated if the drift exceeded a preset criterion.

The normal text initially appeared for each of the 47 pages of text. However, at the onset of the 9th, 10th, or 11th forward saccade on each page, one of the alternative versions appeared, and then at the onset of the following saccade the normal text reappeared. Thus, the alternative version was present on the screen for the period of a single eye fixation, the critical fixation. In the control condition (Orig+) the normal text was replaced by itself, producing no visible stimulus change but marking the data with fixations that were selected in the same manner as in the experimental conditions. This cycle was repeated throughout

---

[2] Four other conditions were also included in the current study, all with interword spacing, that are not discussed further here: alternating case, reversed case, and two conditions in which the locations of the spaces were moved.

the reading of the page, with an alternative version of the page presented about every 12 to 16 fixations, depending on the frequency of regressive eye movements. The control condition and three alternative versions were scheduled for each page and they appeared in a predetermined order that was counterbalanced across pages. The combinations of conditions that appeared together on a page were also counterbalanced as far as possible.

Data analysis

Data analyses examined the effects of the various stimulus conditions on the saccade onset time (fixation duration) and the direction and length of the saccade that ended the critical fixation. The values for these variables were taken from the data analysis program included with the EyeLink system. Data were excluded from analysis if any of the following conditions occurred:

(1) The fixation was preceded or followed by a movement with a vertical component of more than two lines (100 pixels) or horizontal component larger than half of the line (320 pixels)—12% of critical fixations excluded.
(2) The fixation was preceded or terminated by a blink—10% excluded.
(3) A display change occurred during the period of fixation (this occurred because the online algorithm used to detect saccade onsets was in some cases more sensitive than that in the data reduction software provided by the manufacturer, which was used in the final data reduction; this conservative procedure resulted in data being included in final analyses only if both algorithms indicated that the display change occurred during a saccade)—6% excluded.
(4) The fixation occurred before the onset of the first display change on each page, or was the final fixation on the page, terminated by a button press—<1% excluded.

In all, 29% of the critical fixations were excluded.

Data were initially analysed using repeated measures ANOVAs. The tests of statistical significance for main effects and interactions were done by first using Wilks' A. Homogeneity of the samples for each factor was tested using the Mauchly's test of sphericity with the criterion of α .05. When sphericity was violated, the Greenhouse-Geisser method was used to provide an unbiased estimate of the statistical significance, which is accomplished by adjusting the degrees of freedom applied in the F test. For factors with significant results, Bonferroni pairwise comparisons of means for each level were calculated, using a family-wise a level of .05 to take into account the number of comparisons involved. Since the frequency distributions of fixation durations were positively skewed, especially in experimental conditions, base-10 logarithmic transformations were used to transform the data before obtaining means for use with the ANOVA analyses and subsequent comparisons in order to reduce the influence of outliers. While the resulting geometric means are not unbiased estimators, with bias varying with sample size, the large sample sizes of the current data set should keep any such bias small.

For further exploration, saccade onset time hazard curves were derived from fixation duration frequency distribution data. Using specified time intervals or "bins" (here, 25ms), a hazard value is calculated by dividing the proportion of saccades occurring in a bin by the proportion of still surviving fixations at the beginning of that time interval. This was done separately for forward and regressive saccades, as follows:

$$\begin{cases} HF(i) = f_f(i)/(1 - F(i)) \\ HR(i) = f_r(i)/(1 - F(i)) \end{cases}$$

Here HF(i) and HR(i) are the hazard levels for forward and regressive saccades during the ith time period; $f_f(i)$ refers to the proportion of all saccades that are forward and occur during the ith time period; $f_r(i)$ refers to the similar proportion of all saccades that are regressive and occur during this same period; and F(i) refers to the proportion of all saccades that occur prior to the beginning of the ith time period.

Hypotheses derived from different types of theories, as described above, often concern the time at which saccades begin to be delayed or cancelled. This time is observable in the hazard curves as a point at which the experimental curve falls below the control condition curve. This will be referred to as the *effect onset time*. This time is identified here statistically by testing the difference between hazard values as proportions in corresponding time bins in an experimental and the control conditions, with the a level of the comparisons being set at .0005 because of the number of tests performed, and with the requirement that there be a significant difference in the same direction for at least three successive bins. The first of these three bins is taken as the effect onset time.

RESULTS

Overall patterns of eye movements

The overall characteristics of eye movements in the current study are close to normal, in spite of the fact that critical fixations made up about 7% of all fixations (mean duration of all fixations made: 212ms; mean forward saccade length: 6.8 letters; regressive saccade length: 4.9 letters; regression percentage: 28.7%). The regression rate is slightly higher than that obtained by Yang and McConkie (2001) using the same text, procedure, and general type of manipulation, and the performance on comprehension questions slightly lower (66% vs. 72%), suggesting somewhat greater reading difficulty in the current study.

TABLE 2
Means and pairwise comparisons for fixation durations, saccade lengths, and regression frequency

	Display conditions[3]					
Data source	*Orig+*	*Orig−*	*Rand+*	*Rand−*	*X+*	*X−*
Fixation duration						
N	848	912	951	840	767	694
Transformed means[1]	2.333	2.439	2.407	2.477	2.499	2.710
SE	.010	.015	.013	.016	.018	.023
Antitransformed (ms)[2]	215[a]	275[c,d]	255[b,c]	300[d]	316[d]	513[c]
Regression proportion						
N	208	346	350	414	328	432
Transformed percentage	1.165	1.428	1.363	1.686	1.509	1.914
SE	.066	.048	.058	.071	.064	.080
Antitransformed (%)	30.2[a,b]	42.9[c,d]	39.7[b,c]	55.7[d,f]	46.9[c,d]	66.8[f]
Forward saccade length						
N	640	566	601	426	439	262
Transformed means	1.747	1.675	1.653	1.642	1.649	1.492
SE	.018	.019	.016	.019	.019	.041
Anti-transformed (pixels, letters)	55.8 (7.0)[a]	47.3 (5.9)[a,b]	45.0 (5.6)[b]	43.9 (5.5)[b]	44.6 (5.8)[b]	31.0 (3.9)[c]
Regressive saccade length						
N	208	346	350	414	328	432
Transformed means	1.608	1.449	1.394	1.428	1.427	1.440
SE	.043	.033	.034	.031	.031	.046
Antitransformed (pixels, letters)	40.6 (5.1)[b]	28.1 (3.5)[a]	24.8 (3.1)[a]	26.8 (3.3)[a]	26.7 (3.3)[a,b]	27.5 (3.4)[a,b]

[1] Fixation duration and saccade length scores were logarithmically transformed; inverse sine transformation was used to transform regression percentages.
[2] The means here were untransformed from the transformed means above.
[3] The superscript letters here indicate the results of mean comparisons. When two conditions share the same letter, there is no significance difference between them. The family-wise significance level for all comparisons was set at a • .05 for pairwise Bonferroni comparisons.

Table 2 presents the Ns, means, and standard errors for the various dependent variables on critical fixations in the conditions of the study. Means having the same superscript letter are not significantly different. [3]

Fixation duration. A three-way repeated measures ANOVA was carried out on fixation duration data with segmentation (segmented vs. unsegmented marked by "+" or " –") by text anomaly (Orig, Rand, or X) by page group (three sets of pages) as factors. Significant main effects were obtained for segmentation, $F(1, 36)=169.047$, $p<.0005$, text anomaly, $F(1.627,58.571)= 134.566$, $p<.0005$, and page group, $F(2,35)=11.628$, $p<.0005$, as well as for the interaction of segmentation and text anomaly, $F(1.466, 52.768)=16.710$, $p< .0005$. Fixations were longer with unsegmented text (348 ms) than segmented (259ms). Replacing text with Xs (X, 403ms) produced longer fixations than with random letters (Rand, 277 ms), and both were longer than the original text (Orig, 243ms). Fixations were longer on the initial two page groups (310 ms, 310ms) than on the last page group (281 ms). The interactions between page group and the experimental conditions were not significant.

The significant interaction observed above came primarily from the fact that the Orig+ (control) condition had the shortest fixations (see contrasts in Table 3: Orig+ vs. the group of Orig−, Rand−, Rand+ and X+; $t=-3.73$, $p<.0005$), and the X− condition had the longest (contrast: X−vs. the same group; $t= 7.74$, $p<.0005$). A contrast between the nonextreme segmented conditions (Rand+, X+) and unsegmented conditions (Orig−, Rand−) was not significant ($t=-0.17$), but Table 2 indicates that removing word segmentation in the Rand condition did produce significantly longer fixations (Rand−#DXGT# Rand+). While eliminating spaces between words disrupts reading, as expected, it does not produce the massive disruption observed in the X− condition, and, as Table 2 indicates, filling spaces in the text (Orig–) produces a mean fixation duration that is not significantly different from those in the X+ or Rand+ conditions where spaces are preserved.

Figure 1 presents the hazard curves for the forward saccade onset times for the same six conditions through 450ms; low Ns in some of the conditions make the curve quite unstable beyond that point. Data for all conditions are very similar through 125 ms: these are the early saccades that are unaffected by stimulus patterns. The Orig+ (control) condition data show the pattern typically observed during reading, with a low saccadic activity level initially, a period in which the activity level rises

[3] Means in Table 1 are based on two-way repeated measures ANOVAs. These are sometimes not identical to means reported in the text from three-way repeated measures ANOVAs, with the additional stage factor included.

TABLE 3
Contrasts of means for fixation durations, saccade lengths, and regression frequency for selected groups of conditions

	Manipulation					
Dependent variables and contrasts	*t*	*SE*	*df*	*p value*	*Mean 1*	*Mean 2*
Fixation duration (ms)						
Contrast 1	−3.73	.033	37	<.0005	2.33 (214)	2.46 (288)
Contrast 2	7.74	.033	37	<.0005	2.71 (513)	2.46 (288)
Contrast 3	−0.17	.029	37	n.s.	2.46 (288)	2.45 (282)
Regression proportion						
Contrast 1	−4.90	.068	36	<.0005	1.17 (30.5)	1.50 (46.5)
Contrast 2	6.17	.068	36	<.0005	1.91 (66.6)	1.50 (46.5)
Contrast 3	−2.00	.060	36	<.05	1.56 (49.5)	1.44 (43.5)
Forward saccade length (letters)						
Contrast 1	1.76	.052	37	n.s.	1.75 (7.0)	1.65 (5.9)
Contrast 2	−3.11	.052	37	<.0005	1.49 (3.9)	1.65 (5.9)
Contrast 3	0.16	.047	37	n.s.	1.66 (5.7)	1.65 (5.9)
Regressive saccade length (letters)						
Contrast 1	4.59	.040	36	<.0005	1.61 (5.1)	1.42 (3.3)
Contrast 2	0.39	.040	36	n.s.	1.44 (3.4)	1.42 (3.3)
Contrast 3	0.78	.036	36	n.s.	1.44 (3.4)	1.41 (3.2)

[1] Antitransformed values of the means are included in parentheses. The units for these antitransformed means are indicated in parentheses in the first column.
[2]Means included in each contrast: Contrast 1: Mean 1 (Orig+) vs. Mean 2 (mean of Orig–, Rand– , Rand+, X+); Contrast 2: Mean 1 (X–) vs. Mean 2 (mean of Orig–, Rand– , Rand+, X+); Contrast 3: Mean 1 (mean of Orig–, Rand–) vs. Mean 2 (mean of Rand+, X+).

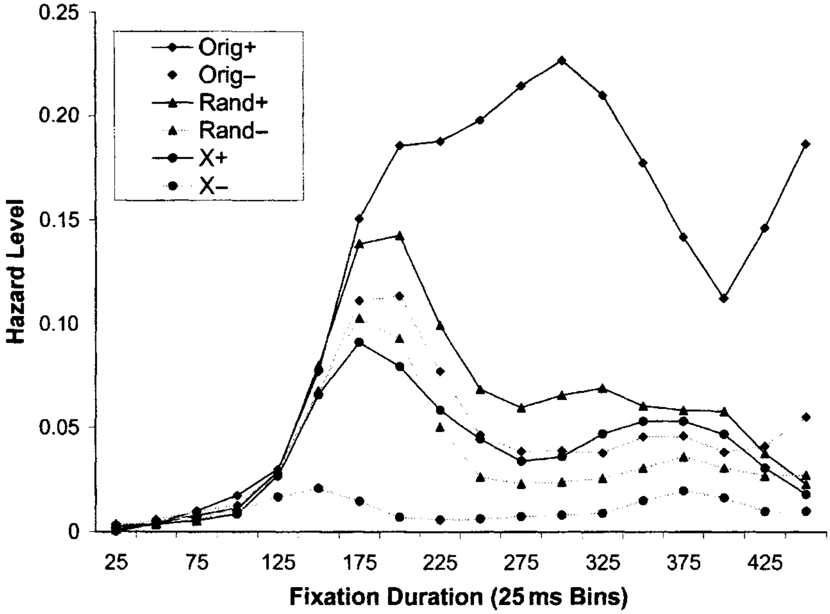

Figure 1. Hazard curves for the onset times of forward saccades that terminate the critical fixations.

quickly to an asymptote, and then a final period of high activity. All other conditions differ substantially from this condition, dropping below it at different times, showing the delay or cancellation of saccades. The X– condition, with unsegmented lines of Xs, is also different from all other conditions, showing the same pattern as observed by Yang and McConkie (2001): initial low saccadic activity similar to the control condition through 125ms, followed by a drop in activity at the time when a steep rise normally occurs, coming near zero by 200 ms, and then gradually rising to a low and rather stable level later. Thus, the current study replicates the X– effect from the earlier study. Importantly, this severe disruption is not observed in the Orig– or Rand– conditions, thus disconfirming the hypothesis that the X–effect is caused by a lack of perceptible word-like units. Neither is that effect observed in the X+ condition. Apparently the X—Effect requires the combination of a lack of segmentation and homogenous stimulus strings.

The X—condition curve shows an initial period, through 175 ms, in which some saccades occur at normal times in spite of the aberrant stimulus pattern. These appear to be the early saccades whose onset times are unaffected by stimulus conditions that are present on the new fixation (Morrison, 1984). Only some time later does the hazard curve begin to rise again, presumably indicating the time at which the delayed or cancelled saccades finally begin to occur.

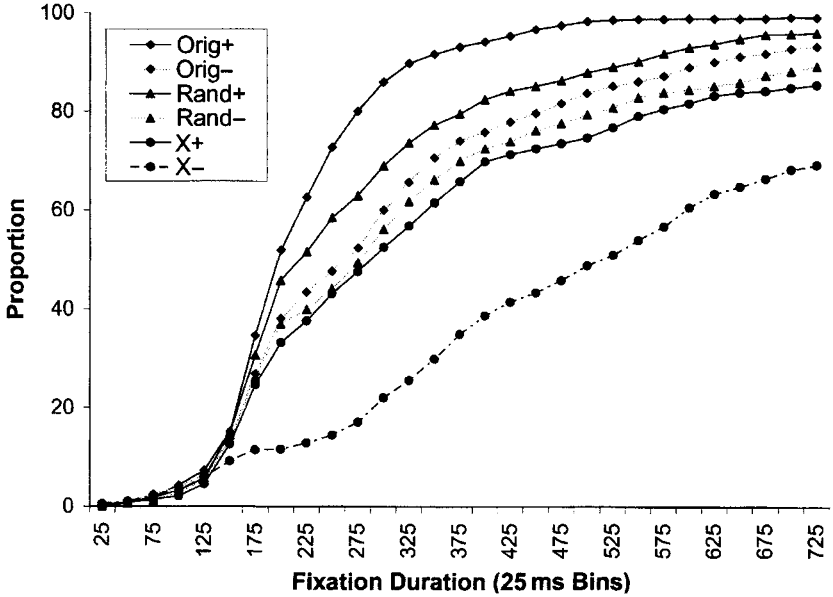

Figure 2. Cumulative frequency distributions of durations of critical fixations ending with forward saccades. Some curves do not reach 100 because not all saccades have occurred within the time represented.

Figure 2 presents cumulative frequency distributions for these same six conditions. The cumulative frequency for saccade onset times in the X−condition in the 175 ms period is about 0.12, indicating that this proportion of all saccades made were early saccades. We assume that this set of saccades exists in all the conditions of the study, including the control condition; their presence is usually hidden by the distribution of normal saccades that typically occurs.

The Rand+ condition (segmented random letters) also replicates results from the earlier study: a hazard curve similar to the control condition through about 175 ms, with a sizeable reduction in saccadic activity (or effect onset time) by 200 ms, reducing further over time until a low, relatively stable level is achieved at about 250 ms. This is the data pattern predicted by the C/I theory, rather than that expected from the cognitive saccade triggering assumption. Furthermore, if the early period of the curve for the X−condition is accepted as revealing the early saccades, then a comparison of this and the Rand+ condition indicates the existence of a large number of non-early saccades (that is, saccades whose onset times are being affected by the present stimulus pattern). Figure 2 indicates that about 46% of all forward saccades are initiated by the effect onset time for these nonwords. These are saccades that occur at their normal time in spite of the violation of orthographic regularities and the lack of identifiable words. These saccades clearly do not require the occurrence of any cognitive triggering event at the orthographic or lexical level; if they are being triggered by some perceptual or cognitive event, it must be at a level below that of involving orthographic constraints.

A comparison of the curves for the Orig−and Rand−conditions with that of the Orig+ (control) condition indicates an effect onset time of 175 ms and the pattern is predicted by the C/I theory. A comparison of these conditions with the X −condition curve again reveals many non-early saccades occurring at their normal times. Figure 2 indicates that by 175 ms about 26% of all forward saccades produced in these conditions have been initiated. These are saccades that are initiated in spite of a lack of word-like units in the text. Thus, their existence indicates that word-like units are not required in order for non-early saccades to be produced at their normal times.

The relative effects of replacing spaces and substituting letters on saccade onset times can be observed in the cumulative frequency curves of Figure 2. Comparing each experimental condition with the control condition at different points in time provides an indication of the general amount of delay in saccade production that occurs in this condition. Considering the curves for the four nonX −experimental conditions, the two segmented conditions consistently show the least (Rand+) and the greatest (X+) delay, with the two unsegmented conditions lying between. This pattern does not argue for a uniquely important role for the presence of word-units in the stimulus pattern; rather, it suggests that various forms of text stimulus abnormality all delay saccades, varying in degree. Segmented nonwords (Rand+) produce the least effect, Xs produce the most even when segmented (X+), filling spaces in normal text (Orig−) produces a greater delaying effect than does replacing words with nonwords (Rand+), and filling spaces between nonwords (Rand−) produces greater delay than do nonwords alone (Rand+). This pattern of results indicates that either removing spaces or having a homogeneous (though segmented) text pattern delays saccades more than simply having unreadable text, suggesting that their effects are occurring at a prelexical processing level. Further evidence for this latter point comes from the fact that the effect onset time of 200 ms for the Rand+ condition, as seen in Figure 1, is later than the 175 ms time for unsegmented and X+ conditions, suggesting interference at a later point in processing.

Figure 3 shows the frequency distributions of fixation durations on the critical fixations in the various conditions. It shows a distribution of early saccades in the X−condition, and the lack of effect on saccade onset times in the other experimental conditions until 175 ms. All four experimental conditions, excluding the X−condition, have distributions that suggest that many saccades that normally occur prior to 225 ms are being cancelled and delayed until after 225 ms, the point at which the decline in number of saccades, relative to the control conditions, is reversed. If the distributions up to 225 ms can be taken as the normal saccades (ignoring the tail of the distribution that is probably hidden in the next bin or two), then their number, as estimated from Figure 2, is about 37–52% of fixations. Subtracting the estimate of early saccades (12%) indicates

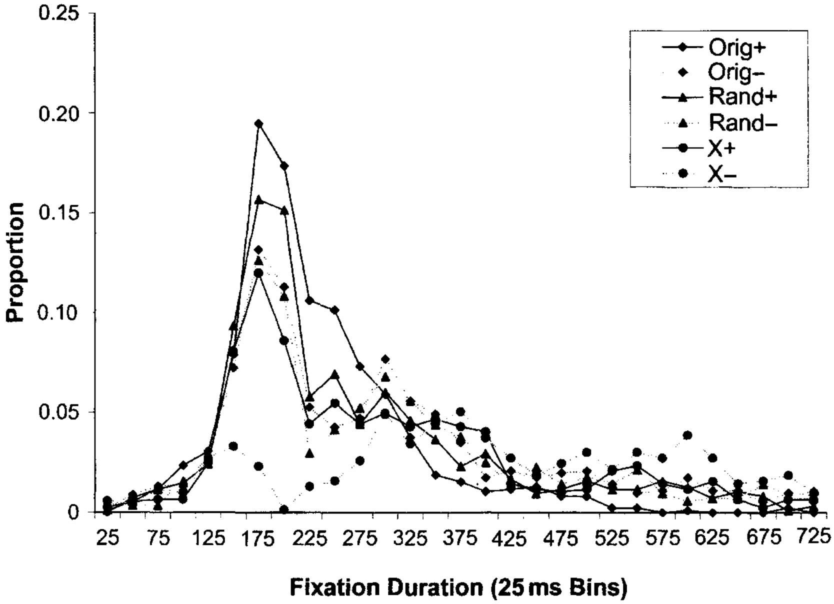

Figure 3. Frequency distributions of durations of critical fixations ending with forward saccades.

that 24% to 40% of all forward saccades are normal (undelayed) saccades in these experimental conditions. Later bins, rising above the control condition, clearly include late saccades. Analyses presented below consider the effects of the experimental variables on saccades in the three categories discussed above early, normal, and late saccades—by considering saccades in these different time ranges (0–125ms consisting of early saccades, 175 to 225ms consisting almost entirely of normal saccades, and those of 250 ms or greater being almost entirely late saccades).

Processing problems that would cause cancellation or delay of saccades undoubtedly also occur during normal reading. It is not surprising, then, to observe that the frequency distribution for the Orig+ (control) condition in Figure 3 shows an irregularity at 250 ms similar to those in the experimental conditions, though not as pronounced, suggesting the occurrence of delayed saccades in this condition as well. This feature was also present in the control condition data from the previous study. Thus, it is proposed that the frequency distribution of the durations of fixations in normal reading is actually a mixture of three distributions (Feng, 2001). Figure 3 indicates that in the control condition, 63% of the saccades occur by 225 ms, suggesting that something less than a third of the saccades made in that condition are late saccades.

Figure 4 presents the saccade onset time hazard curves for regressive saccades for the same six conditions. The Orig+ condition shows bilateral change quite

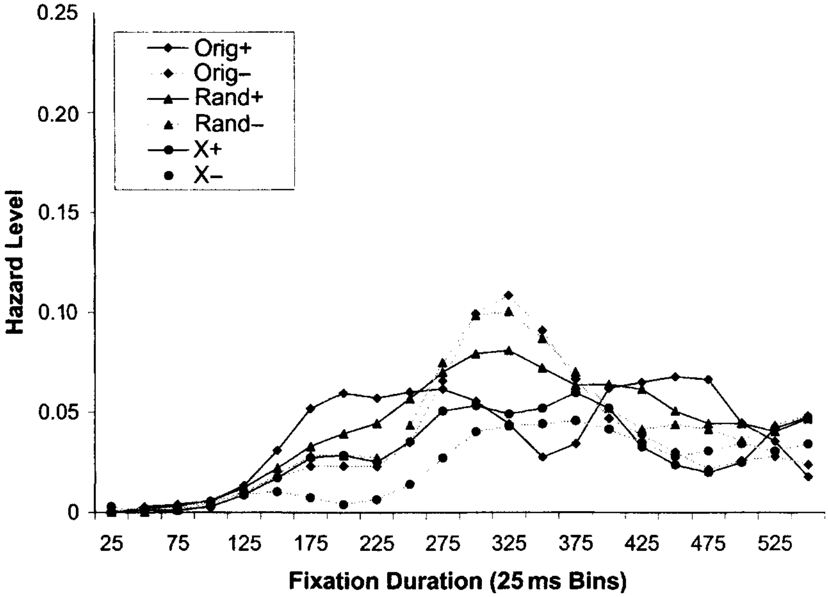

Figure 4. Hazard curves for the onset times of regressive saccades that terminate the critical fixations.

similar to the forward saccade onset hazard curve: initially low, rising, then flattening, and at approximately the same times, though with lower hazard levels reflecting the tendency to make fewer regressive than forward saccades. Regressive saccades are reduced in the experimental conditions relative to the control condition in the period of normal saccades, again bilaterally, with near complete suppression in the X−condition. These patterns indicate a symmetric reduction in saccadic activity between the left and right move centres. However, unlike forward saccades, regressive saccadic activity rises unilaterally for late saccades, with the Orig−, Rand–, and Rand+ conditions being well above the control in the 300–375 ms period and with the X−condition reaching the level of the control condition. This is probably a competitive effect: inhibition in the right move centre accompanied by increased activity in the left. Thus, the onset times of early saccades are unaffected by the experimental conditions, normal saccades are bilaterally inhibited, and late saccades show a competitive change with increased regressive activity and decreased forward activity. This pattern is discussed further below.

Saccade direction

A two-way, repeated measures ANOVA was carried out on proportions of saccades that are regressive with the data collapsed across page group because of empty cells. It found a significant effect for segmentation, $F(1, 35)=55.051$, $p<.0005$, and text anomaly, $F(2,34)=18.180$, $p<.0005$, but not for their interaction, $F(2,34)=1.527$, $p=.232$. Unsegmented text leads to a higher regression rate (55.3%) compared to segmented text (38.9%). Xs produced higher regression rates (57.0%) than random letters (47.7%) and original text (36.5%).

Since the fixation duration data indicate that all experimental conditions disrupt normal saccadic activity, showing large deviations from the control condition, with extreme disruption for the X−condition, further analyses on other dependent variables were conducted using the same four contrasts as were used with fixation duration data: first, comparing the Orig+ condition and the X−condition separately with the other four conditions, then comparing segmented (Rand+, X+) and unsegmented (Orig–, Rand–) conditions, and finally comparing Rand+ and Rand−. For each dependent variable, Table 3 shows the results of testing the first three contrasts and Table 2 shows the results of the fourth. Finally, an examination is made of changes in the dependent variable over time, measured from the beginnings of fixations, in order to understand the timing and nature of the observed effects.

For the regression frequency data, Table 3 indicates that the four critical experimental conditions increase the frequency of regressing relative to the control condition, the X−condition increases this frequency even more, and replacing spaces with letters (Orig–, Rand–) significantly increases the frequency of regressing relative to experimental conditions with nonwords (Rand+, X+). Table 2 indicates that in the Rand conditions, replacing the spaces (Rand–) produced significantly more regressions than did leaving them in (Rand+).

Figure 5 shows the proportions of saccades that are regressions at different times after fixation onset. The X−condition shows an elevated regression rate, compared to the other conditions, beginning at 150ms During the period when most saccades are cancelled, the remaining saccades are more likely than normal to be regressive. The fact that this tendency begins so early in the fixation suggests that it may result, not from a redirection of saccades, but rather from a tendency for forward saccades to be inhibited more strongly than regressive saccades. Changing the directions of saccades typically requires more time than this (Everling & Fischer, 1998). This suggestion is confirmed by Figure 4: The level of regressive saccadic activity in the X–condition does not rise in this period, it is actually reduced substantially relative to the control beginning in the 150 ms period. Thus, the increased proportion of regressions is the result of a reduction in saccadic activity level that is greater for forward than for regressive saccades. At 275 ms and later the tendency to regress rises up to 80%, indicating a strong redirection of saccadic activity among late saccades, also confirmed by data in Figure 4. Thus, unsegmented strings of Xs have no effect on the directions of early saccades, suppress normal saccades almost completely, and increase the frequencies of regressions among late saccades.

The other experimental conditions are quite similar to the control condition through 225 ms, showing little or no effect of the stimulus pattern on the

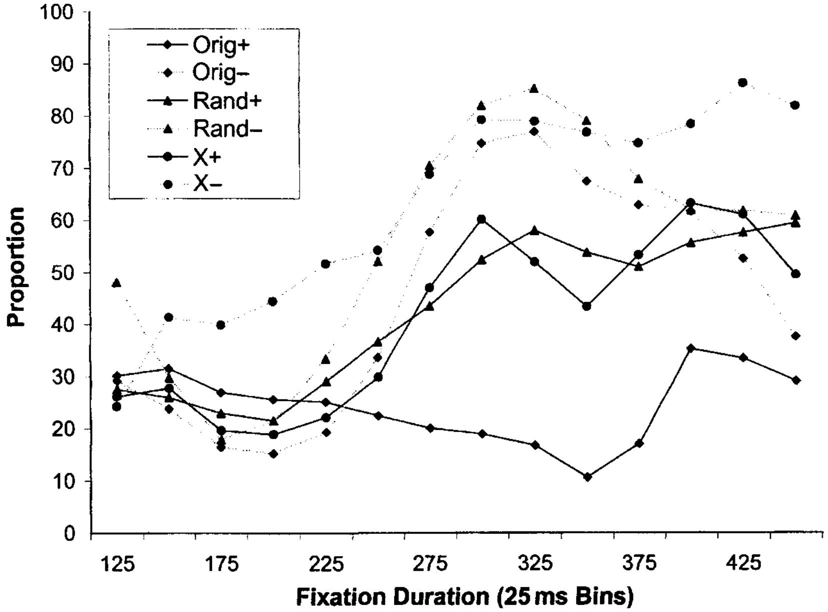

Figure 5. Time-based regression rate curves of the saccades that terminate critical fixations.

direction of either early or normal saccades. These four conditions then show a great elevation in the number of regressions beginning at 250 ms and continuing through the period charted; thus, many late saccades are being redirected left ward. In the 275–375 ms period, this tendency is greater for the unsegmented than the segmented conditions, with the unsegmented conditions reaching the level of the X–condition, while the segmented conditions hover around 50% regressions.

Lengths of forward saccades

A three-way repeated measures ANOVA of forward saccade length data found a significant effect for segmentation, $F(1, 36)=18.685$, $p<.0005$, text anomaly, $F(1.567, 56.409)=16.930$, $p<.0005$, and page group, $F(1.680,60.481)=7.692$, $p=.002$, and for the interaction between segmentation and text anomaly, $F(1.639, 58.998)=9.403$, $p=.001$. Unsegmented text (5.0 letters) resulted in shorter forward saccades than segmented text (6.0 letters), X conditions (4.6 letters) produced shorter saccades than Random letter conditions (5.5 letters), and Original text conditions (6.4 letters), and the third page group (6.1 letters) had longer saccades than the first or second groups (5.1 and 5.3 letters).

Table 3 indicates that only the second contrast is significant (X−vs. the other conditions), but the first (Orig+ vs. other conditions) and the third are not. There is no significant difference between the Rand+ and Rand−conditions. Thus, segmented and unsegmented alternative stimulus patterns both shorten forward saccades, but there is little difference between the effects of these conditions, and only the X−condition differed from the control. Thus, there is no evidence that an unsegmented text pattern, with its lack of word-like units, creates unique effects on where forward saccades are sent. This raises doubts about whether word units play a critical role in the programming of saccades.

Figure 6 shows the frequency distributions of the lengths of forward saccades for all six conditions. In the X−condition there is a great shortening of saccades (Mode=3 letter positions) compared to the control condition (Mode=7 letter positions). The other conditions all lie between these extremes, and again do not show any particular difference between conditions in which spaces are removed (Orig–, Rand–) and those with other manipulations (Rand+, X+).

Figure 7 presents the mean forward saccade length at different times following the onset of the critical fixation for the six conditions being considered. Low sample sizes for the earliest time periods leave the data quite unstable. For all conditions, including the control, the mean saccade length increases until 175 ms, then drops slowly over time. It appears that by 175 ms normal saccades and, later, late saccades are being shortened in the experimental conditions, relative to the control, with no clear difference between the unsegmented conditions, including X− , and the other experimental conditions. Of course, in the normal saccade time period, these means are based only on uncancelled saccades. Thus, the lengths of normal saccades can be modified by stimulus characteristics even when their onset times are unaffected. The observation that the X−condition shows saccade lengths similar to those of other conditions in Figure 7, but shorter than others in Figure 6, results from the fact that saccades following still longer fixations, not included in the time-based figures, tend to be quite short.

Lengths of regressive saccades

A two-way, repeated measures ANOVA of regressive saccade length data, collapsed across page groups, found no significant effect for segmentation, $F(1, 35)=2.487$, $p=.124$; but there are significant effects for text anomaly, $F(1.715, 60.022)=7.356$, $p=.002$, and the interaction, $F(2,34)=7.981$, $p=.001$.

Table 3 indicates that contrast 1 is significant, showing that regressions are shortened by these experimental conditions, but the other contrasts are not significant. Table 2 indicates no significant difference between the Rand+ and Rand–conditions. Thus, all types of alternate stimuli studied here shortened regressive saccades, with no differences among conditions. Of particular interest, a lack of word units produced no unique effect.

Figure 8 presents the frequency distributions of the lengths of regressive saccades. The distributions for experimental conditions show shorter saccades

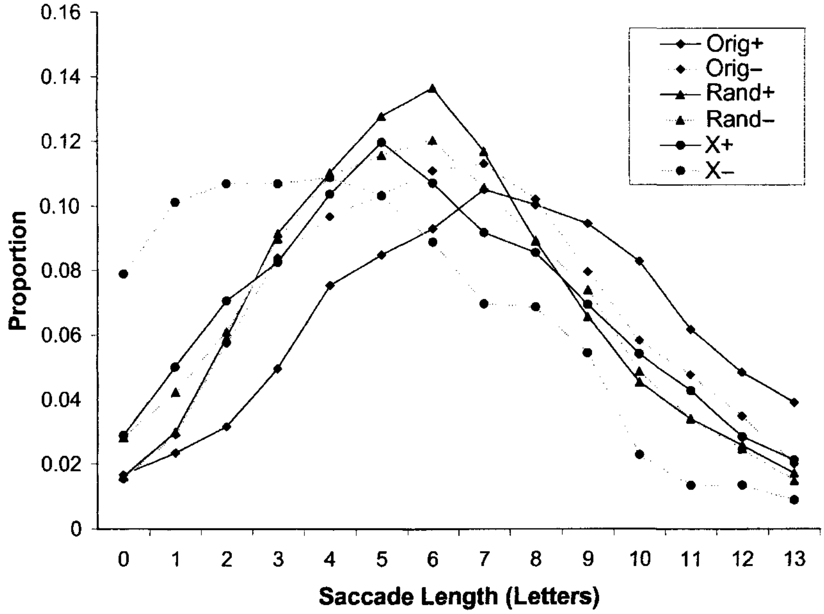

Figure 6. Frequency distributions of lengths of forward saccades that terminate the critical fixations.

than does that of the control condition but do not appear to differ among themselves. It is striking that the distribution for the X–condition is so similar to that of the other experimental conditions, in spite of the great differences in the numbers of regressions that occur and their onset times in these conditions. This replicates the Yang and McConkie (2001) finding of little effect of the various experimental conditions on regressive saccade length.

Figure 9 presents the mean lengths of regressive saccades made at different times following the onsets of critical fixations. In the control condition, the mean length drops over the first 200 ms, then rises beginning about 350 ms. All experimental conditions are similar to the control condition until 250 ms, at which time the X–condition begins producing shorter than normal regressive saccades. Other experimental conditions fall below the control beginning about 350 ms. Thus, unlike forward saccades, only the lengths of late regressive saccades are being shortened.

Summary

The analysis methods used revealed a fairly complex pattern of results that can be summarised as follows. (1) There is a set of early saccades whose onset time, direction, and length are all unaffected by alternative stimuli, and whose distribution can be observed in data from the X–condition. Onset times for these

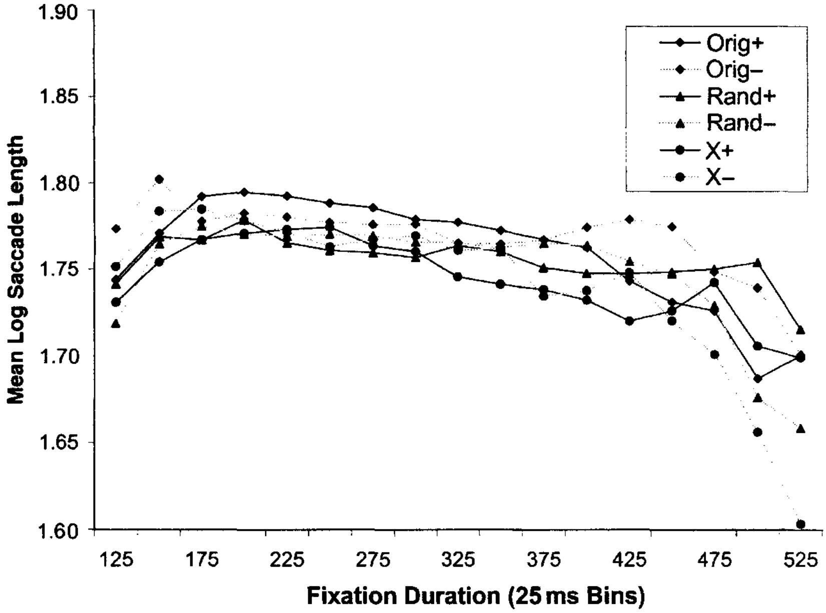

Figure 7. Mean log lengths of forward saccades that terminate the critical fixations at different times.

saccades range up to about 175 ms. (2) Both the effect onset time and the degree of reduction in the level of saccadic activity vary among experimental conditions, with the X–condition showing the earliest onset time (125 ms) and the Rand+ condition the latest (200 ms). (3) Many saccades occur at their normal times when spaces are filled and when words are replaced by nonwords or segmented strings of Xs. The frequency distributions of these normal saccades suggest that the delayed saccades, here called late saccades, begin after 250 ms. The occurrence of the undelayed normal saccades indicates that neither word spaces nor real words are required for normal saccade programming activities to occur during reading. (4) Alternate stimulus patterns produce a bilateral reduction in the level of saccadic activity for normal saccades (e.g., a reduction in both forward and regressive saccades, with a greater proportional reduction in forward saccades than in regressions) but a unilateral rise in regressive late saccades. (5) The mean lengths of normal and late forward saccades in experimental conditions are reduced, with little difference among conditions. (6) Late regressive saccades were shortened, but only after 225 ms in the X– condition and 325 in the other experimental conditions. (7) The effects of alternative stimuli on the saccade onset time hazard curve shows the pattern predicted by C/I theory (an initial rise, followed by a drop) rather than that expected on the basis of the cognitive saccade triggering assumption (a delay or reduction in the rise of the hazard level).

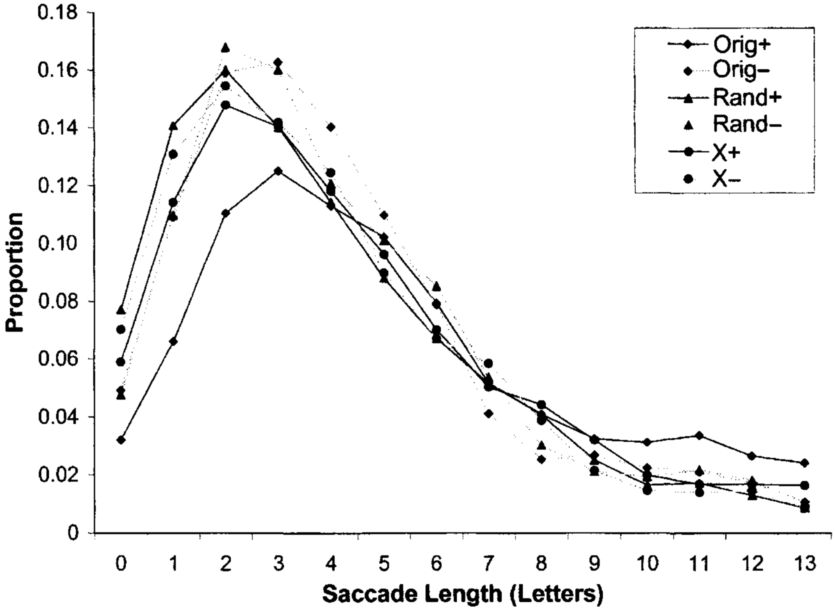

Figure 8. Frequency distributions of lengths of regressive saccades that terminate the critical fixations.

DISCUSSION

Most current theories of eye movement control in reading intimately involve word units in determining both the onset times and spatial properties of saccades. The current study addresses two assumptions that are common in word based control theories: the cognitive saccade triggering assumption and the word-based targeting assumption. The question is whether having alternate stimuli that do not support these word-based activities substantially delays the making of saccades and changes their properties in ways that would be expected from a word-based control position.

The study was stimulated by Yang and McConkie's (2001) observation of the X–effect and their suggestion that this near-complete elimination of saccades right at the time when most saccades are produced, as a result of having unsegmented strings of Xs on that fixation, may be evidence for word-based control. The lack of space- or otherwise-delineated word units could be preventing the normal, word-based saccade-initiating processes from occurring. The current study did not support that interpretation, finding that replacing spaces in normal text and in nonword text-like stimuli with random letters does not produce the X–effect. While filling spaces in text does reduce saccadic activity levels, it does not eliminate saccades in the way that unsegmented strings of Xs do. Further research is needed in order to identify the cause of the X –effect, which Yang and McConkie (2001) also observed in two other

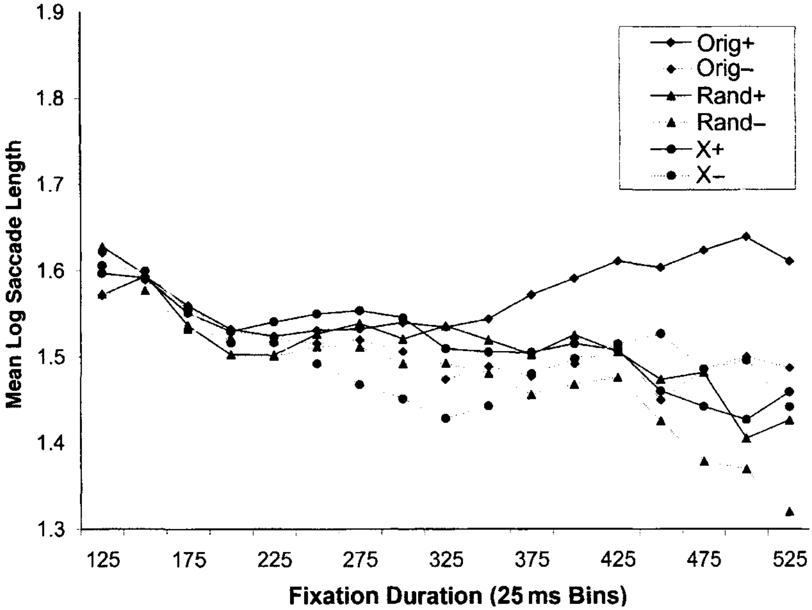

Figure 9. Mean log lengths of regressive saccades that terminate the critical fixations at different times.

conditions: replacing text with a blank page or with underline characters instead of letters. We suspect that the cause lies in the mechanism involved in the capture of visual attention: This probably does not require a salient change in the overall illumination (Theeuwes, Kramer, & Atchley, 2001), nor does it necessarily occur with unfoveated stimulus changes (see change blindness literature: O'Regan & Noë, 2002; Rensink, 2001; Simons, Franconeri, & Reimer, 2000). The X–effect does not occur when letters are replaced by Xs if word spaces remain, nor does it occur with text-like stimuli if word spaces are removed; thus, it seems to require both the removal of spaces and a foveal, nontext-like pattern that quickly alerts the system to the stimulus abnormality.

Data from the X–condition reveals a set of early saccades, consisting of about 12% of all saccades in the current study, whose onset times and other properties are unaffected by the stimulus pattern on the critical fixation (i.e., the fixation that is terminated by that saccade). Part of these may be the saccades that Morrison (1984) and others have suggested are programmed on the basis of information from the prior fixation.

All experimental conditions were found to produce changes in the eye movement patterns, affecting both the saccade onset times and their directions and (for forward saccades) their lengths. This is consistent with all current theories of eye movement control. The matter of concern here is whether the pattern of effects in conditions other than X–are harmonious with word-based control assumptions or with the C/I theory.

Cognitive saccade triggering assumption

If all saccades, other than early saccades, were initiated by the occurrence of some critical event in the processing stream for words, then stimulus changes that delay or prevent that event should delay or prevent normal saccades. If a saccade is not initiated on the normal basis, then the eyes must be moved on some other basis, probably by some type of recovery process. This perspective seems to distinguish between four types of saccades within a SFR study framework: (1) the early saccades described above, (2) timely saccades that occur at their normal time because the usual saccade-relevant processes are unaffected by the alternative stimulus pattern, (3) delayed saccades, which occur through the normal saccade initiation processes but are delayed by a delay in the critical saccade triggering event that is caused by the alternate stimulus pattern, and (4) recovery saccades that are initiated on some other basis when the normal saccade initiation processes fail. Henderson (1992), for example, proposed that if no saccade has occurred by a particular time, a saccade is generated on some other basis.

While the data from the current study clearly show that nonword and unsegmented text delay some saccades, we have argued above that the saccade onset time (fixation duration) distributions do not show the properties that would be expected from the cognitive saccade triggering assumption. If saccades require word-based perceptual and cognitive activities for their normal initiation, then the lack of word units in the text (Orig—and Rand—conditions) should result in only delayed and recovery saccades, beyond the early saccades. However, with the exception of the X–condition there were large numbers of timely saccades in all experimental conditions, as indicated by the lack of any reduction in saccadic activity, relative to the control condition, until 175 ms. Many saccades appear to occur at their normal times even after that. These timely saccades serve as an existence proof that normal saccadic control does not require word-based processes to occur, including those assumed by the cognitive saccade triggering assumption. The fact that these same stimulus conditions produce many delayed and recovery saccades at later times does not invalidate this conclusion, since almost any theory of saccade control would expect such effects to eventually occur in the face of such reading-unfriendly stimulus patterns. The issue here is not whether the lack of a properly segmented stimulus pattern can produce saccade delays, but whether word-based processes are so critical to normal saccade initiation that their absence makes normal saccadic activity impossible. The existence of timely saccades that are not early saccades, in the presence of unsegmented text, argues against this necessity and, hence, against the assumption itself.

Word-based targeting assumption

The decision of where to send the eyes is usually assumed either to result from sequentially attending each word (Henderson, 1993; Reichle et al., 1998), or from a competition among the words (Kliegl & Engbert, 2003). Both of these positions require the presence of word-like units in the stimulus pattern. While it is clear that unsegmented text should prevent these normal processes from occurring, thus requiring the decision of where to move the eyes to be made on some other basis, no precise prediction can be made about where the eyes would be sent in this situation. A lack of word units could cause the eyes to remain in place, producing short saccades and many regressions, to move forward an average distance with a reduction in regressions, or to move rather randomly in direction and length. The least likely alternative would be to find little effect on saccade direction and length. The presence of nonwords (Rand+) and segmented strings of Xs (X+) could be expected to support the word-based attentional or selection activities, though the lack of lexical-level linguistic information would be expected to shorten saccades, as compared to the control (Orig+) condition, representing a reduced likelihood of skipping a word (Ehrlich & Rayner, 1981). While the study found that forward saccades were indeed shortened by the experimental conditions, and the number of regressions increased, these effects were very similar for the segmented and unsegmented experimental conditions, thus providing no evidence for a critical role of the presence of word-like units in determining where to send the eyes. The increase in regressions only appears after 250 ms, indicating that the direction of most saccades is not affected by whether or not words or word-like units are present. These results do not provide support for a critical role of word units in decisions of where the eyes are sent during reading. A proposal for how this conclusion can be reconciled with current evidence that saccadic landing positions are word based is given below.

There is an important caveat to the claims made above. The current study focuses entirely on the effect that the stimulus present during a particular eye fixation has on the onset time, direction, and length of the saccade that ends that fixation. It is clear that carry-over effects exist, in which processes initiated by stimuli present on fixation *N* affect the characteristics of the saccade that terminates fixation *N*+1 (Rayner, 1998). For example, the absence of spaces in the text may have a greater or different effect on the later saccade than does the presence of segmented nonwords or Xs. The current analyses do not address these effects, since, with the exception of the typical account of early saccades (Morrison, 1984; Reichle et al., 1998), they have not been well developed in theories of eye movement control during reading (though some such influences are included in the E-Z Reader model, Reichle et al., 1998).

Alternative word-based explanations

The conclusions reached above conflict with commonly made assumptions about the nature of eye movement control during reading. There are several possible directions that can be taken to account for these results without giving up a word-based control theory. The first approach is to propose that the initial distribution of saccades seen in the X−condition does not represent the full set of early saccades (e.g., saccades that are initiated on the basis of processing initiated by information from the previous fixation). Rather, all fixations whose onset times remain unaffected by, say, the Orig−or the Rand+ condition, might be considered to be early saccades. However, this would mean that many, and perhaps even most, saccades are being triggered by a cognitive event that results from information obtained on the previous fixation. Only by 175 or 200 ms would the critical cognitive event be assumed to occur from current stimulus information, triggering non-early saccades that, therefore, are susceptible to being delayed by alternative stimuli such as those employed in the current study. It should be noted that this proposal would also require invoking another process in the case of the X−condition to account for the cancellation of many early saccades. This would probably involve an emergency reaction process that cancels even early saccades when a very abnormal visual event occurs.

A somewhat related proposal is included in the E-Z Reader model (Reichle et al., 1998). This model assumes that a saccade to the currently attended word is automatically planned, starting at the beginning of every fixation. This saccade is executed unless an estimate of the frequency of the fixated word is accomplished prior to the completion of that planning, enabling its cancellation and the planning of a new saccade to the immediately following word. Thus, it is assumed that there is another set of saccades, beyond the early saccades, whose onset time occurs too early to be determined by the characteristics of the current stimulus configuration. This proposal could accept the initial distribution of saccades in the X−condition as indicating early saccades, while attributing the further set of saccades that are unaffected by the lack of spaces (which should delay or prevent processes involved in word familiarity estimation) to a failure to cancel the initially planned saccade. However, this proposal makes a prediction that is not supported by the present data. It predicts that the initially planned saccades would all be refixations, while early and later saccades should typically not be. Early saccades should take the eyes to the next word, according to the mechanism by which they are assumed to be generated, while later (forward) saccades should go to the next or a following word. This should result in a reduction in the mean saccade length, relative to the control condition, during the period when the initially planned saccades are occurring, probably 150–175 ms. The data presented in Figure 7 do not show this pattern.

Two other explanations of the current results would attribute them to artefacts of the experimental method itself, thus suggesting that they do not call into question the assumptions described as being tested in this study. The first of these would claim that the results are in some way caused by the occurrence of display changes. Previous research has demonstrated that changing the word to which the eyes are going during the saccade results in an increased fixation duration and increased likelihood of refixation (Rayner, 1974) which may be an indication of the general disruption of normal processing under these conditions. From this perspective, the results obtained in the current study would not be assumed to provide information about the normal process of saccade generation during reading. An unfortunate result of taking this position, of course, is that all studies involving this eye movement contingent display control methodology would then be assumed to be irrelevant to the understanding of normal reading. While we agree that such display changes do disrupt the normal processing taking place, it is the conditions, timing, and effects of this disruption that provide useful information about the nature of the normal processes.

The second artefact-oriented explanation suggests that the frequent occurrence of alternative stimuli in the present study, and the processing disturbances that result, may cause readers to change their eye movement control strategies in ways that invalidate the current results as indications of the nature of normal control. In particular, since the alternate stimulus patterns appear until the reader makes a saccade, an optimal strategy would be one in which a short saccade is made as quickly as possible when an alternate pattern appears, as a way of bringing normal text back onto the screen. Again, we agree that the display changes make reading more difficult, which probably causes some change in strategy, though that itself would not mean than an abnormal strategy is being employed. The more important question is whether readers can adopt a strategy by which basic eye movement decisions are altered in a fundamental manner that changes the effect onset times, hazard levels, and saccade directions and lengths of saccades. McConkie and Yang (2003) provide evidence of direct cognitive control of saccade direction, but also indicate that this occurs quite late, in this case only after 325 ms. Thus, we suspect that effects occurring earlier in the fixation can not be directly cognitively controlled based on information obtained from the current fixation. Further research is needed to determine whether it is possible to establish alternate oculomotor control strategies in reading that change such early aspects of eye behaviour, and whether that is taking place under the current circumstances. Our own scepticism that this is the basis for the data patterns reported above comes from the general failure of the subjects to develop the proposed optimal strategy: Saccade onset was very slow in the X–condition where the stimulus abnormality was most obvious, and the proportion of regressions in experimental conditions other than the X–condition did not begin to increase, relative to the control condition, until the 250 ms period.

C/I theory

An alternative framework within which to interpret the results of the present study is found in the theory initially presented by Yang and McConkie (2001). It postulates that saccades during reading are initiated after random waiting times, though the state of the system varies over the fixation period in a manner that is represented as a change in the value of the parameter controlling the momentary mean waiting time. This results from a saccade-related inhibition of further saccadic activity once a saccade has been initiated. This inhibition is lifted at a particular time as part of an eye movement control strategy, allowing the saccadic activity level to rise until it reaches a maximum whose level is determined by another parameter that is defined strategically, controlling the maximum activity level which, as a result, controls the general mean fixation duration during this period. Where the eyes go results from a competition between and within the left and right portions of the saccade generating system, including the superior colliculus. The rightward bias in reading European languages results from strategy-based activation being added to the right portion of the system, with a maximum activation level occurring at a strategically determined retinal location. The stimulus also affects the pattern of activation, with greater activation occurring at the locations of words. Finally, when problems are encountered at some processing level (here produced by the alternate stimulus patterns), a distress call causes a processing-based inhibition of the saccadic system, reducing the saccade onset hazard level. Whether a particular saccade is inhibited depends, first, on whether it is produced before or after the oculomotor system is inhibited, and, second, if after, on the degree of processing-based inhibition.

If processing problems are encountered and the saccadic system is inhibited, the effect onset time and degree of this inhibition varies with the processing centre having difficulty, and the degree of this difficulty. The data from the present study indicate that unsegmented text and segmented strings of Xs produce slightly earlier and stronger inhibition than does the presence of nonwords (Rand+ condition), suggesting that different processing centres are being affected. Early saccades are generated in the same manner as other saccades, but they occur too early to be affected by inhibition from problems in any processing centre that result from the current stimulus pattern. From this perspective, there is no need for word-based processes to be involved in the normal generation of saccades during reading, so the onset times for normal saccades are unrelated to the current momentary language processes. Delayed influences can occur, such as by adjusting the maximum hazard level, which modifies the mean fixation duration, or changing the location of maximum strategy-based activation, which modifies forward saccade length as means of adjusting the reading speed. Processing-based inhibition has the effect of cancelling some proportion of the saccades being programmed. The uncancelled saccades occur at their normal times and directions; it is the properties of the late saccades, those that are delayed because of the earlier cancellation, that are most frequently modified. The current data patterns suggest that these changes result primarily from a second effect of processing based inhibition; that is, the elimination of strategy based activation. Eliminating this activation reduces or eliminates the rightward bias in saccadic activity in reading, leading to an asymmetric increase in the number of regressive eye movements, and also reduces the lengths of forward saccades. Length reduction comes from assuming that in reading the peak of the distribution of strategy-based activation is usually more eccentric than the peak activation would be without it. At the same time, the removal of the added activation has no effect on the lengths of late regressive saccades, since the activation pattern in this left visual region is not changed. At a later point in time (after 325 ms) direct cognitive influences can begin producing a new source of bias in where the eyes go (McConkie & Yang, 2003), in the present case, shortening the saccades.

There is one interesting effect that was not anticipated: Normal forward saccades are shortened by the alternative stimulus patterns, relative to the control condition. Since many of the normal saccades were cancelled and only occurred later, the saccade length data in the normal saccade range comes from saccades that occurred in spite of the inhibition. This leads to the conclusion that the onset of processing-based inhibition must immediately reduce strategy-based activation, resulting here in a shortening of forward saccades, even for saccades whose onset time is unaffected by the inhibition. This shortening is apparent in Figure 7 beginning at 175 ms, when the inhibition is first observed. However, there is no increase in regressions until late saccades begin. Thus, the normal forward saccades appear to be committed to a direction that can only be changed by cancelling the saccade, but the final destination can be modified by a change in the activation pattern.

There are several results that might have been expected but that were not found. The lack of these differences is also compatible with C/I theory. First, although there is considerable evidence that the eyes tend to go to words during reading, the filling of spaces did not produce delays in the saccade onset times that were much greater than having nonwords nor did it produce great changes in the distribution of forward saccade lengths. All experimental conditions, other than the X–condition, showed quite similar data patterns. One might expect that the lack of word units would produce great changes in when and where the eyes are sent, but this was not so in the data. According to the C/I theory, the presence of words in the stimulus pattern does modify somewhat the pattern of activation, such that locations in words are more likely to be selected as target locations than are locations between words. However, the absence of words in no way changes the dynamics of saccade generation, it just modifies the likely winning location. Furthermore, since strategy-based activation biases a given region to be more likely to include the winning location, this will be the case whether or not the stimulus is segmented. Thus, segmented and unsegmented text should produce similar saccade length distributions in SFR studies. Second, in conditions that produce great increases in the number of regressive saccades, the lengths of these saccades appear to be distributed similarly to those of the saccades in the control condition (except for late-developing influences). One might expect that these many “induced” regressions would reflect the processing difficulty that produced them and thus show quite a different distribution from normal regressions. However, the C/I theory assumes that these saccades are redirected leftward because of a reduced attractiveness of locations in the right visual field, due to reduced strategy-based activation, with no changes in the activation pattern of words to the left; it is this pattern that determines where the eyes will go in a regression, hence producing a normal length distribution. This again shows the separation between momentary cognitive activities and eye movement decisions in the C/I theory.

The above discussion highlights the fact that a number of findings from this study, while being unexpected from the perspective of word-based control theories, are quite compatible with a competition/interaction type of theory.

Eye movements and research on language processes

The C/I theory account of eye movement control abandons the common assumption of a rather tight link between cognition and eye movements in the case of reading. This seems to fly in the face of the many studies that have successfully used eye movement recording to investigate language processes during reading, showing longer fixations and gaze durations for word locations where processing difficulties lie, whether lexical, syntactic, or semantic (Rayner, 1998; Rayner & Pollatsek, 1989). The very existence of this literature might be considered to be a major argument against the acceptability of the C/I theory. However, it is important to note two facts. First, the current data deal only with the processes that determine the time and properties of a saccade that ends the fixation on which the information that enables those processes is obtained. From the point of view of language research, the current study deals only with the duration of the first fixation on a stimulus, and then without the benefit of much information that might normally be obtained through peripheral preview. Up to this point we have explicitly excluded any study of the effect of the stimulus on later fixations which would be necessary to illuminate the nature of the gaze duration measure. Second, the reading studies have been conducted using mean fixation times and gaze durations as data. It has been assumed that the durations of the individual fixations are affected in a continuous or graduated manner by the language variables being studied, similar to the differences observed in the means. However, in principle these differences in means could just as easily arise from a more discrete and probabilistic system, such as that proposed in the C/I theory. A system in which saccades are generated at random times, and in which the occurrence of a processing difficulty delays a saccade for 100 ms with a likelihood that varies with the degree of the difficulty, could produce mean fixation durations between conditions that vary continuously with the processing difficulty. Thus, the C/I theory predicts that one would find the processing difficulty-related differences in mean fixation durations that have been observed; it just proposes a different basis on which they are produced.

PrEview proof published online September 2003

REFERENCES

Beauvillain, C., Doré, K., & Baudouin, V. (1996). The “center of gravity” of words: Evidence for an effect of the word-initial letters. *Vision Research, 36*(4), 589–603.

Ehrlich, S.F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20,* 641–655.

Epelboim, J., Booth, J.R., Ashkenazy, R., Taleghani, A., & Steinman, R.M. (1997). Fillers and spaces in text: The importance of word recognition during reading. *Vision Research, 37*(20), 2899–2914.

Epelboim, J., Booth, J.R., & Steinman, R.M. (1994). Reading unspaced text: Implication for theories of reading eye movements. *Vision Research, 34*(13), 1735–1766.

Everling, S., Dorris, M.C., Klein, R.M., & Munoz, D.P. (1999). Role of primate superior colliculus in preparation and execution of anti-saccades and pro-saccades. *Journal of Neuroscience, 19*(1), 2740–2754.

Everling, S., & Fischer, B. (1998). The antisaccade: A review of basic research and clinical studies. *Neuropsychologia, 36*(9), 885–899.

Feng, G. (2001). *Share: A stochastic, hierarchical architecture for reading eye-movement.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, USA.

Henderson, J.M. (1992). Visual attention and eye movement control during reading and picture viewing. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 260–283). New York: Springer-Verlag.

Henderson, J.M. (1993). Visual attention and saccadic eye movements. In G.d’Ydewalle & J.van Rensbergen (Eds.), *Perception and cognition* (pp. 37–50). New York: Elsevier Science.

Henderson, J. M, & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 417–429.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*(2), 1201–1221.

Hsu, S.-H., & Huang, K.-C. (2000). Effects of word spacing on reading Chinese text from a video display terminal. *Perceptual and Motor Skills, 90*(1), 81–92.

Inhoff, A.W., Starr, M,Liu, W., & Wang, J. (1998). Eye-movement-contingent display changes are not compromised by flicker and phosphor persistence. *Psychonomic Bulletin and Review, 5*(1), 1101–1106.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind’s eye: Cognitive and applied aspects of eye movement research* (pp. 391–411). Amsterdam: Elsevier Science.

Liu, I.M., Yeh, J.S., Wang, L.H., & Chang, Y.K. (1974). Effects of arranging Chinese words as units on reading efficiency. *Acta Psychologica Taiwanica, 16,* 25–32.

Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81*(12), 1899–1917.

McConkie, G.W. (1979). On the role and control of eye movements in reading. In P.A.Kolers, M. Wrolstad, & H.Bouma (Eds.), *Processing of visible language* (Vol. I, pp. 37–48). New York: Plenum Press.

McConkie, G.W., & Dyre, B.P. (2000). Eye fixation durations in reading: Models of frequency distributions. In A.Kennedy, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 683–700). Oxford, UK: Elsevier.

McConkie, G.W., Kerr, P.W., Reddix, M.D., & Zola, D. (1988). Eye movement control during reading: (I) The location of initial eye fixations on words. *Vision Research, 28* (10), 1107–1118.

McConkie, G.W., & Yang, S.-N. (2003). How cognition affects eye movements during reading. In J. Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 413–427). Amsterdam: Elsevier Science.

Meredith, M.E., & Ramoa, A.S. (1998). Intrinsic circuitry of the superior colliculus: Pharmacophysiological identification of horizontally oriented inhibitory neurons. *Journal of Neurophysiology, 79,* 1597–1602.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 667–682.

O'Regan, J.K. (1989). Visual acuity, lexical structure, and eye movements in word recognition, In B. G.Elsendoorn & H.Bouma (Eds.), *Working models of human perception* (pp. 261–292). London: Academic Press.

O'Regan, J.K., & Levy-Shoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. In M.Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 363–383). Hove, UK: Lawrence Erlbaum Associates Ltd.

O'Regan, J.K., & Jacobs, A.M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception & Performance, 18*(1) 185–197.

O'Regan, K. (1994). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (Vol. 4, pp. 397–454). Amsterdam: Elsevier.

O'Regan, J.K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 883–917.

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

Rayner, K. (1974). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65–81.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*(1), 21–30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K., Fischer, M.H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice-Hall.

Rayner, K., & Pollatsek, A. (1996). Reading unspaced text is not easy: Comments on the implications of Epelboim et al.'s (1994) study for models of eye movement control in reading. *Vision Research, 36*(3), 461–465.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39*(26), 4403–4411.

Reilly, R.G., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429–455). Amsterdam: Elsevier Science.

Reingold, E.M., & Stampe, D.M. (1999). Saccadic inhibition in complex visual tasks. In W.Becker & T.Mergner (Eds.), *Current oculomotor research: Physiological and psychological aspects* (pp. 249–255). London: Plenum.

Rensink, R.A. (2001). Change detection. *Annual Review of Psychology, 53*(1), 245–277.

Simons, D.J., Franconeri, S.L., & Reimer, R.L. (2000). Change blindness in the absence of a visual disruption. *Perception, 29*(10), 1143–1154.

Sun, F., Morita, M., & Stark, L.W. (1985). Comparative patterns of reading eye movement in Chinese and English. *Perception and Psychophysics, 37,* 502–506.

Theeuwes, J., Kramer, A.F., & Atchley, P. (2001). Spatial attention in early vision. *Acta Psychologica, 108*(1), 1–20.

Tsai, J.-L., & McConkie, G.W. (2003). Where do Chinese readers send their eyes? In J.Hyönä, R. Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 159–176). Amsterdam: Elsevier Science.

Walker, R., Deubel, H., Schneider, W.X., & Findlay, J.M. (1997). Effect of remote distractors on saccade programming: Evidence for an extended fixation zone. *Journal of Neurophysiology, 78*(2), 1108–1119.

Yang, H.-M. (1998). The influence of adjustment in space size of between- and within words on Chinese reading. *Journal of National Taiwan Teachers College,* 37, 303–325.

Yang, H.-M., & McConkie, G.W. (1999). Reading Chinese: Some basic eye-movement characteristics. In J.Wang & A.W.Inhoff (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 207–222). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Yang, S.-N. (2002). *Inhibitory control of saccadic eye movements in reading: A neurophysiologically based interaction-competition theory of saccade programming.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, USA.

Yang, S.-N., & McConkie, G.W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research, 41,* 3567–3585.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 262–284

Length, frequency, and predictability effects of words on eye movements in reading

Reinhold Kliegl, Ellen Grabner, Martin Rolfs, and Ralf Engbert

University of Potsdam, Germany

We tested the effects of word length, frequency, and predictability on inspection durations (first fixation, single fixation, gaze duration, and reading time) and inspection probabilities during first-pass reading (skipped, once, twice) for a corpus of 144 German sentences (1138 words) and a subset of 144 target words uncorrelated in length and frequency, read by 33 young and 32 older adults. For corpus words, length and frequency were reliably related to inspection durations and probabilities, predictability only to inspection probabilities. For first-pass reading of target words all three effects were reliable for inspection durations and probabilities. Low predictability was strongly related to second-pass reading. Older adults read slower than young adults and had a higher frequency of regressive movements. The data are to serve as a benchmark for computational models of eye movement control in reading.

An important goal of reading research is to disentangle visual, lexical, and contextual processing factors. Three word characteristics that have been reliably linked to these three factors are their length, their frequency in written texts, and their predictability from the preceding context of the sentence. These relations have been established for various measures of inspection duration (e.g., first fixation duration, single fixation duration, gaze duration; see Methods for definitions) and inspection probabilities (e.g., of skippings, single fixations, or multiple fixations) but there is still little research looking at these effects in combination within the same data set (see Rayner, 1998, for a general review, and Calvo & Meseguer, 2002, for a review and the most recent specific attempt to examine the combination of these effects). In the present experiment we examined the influence of these variables on the set of standard measures and a few additional measures of processing beyond first-pass reading. Most importantly, we tested the effects (1) for all words of the sentence corpus (facing the problem of correlated effects) and (2) for a subset of target words uncorrelated in length and frequency (facing the problem of generalisability to all words of the corpus) with repeated-measures multiple regressions in both sets of analyses. This way we could compare effects of word length, frequency, and predictability in a statistical control approach with those from an experimental control approach. Finally, we also checked the stability of these effects across the adult life span with samples of young and older adults.

A VIEW FROM THE EXPERIMENTAL CONTROL APPROACH

Given the large correlation between length and frequency of words for all words of a text or sentence corpus (e.g., −.64 in the one used in this experiment), it has been a longstanding problem to determine the independent contributions of these variables to visual and lexical processes.[1] The experimental method of choice to dissociate effects such as word length and word frequency, and the one most frequently used, is to restrict the analyses to a subset of target words which then by experimental design are orthogonal or uncorrelated. Our experiment also afforded such an analysis because one target word per sentence had been specified a priori to contribute to an orthogonal design with frequency and word length as factors.[2]

There is already much information available about the effects of word length and word frequency on inspection durations and inspection probabilities. In general, they covary with processing difficulty, that is they increase with word length and decrease with word frequency (e.g., Rayner, Sereno, & Raney, 1996; Schilling, Rayner, & Chumbley, 1998). However, the effect of word frequency on inspection probabilities has not been observed consistently (e.g., Gautier, O'Regan, & Le Gargasson, 2000; Henderson & Ferreira, 1993; Kennison & Clifton, 1995) and there were also exceptions to effects of word length on early

Correspondence should be addressed to R.Kliegl, Dept of Psychology, University of Potsdam, PO Box 60 15 53, 14451 Potsdam, Germany. Email: Kliegl@rz.uni-potsdam.deWe thank Petra Grüttner and Sabine Kern for research assistance. Antje Nuthmann, Ralph Radach, and an anonymous reviewer provided helpful comments on an earlier version of this paper. Sentences and associated word norms are available from the authors. This research was supported by grant KL 955/3 from Deutsche Forschungsgemeinschaft.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000213

processing stages (i.e., no effect on the duration of first fixations; Hyönä & Olson, 1995).

There are quite a few studies demonstrating the influence of frequency and predictability with word length controlled. For example, Brysbaert and Vitu (1998) identified seven previously published studies for their meta-analysis of frequency effects on word skipping and eight studies for the analogous analysis of predictability effects (see also Rayner, Binder, Ashby, & Pollatsek, 2001, for a recent report). Low-predictability words also attract regressive movements (Rayner & Well, 1996). There are also reports about effects of predictability on inspection durations. Binder, Pollatsek, and Rayner (1999; see also Rayner et al., 2001, Rayner & Well, 1996) reported shorter first-fixation durations for predictable words. Calvo and Meseguer (2002) also examined the time course of effects: Word length was significant for early and late processing stages (i.e., first-fixation and gaze durations), frequency for early stage (i.e., first-fixation durations) and predictability primarily for late stage of processing (i.e., gaze durations). Moreover, word length accounted for more variance (25%) than frequency (4%) and predictability (10%) when averaging across a number of statistically significant eye movement measures.

A VIEW FROM THE STATISTICAL CONTROL APPROACH

Clearly, then, there is already very reliable evidence for the influence of word length, frequency, and predictability on eye movement measures from studies using the experimental control approach. There remains the question of how results based on this approach generalise to all words of a sentence corpus. One can use statistical control techniques such as multiple regression to assess the effects of correlated predictors. Lorch and Myers (1990) described an appropriate procedure following less suitable earlier proposals (Blanchard, 1985; Just & Carpenter, 1980; Kliegl, Olson, & Davidson, 1982), which nevertheless are still in use. Specifically, variables such as word length, frequency, and predictability represent continuous repeated-measures design factors. Therefore, the data need to be analysed with repeated-measures multiple regression analyses removing systematic variance between persons. This can be accomplished by (1) regressing the dependent variable of choice (i.e., various inspection durations or

[1] It could be argued that word length is also a linguistic variable. Specifically, one may assume that knowledge about the length of a to-be-processed word constrains the number of candidates in the lexical processing stage. However, as shown in a recent series of studies, this appears not to be the case (Inhoff & Eiter, 2003; Inhoff, Radach, Eiter, & Juhasz, 2003).

[2] Another strategy is post-hoc selection of words from the reading material for an orthogonal design (e.g., Kliegl, Olson, & Davidson, 1983; Radach & Heller, 2000; Radach & McConkie, 1998).

probabilities) on the predictor set for each participant and (2) testing whether the means of the unstandardised coefficients (averaged across participants) differ significantly from zero. Moreover, one can test whether the means differ significantly between groups such as young and older adults in the present experiment. In ANOVA terminology, a significant group difference on a predictor (such as length, frequency, or predictability) translates into a significant interaction between the group and this factor. The disadvantage associated with the multiple-regression approach (repeated-measures based or not) is that variance shared between predictors cannot be used for statistical tests of effects, leading to much more conservative tests compared to experimental designs with orthogonal factors, especially in the case of highly correlated predictors such as word length and word frequency.

The studies mentioned above examined the effects of word length and word frequency, mostly in combination with a set of other language-related factors (e.g., function vs. content words, number of syllables). In general but not always, results related to word length and frequency were consistent with those reported above but, as the statistics were not appropriate (i.e., between-subject variance was not removed), we do not present a review here. Moreover, as far as we know, there is no research examining word length and frequency together with word predictability for data of all words of sentences or texts. Indeed, apparently only Reichle, Pollatsek, Fisher, and Rayner (1998) provided length, frequency, and predictability information for all words of a corpus of sentences; they supplemented data collected by Schilling et al. (1998) with predictability information for use in the E-Z Reader model. Reichle et al. did not report any specific analyses contrasting length, frequency, and predictability effects.

Thus, there still is no information about differences in effect sizes associated with the experimental control and statistical control methods. Of course, there can be no doubt that, irrespective of generalisability, experimental effects are highly important for theories about eye movement control in reading. However, we often do not know whether these effects are of any practical relevance in normal reading. For example, if there is no reliable unique variance associated with frequency or predictability beyond that accounted for by word length in an analysis of eye movements including all words of a text, one may well ignore these variables for the prediction of individual differences in reading efficiency. More importantly even, knowledge about the generalisability of experimental effects is of high theoretical relevance for the development and evaluation of computational models of eye movement control during reading because these models claim to embody the dynamics of "normal reading" and, consequently, must be evaluated on a variety of eye movement measures for all words of sentences or texts, not only a select subset of them. Ideally, of course, computational models should reproduce differences between results from experimental control and statistical control techniques.

Our selection of dependent variables closely followed the scheme proposed by Reichle et al. (1998). Engbert and Kliegl (2001), Engbert, Longtin, and Kliegl

(2002) and Reichle et al. (1998; Reichle, Rayner, & Pollatsek, 1999, in press) used these benchmark data as target for their computational models. Obviously, a second corpus with similar and additional information will facilitate the further development of computational models. Reichle et al. (1998) reported first fixation durations (FF), single fixation durations (SF), gaze durations (GD), as well as probabilities for skipping (P0), single fixations (Pl), and multiple fixations (P2) on first-pass reading as a function of five word frequency classes. We extended this protocol with information about the effects of word length and predictability and with measures relating to information uptake beyond first-pass reading. Specifically, we added total reading time per word (TT) and probabilities of a word serving as origin (RO) or goal (RG) of a regressive eye movement. The general expectation was that difficult words might attract reinspections to compensate for word identification problems (Vitu & McConkie, 2000). It turned out that low predictability may be the most critical feature triggering such reanalysis.

The main goal of this experiment, then, was to replicate and extend well known effects of word length, frequency, and predictability with a standard set of eye movement measures within a coherent framework of analyses covering both data from all words as well as data from an a priori specified set of target words. Moreover, we hoped to shed new light on some less stable effects (such as the effect of word frequency on inspection probabilities and the effect of word length on first fixation duration). Finally, on another dimension of generalisability, we wanted to check the stability of all these effects across the adult life span.

METHOD

Participants

Thirty-three university students (*M*=21.9, *SD*=2.2, range: 19–28 years) and 32 older adults (*M*=69.9, *SD*=3.9, range: 65–83 years) participated in this study: Young and older adults showed the typical pattern of equivalence in Lehrl's (1977) multiple-choice measure of vocabulary (old: *M*=33.1, *SD*=1.2; young: *M*=32.8, *SD*=1.0) and large differences in Wechsler's (1964) digit symbol substitution (old: *M*=49.3, *SD*=10.0; young: *M*= 67.9, *SD*=8.1). They were all native speakers of German. Sessions lasted about one hour. Participants were paid an equivalent of 5 / hour or received credit in partial fulfillment of study requirements. Data from an additional 5 young and 10 older adults were excluded because of calibration problems during testing or because of low data quality (see *Analyses* below); their psychometric scores did not differ from the final samples.

Apparatus

Sentences were presented on the centre line of a 21-inch EYE-Q 650 monitor (832 x 632 resolution; frame rate 75 Hz; font: regular New Courier 12) controlled by an Apple Power Macintosh G3 computer. Participants were seated 60 cm in front of the monitor with the head positioned on a chin rest. Thus, letters subtended 0.35° of visual angle. Eye movements were recorded with an SR EyeLink System (SMI) with a sampling rate of 250 Hz and an eye position resolution of 20 sec-arc. Calibrated eye position was recorded accurately at the level of letters. Data were collected in two laboratories with identical equipment and setup.

The Potsdam Sentence Corpus

Word length. The Potsdam Sentence Corpus comprises 144 German sentences (1138 words). They were constructed with the goal to represent a large variety of grammatical structures around a set of target words (one or two per sentence; see below) that are uncorrelated in length and frequency. Sentence lengths range from 5 to 11 words with a mean of 7.9 words. Excluding the first word of each sentence which was not used in the analyses, frequencies of word lengths 2–13+ are: 54, 222, 134, 147, 129, 92, 72, 66, 20, 25, 16, and 17. (The category 13+ contains seven words of length 14–20.)

Printed frequency. CELEX Frequency Norms (Baayen, Piepenbrock, & van Rijn, 1993) are available for all 1138 words. Excluding the first word of each sentence, the corpus contains at least 76 words in each of five logarithmic frequency classes: class 1 (1–10 per million): 242 words; class 2 (11–100): 207 words; class 3 (101–1000): 242; class 4 (1001–10,000): 227; class 5 (10, 001max): 76 words. The CELEX corpus is based on approximately 5.4 million words.

Predictability. Predictability of words was collected in an independent norming study from 272 native speakers of German: 116 high-school students (age range: 17–19 years), 76 university students (age range: 19–38 years), and 80 older adults (age range: 66–80 years). Participants differed in the number of sentences they worked through: Twenty older adults generated predictions for all of the sentences; the other participants generated predictions for a quarter of the sentences. Collapsing the complete and partial protocols across participants, there were 83 complete predictability protocols, specifically 29 from highschool students, 19 from university students, and 35 from older adults. Completion of a complete protocol (144 sentences) lasted about 2.5 hours. Sentences were presented in a random order for each participant. Participants guessed the first word of the unknown original sentence and entered it via the keyboard. In return, the computer presented the first word of the original sentence on the screen. Responding to this, participants entered their guess for the second word and so on, until a period indicated the end of the sentence. Correct words stayed on the screen. There were no significant differences between the predictability distributions generated by the three samples. Excluding the first word of each sentence, the corpus contains at least 88 words in each of five logit-based predictability classes: class 1 (−2.553 to−1.5): 506 words; class 2 (−1.5 to −1.0): 111 words; class 3 (−1.0 to −0.5): 114 words; class 4 (−0.5 to 0): 88 words; class 5 (0 to 2.553): 175 words. Logits are defined as 0.5*ln(pred/(1− pred)); predictabilities of zero were replaced with 1/(2*83) and those of the five perfectly predicted words with (2*83–1)/(2*83), where 83 represents the number of complete predictability protocols (Cohen & Cohen, 1975). Upper boundaries of classes correspond to predictabilities of 0.0474, 0.1192, 0.2689, 0. 5, and 1.0. For a word with predictability 0.50, the odds of guessing are 1:1=1, and the log odds of guessing are ln(1)=0. Thus, words with predictability larger than 0.50 yield positive logits, those with predictabilities smaller than 0.50 negative logits. The logit transformation corrects for the dependency of mean probabilities *(p)* and associated standard deviations *(SD)* [i.e., *SD=p(1-p)*] by stretching the tail of the distribution.

Target words. Each sentence contained one target word selected from the CELEX database contributing to a 2x2x3 design with word class (noun vs. verb), printed frequency (high:#DXGT#50 occurrences/million vs. low: 1 to 4 occurrences/million), and word length (short: 3, 4 letters, medium: 5–7 letters, long: 8, 9 letters); there were 12 sentences in each cell of this design. The position of the target word ranged from being the second to the second word from the last word in the sentence; mean word position is 4.9. For a subset of 32 sentences, two directly adjacent target words (a verb-noun or noun-verb sequence) set up a four-factorial minidesign with the frequency of the second target word as a fourth orthogonal factor in addition to word class, frequency, and length of first word; the additional target word of a sentence was of the same length as the first one. There were two sentences contributing to each cell of the 2x2x2x2 design. This subset of 32 sentences is listed in the Appendix; target words and second target words are set in italics. Analyses of the embedded design with additional target words as well as analyses of languagespecific variables beyond frequency and predictability (e.g., word category, syntactic structure of sentence, spillover effects) are not reported here.

Procedure

Participants were calibrated with a standard nine-point grid for both eyes. They were instructed to read the sentence for comprehension and fixate a dot in the lower right corner of the monitor to signal the completion of a trial. After validation of calibration accuracy, a fixation point appeared on the left side of the centre line on the monitor. If the eye tracker identified a fixation on the fixation spot, a sentence was presented so that the midpoint between the beginning and the centre of the first word was positioned at the location of the fixation spot. Therefore, each sentence-initial word was read from a wordspecific optimal viewing position (e.g., O'Regan & Lévy-Schoen, 1987; Rayner, 1979). Sentences were shown until participants looked to the lower right corner of the screen. Then, the sentence was replaced by (1) an easy three-alternative multiple-choice question pertaining to the current sentence on 27% of the trials which the participant answered with a mouse click, (2) a fixation spot indicating the beginning of the next trial which participants then initiated by fixating the fixation point, or (3) a complete recalibration with the nine-point grid after 15 sentences each. In addition, the experimenter carried out an extra calibration if the tracker did not detect the eye at the initial fixation point within 2s.

Analyses

Eye movement data from reading the 144 sentences were screened for loss of measurement and blinks. Data from participants containing less than 100 sentences after data screening were excluded from the database. The remaining participants (33 young, 32 older adults) contributed on average 133 (young: 137, older adults: 127) sentences to the database. Data of sentences without problems were reduced to a fixation format after detecting saccades as rapid binocular eye movements with amplitudes of more than 0.5°, using a velocity-based detection algorithm originally developed for the analyses of microsaccades (Engbert & Kliegl, 2003). In a second level of data screening, words with first-fixation durations shorter than 30 ms or longer than 1 s and words with gaze durations or reading times longer than 1.5 s were removed from the database. This screening amounted to a total loss of 194 words (young: 101, older adults: 93; i.e., three words per participant). Finally, we removed the first word of each sentence. Overall an average of 92% of the words (relative to total number of corpus words minus 144 first words of sentences) remained in the analyses (young: 95%, older adults: 88%).

For each word we determined whether it was fixated once (P1), fixated more than once (P2), or skipped (P0) in first-pass reading. A word contributed to first pass reading statistics as long as the eyes had not moved past it. Note that the three probabilities sum up to 1.0. For first-pass fixations we determined the duration of first fixation on a word (if fixated at least once, FF), the duration of single fixations (if fixated exactly once, SF), and gaze durations (i.e., the sum of fixations if fixated at least once, GD). In addition to first-pass reading, we counted how often words served as the origin (RO) and the goal (RG) of a regression back to a previous word of the sentence. The durations of fixations following a regression were added to the gaze duration for this word and yielded a measure of total reading time (TT). Fixations after the first encounter of the last word (i.e., re-readings) were not included in the analyses.

Statistical analysis followed the procedure described by Lorch and Myers (1990, method 3) for multiple regressions with repeated measures. Specifically, for each participant we regressed each dependent variable mentioned above on the following independent variables: linear and quadratic effects of word length

(2–13+) centred for word length 7 (i.e., range: −5–6), logarithm (base 10) of CELEX frequency, and logit-transformed predictabilities (see above under Predictability). Subsequently, we checked whether means of unstandardised regression coefficients (across persons) were significantly different from zero and whether they differed between age groups.

RESULTS

Results on four inspection durations and five inspection probabilities are presented in two sections. First, we report the results based on multiple regression analyses of repeated measures of word length, frequency, and predictability for all corpus words. These analyses were carried out for all participants. Then, we compare the effects obtained for corpus words with target words representing an orthogonal design of word length and word frequency. The significance level for ANOVA effects was set at .001 in order to protect against isolated chance effects but we report effects that exhibit a meaningful pattern across measures for the .05 level. In general, as expected, first-fixation and single-fixation durations as well as gaze durations and total reading times exhibited very similar trends.

Corpus words

Means (standard deviations) of unstandardised regression coefficients are listed in Table 1 for the entire sample of 65 participants. Values reported under Constant represent the estimate for a seven-letter word with a printed frequency of zero and a 50% predictability. We present results for first-pass inspection probabilities (P0, P1, and P2+), for inspection durations (FF, SF, GD, TT) and for reinspection probabilities (RO, RG). The median R^2 of these multiple regressions was .07.

Inspection probabilities for first-pass reading. Figure 1 displays the main effects of word length, logarithmic word frequency, and logits of predictabilities for the probability of skipping a word (P0), fixating it once (P1), or fixating it two or more times (P2+). The lines are computed from regression equations with the unstandardised regression coefficients listed in Table 1. In general, results were consistent with expectations: Word skipping (P0) decreased with word length and increased with frequency and predictability, whereas the reverse pattern was observed for multiple fixations (P2). Symbols in Figure 1 represent means across words of different lengths as well as frequency and predictability classes (see Methods for definition) and are plotted at the means of their respective classes. There is good agreement between the probabilities computed from regression equations (lines) and the observed means (symbols). The probability of a single fixation (P1) followed an inverted U-shape over word length (i.e., no significant linear trend), $F(1, 63)=1.11$, $p=.296$, but as it is the complement of P0 and P2

TABLE 1

Means (standard deviations) of unstandardised regression coefficients based on 65 participants for corpus words

	Constant	*Length*	*L-sq*	*Freq*	*Pred*
FF	214 (36)	*0.4* (1.9)	0.3 (0.6)	−4.5 (5.0)	*1.2* (7.7)
SF	213 (37)	1.6 (2.7)	0.6 (0.7)	−4.1 (5.9)	−*0.5* (8.0)
GD	247 (50)	9.1 (5.4)	2.0 (1.0)	−8.1 (7.2)	−*3.5* (12.4)
TT	253 (51)	9.1 (5.2)	1.9 (1.0)	−9.5 (7.5)	−7.3 (13.3)
PO	7.1 (10.1)	−4.6 (1.5)	0.8 (0.4)	3.0 (3.0)	1.7 (2.8)
P1	76.2 (9.0)	*0.4* (3.1)	−1.4 (0.4)	−1.8 (2.6)	−*0.1* (2.7)
P2	16.4 (9.2)	4.2 (2.0)	0.5 (0.3)	−1.2 (1.4)	−1.7 (2.3)
RO	7.3 (5.9)	*0.0* (0.6)	*0.0* (0.2)	−0.6 (1.1)	−1.8 (1.6)
RG	*0.2* (3.0)	−1.0 (0.8)	0.1 (0.2)	*0.5* (1.4)	−2.5 (0.2)

Length=−5 to 6+ letters (i.e., centred on seven-letter word); L-sq=quadratic trend of length; Freq(uency) (in 1g10-units); Pred(ictability) (in log-odds units).
FF=first fixation duration, SF=single fixation duration, GD=gaze duration, TT=total reading time; P0, P1, P2=probability of zero (skipping), one, two+ fixations; RO, RG=probability of origin, goal of regressive eye movement (multiplied with 100 in table). Estimates are based on a mean of 913 words per participant. Italicised coefficients were *not* significant (99.9% confidence intervals).

(i.e., $p1=1-p0-p2$), it carries no independent information. Similarly, the effect of predictability was not significant for P1, $F(1, 63)=0.05$, $p=.829$. There were no significant age differences in inspection probabilities. Thus, despite substantial correlations among the predictors, there were statistically independent and reliable effects of word length, frequency, and predictability on inspection probabilities.

Inspection durations. Figure 2 displays the main effects of word length, logarithmic word frequency, and logits of predictabilities for first-fixation duration (FF), single-fixation duration (SF), gaze duration (GD), and total reading time (TT). In general, durations increased with word length and decreased with frequency but predictability was not significantly associated with first-pass inspection durations after controlling for length and frequency (see Table 1); FF: $F(1, 63)=1.64$, $p=.205$; SF: $F(1, 63)=0.27$, $p=.605$; GD: $F(1, 63)=5.0$, $p=.029$. Predictability was significant, however, once regressions to words were added in: TT: $F(1, 63)=19.3$, $p<.001$. The linear effect of word length was not significant for first fixation durations, $F(1, 63)=2.9$, $p= .093$. All other regression coefficients associated with length and frequency were significantly different from zero in the expected direction (all *F*#DXGT#12.8). Again, there was a very good agreement between estimates based on multiple regressions for individual participants (lines) and the observed mean durations computed at the level of words (symbols).

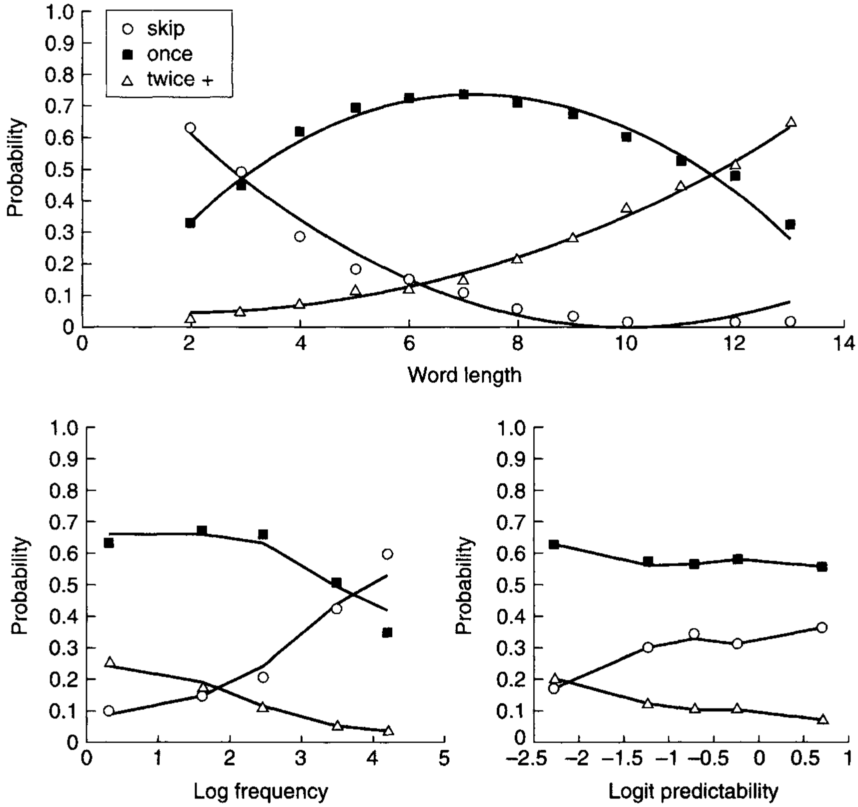

Figure 1. Observed and predicted first-pass inspection probabilities over word length (top), logarithmic word frequency (bottom left), and logit predictability (bottom right) based on all words of the corpus (except first words of sentences). Each curve was computed from the means of five regression coefficients (i.e., intercept, length, length-square, log frequency, and logit predictabilities) determined individually for 65 participants (see Table 1); they were not fitted directly to the observed means shown in symbols.

Age differences in inspection durations. Nominally, there were no age differences at the .001 level of significance for any of the regression coefficients but there was a very consistent pattern at the .01 level across the four inspection measures: Older adults read somewhat more slowly than young adults as reflected in larger constants: FF: 228vs. 201ms, SF: 227 vs. 200ms, GD: 265 vs. 230 ms, TT: 271 vs. 234ms, $F(1, 63)=10.4$, 10.4, 9.0, and 9.5 for the age differences, respectively. Also, they were more sensitive to word frequency: FF: −6.4 vs. −2.8, SF: −6.0 vs. −2.3, GD: −10.3 vs. −5.9, TT: −12.0 vs. −7.2, $F(1, 63)=9.6$, 7.0, 6.5, and 7.1 for the corresponding age differences. However, the frequency effects were significant within both groups.

Reinspection probabilities. In addition to first-pass reading measures we also examined regressive movements (see RG and RO in Table 1). Regressions

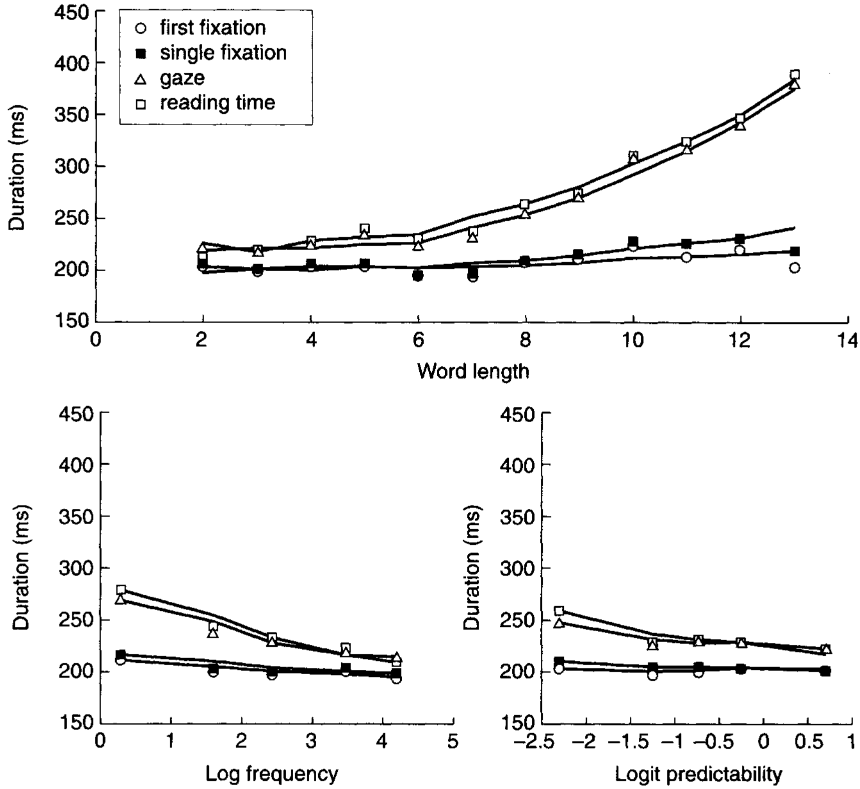

Figure 2. Observed and predicted inspection durations over word length (top), logarithmic word frequency (bottom left), and logit predictability (bottom right) based on all words of the corpus (except first words of sentences). Each curve was computed from the means of five regression coefficients (i.e., intercept, length, length-square, log frequency, and logit predictabilities) determined individually for 65 participants (see Table 1); they were not fitted directly to the observed means shown in symbols.

originated at words of low frequency and low predictability and landed on short words of low predictability (RO and RG in Table 1) (all *F*#DXGT#40). The absence of a frequency effect did not quite match our expectations given Vitu and McConkie's (2000) results of high regression probabilities to low-frequency skipped words. However, the fact that short words of low predictability were probable regression goals makes sense if these words were initially skipped due to their shortness but subsequently inspected due to their low predictability.

Sources of reinspection probabilities. For a better understanding of reinspection probabilities, we searched for regressive movements from word $n+1$ to a previously fixated word n (i.e., an average of 16 refixations per person contributed by 64 participants) and for regressive movements to a previously skipped word n (i.e., an average of 31 first fixations after regression per person contributed by 62 participants). Comparing lengths, frequencies, and predictabilities of word *n* with those of control words (i.e., averages of 528 and 196 cases per person without regressive movement to word *n,* respectively), we found refixated words to be shorter (5.5 vs. 5.8 letters), paired $t(63)=-3.0$, $p=.004$, $N=64$, of lower log frequency (1.8 vs. 2.0), $t(63)=-2.0$, $p=.047$, and of lower logit predictability (−2.0 vs. −1.3), t(63)=−14.2, $p<.001$, $N=64$. Previously skipped words targeted by a regressive movement from the next word differed only in logit predictability (−1.3 vs. −0.8), paired $t(61)=-14.7$, $p<.001$, $N=62$; they were of similar length (3.7 vs. 3.8 letters), $t(61)=-1.1$, $p=.281$, and log frequency (2.92 vs. 2.90), $t(61)=0.293$, $p=.770$.

Age differences in reinspection probabilities. Consistent with results reported for inspection durations, higher reinspection probabilities also contributed to a slower reading of older adults (means of 11% vs. 7% of all fixations), $t(63)=3.5$, p=.018. However, old and young adults did not differ in the length, frequency, and predictability effects reported in the last paragraph (all $t< 1.5$). Thus, the strong predictability and relatively weak frequency and lengths effects replicated across age groups.

Target words

One well-known problem associated with analyses of corpus words is the correlation among predictors. Word length and log frequency of 994 corpus words (i.e., excluding the first word of each sentence) correlated −.64, length and logit predictability −.29, and log frequency and logit predictability .38. For a subset of 144 target words (one word per sentence), the corresponding correlations were −.01, −.19, and .31. Therefore, for this subset we could assess the independent contributions of length and frequency. Means (standard deviations) of unstandardised regression coefficients (averaged across 65 participants) for these analyses are summarised in Table 2. The median R^2 of these multiple regressions was .07.

Inspection probabilities. There were no effects of length (linear trends), frequency, or predictability on single-fixation probabilities (all *F*s<1.9). All other coefficients were in the expected direction and were highly significant (all *F*s#DXGT#7.3). A comparison between corpus and target word analyses (i.e., Table 1 vs. Table 2) revealed mainly a reliable effect of frequency on single-fixation probability of corpus words and a stronger effect of frequency on multiple fixation probability of target words. The overall pattern is quite similar.

Age differences in inspection probabilities. Young and old adults differed in the measure that showed susceptibility to the predictability of words: High predictability reduced young adults' single-fixation probability but increased that of old adults (young: −.023, old: .042), $F(1, 63)=23.2$, $p<.001$; high predictability increased only young adults' probability of word skipping (young: .036, old:−.001), $F(1, 63)=7.9$, $p=.007$; finally, high predictability decreased old adults'

TABLE 2

Means (standard deviations) of unstandardised regression coefficients based on 65 participants for target words

	Constant	*Length*	*L-sq*	*Freq*	*Pred*
FF	207 (36)	*1.5* (5.0)	*0.3* (1.9)	−5.4 (7.4)	*−3.2* (10.0)
SF	210 (38)	3.3 (6.0)	*0.5* (2.1)	−6.3 (8.1)	−5.3 (10.9)
GD	241 (49)	8.5 (9.7)	*0.5* (3.1)	−11.8 (11.4)	−10.3 (15.0)
TT	245 (48)	7.4 (10.4)	*−0.1* (3.4)	−14.5 (13.5)	−17.5 (17.4)
P0	9.1 (13.5)	−3.4 (2.6)	0.8 (0.9)	2.3 (2.8)	*1.8* (5.5)
P1	74.1 (17.2)	*−0.6* (5.6)	−1.0 (1.3)	*0.7* (3.8)	*0.9* (6.3)
P2	17.0 (12.2)	4.1 (4.2)	*0.2* (1.1)	−2.5 (2.9)	−2.6 (4.4)
RO	12.5 (14.7)	*−0.6* (1.8)	*−0.3* (0.8)	*−0.8* (2.9)	*−0.0* (3.9)
RG	0.4 (7.7)	−1.0 (1.9)	*−0.1* (0.6)	*−0.7* (2.5)	−3.7 (3.5)

Length=–5 to 6+ letters (i.e., centred on seven-letter word); L-sq=quadratic trend of length; Freq(uency) in 1g10-units); Pred(ictability) (in log-odds units).
FF=first fixation duration, SF=single fixation duration, GD=gaze duration, TT=total reading time; P0, P1, P2=probability of zero (skipping), one, two+ fixations; RO (N=63), RG (N=63)= probability of origin, goal of regressive eye movement (multiplied with 100 in table).
Estimates are based on a mean of 135 words per participant. Italicised coefficients were *not* significant (99.9% confidence intervals).

probability of multiple fixations (young: −.012, old: −.042, $F(1, 63) = 9.0$, p=. 004. Thus, for both groups predictability allowed faster reading but they differed in how they realised the speed-up.

Inspection durations. With the exception of (1) the quadratic trends of word length, (2) the linear trend of word length for FF, and (3) the predictability for FF, regression coefficients for the prediction of inspection durations were significant at the .001 level (all *F*s#*DXGT*#15.0). (The two nonsignificant effects associated with FF were in the expected direction at the .05 level.) A comparison between the corpus and target word analyses (i.e., Table 1 vs. Table 2) reveals reliable effects of predictability on SF and GD for target words only. The restriction in range of word length for target words (i.e., 3–9 letters) eliminated quadratic trends of word length observed for corpus words. Finally, there were no age differences associated with inspection durations of target words.

Reinspection probabilities. There were reliable effects of linear word length and predictability on the probability of a word to be selected as the goal of a regressive movement (all *F*s#DXGT#19.9). (See Table 2 for corresponding unstandardised regression coefficients in lines RG and RO.) As in the corpus analyses, short words and low predictable words had a higher probability of being selected as a goal; different from the corpus analyses, however, low frequency was also marginally associated with a high selection probability for a regressive movement (RG in Table 2), $F(1, 62)$=4.3, p=.042. Thus, for target words there is a weak effect of low word frequency and predictability on reinspection probability compatible with word identification problems as their source (Vitu & McConkie, 2000). There were no reliable frequency or predictability effects associated with the launch site of regressive movements (RO in Table 2).

DISCUSSION

We replicated and extended well-known effects of word length, frequency, and predictability for a standard set of eye movement measures within a coherent framework of repeated-measures multiple regression analyses covering data from corpus words as well as data from an a priori specified set of target words that were uncorrelated in length and frequency. We summarise our results in relation to first-pass and second-pass reading and the generalisability of effects from target to corpus words. Then we address the quite remarkable stability of these results across the adult life span, compare effect sizes associated with three predictors, and "defend" the selection of predictors for the present set of multiple regressions. Finally, we argue that the set of results presented here, extended with other features of the text material and related key findings of eye movement research, will serve a useful purpose in constraining computational models of eye movement control in reading.

First-pass reading

Inspection probabilities. First-pass reading was influenced by word length, frequency, and predictability of words. The reliability of effects could be established for corpus and target words. Word length decreased the word skipping probability and increased multiple-fixation probability as in all other previous eye movement research. Also, consistent with previous research (e.g., meta-analysis by Brysbaert & Vitu, 1998), word frequency reliably increased skipping probability. Finally, predictability increased word skipping and decreased the associated multiple-fixation probability for corpus and target words.

Inspection durations. The three measures of first-pass inspection duration (FF, SF, GD) increased with word length in corpus and target words, in agreement with most previous research (see introduction). Quadratic trends of word length were reliable only for corpus words; probably this difference reflects the restriction in word length for target words. There were statistically reliable effects of word frequency on these measures in both sets of analyses. Finally, first-pass reading effects of predictability were observed only for single fixation and gaze durations of target words.

Comparison of effects for corpus and target words. Can we generalise results from a select set of target words (i.e., the experimental control approach) to an unselected set of corpus words (i.e., the statistical control approach)? Given the similarity of coefficients in Tables 1 and 2, our answer is an almost unqualified "yes". For both corpus and target words, we obtained reliable effects of word length and word frequency on inspection probabilities (P0, P2+) as well as inspection durations. Also the effects of word predictability on inspection probabilities were very similar for corpus and target words, even in terms of the size of regression coefficients. Aside from the quadratic effect of word length, the main difference between analyses of corpus and target words related to effects of predictability on inspection durations that were reliable only for single-fixation durations and gaze durations of target words. Thus, counter to expectations, the main benefit of an a priori selection of target words was not the dissociation of word length and word frequency but the reliability of word predictability in first-pass reading. Regression coefficients for frequency, however, were consistently larger in the analyses of target compared to corpus words.

Second-pass reading and predictability effects

Second-pass reading effects were primarily linked to the predictability of words. This is quite plausible if we assume that words of low predictability set up small-scale garden-path effects that are known to trigger reanalysis and regressive movements (e.g., Rayner, 1998, for a review). We observed a much larger effect of predictability on total reading time than on gaze duration (–7.3 vs.–3.5 for corpus words and −17.5 vs.−10.3 for target words, respectively). These results are consistent with the expectation that low-predictable words were a primary target of regressive movements because differences between gaze duration and total reading time reflect the time spent in refixations. Vitu and McConkie (2000) reported their results for low-frequency words. In our data, after statistical control of word length and predictability, the frequency effect was not significant but it showed a trend in the expected direction for target words.

We attempted to extend the set of dependent measures with reinspection probabilities after first-pass reading and distinguished between the probabilities of a word for serving as the origin or the goal of a regressive eye movement. For the goal probability we had expected a tendency towards low-frequency and low-predictable words (Vitu & McConkie, 2000). The strongest effect was associated with low predictability, which showed the expected effect both for corpus and target words. Moreover, short words were more likely to be selected as a goal for a regressive movement in both analyses. Obviously, one may easily and inadvertently skip a short and low-predictable word. As a first attempt to look at the dynamics of this situation, we checked predictability, length, and frequency effects for regressions to the immediately preceding word, given that this word was fixated or skipped prior to the regression. Refixated words were less predictable, shorter, and of lower frequency than a set of control words; effect sizes were largest for predictability and weakest for frequency. In contrast, previously skipped words differed only in (low) predictability from their control words.

Finally and in general, regressions originated from longer, less frequent, and less predictable words relative to the means of all corpus words.

Adult age differences

Easy reading yielded only minor effects of age. Older adults read somewhat slower than young adults as reflected in higher intercepts for inspection durations compatible with earlier reports of age-related slowing of saccadic latency (Abel, Troost, & Dell'Osso, 1983) and age-related slower reading (Solan, Feldman, & Tujak, 1995). Older adults were also more likely to reinspect an earlier word than young adults, again in agreement with results reported by Solan et al. (1995).

Aside from this general age-related speed effect, there were a few specific age differences related to effects of word frequency and predictability: Older adults responded earlier (i.e., in first and single fixation durations) and more consistently to word frequency. This second age difference related to predict ability effects on inspection probabilities of target words: Young adults increased skipping probability (P0) with word predictability, whereas old adults reduced the probability of multiple fixations (P2+) for such words more than young adults. Obviously, both effects reflect a utilisation of predictability for an increase of reading speed.

Overall the similarities in the reading profiles, especially the similarity of first-pass inspection probabilities, of the two samples of adult readers, which differed an average of 47 years in age, are much more impressive than the differences between them. Obviously, it remains to be seen how age affects eye movement control when difficult sentences must be read, when difficult questions need to be answered, or when working memory load is manipulated simultaneously. The similarity in profile of young and older adults under the easy-reading conditions of the present experiment could serve as a very useful baseline against which specific effects of adult age (e.g., deficits suspected in executive control processes; Mayr, Spieler, & Kliegl, 2001) could manifest themselves.

Effect sizes

Figures 1 and 2 provide direct information about absolute effect sizes associated with word length, frequency, and predictability in units of the dependent variables relative to the range of predictor values. Obviously, word length affects inspection durations and probabilities much more than frequency and predict ability. For example, for gaze durations the difference between a 13-letter and a 2-letter word was 373ms–252ms =121ms (for log frequency=0 and logit predictability=0), whereas the difference between words of the lowest and highest frequency was 36ms (for word length=7 and logit predictability=0) and the difference between words of zero and perfect predictability was 18 ms (for word length=7 and log frequency=0). In addition, unstandardised regression coefficients can be compared directly to this end. For example, from Tables 1 and 2 it is quite clear that the effects of frequency and predictability were about twice as large for aggregated measures (i.e., gaze duration, total reading time) than for first-fixation and single-fixation durations.[3]

We want to point out a limitation of the present set of analyses. We carried out separate multiple regression analyses for several measures of inspection probabilities and inspection durations (see Figures 1 and 2). Evidently, these measures are highly redundant. Most blatantly, the three inspection probabilities sum to one. There is also an obvious dependency between the probability of multiple fixations (P2+) and gaze duration. Finally, all inspection measures are correlated given their cumulative operational definitions—single fixations are a subset of first fixation durations, both are part of gaze duration, and gaze durations are part of total reading time. Strictly speaking, such technical dependencies prevent statistical tests of differences between the measures. We opted for the present analysis framework because it represents most of the measures used in the research community. In perspective, a coherent framework of nonredundant, independently defined dependent measures would be highly desirable.

Higher-order (multiplicative) predictor terms

Aside from linear terms of word length, log word frequency, and logit predictability, we included a quadratic term for word length as an additional higher order predictor in the regression equation. The need for this predictor was motivated by the visual inspection of word-length functions and clearly needed for an adequate account of the observed means (see Figures 1 and 2). In additional analyses we also included quadratic terms for the other two predictors as well as terms coding multiplicative interactions between them. With an average of 913 words (i.e., cases) in the multiple regressions, all of these terms were significant for at least one of the dependent variables but they did not lead to a qualitative improvement of the reproduction of the observed means as shown in Figures 1 and 2; mean incremental R^2s were always smaller than 1%. In addition, the collinearity between linear and higher order predictor terms greatly reduced the interpretability of associated unstandardised regression coefficients given in Tables 1 and 2. Thus, although based on a qualitative rather than a quantitative judgement of goodness of fit, we restricted the report of results to the four-predictor model.

There is a second argument for a small set of predictors. We did not expect major theoretical advances in our understanding of eye movement control during reading from fitting higher order polynomial regression models because, at this

[3] We restricted our comparison of effect sizes to unstandardised regression coefficients; for arguments against the interpretation of unique amounts of variance see Duncan (1975, pp. 63–66).

point in time, we simply have no good theoretical reasons to postulate specific quadratic or multiplicative effects. If these higher order predictors are significant, they mainly refine the description of our particular set of corpus words. Rather, we expect theoretical advances from the further development of inherently nonlinear computational models (see next section).

Finally, although the regression model comprised only linear and quadratic effects of word length as well as linear effects of word frequency and predict ability, the regression lines in Figures 1 and 2 exhibit trends that are suggestive of higher order terms. For example, word-predictability curves for skipping probabilities are suggestive of a cubic term (i.e., they contain nonmonotonic segments), which actually traces the observed pattern of means. These effects were due to nonlinear relations between the predictors (e.g., words of low and high predictability were of lower frequency than words of medium predict ability). Such nonlinear relations among predictors can generate disordinal trends.

Implications for computational models of eye movement control

Descriptive accounts of eye movement data from reading are needed as simulation targets for computational models of eye movement control. The data reported here are to serve as a complement to the corpus of Schilling et al. (1998, supplemented by Reichle et al., 1998). The present corpus represents a more comprehensive benchmark for the evaluation of computational models of eye movement control during reading because we report eye movements not only as a function of word frequency but also as a function of word length and as a function of word predictability. Moreover, there is information about independent effects of word length and word frequency for a subset of target words. In principle, the data pattern reported for this experiment is within reach of current theoretical proposals (see Reichle et al., in press, for a comparison of models).

There are a few challenges for future computational models. For example, it might be very difficult to reproduce the different pattern of word-frequency and word-predictability effects on inspection probabilities and inspection durations during first-pass reading of corpus and target words (see above). At the same time, to avoid overfitting a specific set of corpus words, it may be advantageous to fit models to length, frequency, and predictability functions based on a limited number of unstandardised regression coefficients rather than to means derived from post-hoc categorisations of frequency and predictability classes. Moreover, these functions should be based on a less redundant and technically dependent set of measures than the suite of inspection probabilities and durations presented here.

The present set of data represents only a small segment of eye movement measures distilled from the reading protocols for a select set of word characteristics. There are other features that can be determined for the present corpus of words, among them neighbourhood effects (Grainger, O'Regan, Jacobs, & Sequi, 1992) as well as initial trigram frequency and informativeness (Kennedy, Pynte, & Ducrot, 2002). Some of them appear to be relevant for a better understanding of dynamics of eye movement control and hypotheses about distributed lexical processing (e.g., parafoveal preview and spillover effects). To this end we also need to integrate knowledge about the distributions of landing sites within words, that is the observations that landing sites are normally distributed with a dependency of the mean on the launch distance of the last saccade (McConkie, Kerr, Reddix, & Zola, 1988; Radach & McConkie, 1998) and the associated effects on fixation durations (Vitu, McConkie, Kerr, & O'Regan, 2001). According to Rayner et al. (2001) these eye-position effects are dissociated from lexical processing, that is they are not influenced by word frequency or predictability. In summary, eye movement research has produced a rich and reliable set of results since the 1980s. It is time to document that harvest in a single database because we will need such a corpus for constraining the parameter space of extant and future computational models of eye guidance during reading.

PrEview proof published online September 2003

REFERENCES

Abel, L.A., Troost, B.T., & Dell'Osso, L.F. (1983). The effects of age on normal saccade characteristics and their variability. *Vision Research, 23,* 33–37.

Baayen, R.H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (Release 1)* [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Binder, K.S., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1162–1172.

Blanchard, H.E. (1985). A comparison of some processing measures based on eye movements. *Acta Psychologica, 58,* 1–15.

Brysbaert, M, & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–147). Amsterdam: Elsevier.

Calvo, M.G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *Spanish Journal of Psychology, 5,* 66–77.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Duncan, O.D. (1975). *Introduction to structural equation models.* New York: Academic Press.

Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics, 85,* 77–87.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43,* 1035–1045.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636.

Gautier, V., O'Regan, J.K., & Le Gargasson, J.F. (2000). "The-skipping" revisited in French: Programming saccades to skip the article "les". *Vision Research, 40,* 2517–2531.

Grainger, J., O'Regan, K., Jacobs, A.M., & Sequi, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception and Psychophysics, 51,* 49–56.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47,* 201–222.

Hyönä', J., & Olson, R.K. (1995). Eye movement patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1430–1440.

Inhoff, A.W., Radach, R., Eiter, B.M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *Quarterly Journal of Experimental Psychology, 56A,* 803–827.

Inhoff, A.W., & Eiter, B. (2003). Knowledge of word length does not constrain word identification. *Psychological Research, 67,* 1–9.

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

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Kennison, S. M, & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 68–81.

Kliegl, R., Olson, R.K., & Davidson, B.J. (1982). Regression analyses as a tool for studying reading processes: Comments on Just and Carpenter's eye fixation theory. *Memory and Cognition, 10,* 287–296.

Kliegl, R., Olson, R.K., & Davidson, B.J. (1983). On problems of unconfounding perceptual and language processes. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 333–343). New York: Academic Press.

Lehrl, S. (1977). *Mehrfach-Wortschatz-Test Form B* [Multiple-choice vocabulary test]. Erlangen, Germany: Perimed.

Lorch, R.F., & Myers, J.L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Language, Memory, and Cognition, 16,* 149–157.

Mayr, U., Spieler, D., & Kliegl, R. (Eds.). (2001). Ageing and executive control. *European Journal of Cognitive Psychology, 13* (Whole no. 1/2).

McConkie, G.W., Kerr, P.W., Reddix, M.D., & Zola, D. (1988). Eye movement control during reading: 1 The location of initial eye fixations on words. *Vision Research, 28,* 1107–1118.

O'Regan, J.K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. *Attention and Performance, 12,* 363–383.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 165–192). Oxford, UK: Elsevier.

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

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

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

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41,* 943–954.

Rayner, K., Sereno, S.C., & Raney, G.E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1188–1200.

Rayner, K., & Well, A.D. (1996). Effects of contextual constraints on eye movements in reading: A further examination. *Psychonomic Bulletin and Review, 3,* 504–509.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39,* 4403–4411.

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

Schilling, H.E. H., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26,* 1270–1281.

Solan, H.A., Feldman, J., & Tujak, L. (1995). Developing visual and reading efficiency in older adults. *Optometry and Vision Science, 72,* 139–145.

Vitu, F., & McConkie, G.W. (2000). Regressive saccades and word perception in adult reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 301–326). Amsterdam: Elsevier.

Vitu, F., McConkie, G.W., Kerr, P.W., & O'Regan, J.K. (2001). Fixation location effects on fixation duration during reading: An inverted optimal viewing position effect. *Vision Research, 41,* 3513–3533.

Wechsler, D. (1964). *Der Hamburg Wechsler Intelligenztest für Erwachsene (HAWIE).* Bern, Switzerland: Huber.

APPENDIX

1 Den *Ton gab* der Künstler seinem Gehilfen.

2 Der *Hof lag* weit außerhalb des eigentlichen Dorfes.

3 Die Wanderer sahen Rehe auf einer Lichtung im *Wald äsen.*

4 Den *Kopf hieb* man früher nur Mördern und Verrätern ab.

5 Vorne am *Bug sah* man eine prächtige Galionsfigur.

6 Sogar aus *Raps läßt* sich Kraftstoff herstellen.

7 Torsten beobachtete gestern eine Maus, die *Efeu fraß.*

8 Der schüchterne kleine *Gnom mied* die Nähe der Elfen.

9 Claudia hatte ihr Fahrrad auf der *Straße stehen* lassen.
10 Wir hätten schon vor einer *Stunde wissen* sollen, ob ihr kommt.
11 Die Eltern konnten ihre Kinder im *Garten raufen* hören.
12 Er hätte nicht auch noch am *Telefon nörgeln* sollen.
13 Wegen ihrer Diät hatte die Gräfin leider keine *Auster nehmen* dürfen.
14 Die meisten *Hamster bleiben* bei Tag in ihrem Häuschen.
15 Man sollte nie Geschirr mit einem dreckigen *Lappen spülen* müssen.
16 Man kann *Spargel dämpfen* oder in viel Wasser kochen.
17 Manchmal *sagen Opfer* vor Gericht nicht die volle Wahrheit.
18 Die meisten Befragten *hören Musik* zur Entspannung.
19 Kinder *essen Quark* am Liebsten mit Früchten.
20 Bei Wölfen *leben Rudel* nicht verwandter Tiere in getrennten Revieren.
21 Die Frauen in den Andendörfern *weben Stoff noch* auf traditionellen Webstühlen.
22 Die Platzwarte *ebnen Stück* für Stück den Rasen nach dem Spiel.
23 In den Fässern *gären Beize* und Lauge.
24 Die Förster *küren Ahorn* zum Baum des Jahres.
25 Wolfgangs Töchter *studieren Literatur* und Maschinenbau.
26 In der Klosterschule *herrschen Schwester* Agathe und Schwester Maria.
27 Hier *scheinen Klempner* am Werk zu sein.
28 Im Aussehen *gleichen Bratsche* und Geige sich sehr.
29 Angeblich *flunkern Künstler* oft bezüglich ihrer Einnahmen.
30 Manchmal *krakeelen Politiker* genauso wie Demonstranten.
31 Die Armen *plündern Speicher* und Vorratskeller der reichen Bauern.
32 Die Richter der Landwirtschaftsschau *prämieren Rhabarber* und Mangold.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 285–311

Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap

Sally Andrews

University of Sydney, Australia

Brett Miller and Keith Rayner

University of Massachusetts, Amherst, USA

In two experiments, readers' eye movements were monitored as they read sentences containing compound words. In Experiment 1, the frequency of the first and second morpheme was manipulated in compound words of low whole word frequency. Experiment 2 compared pairs of low frequency compounds with high and low frequency first morphemes but identical second morphemes that were embedded in the same sentence frames. The results showed significant effects of the frequency of both morphemes on gaze duration and total fixation time on the compound words. Regression analyses revealed an influence of whole word frequency on the same measures. The results suggest that morphemic constituents of compound words are activated in the course of retrieving the representation of the whole compound word. The fact that the frequency effects were not confined to fixations on the morphemic constituents themselves implies that saccadic eye movements are implemented before morphemic retrieval has been completed. The results highlight the importance of developing more precise models of the perceptual processes underlying reading and how they interact with the processes involved in lexical retrieval and comprehension.

Visual word recognition is a very active area of current research in experimental psychology. Research on the processes involved in lexical retrieval has generated a plethora of theoretical accounts couched within a variety of conceptual frameworks. Lexical retrieval has also been the focus of developments in computat ional modelling. Models of visual word recognition reflecting conceptually distinct frameworks have been made completely explicit through computational implementation (e.g., Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger & Jacobs, 1996; Plaut, McClelland, Seidenberg, & Patterson, 1996). It has therefore become possible to directly compare the predictions of different theoretical accounts with measures of human performance (e.g., Andrews & Scarratt, 1998; Coltheart et al., 2001; Kello & Plaut, 2000).

The theoretical precision implied by computational models of lexical retrieval can, however, be challenged because of the relatively impoverished body of evidence that has provided the basis for their development. There are two major limitations associated with the data on which many current models of word identification have been based.

First, virtually all computationally implemented models are restricted to monosyllabic words and research on lexical retrieval has been dominated by investigations of monosyllables. Whatever their level of precision, models that only simulate processing of monosyllabic words remain an inadequate account of visual word identification until it is established how they can be generalised to the processing of multisyllabic words. It has been estimated that over 80% of the words in spoken discourse have more than one syllable (Gimson, 1980) and the proportion is probably higher in written text (Henderson, 1982). Systematic evidence about the processes underlying recognition of multisyllabic words is necessary to provide constraints that can guide the elaboration of theoretical and computational models to the full vocabulary of written English.

The second limitation of both current computational models and of conceptual models developed to account for processing of multisyllabic words is that they have been based primarily on investigations of single words presented in tasks requiring speeded classification or naming. As well as raising questions about the generalisability of such data to natural reading, the models fail to take account of perceptual characteristics of the reading process such as the multiple fixations directed towards many monosyllabic words.

Correspondence should be addressed to S. Andrews, School of Psychology, University of Sydney, Griffith Taylor Building (A19), Sydney, 2006, Australia. Email: sallya@psych.usyd.edu.auThis research originated when the first author was on sabbatical leave at the University of Massachusetts. The research was supported by Grants HD17246 and HD26765 from the National Institute of Health, and by Grant 79917228 from the Australian Research Council. The third author was also supported by a Research Scientist Award from the National Institute of Mental Health (MH01255). We thank Barbara Juhasz, Wayne Murray, Sandy Pollatsek, Matt Starr, and two anonymous reviewers for their helpful comments.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000123

The present research addresses these two broad limitations of the existing research on lexical retrieval by recording eye movements to multimorphemic compound words during a sentence reading task. To provide the theoretical and empirical background to the experiments, the sections below first overview theories and research on lexical retrieval for morphologically complex words and then consider relevant eye movement research.

LEXICAL RETRIEVAL FOR MORPHOLOGICALLY COMPLEX WORDS

Two issues are central to understanding processing of multimorphemic words. The first concerns the form of lexical representation of these items. Most long words are morphologically complex. A major question driving psycho linguistically oriented research on multimorphemic words concerns whether the whole word forms of morphologically complex words are fully listed in the lexicon (e.g., Fowler, Napps, & Feldman, 1985); or whether morphologically complex words are represented in decomposed form—the root form along with a listing of its permissible affixes (e.g., Taft & Forster, 1976). The second central issue follows from the first. If multimorphemic words are represented as root and affix units, how are these morphemes extracted from the complete word string to provide the basis for access?

Most current approaches to the representational question adopt some form of hybrid account that assumes representations of both whole forms and morphemic constituents. Debate now centres around the organisation of the different forms of lexical representation. Taft (1994) proposed a hierarchical architecture in which common prefixes and suffixes are represented at a lower level in the hierarchy than lexical roots. Other models assume that whole word and decomposed morphemes constitute two independent routes to meaning (e.g., Baayen, Dijkstra, & Schreuder, 1997; Caramazza, Laudanna, & Romani, 1988; Niemi, Laine, & Tuominen, 1994). Such dual route accounts differ in their assumptions about what determines the selection of processing routes. For example, Caramazza et al.'s (1988) Augmented Addressed Morphology model assumes that the constituent route is only used when retrieval of whole word forms is unsuccessful as is the case for nonwords, while other frameworks assume that the two forms of knowledge are activated in parallel (Baayen et al., 1997; Niemi et al., 1994).

Early approaches to the second question about how morphemic segments are parsed from whole words adopted a "segmentation-*then*-recognition" account of morphological parsing (Andrews & Davis, 1999) which assumes that a segmentation procedure is applied to decompose words into access units before lexical retrieval is attempted. The most precise specification of such a procedure was provided by Taft (1979b, 1985; Taft & Forster, 1975, 1976), who argued that words are located in lexical memory via a unit corresponding roughly to an orthographically defined first syllable. Although vestiges of this prelexical decomposition account remain in more recent formulations (e.g., Taft, 1991), most recent theories rely on activation-based accounts of morphological influences on word recognition which assume a "segmentation-*through*-recognition" process (Andrews & Davis, 1999). Morphological segments are not prelexically parsed but rather extracted by virtue of their match with lexical representations. Such views are usually expressed within some form of interactive activation (IA) framework (e.g., McClelland & Rumelhart, 1981; Taft, 1991), which assumes that multimorphemic words simultaneously activate representations of both the whole word and those constituent morphemes that are represented in lexical memory (e.g., Fowler et al., 1985; Taft, 1991, 1994). Some versions of this framework assume that only whole words are lexically represented, and therefore attribute morphological influences to patterns of orthographic and semantic cross-activation (Henderson, 1985; Marslen-Wilson, Tyler, Waksler, & Older, 1994). Others assume that both bound and unbound root morphemes are explicitly represented along with representations of their affixed variants (e.g., Caramazza et al., 1988; Dell, 1986; Taft, 1991, 1994).

Despite differences between the precise representational and processing assumptions of the various IA accounts of morphological influences on word recognition, they share the common assumption that these effects reflect interactions between the activation of multiple lexical representations. At a general level, they assume that subword constituents that are lexically represented, and of higher frequency than the whole word, achieve high levels of activation early in processing and contribute to the speed with which the detector for the complete word reaches threshold. In principle, this provides an appealing and general account of a variety of results demonstrating that morphological constituents contribute to identification of the whole word forms that contain them. However, many of the precise details of the nature of the activation process and the units it operates on remain unclear (e.g., Andrews & Davis, 1999).

The present research uses compound word stimuli to provide further evidence relevant to evaluating and refining the general IA account. The broad goal is to determine whether lexical constituents of compound words are retrieved when compound words are presented in sentence contexts under conditions more closely approximating natural reading. The use of eye movement measures provides more precise information about the timecourse of the processes underlying activation of constituents and whole word forms than is usually available from behavioural methods.

Compound words provide a stronger test of the validity and generality of the segmentation-through-recognition process that is central to IA frameworks than do affixed forms. Affixed words contain a limited set of very frequent bound morphemes (e.g., pre-, re-, -ed, -ing) that, in English, occur in predictable word initial or word-final positions. Evidence for extraction of such constituents might reflect an affix-stripping procedure that is implemented either prelexically or through activation of a special purpose "affix dictionary". Compound words consist of two unbound morphemes and contain a variety of different morphemic constituents. They therefore provide stronger evidence of a general segmentation-through-recognition strategy that is available for all multimorphemic words.

Previous investigations of compound words using standard word identification paradigms such as the lexical decision and naming tasks have provided evidence consistent with the segmentation-through-recognition account although a number of contradictions and ambiguities remain. One issue concerns the relative influence of first and second compound constituents. Taft and Forster's (1976) early investigations of such stimuli led them to conclude that only the first constituent of the compound contributes to accessing the whole word form because nonword classification time was affected by the lexical status of the first but not the second constituent of the compound (e.g., response times to *toast pull* and *spellcung* were equal, and longer than those *to flurbpair* and *thrimnade*) and compound words with high frequency first syllables (e.g., *headstand*) were classified more quickly than those with low frequency first syllables (e.g., *loincloth*). However, subsequent evidence has challenged this claim. Lima and Pollatsek (1983) found effects of the frequency of the second constituent of a nonword compound that were significant over subjects but not items, and Andrews' (1986) investigation of compound words showed that the frequency of *both* first and second constituents were correlated with reaction time, suggesting that they play an equivalent role in the access procedure. Sandra (1990) also found equivalent semantic priming of the first and second constituents of Dutch compound words but the effects were restricted to semantically transparent compounds (e.g., priming occurred for *loose—tightrope* but not for *bread—butterfly*).

A more general problem with all of the results reviewed is that they have been obtained in word judgement tasks that require task-specific decision processes that may, themselves, be sensitive to morphological complexity. Decomposition effects in the lexical decision task, one of the most frequently used paradigms, might arise from postlexical decision processes invoked to deal with the specific decision requirements of the task (Balota & Chumbley, 1984) and the stimulus context in which they are presented (Andrews, 1986). For example, in the same way that De Groot (1984) argued that the semantic relationship between prime and target items might bias towards classifying the target as a word, the association between the elements of compound words might contribute to lexical classification independently of retrieval of the complete compound.

EYE MOVEMENT INVESTIGATIONS OF MORPHOLOGICAL EFFECTS ON LEXICAL RETRIEVAL

Measurements of eye movements during the reading of sentences or text provide a means of assessing sensitivity to the lexical constituents of compound words in a more naturalistic reading situation that does not draw participants' attention to morphological structure or encourage a decision strategy based on morphological decomposition.

Indirect evidence about the segmentation of constituents during normal reading is provided by comparisons of eye fixation measures on compound items with matched pseudomorphemic words (e.g., *cowboy* vs. *napkin*) in a sentence reading task. However, the available data are contradictory. Inhoff (1989) found no differences between gaze durations (the sum of all fixations on a word prior to moving to another word) or amount of parafoveal preview benefit for the two item classes. However, Inhoff, Briihl, and Schwartz's (1996) investigation of longer compounds and matched suffixed and monomorphemic words (e.g., *blackberry* vs. *sainthood* vs. *arthritis*) in a sentence reading task revealed that first fixation durations—but not gaze durations—were longer for compound words than for the other word types. This finding was attributed to differences between the morphological structure of suffixed and compound words. In contrast with suffixed words, the meaning of English compound words is usually conveyed primarily by their second constituent. Inhoff et al. argued that the "atypical distribution of meaning-defining information in compound words… could hamper the derivation of conventional word meanings, accounting for the increased first-fixation durations", but that this cost might be "offset by savings during target re-fixations" accounting for the lack of differences between conditions in gaze duration (p. 474). These results are generally consistent with the view that the constituents of morphologically complex words influence sentence reading, but the complex form of the effects makes it difficult to determine their locus.

A recent investigation of suffixed words using eye movement data also produced ambiguous evidence. Niswander, Pollatsek, and Rayner (2000) presented inflected and derived words that independently manipulated root frequency and whole word frequency in a sentence reading task. Derived words showed effects of root morpheme frequency that occurred earlier in processing than the effects of whole word frequency. However, inflected words showed early effects of whole word frequency and only plurals, not inflected verbs, showed root frequency effects. The apparent absence of root morpheme effects for the regular inflected verbs is surprising given that effects were evident for derived words in which the relationships between constituent and whole forms are less direct than for regular inflections. Niswander et al. suggested that the anomaly may occur because root morpheme effects can be obscured by conflicts between the meaning and/or grammatical category of the most common form of the root of the inflected word and that of the whole form (e.g., *hand*|*handed*).

Direct evidence of the contribution of morphemic constituents is provided by recent investigations of eye-movements on bimorphemic Finnish compound words in sentence reading tasks. Hyönä and Pollatsek (1998) found effects of both the length and the frequency of the initial morpheme on the pattern of eye movements. The length of the initial morpheme influenced the location of the second fixation on the word, while first morpheme frequency influenced measures of first fixation duration, gaze duration, and the location of first and second fixations. Post-hoc analyses showed an effect of second morpheme frequency on gaze duration implying that activation of this constituent influences later processing of the word, but the effects were more restricted than those of first morpheme frequency. Subsequent research by Pollatsek, Hyönä, and Bertram (2000) directly investigated the independent effects of second constituent and whole word frequency on eye movements during sentence reading. Both frequency manipulations significantly affected gaze duration, but the effects occurred relatively late in processing and were most clearly evident in the probability of making a third fixation on the word. The fact that whole word frequency influenced processing at least as early as constituent frequency was taken as evidence that "identification of…compound words involves parallel processing of both morphological constituents and whole-word representations" (Pollatsek et al., 2000, p. 820). The time course of effects for different units is argued to be most compatible with models that assume a "race" between retrieval processes for constituent and whole words.

These results appear to confirm the general assumptions of IA frameworks by showing that the constituent elements of compounds are activated by presentation of the complete compound word. The fact that constituent effects can be demonstrated in the eye movement record of sentence reading tasks shows that sensitivity to the constituent structure of compound words is not an artefact of the decision requirements of tasks like lexical decision. However, questions remain about the implications of the relatively late manifestation of the constituent effects and about the generality of the results beyond Finnish. Productive compounding is very frequent in Finnish and Finnish compound words tend to be long and to require multiple fixations. These two factors may increase the likelihood that morphemes are treated as independent units. Hyönä and Pollatsek (1998) therefore "hesitate[d] to generalize either to situations where compounds are shorter or to languages in which compounding is not as productive" (p. 1625).

THE EXPERIMENTS

The present experiments use English compound word stimuli in paradigms similar to those of Hyönä and Pollatsek (1998) and Pollatsek et al. (2000). Two experiments were run in a single session. In Experiment 1, we orthogonally manipulated the frequency of the first and second morpheme of compound words while holding the frequency of the whole compound constant. The compounds were embedded in congruent sentence contexts and participants were asked to read the sentences for comprehension. Measures of the location and duration of eye fixations on the target compound words were analysed to provide insight into the contribution of the morphemic constituents of the compound to word identification. To provide an even more stringent test of the effects of the first morpheme, Experiment 2 used pairs of compound words with the same second constituent and different first constituents that were either high or low frequency (e.g., *raindrop/dewdrop*). This allowed the effect of first constituent frequency on the pattern of eye movements to be assessed in identical contexts (e.g., *The girl thought that the raindrop/dewdrop in the middle of the rose petal looked like a jewel*).

English compound words are virtually all low frequency. For example, the present items were selected from a set of 315 English compound words listed as whole forms in the Standard Oxford Dictionary. Over one third of these words are not listed in the CELEX database (derived from a corpus of 17 million words) and the average frequency of the 191 words that are listed is only 3.3 per million. Only 15 have frequencies of higher than 10 per million and only 5 are higher than 20. Given the almost complete absence of high frequency English compound words, it is not feasible to directly manipulate whole word frequency in combination with morpheme frequency, but regression analyses were conducted to determine its influence on performance.

METHOD

Participants

Forty-two members of the University of Massachusetts community participated in the experiments. All were native speakers of American English and received course credit or were paid for their participation.

Apparatus

Participants were asked to read individual sentences on a 17-inch ViewSonic 17PS monitor attached to a Pentium 166 MHz Compaq DeskPro 2000. During the experiment, their eye movements were monitored using a SR Research Eyelink 1 video-based eyetracking unit which sampled the reader's eye position every 4 ms (250 Hz) by means of infrared (IR) cameras attached to a lightweight helmet. Viewing was binocular but the position and duration of each fixation was stored for only the right eye. Another camera monitored the position of four IR markers positioned on the monitor allowing the eyetracker to compensate for head movements. Participants were seated in a recliner style chair approximately 85 cm from the monitor and allowed unrestricted head movement, but encouraged to remain relatively still. At this viewing distance, three letters occupied approximately 1 ° of visual angle. All sentences were displayed on a single line with a maximum length of 80 characters.

Stimuli

The characteristics of the target words used in each of the experiments will be described separately. The Appendix contains a complete list of the target stimuli and sentences presented in each experiment.

Experiment 1. The target words initially consisted of 80 compound words selected to complete a factorial manipulation of the first and second morpheme: high frequency first morpheme/high frequency second morpheme (HH), high frequency first morpheme/low frequency second morpheme (HL), low frequency first morpheme/high frequency second morpheme (LH), and low frequency first morpheme/low frequency second morpheme (LL). These words varied in length from 6 to 11 characters. Average length, the length of each morpheme and average whole-word frequency (Kučera & Francis, 1967), were matched across the four conditions (see Table 1). Comparison with frequency estimates from the larger CELEX corpus (Baayen, Piepenbrock, & van Rijn, 1993) confirmed that the items were well matched on word and morpheme frequency.

The target words were embedded in sentence frames of between 7 and 17 words in length in which at least three words preceded the target. The ideal way to ensure that eye movement differences between conditions were due to the critical targets rather than the sentence frames would have been to use the same sentence frame for one item from each condition. However, the difficulty in fitting four different words into the same sentence frame while maintaining neutral and plausible sentences was insurmountable. We therefore adopted the best approximation to this that could be achieved. The 80 words were assigned to 40 pairs such that both words could be embedded into exactly the same sentence frame. Four sets of 10 pairs were constructed that systematically combined the items from different conditions: (1) HH with HL words, (2) HH with LH words, (3) HL with LL words, and (4) LH with LL words. This assignment ensured that identical sentences were used to assess the simple effect of each constituent (first constituent: HH/LH, HL/LL; second constituent: HH/ HL, LH/LL). The members of each pair were embedded in 40 sentence frames that were constructed so that each member of the pair was used as a noun and was plausible in the sentence but unpredictable from the words that preceded it. [1] Two lists were constructed, each containing one of the pair members, such that 10 target items from each condition were included. Each participant therefore saw each sentence frame only once with one of its possible target completions. Each list was presented to equal numbers of participants.

After the experiment was run, the stimuli were subjected to a norming task in which a group of 18 participants (who had not participated in the actual experiment) were asked to rate whether each pair of target words fit equally naturally into a sentence frame. For example, a participant might see the following two sentence fragments: 1. *The exact location of the* ***battlefront*** and 2. *The exact location of the* ***shipwreck.*** Participants were asked to rate the

---

[1] We tried to avoid repeating morphemic constituents but it was not entirely possible. On the few occasions that words were repeated they appeared as different constituents (e.g., *hacksaw, sawdust*).

TABLE 1

Frequency of the whole compound word, and frequency and length of the first and second morpheme of the stimuli in each condition of Experiments 1 and 2

	Frequency			*Length*	
	Whole word	*Morpheme 1*	*Morpheme 2*	*Morpheme 1*	*Morpheme 2*
Experiment 1					
HH	1.8	192.5	200.2	4.2	4.4
HL	1.3	188.1	11.1	4.3	4.6
LH	0.7	8.6	193.7	4.4	4.2
LL	1.1	10.5	9.3	4.6	4.4
Experiment 2					
HH	2.6	234.0	316.0	4.2	4.2
LH	2.1	11.0	316.0	4.2	4.2

fragments on a 3-point scale coded so that –1 indicated that fragment 1 was more plausible than fragment 2, 0 that they were equivalently plausible, and +1 that fragment 2 was more plausible. For all but four sentence frames, the two completions were judged to fit equally well on this norming task or there was no systematic preference for one fragment over the other (mean rating=0.07). The four item pairs for which one fragment was systematically rated as more plausible than the other (mean rating above +0.75 or below –0.75) were removed from analyses based on 72 sentences –36 sentence frames each presented with two different completions. Another set of 15 individuals were given the target sentence up to the target word and asked to write down a noun that could complete the sentence fragment. None of the target words were generated as completions of the sentence fragment. The final set of target words was therefore plausible, but not predictable from the preceding sentence context.

Experiment 2. The target words in Experiment 2 initially consisted of 30 pairs of compound words that differed in the frequency of their first morpheme but had the same high frequency second morpheme, corresponding to the HH and LH conditions of Experiment 1. The words varied in length from 7 to 10 characters. The average length of first morpheme and the average whole-word frequency was controlled on average over the two lists (Kučera & Francis, 1967; see Table 1). Sentence frames were constructed in which both pair members were used as nouns, and were plausible but not predictable from the preceding words. Two lists of 30 sentences containing one member of the pair and 15 of each target condition were presented to equal numbers of participants.

The stimuli from Experiment 2 were subjected to the same norming pro cedures as those from Experiment 1 (the same participants evaluated both sets of stimuli). Two of the target words were predicted from the sentence fragment and five sentences, including one of those for which the target was predicted, yielded plausibility ratings favouring one completion (mean rating outside +/–0.75). Six sentences were therefore eliminated from analyses leaving 48 sentences (24 frames with two different completions).

Procedure

Upon arrival, participants were familiarised with the eyetracker and directions concerning the experiment were given. Participants were informed that single sentences would appear on the screen in front of them and they were instructed to simply read each sentence as they normally read. To ensure that they were reading the sentences for meaning, comprehension questions requiring a simple "yes" or "no" vocal response were presented after approximately 15% of the items.

RESULTS

Consistent with standard practices in eye movement research (see Rayner, 1998), fixations shorter than 100 ms and longer than 800 ms were eliminated from the data analyses for both experiments (2.36% and 2.57% of the total number of fixations on the target for Experiments 1 and 2, respectively). Fixations shorter than 100 ms are thought to reflect oculomotor programming time as opposed to cognitive processes, while long fixations are often due to track losses.

The target words were fairly long, so readers rarely failed to fixate on them at all. Skipping rates in first pass fixations were only 1.9% and 2.5% in Experiments 1 and 2 respectively, and even fewer words were never fixated at all (1.2% and 1.8%). There were no differences across conditions in either first pass or total skipping rates in either experiment (all F_s<2.18). Analyses therefore focused principally on measures of fixation time. For all measures, separate analyses of variance (ANOVAs) treating subjects and items as random factors (Clark, 1973; Wickens & Keppel, 1973) are reported as F_s and F_i, respectively.

The major variables reported are *first fixation duration, gaze duration,* and *total time* on the word. First fixation duration is the duration of the first fixation on a word independent of the number of fixations that are made, while gaze duration is the sum of all fixations on a word prior to moving to another word. Both of these measures are computed contingent on the word being fixated on a first pass fixation. The total time measure is the sum of all fixations on a word including both first pass and regressive fixations. In addition to these measures of processing time, a number of other eye movement measures (landing position in the target word, number of fixations on the target word, regressions to the target word and a "spillover" measure of duration of the first forward-moving fixation into a post-target region consisting of the next two-three words in the sentence) were examined and are reported when appropriate.

TABLE 2

Processing time measures (ms) for each condition of Experiment 1, standard errors in parentheses

	Morpheme frequency							
	HH		*HL*		*LH*		*LL*	
First fixation	270	(4.8)	273	(5.5)	277	(5.4)	282	(6.3)
Gaze duration	383	(12.1)	390	(16.5)	397	(16.5)	419	(14.7)
Total time	431	(15.3)	449	(22.3)	453	(20.7)	486	(22.3)

Experiment 1

Table 2 shows the processing time measures from Experiment 1. For first fixation duration, there was a marginally significant effect of the frequency of the first morpheme in the participants' analysis, $F_s(1, 41)=3.18$, $p=.08$; $F_i(1, 76)=1.94$, $p=.17$. First fixations on words with high frequency first morphemes were 8 ms shorter than those on words with low frequency first morphemes. There were significant effects of the frequency of both morphemes in the subjects' analyses of gaze and total duration that were not reliable in the item-based analyses. Gaze durations were 21 ms shorter for items with high than low frequency first constituents, $F_s(1, 41)=8.74$, $p<.005$, $F_i(1, 76)=1.56$, $p= .21$; and 15 ms shorter for high than low frequency second constituents, $F_s(1, 41)=2.86$, $p=.10$, $F1(1, 76)=1.16$, $p=.29$.[2] Similarly, average total fixation times on words with high frequency first and second morphemes were shorter by 29 and 25 ms, respectively, than when the constituent morpheme was low frequency: first: $Fs(1, 41)=11.60$, $p<.001$, $F1(1, 76)=2.00$, $p=.16$; second: $F_s(1, 41)=5.06$, $p<.05$, $F1(1, 76)=1.83$, $p=.18$. The interaction between first and second morpheme frequency was not significant in the analyses of any of the fixation measures indicating that the two constituents exerted statistically equivalent effects.

Measures of the mean landing position in the target word (i.e., the location of the first fixation on the target word) varied from 3.0 to 3.3 letter positions from the beginning of the word and did not significantly differ between conditions. Readers made more fixations overall on the target word when the first morpheme was low frequency (1.76 fixations) than when it was high frequency (1.68 fixations), $Fs(1, 41)=5.105$, $p<.05$, $F1(1, 68)=1.16$, $p=.29$, but this effect was not significant in first-pass fixations $F_s(1, 41)=1.476$, $p=.23$, Fi(1, 68)<1, suggesting that it is due, in part to regressive fixations. However, the only significant finding in the direct analysis of the regression rates was an effect of second morpheme frequency, $F_s(1, 41)=6.363$, $p=.016$, $F_i(1, 76)= 3.496$, $p=.07$, that occurred because the average regression rate was higher when the second morpheme was low frequency (11.4%) than when it was high frequency (8.0%).

Inspection of the patterns of fixations on the two morphemic constituents of the compound words showed that the majority of compound words received only one or two forward fixations: Over half the compound words were fixated once on either the first (37.8%) or second morpheme (19.0%) and a further 28.6% of compound words received a single fixation on each morpheme. The patterns of fixations were very similar for the four-item groups suggesting that the effects of morpheme frequency on fixation times are unlikely to be due to differences in the probability of fixating the two constituents.

Experiment 2

Table 3 shows the measures of average processing time for each condition of Experiment 2. The results obtained in these more controlled contexts paralleled those for first morpheme frequency in Experiment 1 but showed more robust effects in the item-based analyses, presumably reflecting the greater control exerted over the sentence and item context in which the initial morphemes were presented. The effect of frequency of the first morpheme was only marginally significant in the subjects' analysis for first fixation duration, F_s(l, 41)=2.99, p =. 09, F_i<1, but there were significant effects in the subjects' analysis of gaze duration, F_s(l, 41)=9.52, p<1.0101, F_i(l, 46)=3.23, p=.08, and total fixation time, F_s(1, 41)=16.31, p<.001, F_i(l, 46)=3.76, p=.06, which were 27ms and 42ms shorter, respectively, for words with high frequency compared to low frequency first morphemes.[3]

The difference in overall number of fixations on items with high rather than low frequency first constituents (1.49 vs. 1.59) was marginally significant in the subjects' analysis, F_s(l, 41)=3.73, p=.06, F_i(l, 46)=1.29, p=.26, but first constituent frequency did not significantly affect the number of first pass fixations, F_s(l, 41)=1.41, p=.24, F_i<1, the probability of regressing to the target word (F_s<1) the first fixation duration in the spillover region, F_s(1, 41)=1.74, p=. 19, F_i(1, 46)=1.173, p=.284, or the landing position on the target word (F_s< 1.2).

[2]The effect of the first and second constituent frequency in gaze duration appears to be due to the increased likelihood of refixation on the target in Experiment 1. When examining the duration of the second fixation on the target, no effect for second fixation duration was found. However, a nonsignificant 12 ms effect was present with the high frequency second morphemes yielding slightly shorter second fixation durations than low frequency second morphemes. This trend remained whether one analyses the second fixation durations contingent upon the initial fixation landing on the first morpheme or contingent upon the second fixation landing on the second morpheme.

[3] In Experiment 2, high frequency first morphemes resulted in shorter second fixation durations on the target than words with low frequency initial morphemes, F_s(l, 34)=3.05, p=.09, $F1$(1, 46) = 4.21, p<.05.

TABLE 3

Processing time measures (ms) for both conditions of Experiment 2, standard errors in parentheses

	Morpheme frequency	
	HH	*LH*
First fixation	267 (5.6)	274 (5.8)
Gaze duration	343 (11.0)	370 (10.8)
Total time	380 (12.8)	422 (15.0)

The pattern of fixations on each morpheme was similar to that observed in Experiment 1. Most items received one or two fixations. As in Experiment 1, items that received only one fixation were much more likely to be fixated on the first (41.2%) than the second constituent (23.0%) while 24.4% were fixated once on each morpheme. The fixation patterns were similar for the two item groups, although words with a low frequency first morpheme were slightly less likely to receive a single fixation on the first morpheme than words with a high frequency first morpheme (38.1% vs. 44.4%).

Regression analyses

The ANOVA analyses of both experiments demonstrate effects of the frequency of both morphemic constituents on processing time for compounds matched on whole word frequency. However, these effects were not reliable by conventional criteria in Experiment 1 and only marginally so in Experiment 2. Applying item based analyses is, in some senses, an unnecessarily stringent analytical approach because the items were highly selected and, in at least some conditions, virtually exhausted the population of available words. Nevertheless, establishing generality across items is important to establishing the internal consistency of the experimental manipulation independently so it is relevant to consider the factors determining variability between items.

It is also important to determine whether there is an influence of whole word frequency on fixation times. The uniformly low frequency of compound words makes it impossible to manipulate this variable orthogonally with constituent frequency. Instead, we conducted simultaneous regression analyses predicting the three main processing time measures using measures of the frequency of the whole compound and of each constituent (see Table 4).

Neither the overall regression equation nor any individual predictor significantly contributed to explaining first fixation time. The three predictors accounted for around 14% of variance in gaze and total fixation duration in Experiment 1 and, for both dependent measures, whole word frequency uniquely accounted for a little over half of the predictable variance while the frequency of

TABLE 4

Results of simple regression analyses using whole word, morpheme 1, and morpheme 2 frequency to predict processing time measures in Experiment 1. The total R^2 accounted for by the regression equation as well as the change in R^2 resulting from removing each variable from the equation is reported

	First fixation		*Gaze duration*		*Total duration*	
	β	*R^2 change*	β	*% R^2 change*	β	*% R^2 change*
Experiment 1						
Frequency Morpheme 1	−0.015	.006	−0.034	.003	−0.062	.005
Frequency Morpheme 2	−0.024	.023	−0.096	.037*	−0.142	.044*
Frequency Whole word	−1.330	.026	−7.754	.088**	−9.700	.074**
R^2	.063		.146**		.141**	
Experiment 2						
Frequency Morpheme 1	0.007	0.007	−0.054	.052*	−0.070	.041
Frequency Morpheme 2	−0.012	.038	−0.031	.041	−0.036	.026
Frequency Whole word	−2.180	.063*	−6.840	.099**	−9.570	.090**
R^2	.127		.243**		.198**	

*p<.1; ** p<05.

the second but not the first morpheme made a marginally significant unique contribution. Larger proportions of overall variance were accounted for in the more controlled context of Experiment 2, but whole word frequency remained the strongest predictor, even making a marginally significant contribution to first fixation times. In contrast to Experiment 1, which only showed significant unique effects of second morpheme frequency, in Experiment 2 which held the second morpheme constant, first morpheme frequency accounted for larger proportions of unique variance in gaze and total duration than second morpheme frequency, but its contribution only approached significance in the regression equation for gaze duration. Thus, the regression analyses for both experiments show strong effects of whole word frequency and weaker more variable effects of the frequency of the constituent morphemes.

DISCUSSION

The results confirm those of earlier studies of behavioural performance in isolated word identification tasks by showing reliable effects of the frequency of both the first and second constituent of compound words on various measures of eye fixations during sentence reading. They also demonstrate that Hyönä and Pollatsek (1998) and Pollatsek et al.'s (2000) findings for Finnish compound words generalise to English. Given that all of the whole compound words were of low frequency, these results clearly establish that the morphemic constituents of compound words exert an influence on processing independently of the frequency of the whole compounds. Regression analyses showed that fixation times are also influenced by whole word frequency. At a general level, these findings are consistent with the assumption that lexical representations of the whole word and both morphemic constituents are activated in parallel. The pattern of effects on different eye movement measures also provides evidence about the timecourse of constituent processing.

The patterns of fixation across the two constituents showed a bias to fixate on the first constituent. A single fixation on the first morpheme was the most common pattern in both experiments and over 40% of items were only ever fixated on the first constituent, while only half that number were fixated only on the second morpheme. These fixation patterns were minimally influenced by constituent frequency. Neither did frequency influence landing position in either experiment. This is consistent with the accumulated evidence demonstrating that the information extracted from parafoveal vision is limited to gross features such as interword spaces and letter shape (see Rayner & Pollatsek, 1989; Rayner, 1998, for reviews). This information contributes to determining the length and landing position of saccades, but is insufficient to trigger lexical retrieval. The target items all had similar overall orthographic structure regardless of their constituent frequency and were selected to be both unable to be explicitly predicted from context, and equivalently plausible in the sentence. This combination of contextual and target factors determines where participants land their eyes when they reached the region of the target word, and are independent of constituent morpheme frequency.

Recognising that a bias towards the initial constituents of long words may reflect general saccadic strategies that are independent of the processes involved in lexical retrieval provides a possible alternative account of the first constituent primacy effects that Taft (1991) has attributed to morphological structure. Early constituents of long words may play a dominant role in processing simply because this is where people tend to land their eyes rather than because the first segment has an inherently special role in the retrieval process. Lexical retrieval processes require the high spatial frequency information provided by foveal vision and are therefore reflected in fixation times.

The measures of gaze duration and total fixation duration in Experiment 1 showed significant frequency effects for both morphemes. These were of essentially equivalent magnitude despite the bias to fixate on the first morpheme. There was also a small effect of first constituent frequency on first fixation duration that was marginally significant in both experiments. This pattern of findings is consistent with the "sequential decomposition" model that motivated

Hyönä and Pollatsek's (1998) research because the first morpheme appears to influence the early processing indexed by first fixation duration as well as the later processing reflected in gaze duration, while second morpheme frequency influences both the amount of time spent initially processing the compound and also the probability of regressing to the word.

The present items were between seven and eleven letters long so they usually attracted two fixations. Most first fixations fell about three characters into the word. Initial morphemes were between three and five letters long so this landing position would have ensured that the first morpheme was in foveal vision but allowed only partial foveal processing of the second morpheme. It is therefore not surprising that only first morpheme frequency influenced first fixation duration. What is more interesting is that the effect of first morpheme frequency on first fixation duration was so much smaller than the effects on gaze duration and total fixation time given that most first fixations fell on the first morpheme and less than 10% of items received more than one fixation on the first morpheme. This implies that the first morpheme effects are due, at least in part, to processing that occurs after participants have stopped directly fixating on the first constituent (see also Hyönä & Pollatsek, 1998).

This is confirmed by the results of Experiment 2 in which the stimulus context was more tightly controlled than had been possible in Experiment 1. In Experiment 2, the stimuli that participants had to read were identical except for the first morpheme of the target word. In particular, the second morpheme was identical so differences between fixations on words containing high and low first morphemes cannot be attributed to the influence of the second morpheme. The results were similar to those for the same conditions of Experiment 1 in that effects of first morpheme frequency were larger and more robust when the dependent measure included fixations beyond those on the first morpheme. Thus, as for Experiment 1, the influence of the frequency of the first morpheme persisted after participants had stopped fixating on that constituent. Further, in contrast to Experiment 1, first morpheme effects on gaze and total fixation time generalised across items suggesting that the item-related variance in Experiment 1 may have been due, in part, to interactions between constituent attributes and the compound word and sentence context in which they occurred.

The lack of a direct relationship between the focus of fixation and the effects of first constituent frequency is consistent with evidence suggesting that eye movements can be triggered before lexical retrieval for the fixated constituent has completed. Such evidence is inconsistent with early models of eye movement control (e.g., Morrison, 1984). Reichle, Pollatsek, Fisher, and Rayner (1998) proposed an alternative model in which different mechanisms underlie the programming of saccadic eye movements to a new position in the text and the shift of covert attention to that position. The model distinguishes between the early computation of a word's familiarity and the ongoing lexical activity generated by the fixated stimulus that ultimately leads to retrieval of a single lexical representation. The level of familiarity is "related to the proficiency and probability of successful resolution…[and therefore serves as a] signal to the eye movement system that lexical access is imminent and that a saccade should be planned to the subsequent word" (p. 133). The subsequent "completion of lexical access [is] the signal for a shift of covert attention" (p. 133). Because there is a delay of around 150–175 ms between the time of initiating an eye movement and the actual execution of the movement (Rayner, Slowiaczek, Clifton, & Bertera, 1983), the shift of covert attention triggered by lexical access can occur before the planned eye movement has been executed. The information obtained while covert attention is directed to a word before it has actually been fixated gives rise to the "preview benefits" that have been observed from parafoveal words (e.g., Rayner, 1975). If familiarity is computed quickly enough, it may allow a new eye movement programme to be initiated before the previous one has been executed. This can, in turn, abort the original movement and lead to the skipping that sometimes occurs for short or highly predictable words.

The present results for English compound words, like those obtained in Finnish, are consistent with this view of the relationship between eye movements and lexical retrieval. Hyönä and Pollatsek (1998) also found larger effects of first morpheme frequency on gaze duration than first fixation duration suggesting that "the putative first stage of processing must often extend past the first fixation" (Pollatsek et al., 2000, p. 821) and Pollatsek et al. also found substantial effects of second morpheme frequency on gaze duration. Our English compounds are shorter than the Finnish stimuli so the effects of the second morpheme occur somewhat earlier in processing. In Pollatsek et al.'s data, morpheme frequency exerted only a small effect on second fixation duration and the effects were primarily due to the probability of a third fixation. Nearly all of the items in the present Experiment 1 were only fixated twice so the effects on gaze duration were primarily due to the second fixation, but multiple fixations and regressions were also more common to low frequency second morphemes. Reichle et al.'s (1998) distinction between the shifting of covert attention and gaze may also contribute to explaining why second morpheme frequency exerts as strong an effect on gaze and total fixation duration as first morpheme frequency even though the fixations for many items were restricted to the first morpheme region. For at least some of the occasions on which a single fixation on the first morpheme was recorded, a planned saccade to the second morpheme region may have been initiated, but aborted, because the information extracted after the shift of covert attention was sufficient to allow retrieval of the second constituent or the whole word.

Pollatsek et al. (2000) also found that whole word frequency influenced gaze duration "at least as rapidly as did the frequency of the second constituent" (p. 820) and that the whole word frequency effect was larger than for monomorphemic words. They therefore concluded that the Finnish data support a "parallel version of a dual-route model" (p. 831) in which direct lookup and compositional processes race to identify the word. The fact that there was an effect of whole word frequency over and above the effects of component frequency demonstrates that identification of transparent Finnish compounds does not rely solely on a compositional process and, because the effects of whole word frequency occur relatively early in processing, they do not appear to be solely due to postlexical processes related to determining the meaning of compounds (Shoolman & Andrews, 2003). Pollatsek et al. argued that their data were also incompatible with various “either/or” versions of the dual route framework that assume that “some compounds are always identified by a whole-word process, whereas others are always identified by a compositional process” (p. 831).

We were not able to directly manipulate whole word frequency because English compound words are almost uniformly of low frequency. However, word frequency was included in regression analyses that showed it exerted a significant effect on both gaze and total duration independent of the variability accounted for by constituent frequency, and even accounted for a small portion of unique variance in first fixation data for Experiment 2. The fact that whole word frequency influenced the same measures of processing time as constituent frequency is compatible with the Finnish data in suggesting parallel influences of constituent and compound activation. Indeed, the lack of robust unique effects of first morpheme frequency in the regression analyses of Experiment 1 suggests that variability due to whole word frequency may obscure the effects of constituent frequency that can be observed in comparisons of item sets matched for whole word frequency.[4]

Given that the constituents were generally of higher frequency than the complete compounds, the present data are consistent with the general claim of “segmentation through recognition” accounts in showing that lexical constituents of a compound word that are higher in frequency than the complete word are activated in the course of identifying the complete compound. How ever, the fact that whole word frequency effects are evident at least as early as the effects of constituent frequency implies that whole word representations achieve activation early in processing, even for these long, relatively uncommon words, at least when they are presented in coherent sentences. This seems more compatible with dual route frameworks that assume parallel activation of representations of both whole words and their constituents (e.g., Baayen et al., 1997; Niemi et al., 1994) than with models that assume that constituent morphemes are represented at a lower level of the representational hierarchy than whole words (e.g., Taft, 1994) because the latter models imply that constituent activation should be evident earlier in processing than activation of the whole word (Giraudo & Grainger, 2001).

The early influence of whole word frequency and the strong evidence of second constituent frequency effects in the present results may reflect the meaning-focused processing engaged by the sentence reading task. The early activation of the whole word representation may be due, in part, to contributions from contextually derived expectancies, although these would have to be general rather than specific because the rating data showed that the target items were not predictable from the preceding sentence context. The fact that the second constituent exerted a stronger unique effect on fixation time than the first constituent in the regression analyses of Experiment 1 is consistent with recent findings reported by Juhasz, Starr, Inhoff, and Placke (2003). In an experiment quite similar to our first experiment, they found significant gaze duration effects due to second morpheme frequency, but not first morpheme frequency, that conflict with our ANOVA effects of first morpheme frequency. It is possible that the second constituent effects reflect the fact that the meaning of compound words is usually determined more by their second than their first constituent. More evidence about the relationship between the patterns of performance observed in isolated word and sentence reading tasks is necessary to evaluate how lexical retrieval is modulated by context. This evidence will contribute to determining the validity of isolated word identification tasks as indices of lexical retrieval during online comprehension.

It is important to acknowledge that the results do not allow definitive conclusions as to precisely what level of subword representations underlie the observed effects. The constituents of compound words are lexical units and the manipulations of constituent frequency were designed to detect effects at that level. It is possible that parallel manipulations of syllable frequency might reveal similar sensitivity to this unit. Alternatively, the processes that define the units that serve as inputs to the activation process might rely on processing directed towards syllabic parsing rather than being specific to lexical or morphological units. Apparent morphological effects may be a by-product of the overlap between syllabic units and stem morphemes (Taft, 1979a, 1994, 2003). Neither do the present results discriminate between models that assume explicit morphological levels of representation (e.g., Taft, 1994) and those that attribute morphological effects to the convergence of orthographic and semantic activation (e.g., Marslen-Wilson et al., 1994; Plaut & Gonnerman, 2000). Further insight into the level of the effects could be obtained by comparing constituent effects for compound items with those for pseudomorphemic items (e.g., *carpet, napkin*) to determine whether lexical constituents either facilitate performance for compounds, or interfere with performance for pseudomorphemes, as might be expected if top-down semantic information were implicated in the effects (Pillon, 1998). Inhoff s (1989) comparison of eye fixations on these items revealed no differences in overall fixation duration between the two item classes but it is possible that a more refined investigation of constituent frequency effects would yield a different outcome.

---

[4] This conflict between the results of the ANOVA and regression analyses regarding the contribution of first morpheme frequency may also reflect the non-normal distribution of first constituent frequency resulting from selection of items with low and high first morphemes.

How is morphemic segmentation achieved?

The results are consistent with Pollatsek et al.'s (2000) description of the "unfolding" of the processes involved in reading morphologically complex words. They confirm the findings of less direct paradigms like lexical decision by providing evidence that morphemic constituents are activated during natural reading and show that this online decomposition is not restricted to languages like Finnish in which compounding is much more frequent and productive than in English. The evidence is consistent with the assumptions attributed to the IA framework in the sense that it appears to demonstrate parallel activation of lexical representations corresponding to both constituents and whole forms. But the data also challenge a number of the specific assumptions of computational implementations of the IA model. Current implementations of the model (e.g., Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981) can only process words of a limited range of lengths, and are incapable of demonstrating the position and context invariance that is necessary to allow constituents to be recognised in the context of longer words (Andrews & Davis, 1999).

Importantly, these limitations do not arise from implementation-specific details but from fundamental and inherent attributes of the framework. The processes underlying activation of lexical representations rely on matches and mismatches between position-specific letter detectors for words of a particular length. For example, the identification of the word *bird* requires activation of detectors for *b, i, r* and *d* in the first to fourth position. There is therefore no mechanism by which the last four letters *of blackbird* can activate the word node for *bird*. Indeed, because current implementations assume separate letter nodes for each word length (e.g., Grainger & Jacobs, 1996), even if the model was extended to include nine-letter words, the first five letters of *blackbird* would not activate the node for *black*. These problems can be addressed by modifying the assumption of separate letter nodes for words of different length. For example, Coltheart et al. (1993) implemented a different approach to generalising across word lengths in the IA network that they incorporated into the Dual Route Cascade (DRC) model. This implementation includes a single set of positionspecified letter input nodes that are activated by words of all lengths: A fourletter word activates position 1, 2, 3 and 4; a five-letter word activates positions 1, 2, 3, 4, 5; and so on. Although this approach, in principle, allows *black* to be activated by *blackbird,* it still prevents *blackbird* from activating *bird*. This could obviously be solved by lining words up from the final rather than the initial letter position, but this would create the reverse problem. A further problem derives from the IA model's assumption that mismatching letters inhibit words that do not contain them. That is, even if the letter-coding assumptions allowed *blackbird* to activate *black,* the mismatching letters of *bird* would create inhibition.

Thus, the assumption of position-specific letter codes that is inherent to all IA models prevents the recognition of embedded words that is crucial for segmentation through recognition. Recognition of embedded segments regardless of their within-word position requires context and position-invariance: *Bird* must be coded in such a way that it can be activated regardless of the presence of other letters and where it occurs in a word. This is not possible if activation of word nodes depends on excitation and inhibition from position-specific letter nodes (Andrews & Davis, 1999). Other input coding schemes that are sensitive to the spatial ordering of letters are capable of demonstrating the position and context invariance necessary to achieve segmentation-through-recognition (Mozer, 1987) and, when combined with the learning and retrieval assumptions of self-organising models, can effectively simulate the influence of morphological constituents on word identification (Davis, 2000).

CONCLUSIONS

The present data show that the separate morphemic constituents of compound words are extracted and activated in the course of lexical retrieval during a sentence reading task and therefore confirm that such effects are not solely a function of the decision requirements of tasks like lexical classification. Such evidence is compatible with the assumption that morphemic constituents and whole word forms are extracted and processed in parallel. This view has usually been identified with IA models on the implicit assumption that such frameworks allow morphemic segments to be matched with their lexical representations through processing of the whole word. However, more detailed consideration of exactly how word identification is implemented in IA models reveals major problems in their account of how subword constituents are extracted from whole word forms (Andrews & Davis, 1999; Davis, 2000).

Eye movement data like those reported here demonstrate the importance of taking account of the perceptual mechanisms contributing to word identification. Long compound words usually require more than one fixation and the sequential nature of the observed morphemic effects is presumably at least partly a reflection of this. Processing is, therefore, both serial and parallel. Information intake operates serially on the components of the stimulus pattern available during a fixation, but this occurs in parallel with continuing processing of the information from earlier fixations. The present results suggest that lexical representations of the components and of the whole word are activated simultaneously and influence the time spent fixating on each component and the likelihood of regressing to the word.

Acknowledging these characteristics of the processes involved in acquiring the sensory features of the text highlights the limitations of current models of word identification. Despite the explicit specification required to implement these models in computational form, they remain impoverished accounts of many of the processes involved in reading. In particular, they provide an extremely simplistic account of how sensory evidence is extracted and matched with lexical representations. The further implications of the parallel processing of different word constituents in foveal and parafoveal vision raises a range of questions about how information from different sources and time points in processing are combined that are completely ignored by current models. Such issues are crucial if we are to generalise them even to the processing of words that require more than one fixation, let alone to the processing of text. To address them, we need to develop an empirically validated account of exactly how the perceptual processes that extract information from printed text interact with the retrieval processes that match that information with representations in lexical memory, and of how these processes contribute to online comprehension. Eye movement data provide an important tool for achieving this goal.

PrEview proof published online September 2003

REFERENCES

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language, 25,* 726–740.

Andrews, S., & Davis, C. (1999). Interactive activation accounts of morphological decomposition: Finding the trap in mousetrap? *Brain and Language, 68,* 355–361.

Andrews, S., & Scarratt, D.R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1052–1086.

Baayen, R.H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel route model. *Journal of Memory and Language, 37,* 94–117.

Baayen, R.H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania.

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297–332.

Clark, H. (1973). The language-as-fixed fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel distributed processing approaches. *Psychological Review, 100,* 589–608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Davis, C. (2000). *The self organising lexical acquisition and recognition (SOLAR) model of visual word recognition.* Unpublished doctoral dissertation, University of New South Wales, Australia.

De Groot, A. (1984). Primed lexical decision: Combined effects of the proportion of related primetarget pairs and the stimulus-onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology, 36A,* 253–280.

Dell, G. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Fowler, C.A., Napps, S.E., & Feldman, L.B. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition, 13,* 241–255.

Gimson, A.C. (1980). *An introduction to the pronunciation of English* (3rd ed.). London: Arnold.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence of supralexical representations of morphology. *Psychonomic Bulletin and Review, 8,* 127–131.

Grainger, J., & Jacobs A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

Henderson, L. (1985). Towards a psychology of morphemes. In A.W.Ellis (Ed.), *Progress in the psychology of language, Vol.1* Hove, UK: Lawrence Erlbaum Associates Ltd.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Inhoff, A.W. (1989). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language, 28,* 444–461.

Inhoff, A.W., Briihl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming, and delayed naming tasks. *Memory and Cognition, 24,* 466–416.

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

Kello, C.T., & Plaut, D.C. (2000). Strategic control in word reading: Evidence from speeded responding in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 719–750.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Lima, S.D., & Pollatsek, A. (1983). Lexical access via an orthographic code? The basic orthographic syllabic structure reconsidered. *Journal of Verbal Learning and Verbal Behavior, 22,* 310–333.

Marslen-Wilson, W., Tyler, L., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101,* 3–33.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Mozer, M.C. (1987). Early parallel processes in reading: A connectionist approach. In M.Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Niemi, J., Laine, M, & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes, 9,* 423–446.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15,* 389–420.

Pillon, A. (1998). A pseudo-prefixation effect in visual word recognition: A true—neither strategic nor orthographic—morphemic effect. *Quarterly Journal of Experimental Psychology, 51A,* 85–120.

Plaut, D.C., & Gonnerman, L.M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15,* 445–485.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

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

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 95,* 618–660.

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

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., Slowiaczek, M.L., Clifton, C., & Bertera, J.H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 912–922.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *Quarterly Journal of Experimental Psychology, 42 A,* 529–567.

Shoolman, N., & Andrews, S. (2003). Racehorses, reindeers and sparrows: Using masked priming to investigate morphological influences on compound word identification. In S.Kinoshita & S.J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 241–278). New York: Psychology Press.

Taft, M. (1979a). Lexical access via an orthographic code: The basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior, 18,* 21–39.

Taft, M. (1979b). Recognition of affixed words and the word frequency effect. *Memory and Cognition, 7,* 263–272.

Taft, M. (1985). The decoding of words in lexical access: A review of the morphographic approach. In D.Besner, T.G.Waller, & G.E. MacKinnon (Eds.), *Reading research: Advances in theory and practice, Vol. 5.* New York: Academic Press.

Taft, M. (1991). *Reading and the mental lexicon.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Taft, M. (1994). Interactive activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9,* 271–294.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink & D.Sandra (Eds.), *Reading complex words.* Amsterdam: Kluwer.

Taft, M, & Forster, K.I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Taft, M, & Forster, K.I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15,* 607–620.

Wickens, T., & Keppel, G. (1973). On the choice of design and test statistic in the analysis of experiments with sampled materials. *Journal of Verbal Learning and Verbal Behavior, 22,* 296– 309.

APPENDIX

Experiment 1 stimuli

Set 1: HH/HL

The exact location of the **battlefront|shipwreck** was not disclosed for several days.

*The end of the race was all **downhill|backstroke** so Bill thought he could make it.

When he realised it was a **minefield|spacecraft** Walter's heart began to pound.

Julie tried to forget her **daydream|heartache** and concentrate on her work.

Liz could not find the **checklist|footstool** anywhere in the house.

When Amy saw the **headstone|deathtrap** she stared in shock.

*It seemed to take ages for the **playoff|nosebleed** to completely finish.

Matt accidentally dropped the **seafood|tablespoon** as he walked towards the stove.

Joe mistakenly left his **mouthpiece|billfold** in his locker after practice.

The great expanse of the **farmland|airport** amazed Jill even though she saw it daily.

Set 2: LH/LL

Remembering the day of the **bullfight|thunderbolt** made John feel very anxious.

After inspecting the **cartwheel|crankshaft** the expert found the cause of the problem.

Peter did not realise it was a **trapdoor|tollbooth** until he was very close.

The smell of the **chalkdust|peppermint** was almost overpowering.

Sandy picked up the **clipboard|corkscrew** and put it away in the drawer.

Claire decided to move the **palmtree|beehive** down to the end of the garden.

John thought he stepped on the **anthill|peanut** next to the porch on his way inside.

Paul needed to find a **hacksaw|pawnbroker** before he went home.

The dog sniffed at the **mincemeat|toadstool** but did not try to eat it.

The contractor installed the **floodlight|flagpole** outside the high school gymnasium.

Set 3:
HH/LH

David pulled on the **drawbridge|slingshot** as hard as he could.

When she finally completed the **crossword|scrapbook** Susan felt very satisfied.

It was such a big **snowball|speargun** they were sure it would cause damage.

Roger was sure they would find a **signpost|cloakroom** if they kept looking.

The mill worker pushed the **sawdust|driftwood** over to the nearby corner.

Looking from the **hilltop|stairway** they could see the whole house.

Mary asked for the **songbook|spyglass** that was displayed in the window.

Bob was saving for a **racehorse|surfboard** but he knew it would take a long time.

Liz could not see the **floorshow|tombstone** because someone was blocking her view.

*One of the privileges granted by his **birthright|sailboat** is the mobility of the elite.

Set 4:
HL/LL

It looked too small to be a seagull|hedgehog but the guide said that was what it was.

Maria hoped that her **flowchart|witchcraft** would help them finish the task.

When they saw the **bloodhound|scarecrow** the children began to scream.

The little boy had a **rosebud|silkworm** he had found in the garden.

Pam hoped that they had the **hairspray|grapevine** that she had been searching for.

In the darkness, the **moonbeam|stingray** glowed as if it was luminous.

*Andrew released the **handbrake|gangplank** and they slowly began to move.

No one noticed as the **firebug|bridegroom** left the burning church.

The conductor could feel the **kettledrum|stagefright** as the concerto's climax approached.

The artist said the **eyebrow|greyhound** was not quite right but the composition was good.

Experiment 2
stimuli (HH/LH)

*Stay away from the **rockpool/whirlpool** the mother warned her children.

*The dog walked quietly on the **sidewalk/bushwalk** until another dog came towards him.

It was only a small **manhole/loophole** but Tom thought that it would serve their purpose.

*By the end of the **countdown/knockdown** the whole crowd was standing and cheering.

The dealer said that the old **matchbox/snuffbox** was a very precious object.

Thinking about the **feedback/hunchback** made Mary start to cry.

The man standing in the **campground/foreground** seemed to be very tall.

The girl thought that the **raindrop/dewdrop** on the rose petal looked like a jewel.

You cannot use this **joystick/lipstick** until you are a little bit older said the mother.

The kids were waiting to go in the **spaceship/steamship** that was part of the new exhibit.

After so many hours in the greenhouse/jailhouse the reporter felt relieved to be outside.

Jill was sure that the **crossbar/crowbar** was too heavy for her to lift.

The children were told to draw a picture of a **milkman/caveman** like the one in the story.

The warning was more visible with the **highlight/floodlight** but it was still easy to miss.

I can assure you it is completely **heatproof/rustproof** the salesman told the customer.

There was a long line at the **doorway/archway** in front of the famous palace.

As soon as they finished their **homework/patchwork** the children were allowed to leave.

Jane had not wanted to be a **housewife/midwife** but discovered that she quite enjoyed it.

*The nurse used cotton wool to dry her **blackhead/forehead** before applying the cream.

#They took turns in cleaning the **classroom/bathroom** thoroughly at the end of each week.

Tim knew it was a **foothill/molehill** rather than a mountain, but he was still worried.

The people standing near the **sideline/pipeline** turned when they heard the roar.

John stared intently at the **billboard/chessboard** trying to decide what he should do.

It looked like an **eyeball/mont eball/mothball** but Peter knew it was really a marble.

*By the light of the **campfire/bushfire** they could see lots of animals.

Even though it was **roleplay/foulplay** it had impressed the large audience.

By the time they got to the **turnoff/kickoff** it had already started to rain.

Jim reached for his **penknife/jackknife** as he entered the dark alleyway.

The moss covering the **milestone/limestone** made it difficult to read the inscription.

The guard verified that the **nightwatch/wristwatch** was on time before leaving to go home.

* Sentences removed from analysis because of unequal plausibility ratings for two completions.

Sentences that participants generated the target in a completion task.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16* (1/2), 312–339

Eye movements, word familiarity, and vocabulary acquisition

Rihana S.Williams and Robin K.Morris

University of South Carolina, Columbia, USA

Two experiments examined the effects of word familiarity on word recognition and text comprehension during silent reading. Readers' eye movements were monitored as they read sentences containing words that varied in familiarity as assessed by printed estimates of word frequency, subjective ratings of familiarity, and a multiple-choice test of meaning knowledge. Effects of word frequency were unaffected by differences in subjective familiarity rating for high frequency words. Differential effects of familiarity rating were observed in low frequency conditions. In addition, processing time on high and low frequency words did not differ when familiarity was held constant for moderately familiar words. Readers spent more initial processing time on novel words than familiar words. Performance on a vocabulary test administered after the reading session demonstrated that readers successfully acquired and retained new word meanings. Finally, reanalysis of word processing time as a function of vocabulary test performance demonstrated a systematic relationship between online processing patterns and memory for novel word meaning.

Readers routinely encounter words that vary widely in familiarity from extremely common function words to words that they have never seen before. In fact, encounters with unknown words are not uncommon. Estimates of vocabulary growth suggest that skilled readers 18–25 years' old learn more than five new words a day and that most of these words are learned from text (Landauer & Dumais, 1997) without benefit of explicit instruction (e.g., Nagy, Herman, & Anderson, 1985). There is also a substantial body of evidence demonstrating that readers successfully acquire word meaning from silent reading. This evidence comes primarily in the form of performance on vocabulary tests administered after the reading task is completed (see Morris & Williams, 2003, for review). In contrast to this approach, a recent eye movement study (Chaffin, Morris, & Seely, 2001) examined online processing of texts containing unfamiliar words and demonstrated that following an encounter with a new word, readers spent more time on context that was relevant to establishing word meaning than on context that was not relevant and that readers spent more time on semantically related context following novel words than following familiar words. In addition, readers were more likely to reread a novel than a familiar target word. The experiments presented in this paper examine vocabulary acquisition in silent reading by merging these two approaches to look at the relationship between online processing patterns gleaned from the eye movement records of skilled readers and the performance of the same readers on a test of memory for the new word meanings.

Within the reading literature there is widespread consensus that a reader's familiarity with a word influences the time to recognise the word, at least within samples of words that are known to the reader. Considering novel words as representing the nadir of familiarity one might then expect that readers would spend more initial processing time on novel words than on familiar words. However, Chaffin et al. (2001) reported no initial processing time differences between words that the reader had never encountered before and known words that had received low subjective familiarity ratings (Chaffin et al., 2001). Given this seemingly anomalous finding, we examined the effect of word familiarity in silent reading more closely in our present experiments of vocabulary acquisition.

In eye movement studies of silent reading printed word frequency has been the most widely used operational measure of word familiarity. A word's printed frequency is typically derived from the number of times it appears within a large sample of written text. For example, the Brown Corpus (Francis & Kucera, 1982) consists of over a million words gleaned from 15 genres of edited American English text. Other corpora have also sampled classic literature, children's literature, and newspaper databases (e.g., Burgess & Livesay, 1998; Carroll, Davies, & Richman, 1971; Thorndike & Lorge, 1944). Even in the

---

Correspondence should be addressed to R.K.Morris, Department of Psychology, University of South Carolina, Columbia, SC 29208, USA. Email: morrisr@sc.eduThis work was supported by a Ford Foundation Predoctoral Fellowship for Minorities administered by the National Research Council awarded to R.Williams. This research was conducted in partial fulfilment of the Master's degree requirement for the first author. The authors are indebted to Gary Allen, Matthew Traxler, and Douglas Wedell for their input on this project.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI:10.1080/09541440340000196

CELEX corpus (Dutch Center for Lexical Information, 1995), which samples both written and spoken usage, the majority of the entries were obtained from written sources.

Studies of skilled reading behaviour have consistently demonstrated that readers spend more initial processing time on low frequency words (henceforth, LF) than on words of higher frequency (henceforth, HF) that are matched for word length and part of speech. Effects of printed word frequency have been demonstrated in single sentences (e.g., Rayner&Duffy, 1986; Rayner, Sereno, & Raney, 1996; Schilling, Rayner, & Chumbley, 1998), across multiple frequency categories (Hyönä & Olson, 1995), regardless of word class or syntactic category (Rayner & Duffy, 1986; Schmauder, Morris, & Poynor, 2000), and within discourse that may or may not provide the reader with expectancies of the meaning of the target word (e.g., Inhoff, 1984; Just & Carpenter, 1980). These effects have been observed in first fixation duration and gaze duration. In addition, the increased processing time in the LF word condition has often been shown to spillover onto the next fixation after readers leave the critical word (e.g., Rayner & Duffy, 1986).

In spite of the consistency with which word frequency effects have been reported, some researchers have raised concerns regarding the use of objective frequency counts as the sole metric of word familiarity. For example, there is concern that frequency counts derived from text-based corpora do not provide a representative sample of frequency of exposure as they do not take into account exposure to words through speaking and listening. Perhaps related to this, discrepancies between corpus counts and our intuitive sense of familiarity are easy to find. For example, the word “pizza” and the word “assay” have equivalent frequency counts in the Francis and Kucera (1982) corpus, although we have the sense that most American college students could convey detailed knowledge of pizza while not nearly as many could say much about the meaning of the word “assay”.

Numerous naming and lexical decision latency studies of visual word recognition have demonstrated that participants are sensitive to word familiarity as assessed either by printed word frequency estimates (see Monsell, 1991, for a review) or subjective ratings of word familiarity (e.g., Balota, Pilotti, & Cortese, 2001; Connine, Mullenix, Shernoff, & Yelen, 1990; Ferraro & Sturgill, 1998; Gaygen & Luce, 1998; Gernsbacher, 1984; Gilhooly & Logie, 1980; Gordon, 1985; Whalen & Zsiga, 1994). These studies have produced similar general results in that more frequent words are processed more quickly than less frequent words and, likewise, more familiar words are processed more quickly than less familiar words. Differences between the patterns of response times observed based on estimates of printed word frequency and those based on subjective familiarity ratings emerge within the low printed word frequency category (Connine et al., 1990; Gernsbacher, 1984; Gordon, 1985). Of course, subjective familiarity ratings may be based on additional sources of information beyond a pure frequency count. Some effects based on subjective familiarity ratings have differed depending on how the ratings were obtained; however, it has been the case that regardless of whether participants were asked to rate how familiar they were with a particular word (Gernsbacher, 1984), how frequently a word appeared in the English language (Gordon, 1985), or how well a word's meaning was known (Connine et al, 1990), lexical decision latencies were longer in response to less familiar LF words than they were to more familiar LF words.

In a more naturalistic reading task, Chaffin et al. (2001) monitored readers' eye movements as they read pairs of sentences containing a target word from one of three subjective familiarity conditions: high familiar, low familiar, or novel. The novel words condition consisted of orthographically regular and pronounceable pseudowords (Chaffin, 1997) to ensure that participants did not have any prior knowledge of the word. The comparison of high and low familiar words revealed first fixation duration, gaze duration, and spillover effects that mirror the effects of printed word frequency estimates. Readers spent more time on less familiar words than on more familiar words. However, when readers encountered a novel word, first fixation and gaze duration did not differ in comparison to an existing low familiar word.

On the basis of these data alone one might come to the simple conclusion that readers were equally unfamiliar with the novel and the low familiar words that were selected for use in this study. However, the story is more complicated than that. Processing time on semantically related context that followed the target word and patterns of rereading behaviour revealed greater processing difficulty in the novel word than in the low familiar word condition. Furthermore, in these measures there were no such differences between high and low familiar existing word conditions. These data were taken as evidence that readers were using the information conveyed by the sentence context to establish a meaning for the novel word in ways that were not required for comprehension in the two familiar word conditions.

The Chaffin et al. (2001) study replicated the basic word familiarity effects observed in speeded response paradigms. But it did not include a direct comparison of word frequency and word familiarity effects, leaving open the question of the extent to which these two selection criteria yield similar results in silent reading. Furthermore, the results of the speeded response studies suggest familiarity ratings are particularly sensitive to differences among LF words. Given that, it would seem likely that there would be differences in readers' initial processing time on words that they had never seen before compared to known words that received low familiarity ratings. Yet, as mentioned earlier, Chaffin et al. report no difference in initial processing time for novel words compared to low familiar words.

Finally, reading studies that have examined word familiarity effects (as measured by corpus-based frequency counts) have focused almost exclusively on issues of lexical access and thus looked primarily at initial processing of the target word, and not at reanalysis. In the current study we were interested in the role of word familiarity in both lexical access and in text comprehension. We were particularly interested in how readers comprehend text when that comprehension depends upon understanding the meaning of a previously unfamiliar word. Chaffin et al.'s (2001) results point to the need to look at reanalysis measures in conjunction with the initial processing time measures in order to gain a fuller understanding of vocabulary acquisition, and the effects of word familiarity more generally, in reading.

In light of these issues we conducted two reading experiments. In Experiment 1 readers' eye movements were monitored as they read sentences that contained target words that varied in corpus-based word frequency counts and in subjective familiarity ratings. Experiment 2 compared the online processing of high and low familiar LF words to that of novel words. In addition, eye movement measures from the novel word condition were reanalysed as a function of performance on a vocabulary test administered after the reading session was complete.

EXPERIMENT 1

In this experiment effects of printed word frequency and subjective familiarity ratings on initial processing time were directly compared in a silent reading task. In lexical decision tasks (e.g., Connine et al., 1990; Gernsbacher, 1984), the time to recognise LF words that received low subjective familiarity ratings was greater than the time to recognise LF words that received high subjective familiarity ratings, suggesting that familiarity ratings provide important additional information about LF words. It is not yet clear whether these effects generalise to fluent silent reading. Readers' eye movements were monitored as they read sentences like those displayed in Table 1. The sentences contained an HF-very familiar, HF-familiar, LF-familiar, or LF-less familiar target word, respectively. Only words with meanings that were known to the majority of the participant population were used as target words. The context preceding each target word was neutral with respect to the target word meaning. The context following the target word was consistent with the meaning of the target word. This design allowed us to compare initial processing time on the target word as a function of printed word frequency when familiarity rating was held constant and vice versa, uncontaminated by sentence context effects.

It was expected that readers would spend more processing time (as measured by first fixation duration, gaze duration, and spill over) on LF words than on HF words (replicating Hyönä & Olson, 1995; Inhoff, 1984; Rayner & Duffy, 1986; Schmauder et al., 2000). We also expected to see a familiarity effect primarily within the low word frequency condition such that the condition with the lower rating of familiarity would yield greater processing time. This was based on the findings observed in earlier lexical decision and naming tasks (e.g., Balota et al., 2001; Connine et al., 1990; Gernsbacher, 1984; Gordon, 1985). Given the robust and pervasive history of word frequency effects in the eye movement literature (see Rayner, 1998, for a review), we expected that when familiarity ratings were

TABLE 1

Example sentences for each condition in Experiment 1 (target word in bold)

	Condition	*Example sentence*
Example 1	HF–very familiar	Kate chose the **knife** as her weapon.
	HF–familiar	Kate chose the **rifle** as her weapon.
	LF–familiar	Kate chose the **dagger** as her weapon.
	LF–less familiar	Kate chose the **lance** as her weapon.
Example 2	HF–very familiar	Billy used the **truck** to haul his supplies.
	HF–familiar	Billy used the **wagon** to haul his supplies.
	LF–familiar	Billy used the **cart** to haul his supplies.
	LF–less familiar	Billy used the **sedan** to haul his supplies.
Example 3	HF–very familiar	Sue loved the **bird** that she saw on the farm.
	HF–familiar	Sue loved the **cattle** that she saw on the farm.
	LF–familiar	Sue loved the **goat** that she saw on the farm.
	LF–less familiar	Sue loved the **bison** that she saw on the farm.

A complete set of stimuli are available upon request.

held constant and frequency estimates were either high or low we would see the typical word frequency effects.

We also considered reanalysis effects of word familiarity by examining second pass time and frequency of regressions to the target word. In general, word frequency studies that have reported reanalysis measures have shown that readers engage in more rereading behaviour in low word frequency conditions than in HF conditions (e.g., Hyönä & Olson, 1995; Rayner & Raney, 1996; Schmauder et al., 2000). In addition, Chaffin et al. (2001) reported that readers engaged in more reanalysis of low familiar words than of high familiar words, even in the absence of initial processing time differences.

Method

Participants

Twenty-four students from the University of South Carolina were given either psychology course credit or $5.00 for participating in this experiment. All participants were native speakers of English and had normal uncorrected vision.

Stimuli

The stimuli for the four target word conditions (HF-very familiar, HFfamiliar, LF-familiar, and LF-less familiar) were selected based on ratings of subjective familiarity, printed estimates of word frequency, and vocabulary test responses.

TABLE 2

Target word selection criteria in Experiment 1

Condition	*Example item*	*Printed word frequency*	*Familiarity rating*	*Word length*
HF–very familiar	Knife	109 (78–149)	6.9 (6.8–7.0)	5.5 (4–7)
HF–familiar	Rifle	96 (75–190)	6.6 (6.5–6.7)	5.7 (4–7)
LF–familiar	Dagger	4 (1–10)	6.7 (6.6–6.8)	5.5 (4–7)
LF–less familiar	Lance	1.5 (1–10)	5.7 (5.2–6.1)	5.9 (4–7)

The range of values contributing to each mean are included in parentheses.

Twelve sets of four semantically related target words (one from each condition) were formed. The four words within a set were matched for word length within three character positions. Table 2 displays an example from each target word condition and the mean value for each condition for each of the selection criteria.

Familiarity ratings. One thousand and fifty native English-speaking volunteers from the University of South Carolina community provided subjective ratings of word familiarity via the internet. The words were HF and LF nouns listed in the Francis and Kucera (1982) word frequency estimates (see Table 2). Each word was rated twice, once in a list of other words of similar frequency and once in a mixed list of HF and LF words. Each participant saw only one list. Our instructions asked each participant to rate how often they had seen a given word. Responses were indicated on a 7-point Likert scale with 1 labelled as very unfamiliar and 7 as very familiar.

Ratings were averaged across subjects and lists. Words were sorted by printed word frequency and then familiarity rating quartiles were calculated for each frequency condition. Ratings for HF words were skewed to the higher end of the Likert scale and were generally 6 or higher. In contrast, LF words received a much broader range of ratings including some words that received very low familiarity ratings. Words with very low ratings (in the lowest quartile) were eliminated from further consideration as they were unlikely to be known by most of the participants in the reading study. The LF words that received ratings that fell within the upper three quartiles were tested to ensure that the participant population had knowledge of each word's meaning.

Word meaning assessment. A multiple-choice test was administered to 200 participants in large group sessions to assess their knowledge of word meanings. This test contained the HF and LF nouns for which familiarity ratings had been generated (omitting the lowest quartile from the LF words). Each test consisted of

120 test words with three answer choices for each one. The correct answer was either a synonym, or a description of the function of the test word. One of the other alternatives was unrelated and one was related to the correct response. The correct answer was chosen more than 75% of the time for all words that were included in the final stimulus set (see Table 2).

Sentence frames. Sentence frames were created so that any one of the four target words fit into a given sentence. All sentence frames were generated so that neutral context preceded each target word and supporting context followed. Each subject saw each of the 48 experimental sentence frames with one of the four possible target words. Selection of target word condition by sentence frame was counterbalanced across subjects in a Latin square design. (Refer to Table 1 for examples.)

Procedure

When a participant arrived for the experiment, a bite bar was prepared which served to eliminate head movements during the experiment. The participants were informed that they were participating in a reading experiment and were encouraged to read normally for comprehension. Yes/no comprehension questions were asked by the experimenter only on filler sentences to avoid drawing attention to the target words. The eye-tracking system was calibrated for each participant. Each participant read six sentences as part of a practice session.

At the beginning of each trial, a row of 11 boxes appeared on the screen to allow the experimenter to check the calibration. The participant was instructed to look to the far left box when they were ready to begin reading. The experimenter controlled the sentence presentation. The participants were instructed to push a button to end the trial when they had finished reading the sentence. The sentences were presented in a random order. The eye-tracking session lasted between 20 and 30 min. Each participant read 48 experimental sentences and 54 filler sentences. Every participant performed at 90% accuracy or better on the comprehension questions.

Apparatus

Eye movements of the participants were recorded using a Fourward Technologies Dual Purkinje Image eye movement monitoring system. Viewing was binocular, with eye location recorded from the right eye. The eye-tracking system was interfaced with a Dell computer that controlled stimulus display and data storage. The sentences were presented on a single line with up to 72 characters per line and 4 character spaces per degree of visual angle on a 17-inch Viewsonic monitor. The eye tracker sampled the position of the participant's eye every millisecond with a resolution of 10 min of arc. All characters with the exception of the first letter of proper nouns and the first letter of the first word in a sentence were presented in lower case. Fixations and saccades were computed online and stored on disk for analysis.

Results

Individual fixations shorter than 120 ms or longer than 800 ms were excluded from the analyses. If participants did not fixate directly on the target word, a fixation occurring within three characters to the left of the word was counted as a fixation for the target word. Trials in which there were no fixations on the target word or within three character spaces to the left of the target word were eliminated from analysis. Approximately 2% of the data were lost due to track loss. The data loss was equally distributed across conditions.

The experimental conditions were not orthogonal with respect to frequency and familiarity (see Table 2) so planned comparisons were performed. To test for processing time differences due to estimates of printed word frequency, the two HF conditions were compared to the two LF conditions. To test for differences due to subjective familiarity, the very familiar and familiar HF conditions were compared to one another and the familiar and less familiar LF conditions were compared to one other. In order to test for possible interactions between printed estimates of word frequency and subjective ratings of familiarity, the HF familiar condition was compared to the LF familiar condition.

Initial processing time on the target word. Single fixation duration, first fixation, gaze duration, and spill over were used to measure initial processing time on the target word. The single fixation duration is the amount of time spent when readers make only one fixation on the target word. Single fixation duration data are included in each table for completeness, but no analyses will be discussed because the results are entirely consistent with first fixation duration data. The first fixation duration is the amount of time spent in the reader's first fixation on the target word, independent of the total number of fixations made on the word. Gaze duration is calculated as the sum of all consecutive fixation durations on a target word from the first fixation until the first time that the reader leaves the word. Spill over is calculated as the duration of the fixation immediately following a reader's first pass fixations on the target word (e.g., the duration of fixation n+1). The number of first pass fixations and the landing position of the first fixation were also calculated.

Table 3 shows that readers spent more initial processing time on the LF words than on the HF words. This was tested by comparing the mean of the two HF word conditions to the mean of the two LF conditions. First fixation duration on LF words (313 ms) was also longer than on HF words (286 ms), $F_1(1,23)$= 16. 10, MSE=1152, p=.001; $F2(1,47)$=14.57, MSE=2119, p=.001. The same pattern held for gaze duration. Gaze durations were longer for LF words (388 ms) than for HF words (321ms), $F_1(1,23)$=18.81, MSE=5636, p=.001; F_2 (1, 47)=23.87, MSE=7069, p=.001. There was no frequency effect in the spillover data (Fs<1). Readers made more first pass fixations on LF words (1.28) than on HF words (1.

10), $F1(1,23)=12.96$, $MSE=0.05$, $p=.002$; F_2 $(1,47)=12.13$, $MSE=0.12$, $p=.001$. Initial landing position did not differ as a function of word frequency (all $F_s<1$).

There were effects due to differences in subjective familiarity ratings only in the LF word conditions. First fixation durations on LF-less familiar words were longer than LF-familiar words, $F_1(1,23)=10.00$, $MSE=1935$, $p=.004$; F_2 $(1, 47)=7.94$, $MSE=4526$, $p=.01$. This pattern was also consistent in gaze duration, $F_1(1, 23)=23.28$, $MSE=4097$, $p=.001$; F_2 $(1, 47)=20.46$, $MSE=$ 9759, $p=.001$, and in spill over, $F_1(1,23)=4.72$, $MSE=1731$, $p=.04$; F_2 $(1,47)=3.81$, $MSE=4130$, $p=.06$. The number of first pass fixations on the target word was greater for LF—less familiar words than LF-familiar words, $F_1(1, 23)=11.65$, $MSE=0.08$, $p=.002$; F_2 $(1,47)=9.74$, $MSE=0.20$, $p=.003$. In addition, an analysis of covariance was performed on all low frequency items, with word frequency as a covariate. The effect of familiarity remained in first fixation, $F(1, 23)=12.95$, $p<.001$, gaze duration, $F(1, 23)=18.65$, $p<.000$, spill over, $F(1, 23)=5.10$, $p<.03$, and number of first pass fixations, $F(1, 23)= 16.92$, $p<.000$, even when word frequency was factored out.

Comparisons between HF-very familiar and HF-familiar words did not reach statistical significance. This held true in first fixation duration, $F_1(1,23)=1.15$, $MSE=547$, $p=.29$; $F_2<1$, gaze duration, $F_1(1,23)=3.31$, $MSE=5250$, $p=.08$; $F_2<13$, spill over, $F_1(1,23)=2.71$, $MSE=3960$, $p=.11$; $F_2<1$, and the number of first pass fixations, $F_1(1,23)=1.33$, $MSE=0.06$; $p=.26$; $F_2<1$. There were no significant differences in comparisons using landing position as a dependent measure (all Fs<1).

Surprisingly, amidst the great differences in printed word frequency (MHF= 102.5, MLF=2.75) there was no difference in initial processing time on HF versus LF words when familiarity was held constant. That is, the time that readers spent on HF-familiar words compared to LF-familiar words did not differ in single fixation, first fixation, or gaze duration (all Fs<1). Given that this result was unexpected, we looked for corroborating evidence from beyond our laboratory. In order to see if there was any evidence of similar results in speeded response tasks we went to the English Lexicon Project database (http:// elexicon.wustl.edu). This database contains behavioural data collected from lexical decision and naming time studies at six colleges and universities. Reaction times from 687 participants and a large set of lexical characteristics corresponding to each word have been catalogued for over 40,000 words. First, we verified that participant's reaction times from this database demonstrated a word frequency effect. There was a significant difference between our HF words and our LF words, $t(46)=-3.715$, $p=.001$. Then we looked up lexical decision times for the HF—familiar words and LF-familiar words from our experiment (see Balota et al., 2001, for a more complete description of the database). Familiarity ratings from this database were consistent with those that we had collected and there was no reliable difference in lexical decision times for the two conditions, $t(22)=1.93$, $p<.18$.

Rereading of the target word. Second pass time and the number of regressions to the target word were used as measures of reanalysis. Second pass time is

TABLE 3

Means for eye movement measures for the target word in Experiment

Condition	Initial processing measures				Reanalysis measures		
	Single fixation duration[a b]	*First fixation*[a]	*Gaze duration*[a]	*Spill over*[a]	*Landing position (character position)*	*Second pass time*[a]	*Regressions in (proportion)*
HF–very familiar	286 (1.06)	282	310	304	2.3	30	9
HF–familiar	294 (1.13)	289	331	322	2.6	36	9
LF–familiar	293 (1.14)	293	343	301	2.4	39	13
LF–less familiar	347 (1.41)	333	432	327	2.5	77	18

[a] Means are expressed in milliseconds.
[b] The number of first pass fixations are included in parentheses.

calculated as the amount of processing time spent on the target word after exiting from that word and returning. Regressions in are measured as the number of looks back to the target word after the reader's initial encounter with the word had ended.

The rereading measures in Table 3 yielded the same pattern of results as the initial processing time measures. Readers reread LF words (16) more often than HF words (9), $F_1(1, 23)=6.40$, $MSE=143$, $p=.02$; $F_2(1, 47)=14.18$, $MSE= 131$, $p=.001$, and they spent more time rereading LF words (40 ms) than HF words (33 ms), $F_1(1, 23)=10.15$, $MSE=1518$, $p=.01$; $F_2(1, 47)=14.23$, $MSE =3267$, $p=.001$. There were no subjective familiarity effects in the HF conditions, in number of regressions to the target word, $F_1(1, 23)=1.75$, $MSE= 188$, $p=.20$; $F_2(1, 47)=1.51$, $MSE=270$, $p=.23$, or in second pass time (both Fs < 1).

Second pass times were greater on LF-less familiar words than on LFfamiliar words, $F_1(1, 23)=7.67$, $MSE=2112$, $p=.01$; $F_2(1, 47)=2.95$, $MSE= 8377$, $p=.09$. Readers did not make more regressions into LF-less familiar words than LF-familiar words, $F_1(1, 23)=1.83$, $MSE=121$, $p=.19$; $F_2(1, 47) =2.08$, $MSE=367$, $p=.16$. And, there were no differences in the amount of time or the frequency of rereading in the HF-familiar word condition compared to the LF-familiar word condition on any measure (all Fs<1).

In order to control for any differences due to word frequency in the comparisons between LF-familiar and LF-less familiar words an analysis of covariance was performed on all low frequency items, with word frequency as a covariate. Again, there was a main effect of familiarity in second pass time, $F(1, 23)=9.03$, $p<.004$. The main effect observed in regressions in was marginal, $F(1, 23)=2.89$, $p<.09$.

Discussion

The overall pattern of results from this experiment was similar whether we looked at objective word frequency counts or subjective familiarity ratings. That is, HF words were processed more quickly than LF words and more familiar words were processed more quickly than less familiar words. These patterns hold whether one looks at initial processing time or rereading measures. However, the differences that emerge in direct comparisons between frequency and familiarity effects are quite striking and suggest that familiarity ratings may reflect something more than a pure measure of frequency of occurrence. This issue will be addressed in the general discussion.

EXPERIMENT 2

In this experiment we explored the influence of the quality of word meaning information on word processing and text comprehension during reading. We focused our investigation on LF words that varied in subjective familiarity rating and in the extent of participants' knowledge of word meaning. Since we were particularly interested in how readers acquire meaning for unfamiliar words in reading, a novel word condition was included in which readers had no prior knowledge of target word meanings.

Experiment 1 revealed unique effects of subjective familiarity ratings in silent reading. Balota et al. (2001) suggested that subjective familiarity ratings are based on a different composition of sources of information depending on the frequency of the word under consideration. Their regression analysis of a large sample of subjective familiarity ratings indicated that while word frequency was the best predictor of high familiarity ratings, meaningfulness was a better predictor of low familiarity ratings. Yet, Chaffin et al. (2001) found that novel words were read similarly at initial encounter to actual words that had received low ratings on a familiarity-rating task.

There are several reasons why readers may not have distinguished between low familiar words and novel words at initial encounter in the Chaffin et al. (2001) study. It is possible that the task demands of lexical decision and of reading comprehension are different enough to yield different effects. But, it is also possible that the low familiar and novel words were too similar to each other. The low familiar words used in that study received very low familiarity ratings (M=2.95) and there was no test of participant's level of meaning knowledge for the low familiar words. Connine et al. (1990) suggested that participants are not likely to have well-developed knowledge of the meanings for words at this level of familiarity. However, it is important to remember that readers also made more regressions to, and spent more total time on, novel words than low familiar words in the Chaffin et al. study. The differences demonstrated in the rereading measures suggest that the low familiar words were not completely unknown to the readers, but that more effort was required to generate a new word meaning than to elaborate upon an existing meaning.

Chaffin et al. (2001) presented converging data from three different eye movement measures obtained from three critical regions of the text that suggest that their readers were generating meanings for the novel words on-line. Although these data are quite compelling, they do not provide any evidence of readers' retention of new word meaning beyond the reading experience, nor was there direct evidence regarding how the observed reading patterns were related to successful development of word meaning.

In Experiment 2, we explored the effect of the quality of the meaning representation of individual words in reading. Familiar and less familiar words were selected on the basis of printed word frequency, familiarity ratings, and a vocabulary test score. Scores from a multiple-choice vocabulary test allowed us to separate our words into the three levels of word knowledge. If readers were able to respond accurately to vocabulary questions that tested detailed meaning, a word was placed in the high knowledge category. In contrast, if readers were only able to respond accurately to vocabulary questions that tested general semantic category knowledge, words were placed in the partial knowledge category. All words placed in the unknown category lacked meaning representations. These were words created by the experimenters.

Participants' eye movements were monitored as they read sentences like those displayed in Table 4. Each sentence contained one of four types of target words (in bold). Target words were categorised either as familiar, less familiarhigh knowledge, less familiar-partial knowledge, or unfamiliar (novel) words, respectively. Target words were preceded by neutral context and followed by context that was highly informative as to the meaning of the target.

Previous studies of vocabulary acquisition in reading have included tests of word meaning knowledge that were administered after participants' read narrative passages that contained unfamiliar words (e.g., Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987; Nagy et al., 1985). Performance on these posttests has revealed that readers are able to infer meanings of unfamiliar words from context. However, these measures provide no information about the processing patterns that readers employ to associate unfamiliar words with meaningful information within a text. In order to gain a better understanding of this aspect of vocabulary acquisition we examined the relationship between the pattern of processing time on the target word and the informative context region and later retention of new word meaning.

Method

Participants

Twenty-four students from the University of South Carolina were given either psychology course credit or $5.00 for participating in this experiment. All participants were native speakers of English and had normal uncorrected vision.

Stimuli

There were four target word conditions in this experiment: familiar, less familiar-high knowledge, less familiar-partial knowledge, and unfamiliar (novel). Target words were selected on the basis of subjective familiarity rating, printed word frequency, and quality of meaning knowledge (see Table 5 for examples and the means for each of the selection criteria for each condition).

Only nouns with frequency counts less than or equal to 10 counts per million were included (Francis & Kucera, 1982). The familiarity ratings that were collected in Experiment 1 were used to further select 12 appropriate target words for each condition (see Table 5). The familiarity ratings of the words in our less familiar-high knowledge and less familiar—partial knowledge conditions are comparable to conditions used by Connine et al. (1990) and higher than those in the low familiar condition of Chaffin et al. (2001). In addition, 12 pseudowords

TABLE 4

Example sentences for each condition in Experiment 2 (target word in bold)

	Condition	*Example sentence*
Example 1	Familiar	Sheila selected the **blossom** for her bouquet.
	Less familiar–high knowledge	Sheila selected the **azalea** for her bouquet.
	Less familiar–partial knowledge	Sheila selected the **poppy** for her bouquet.
	Unfamiliar	Sheila selected the **masdor** for her bouquet.
Example 2	Familiar	After recess the **diary** will be read to the children.
	Less familiar–high knowledge	After recess the **saga** will be read to the children.
	Less familiar–partial knowledge	After recess the **parable** will be read to the children.
	Unfamiliar	After recess the **nineer** will be read to the children.
Example 3	Familiar	Jim said that the **wolf** was killed for its fur.
	Less familiar–high knowledge	Jim said that the **camel** was killed for its fur.
	Less familiar–partial knowledge	Jim said that the **stag** was killed for its fur.
	Unfamiliar	Jim said that the **boser** was killed for its fur.

A complete set of stimuli are available upon request.

were constructed to serve as unfamiliar words. These novel words were orthographically regular and pronounceable letter strings.

With the exception of unfamiliar words, all words were subjected to the meaning assessment test described in Experiment 1 (see Table 5 for mean scores). Words in the less familiar-partial knowledge condition received mean scores below 75%. They were then subjected to a second multiple-choice test to test whether participants could place these words in a semantic category. The composition of the alternative answer choices provided on this test differed from the first meaning assessment test. In the earlier test, participants selected the answer that best represented the meaning of the word from a set that included the correct answer, a related but incorrect answer, or an unrelated incorrect answer. In the second test, the alternatives were the correct answer and two unrelated incorrect answers. In all cases, the relationship between the word being tested and the correct answer was the name of the semantic category in which the word is a member. The test was administered to a group of 50 participants. Even though participants could not select a very specified meaning for less familiar-partial knowledge words on the vocabulary test described in Experiment 1, we took the fact that they could classify less familiar-partial knowledge words into a semantic category (see Table 5 for mean score) as evidence that readers would activate some existing meaning knowledge when they encountered these words in sentences.

Sentence frames. Twelve sets of four target words were formed. One word came from each of the four familiarity conditions and all words within a set were semantically related to each other. Four sentence frames were created for each set of four words so that any one of the words would make sense in any of the sentences. Each subject saw each target word in a different sentence frame. Combinations of target word and sentence frame were counterbalanced across subjects. All sentence frames were generated so that each target word was preceded by neutral context and followed by context that was highly informative as to the intended meaning of the word. The sentence frames were normed for target word fit by asking a group of 40 participants to complete the sentence by filling in a single word on a blank line in the target word position. A sentence frame was used as an experimental item if the majority of the responses on this test were semantically related to the target words and no single word constituted the majority of the responses. That is, the sentence frame was not predictive of a particular response. There were 48 experimental items. (Refer to Table 4 for examples.)

Procedure

Each participant read 48 experimental sentences and 32 filler sentences. All other aspects of the procedure were identical to that employed in Experiment 1, with the following exception. After the reading session, each participant was asked to stay to complete a vocabulary test. All participants complied with this request. At that point the participant was moved to another room with a flat writing surface and a 30-item synonym test was administered in order to assess the readers' knowledge of novel word meanings after the reading session. The test included the 12 novel words seen in the experiment and 18 high familiar words that were not included in the reading portion of the experiment. There were two answer choices for each novel word item. For example, the novel word "masdor", supported by context that conveyed information related to flowers, would have corresponding answer choices like choice A and choice B below. Answer choice A was a word that was similar to the item in meaning. The alternative, answer choice B, was constructed to be distinct from the meaning of the item.

Choice A: flower
Choice B: jewel

The word that was provided as a correct answer choice had not been presented in conjunction with the target word during the reading portion of the experiment.

Apparatus

The apparatus was the same as that used in Experiment 1.

TABLE 5

Target word selection criteria in Experiment 2

Condition	*Example item*	*Printed word frequency*	*Familiarity rating*	*Word length*	*Experiment 1 meaning assessment*	*Semantic category knowledge*
Familiar	Blossom	4.3 (1–10)	6.7 (6.6–6.9)	5.5 (4–7)	97% (91–100%)	—
Less familiar–high knowledge	Azalea	4.6 (1–10)	5.9 (5.3–6.2)	5.6 (4–7)	89% (78–100%)	—
Less familiar–partial knowledge	Poppy	5.3 (1–9)	5.4 (4.9–5.9)	5.2 (4–7)	67% (39–73%)	89% (77–96%)
Unfamiliar	Masdor	—	1.8 (1.4–2.3)	5.2 (4–7)	—	—

The range of values contributing to each mean are reported parenthetically.

Results

Individual fixations shorter than 120 ms or longer than 1200 ms were excluded from the analyses. If participants did not fixate directly on the target word, a fixation occurring within three characters to the left of the word was counted as a fixation for the target word. Approximately 2% of the data were lost due to track loss. The data loss was equally distributed across conditions.

Initial processing time on the target word. Single fixation duration, first fixation, gaze duration, spill over, number of first pass fixations, and landing position were used to measure initial processing on the target word. All measures were computed in the same manner as in Experiment 1. Single fixation duration data are included in each table for completeness, but no analyses will be discussed because the results are entirely consistent with first fixation duration data.

The data for first fixation, gaze duration, and spill over in Table 6 show that readers could distinguish between unfamiliar and less familiar words at initial encounter. Unfamiliar words were read more slowly than less familiar-high knowledge words, in first fixation duration, $F_1(1, 23)=27.84$, $MSE=1407$, $p< .0001$; $F_2(1, 47)=8.88$, $MSE=5957$, $p<.005$, and also more slowly than less familiar-partial knowledge words, $F_1(1, 23)=9.43$, $MSE=3738$, $p<.004$; $F_2(1, 47)=15.52$, $MSE=1801$, $p<.0007$. Gaze durations were longer for unfamiliar words compared to less familiar-high knowledge words, $F_1(1, 23)= 32.93$, $MSE=4467$, $p<.0001$; $F_2(1, 47)=18.95$, $MSE=12, 110$, $p<.0001$, and also compared to less familiar-partial knowledge words $F_1(1, 23)=26.43$, $MSE =4139$, $p<.0001$; $F_2(1, 47)=20.72$, $MSE=14,977$, $p<.0001$. This pattern held in spill over when unfamiliar words were compared to less familiar-high knowledge words, $F_1(1,23)=5.43$, $MSE=6373$, $p<.02$; $F_2(1, 47)=11.44$, $MSE=2156$, $p<.003$, and compared to less familiar-partial knowledge words $F_1(1, 23)=5.84$, $MSE=6872$, $p<.02$; $F_2(1, 47)=7.37$, $MSE=3691$, $p<.01$.

Readers were also sensitive to familiarity within the familiar target word conditions. The data for first fixation duration, gaze duration, and spill over presented in Table 6 show that readers spent less time on familiar words than either of the less familiar conditions. This pattern was most clearly expressed in gaze duration such that processing time on familiar words was less than less familiar-high knowledge, $F_1(1, 23)=6.04$, $MSE=3017$, $p<.02$; $F_2(1, 47)= 10.28$, $MSE=6579$, $p<.002$, or less familiar-partial knowledge words, $F_1(1, 23)=4.77$, $MSE=5677$, $p<.04$; F_2 $(1, 47)=3.61$, $MSE=9176$, $p<.06$. The difference between means were numerically different in first fixation duration but did not reach statistical significance when familiar words were compared and less familiar high-knowledge words, $F_1(1, 23)=1.88$, $MSE= 2192$, $p<.18$; $F_2(1, 47)=2.26$, $MSE= 3951$, $p<.14$, or less familiar-partial knowledge words, $F_1(1, 23)=3.83$, $MSE=2357$, $p<.06$; $F_2(1, 47)=5.36$, MSE $=3486$, $p<.03$. This was true for differences in spill over between familiar words and less familar-high knowledge words, $F_1(1, 23)=3.48$, $MSE=1271$, $p <.07$; $F_2(1, 47)=1.53$, $MSE=2567$, $p<.22$, and for less familiar-partial knowledge words as well, $F_1(1, 23)=2.41$, $MSE=1424$, $p<.13$; $F_2<1$.

An additional measure of initial processing of the target word, number of first pass fixations, reflected the same pattern of differences between target words as the measures described above. Unfamiliar words received more first pass fixations than less familiar-high knowledge, $F_1(1, 23)=6.20$, $MSE=0.02$, $p<.02$;F_2 $(1, 47)=3.39$, $MSE=0.09$, $p<.07$, and less familiar-partial knowledge words, $F_1(1, 23)=13.76$, $MSE=0.04$, $p<.001$; $F_2(1,47)=10.27$, $MSE=0.10$, $p<.002$. Readers made fewer fixations on familiar words than on less familiarhigh knowledge words, $F_1(1, 23)=17.15$, $MSE=0.02$, $p<.0004$; $F_2(1, 47)= 10.13$, $MSE=0.06$, $p<.003$, and unfamiliar words, $F_1(1, 23)=22.23$, $MSE= 0.04$, $p<.0001$; $F_2(1, 47)=21.83$, $MSE=0.08$, $p<.0001$. There were no differences in landing position (all Fs<1). There were no significant differences in initial processing time between the less familiar-high knowledge and less familiar-partial knowledge word conditions.

In addition, we entered word frequency as a covariate into analyses that compared our high familiar condition and our less familiar conditions. The observed effect of familiarity remained in first fixation duration, $F(1, 35)=4.66$, $p=.03$ and gaze duration, $F(1, 35)=1.919$, $p=.03$, and the number of fixations, $F(1, 35)=5.297$, p=.02. The main effect of familiarity in spill over was marginal, $F(1, 35)=3.21$, $p=.08$.

Reanalysis of the target word. Second pass time and regressions were used to assess reanalysis of the target word. The degree to which patterns of processing seen in unfamiliar vs. familiar word conditions diverged became more pronounced in reanalysis measures (see Table 6). More rereading time was spent on unfamiliar words than less familiar-high knowledge, $F_1(1, 23)=18.35$, $MSE=5738$, $p<.0003$; $F_2(1,47)=18.63$, $MSE=11{,}512$, $p<.0001$, or less familiarpartial knowledge words, $F_1(1,23)=13.09$, $MSE=6248$, $p<.001$; $F_2(1, 47)= 19.81$, $MSE=8410$, $p<.0001$. In addition, there were more regressions into the unfamiliar target word region than less familiar—high knowledge words, $F_1(1, 23)=24.85$, $MSE=109$, $p<.0001$; $F_2(1, 47)=23.90$, $MSE=200$, $p< .0001$, and less familiar-partial knowledge words, $F_1(1, 23)=6.71$, $MSE=169$, $p<.02$; $F_2(1,47)=9.68$, $MSE=221$, $p<.003$, replicating the result obtained in Experiment 1. Less familiar-high knowledge and less familiar-partial knowledge words did not differ from each other in either measure (Fs<1). Again, we verified that word frequency differences did not account for the observed familiarity effects by entering word frequency as a covariate in comparisons between high familiar words and familiar words in second pass, $F(1, 35) = 11.00$, $p=.002$, and regressions in, $F(1, 35)=2.956$, $p=.09$.

Conversely, there was less rereading time spent on familiar words than on less familiar-high knowledge, $F_1(1,23)=6.80$, $MSE=1438$, $p<.02$; F_2 $(1, 47)= 4.64$, $MSE=4051$, $p<.04$, or less familiar-partial knowledge words, $F_1(1,23)= 12.39$, $MSE=1524$, $p<.002$; $F_2(1, 47)=12.18$, $MSE=3024$, $p<.001$; there were fewer

TABLE 6

Means for eye movement measures on the target word in Experiment 2

	Initial processing measures				*Reanalysis measures*		
Condition	*Single fixation duration*[a]	*First fixation*	*Gaze duration*	*Spill over*	*Landing position*	*Second pass time*	*Regressions in*
Familiar	294 (0.89)	291	326	285	2.3	29	10
Less familiar–high knowledge	320 (1.05)	309	365	304	2.5	57	12
Less familiar–partial knowledge	318 (0.96)	318	374	302	2.4	68	17
Unfamiliar	362 (1.16)	366	476	350	2.4	150	27

[a]The number of first pass fixations are included in parentheses.

TABLE 7

Means for eye movement measures in the context region in Experiment 2

	Initial processing measures		*Reanalysis measures*
Condition	*First pass time*[a]	*Regressions out (proportion)*	*Total time*[a]
Familiar	34	9	40
Less familiar–high knowledge	39	13	44
Less familiar–partial knowledge	39	15	46
Unfamiliar	39	21	49

[a]Means are expressed in milliseconds per character to control differences in the size of the region.

regressions made into familiar words than less familiar-partial knowledge words, $F_1(1,23)=4.44$, $MSE=134$, $p<0.5$; $F_2(1,47)=6.11$, $MSE=193$, $p<.02$

Initial use of context. First pass time and the proportion of regressions out of the context region were used as initial processing time measures. First pass time is calculated as the cumulative processing time on the context region from the first entry to the first exit from the region. This measure is equivalent to the initial processing time measure "gaze duration" but applied to a multiword region. The context region began at the first word following the target word and extended until the last letter of the sentence final word. Approximately 1% of the data were lost due to track loss.

As can be seen in Table 7, there was shorter first pass time on the context following familiar words than after less familiar-high knowledge, $F_1(1,23) = 6.88$, $MSE=194$, $p<.02$; $F_2(1, 47)=4.12$, $MSE=56$, $p<.05$, or less familiar-partial knowledge words, $F_1(1, 23)=4.28$, $MSE=305$, $p<.05$; F_2 (1, 47) =5.35, $MSE=41$, $p<.03$. However, readers did not differ in the amount of first pass time on the context region following unfamiliar words and either less familiar-high knowledge or less familiar-partial knowledge words (Fs<1).

The most notable initial effect of informative context was the extent to which this information prompted readers to make regressions (see Table 7). Readers made more regressions out of this region following unfamiliar words than less familiar-high knowledge words, $F_1(1, 23)=4.49$, $MSE=1748$, $p<.05$; $F_2(1, 47)=5.38$, $MSE=262$, $p<.02$, or less familiar-partial knowledge words (although this effect was only marginal by participants), $F_1(1, 23)=3.07$, $MSE=1684$, $p<.09$; F_2 (1, 47)=4.00, $MSE=219$, $p<.05$. In addition, there were fewer regressions out of the context region following familiar words than less familiar-high knowledge, F_1 (1, 23)=4.74, $MSE=702$, $p<.04$; $F_2(1, 47)= 2.27$, $MSE=239$, $p<.14$, or less familiar-partial knowledge words, $F_1(1,23)= 4.73$, $MSE=1169$, $p<.04$; $F_2(1, 47)=4.50$, $MSE=211$, $p<.04$. *A main effect* of familiarity remained when word frequency was entered as a covariate in comparisons between high familiar words and familiar words in first pass, $F(1, 35)=7.516$, $p=.007$.

Reanalysis of the context. Total time was used to measure rereading of the context region. Total time reflected the amount of time that readers spent in the context including initial processing and rereading. Table 7 shows that readers spent more total time in the context region following unfamiliar words than following less familiar-high knowledge, $F_1(1, 23)=11.38$, $MSE=361$, $p<.003$; $F_2(1, 47)=12.17$, $MSE=57$, $p<.001$, and less familiar-partial knowledge words, $F_1(1, 23)=4.35$, $MSE=356$, $p<.05$; $F_2(1, 47)=4.84$, $MSE=55$, $p<.03$. There was no difference in the total time spent on the context following familiar-high knowledge and familiar-partial knowledge words, $F_1(1, 23)=1.85$, $MSE=356$, $p<.18$; $F_2(1,47)=2.79$, $MSE=55$, $p<.11$. Readers spent less total time following familiar words than less familiar-high knowledge, $F_1(1, 23)=15.42$, $MSE=146$, $p<.0001$; $F_2(1, 47)=8.94$, $MSE=43$, $p< 004$, less familiar-partial knowledge, $F_1(1, 23)= 18.30$, $MSE=288$, $p<.0003$; $F_2(1, 47)= 18.69$, $MSE=47$, $p<.0001$, and unfamiliar words, $F_1(1, 23)=42.52$, $MSE=294$, $p<.0001$; $F_2(1, 47)= 38.55$, $MSE=55$, $p<.0001$. A main effect of familiarity remained when word frequency was entered as a covariate in comparisons between high familiar words and familiar words in total time, $F(1, 35)=11.085$, $p=.001$.

Synonym test. Twenty-one, out of twenty-four participants, scored above chance on the unfamiliar word items on a test of memory for the intended meaning of the novel words. When performance on individual items was averaged across readers, the mean accuracy for unfamiliar words was 62%. Results of a t-test revealed that the accuracy of unfamiliar words was significantly different from chance, $t(11)=2.206$, $p=.05$. This is consistent with the level of retention of word meaning from a single reading encounter that was observed by Nagy et al. (1985, 1987) for school-aged readers. We also observed a relationship between online reading patterns and posttest performance requiring memory for meanings of the unfamiliar words. A separate one-way analysis of variance was performed on gaze duration and on second pass time on the unfamiliar words using accuracy on the synonym test as the independent measure. Readers spent more initial processing time on unfamiliar words which they later answered incorrectly on the synonym test than on unfamiliar words which they answered correctly, $F(1, 248) = 6.65$, $p< .01$. The mean gaze duration for inaccurate responses was 533 ms and the mean for accurate responses was 433 ms. The reverse pattern occurred in second pass time such that readers spent more rereading time on unfamiliar words that were answered accurately than on inaccurate unfamiliar words, $F(1, 285) = 4.01$, $p<.05$. The mean second pass time for accurate responses was 177 ms and the mean for inaccurate responses was 108 ms.

Discussion

The basic familiarity effects observed in Experiment 1 were replicated. Familiar words were processed more quickly than less familiar words, which were read more quickly than novel words. Although the norming data from the meaning assessment task indicated that the quality of meaning representations for the less familiar-partial knowledge condition and less familiar-high knowledge condition differed, these differences did not influence initial processing or rereading time.

Readers also spent less time processing context following high familiar words than moderately familiar words.[1] Familiarity continued to influence readers' behaviour in incorporating words into the sentence. Readers spent the greatest amount of time rereading the informative context that followed unfamiliar words. More total time was spent in the context and a greater proportion of regressions were made out of the context following unfamiliar words than moderately familiar words and familiar words.

Accuracy on the synonym task indicated that readers successfully acquired the intended meaning for the novel words and that they were able to retain that knowledge beyond the duration of the reading session. In conjunction with the reading data, the performance on the synonym task provides evidence that reading patterns are directly related to the development of the novel word's meaning.

GENERAL DISCUSSION

In Experiment 1 familiarity effects were assessed with corpus-based estimates of printed word frequency, subjective ratings of word familiarity, and a vocabulary test. Overall, there was a strong relationship between the effects of word frequency and subjective word familiarity on word processing in reading. That is, readers spent less time on HF words than on LF words and, likewise, spent less time on words that were rated as more familiar than on words that were rated as less familiar. These findings replicate the results of earlier reading studies that examined effects of word frequency (e.g., Inhoff, 1984; Rayner & Duffy, 1986) and word familiarity (Chaffin et al, 2001). In addition, the direct comparison of word frequency and subjective familiarity revealed that readers spent less time on a more familiar LF word than on a less familiar word of equal frequency. This extends previous findings (e.g., Connine et al., 1990; Gernsbacher, 1984) from speeded naming tasks to fluent silent reading. However, direct comparisons of the influence of estimates of printed word frequency and subjective familiarity also revealed the unique finding that processing time did not differ on HF versus LF words when familiarity ratings were held constant in the mid range.

The overall relation between effects of subjective familiarity rating and objective frequency estimates suggests that participants have access to relatively pure information about frequency of occurrence and reflect that in their fami liarity ratings. However, the differences observed among LF words as a function of familiarity rating suggests that these ratings are influenced by other factors, at least in the case of LF words.[2] Balota et al. (2001) demonstrated that meaningfulness was a better predictor of participants' subjective ratings of word frequency for less common words, while objective frequency estimates were a better predictor of participants' ratings of more common words. Other research has demonstrated that meaning-level characteristics of a word can affect processing time on words in silent reading. For example, readers spend more initial processing time on lexically ambiguous words with two relatively equally likely meanings than on words with a highly dominant meaning when the context preceding the word is unbiased (see Rayner & Morris, 1991, for a review).

The data reported here is consistent with this explanation with one possible exception. If meaningfulness influences processing time, then why did we not observe any processing time differences between low familiar-high knowledge words and low familiar-partial knowledge words? These words were equated in word frequency and subjective familiarity ratings and were moderately familiar. Balota et al. (2001) showed that as familiarity decreases the relative weight of influence of frequency and meaningfulness shifts. That is, as familiarity decreases the influence of meaningfulness increases. So it may be that at the moderate level of familiarity employed in Experiment 2, the influence of meaningfulness is too weak to detect.

An alternative (or perhaps additional) explanation is suggested in a recent report by Paap, Johansen, Chun, and Vonnahme (2000). Their work suggests that differences in task demands between the forced choice vocabulary test and silent reading for comprehension may have played a role here. Paap et al. demonstrated that semantic characteristics become particularly important in word recognition when LF words are presented out of context. According to this account, subtle differences in meaningfulness were revealed when the words were presented in isolation in our vocabulary tests, but these effects did not translate into processing time differences in reading the words in sentence contexts.

The second experiment focused on LF words that varied in subjective familiarity: Familiar, less familiar, and unfamiliar or novel words were used. In this experiment readers spent more initial processing time on novel words than on familiar words and they spent less initial processing time on familiar than on less familiar known words. These results differed from those reported by Chaffin et al. (2001), in that Chaffin and colleagues failed to observe initial processing time differences between low familiar and novel words. However, the current experiment differed from their study in several critical ways. By using both familiarity rating data and vocabulary test performance we were assured that our

---

[1] Throughout the remaining discussion "moderately familiar" refers to both our less familiarhigh knowledge and less familiar-partial knowledge conditions since there were no statistically significant differences between the two conditions.

readers had some knowledge of the meaning of the low familiar words. In contrast, Chaffin et al. relied solely on familiarity ratings for selection of low familiar words and the familiarity ratings were considerably lower than those employed in the current study. Our vocabulary test data indicated that the majority of the college-age readers tested did not have basic knowledge of word meaning for words that received familiarity ratings comparable to Chaffin et al. in our norming task.

There was no evidence in the present experiments that readers abandoned normal word processing strategies in their initial processing of unfamiliar words as compared to familiar words. Although the number of first pass fixations is greater for unfamiliar word than for familiar word conditions, there were no differences in initial landing position between these conditions. This is consistent with comparisons of the processing patterns of skilled readers looking at HF versus LF words (Rayner et al., 1996) and with comparisons of beginning versus skilled readers (McConkie, Zola, Blanchard, & Wolverton, 1991). Our data pattern contrasts with data patterns found by Just and Carpenter (1980). They observed the eye movement patterns of skilled readers as they encountered words they were unlikely to know, such as “thermoluminence” and “staphylococci” in sentence contexts and obtained gaze durations that ranged from 913 ms to well over 2 s. The difference in the magnitude of the effects obtained in the two studies may be due in part to differences in word length. However, given that the average gaze durations for their unfamiliar words were more than three times as long as the average gaze duration in our novel word condition, it is doubtful that word length can fully account for the difference. There were also potentially significant morphological differences between the stimuli used in the two studies. The words in the Just and Carpenter study combined familiar morphological constituents to create complex, morphologically transparent words. Thus, it may be that readers in the Just and Carpenter study were parsing the unfamiliar words into morphological constituents and then essentially problem solving their way to a meaning for the word (see Sternberg & Powell, 1983, for a similar view). Consistent with this explanation others have shown that school-aged children make use of morphological information in determining word meaning during the course of reading (e.g., Bertram, Laine, & Virkkala, 2000; Mori & Nagy, 1999). In contrast, none of the words used in the experiment reported here were

---

[2] We draw conclusions based only upon the processing differences observed in our LF conditions. However, the data do not rule out the possibility of familiarity effects within HF words as well. It should be noted that the difference in familiarity rating between our HF word conditions was smaller than in our LF conditions due to constraints imposed by other selection criteria. It is possible that if this difference were increased effects of familiarity would be observed in the HF conditions as well. Although this is theoretically conceivable, it is difficult to implement due to the compressed use of the familiarity rating scale typically seen in response to HF words (see also Connine et al., 1990).

morphologically transparent. Therefore, it is unlikely that our readers would resort to such a strategy.

There are several sources of converging evidence that suggest that readers were making connections between the informative context and the unfamiliar word. The first comes from examining reading patterns in the target word and informative context regions. Readers spent more total time in the informative context following a novel word than following a familiar word. They were also prompted to make more regressions out of the informative context region following novel words than familiar words. In addition, readers were more likely to make regressions to the target word in the novel word condition than in any of the other three conditions. Finally, given that readers did go back to the target word, they spent more time there if the word was a novel word than if it was a known word. This pattern of results suggests that readers are using the contextual information provided to them in order to develop word meaning online as they read (see Chaffin et al., 2001, for a similar pattern of results).

The vocabulary test data provided evidence that the meanings that readers inferred were in fact the intended meanings for the novel words and that they were able to retain this knowledge for a period of time beyond the completion of the reading session. On average the time elapsed between reading a word and seeing the word as a test item was approximately 20–30 min. Readers were not told in advance that the experiment was related to vocabulary acquisition, nor were they told that there would be a test after the reading session. Performance was greater than chance for eight out of the twelve novel word items. This level of retention is consistent with studies that looked at retention for the meaning of novel words after one encounter in adolescent readers using a similar testing procedure (Nagy et al., 1985). It is important to note that there was nothing in the test to indicate to the participant which interpretation was correct unless they had some recollection of the information provided in the sentence reading. The answer choices in the vocabulary test were never words included in the sentence contexts, thereby ensuring that successful choices could not be made on the basis of surface-level matching alone.

In order to investigate the relationship between reading performance and successful acquisition of new word meaning we looked at how the reading patterns coincided with accuracy on our synonym test. Gaze duration and second pass reading time data were reanalysed based on performance on the vocabulary test. Interestingly, readers spent less initial processing time (as measured by gaze duration) on novel words that they later answered correctly than on novel words that they missed on the vocabulary test. This makes sense given that the context that preceded the novel word provided very little information about what the word might mean. Spending additional time on the word at that point would do little to determine its meaning. The context following the word provided information about what the word meant and readers made more regressions to the novel word and devoted more second pass reading time to the novel words that they answered correctly than to the novel words that they answered incorrectly. We take this as evidence that the second pass time on the novel word reflects the process of connecting the newly acquired contextual information to the unfamiliar word. Looking at initial processing time alone, it is not clear why initial processing time on the novel word would differ as a function of vocabulary test performance, and in particular why they spent less time on the words for which they later remembered the intended meaning. The shorter gaze duration on words that were correctly remembered may reflect more efficient metacognitive analysis that the word is unfamiliar prompting the reader to seek informative context. The overall pattern of reading behaviour observed by looking at initial processing and at rereading measures together, suggests that readers were quite flexible in allocating their processing resources to the most informative regions of the text to establish a meaning for the new word.

PrEview proof published online September 2003

REFERENCES

Balota, D.A., Pilotti, M, & Cortese, M.J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition, 29,* 639–647.

Bertram, R., Laine, M., & Virkkala, M.M. (2000). The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology, 41,* 287–296.

Burgess, C., & Livesay, K. (1998). The effects of corpus size in predicting reaction times in basic word recognition tasks: Moving on from Kucera & Francis. *Behavior Research Methods, Instruments, and Computers, 30,* 212–211.

Carroll, J., Davies, P., & Richman, B. (1971). *Word frequency book.* Boston: Houghton Mifflin Company.

Chaffin, R. (1997). Associations to unfamiliar words: Learning the meanings of new words. *Memory and Cognition, 25,* 203–226.

Chaffin, R., Morris, R.K., & Seely, R.E. (2001). Learning new words in context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 225–235.

Connine, C. M, Mullenix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1084–1096.

Dutch Center for Lexical Information. (1995). *The Celex lexical database.* Nijmegen, The Netherlands: Dutch Center for Lexical Information.

Ferraro, F., & Sturgill, D. (1998). Lexical effects and lexical properties associated with National Adult Reading Test (NART) stimuli in healthy young adults and healthy elderly adults. *Journal of Clinical Psychology, 54,* 577–584.

Francis, W., & Kucera, H. (1982). *Frequency analysis of English usage.* Boston: Houghton Mifflin Company.

Gaygen, D., & Luce, P. (1998). Effects of modality on subjective frequency estimates and processing of spoken and printed words. *Perception and Psychophysics, 60,* 465–483.

Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and, polysemy. *Journal of Experimental Psychology: General, 113,* 256–281.

Gilhooly, K.J., & Logie, R.H. (1980). Age of acquisition, concreteness, familiarity, and ambiguity norms for 1,944 words. *Behavior Research Methods and Instrumentation, 12,* 395–427.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language, 24,* 631–645.

Hyönä', J., & Olson, R.K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1430–1440.

Inhoff, A.W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior, 23,* 612–624.

Jenkins, J., Stein, M.L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal, 21,* 767–787.

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

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

McConkie, G.W., Zola, D., Blanchard, H.E., & Wolverton, G.S. (1991). Children's eye movements during reading. In J.F.Stein (Ed.), *Vision and visual dyslexia* (pp. 251–262). London: Macmillan Press.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D.Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Mori, Y., & Nagy, W.E. (1999). Integration of information from context and word elements in interpreting novel kanji compounds. *Reading Research Quarterly, 34,* 80–101.

Morris, R.K., & Williams, R.S. (2003). Bridging the gap between old and new: Eye movements and vocabulary acquisition in reading. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 235–252). Amsterdam: Elsevier Science.

Nagy, W.E., Anderson, R.C., & Herman, P. (1987). Learning new words from context during normal reading. *American Educational Research Journal, 24,* 237–270.

Nagy, W.E., Herman, P., & Anderson, R.C. (1985). Learning new words from context. *Reading Research Quarterly, 19,* 304–330.

Paap, K.R., Johansen, L.S., Chun, E., & Vonnahme, P. (2000). Neighborhood frequency does effect performance in the Reicher Task: Encoding or decision? *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1691–1720.

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

Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14,* 191–201.

Rayner, K., & Morris, R.K. (1991). Comprehension processes in reading ambiguous sentences: Reflections from eye movements. In G. Simpson (Ed.), *Understanding word and sentence* (pp. 175–198). New York: Elsevier Science Publishers.

Rayner, K., & Raney, G.E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin and Review, 3,* 254–248.

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

Schilling, H.E. H., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26,* 1270–1281.

Schmauder, A.R., Morris, R.K., & Poynor, D.V. (2000). Lexical processing and text integration of function and content words: Evidence from priming and eye fixations. *Memory and Cognition, 28,* 1098–1108.

Sternberg, R.J., & Powell, J.S. (1983). Comprehending verbal comprehension. *American Psychologist, 38,* 878–893.

Thorndike, E., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University, Teachers College, Bureau of Publications.

Whalen, D.H., & Zsiga, E. (1994). Subjective familiarity of English word/name homophones. *Behavior Research Methods, Instruments, and Computers, 26,* 402–408.

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2004, *16*(1/2), 340–352

Eye movements in reading: Old questions and new directions

Keith Rayner and Barbara J.Juhasz

University of Massachusetts, Amherst, USA

Trends in the articles reported in this volume are identified (1) landing position effects, (2) word skipping, (3) parafoveal-on-foveal effects, (4) eye movement control, and (5) eye movements and word identification. Each of these issues is discussed in the context of prior research on the issue. We also identify some issues that are not included in the present set of articles, as well as some research questions that need further attention.

Since the mid-1970s there has been an incredible amount of research conducted examining the characteristics of eye movements in reading. Since the publication in 1975 of two classic articles (McConkie & Rayner, 1975; Rayner, 1975) using eye-contingent display change techniques to examine the perceptual span, numerous other questions about reading have been addressed utilising eye movement measures (see Rayner, 1978, 1998; Starr & Rayner, 2001, for reviews). When Tinker (1958) published the last of his reviews of eye movement research on reading, he felt it was the case that all that could be learned about reading from examining eye movements had been learned. Yet, he was clearly wrong, and he was wrong for two reasons: He did not anticipate (1) the technological advances that yielded (a) better (and more accurate) equipment for monitoring eye movements and (b) high speed computers for stimulus presentation and data analysis; or (2) the theoretical advances in psychology and linguistics that would lead to far better understanding of language. Many laboratories throughout the world are now equipped with eye-tracking equipment and the amount of effort devoted to eye movement research on reading (and other cognitive processing activities) continues to increase.

European researchers have been particularly interested in using eye movements to study reading. Furthermore, it is interesting to note that there is nothing in the

US (or other parts of the world, for that matter) equivalent to the European Conference on Eye Movements (ECEM, which is held every other year). This particular volume is further evidence of the fact that eye movement research on reading is alive and well in Europe. It has been argued elsewhere (see Liversedge & Findlay, 2000; Rayner, 1995; Rayner & Liversedge, in press) that there are two extremes with respect to research on reading using eye movements: At one extreme are those researchers who are primarily interested in studying either eye movements *per se* or questions that are related to perceptual processing during reading, while at the other extreme are those researchers who are primarily interested in using eye movements as a tool to study some aspect of the reading process (such as syntactic parsing or comprehension). In mainland Europe, it has been the case that there has been more of the former type of research than the latter. In the United Kingdom, on the other hand, there is probably a more even split between the two types of research. No value judgement is being made with respect to which type of research is more valuable. Indeed, it is clearly the case that researchers from both extremes need to pay attention to the findings from each other (Rayner & Liversedge, in press).

The articles in the present volume therefore reflect the reality of current research activities in Europe which use eye movements to study reading. As such, there are no articles using eye movements to study higher level language processing, such as sentence parsing or discourse comprehension (which would be more likely to emerge from the UK than mainland Europe). Rather, all of the articles deal with aspects of visual and linguistic word processing, in many cases combined with issues concerning the control of eye movements in continuous reading. As such, they provide an accurate overview of current research activities in much of Europe. Indeed, it is also interesting to note that of the 30 comments on the Reichle, Rayner, and Pollatsek (in press) target article (which compared the E-Z Reader model to other models) in *Behavioral and Brain Sciences,* 21 were from Europe. In the remainder of this article, we will not comment extensively on individual articles. Rather, we will focus our comments more generally on trends evident in the articles. Then, we will discuss a bit about what is missing from the current set of articles and also discuss the issue of old questions and new directions.

Correspondence should be addressed to K.Rayner, Dept of Psychology, University of Massachusetts, Amherst, MA 01003, USA. Email:rayner@psych.umass.eduThe first author was supported by Grant HD 26765 from the National Institute of Health and the second author was supported by Grant MH 16745 from the National Institute of Mental Health. We thank Alan Kennedy and Ralph Radach for their helpful comments.

http://www.tandf.co.uk/journals/pp/09541446.html
DOI: 10.1080/09541440340000385

The articles in the present volume can be divided into five general topic areas: (1) landing position effects (Radach, Inhoff, & Heller; White & Liversedge), (2) word skipping (Drieghe, Brysbaert, Desmet, & DeBaecke), (3) parafoveal-on-foveal effects (Hyönä & Bertram; Kennedy, Murray, & Boissiere; Pynte, Kennedy, & Ducrot; Starr & Inhoff; Vitu, Brysbaert, & Lancelin), (4) eye movement control (Yang & McConkie), (5) eye movements and word processing/identification (Andrews, Miller, & Rayner; Kliegl, Grabner, Rolfs, & Engbert; Williams & Morris). All of these topics have been the focus of a considerable amount of recent research.

Before we discuss each topic, we will make a few general points on which we think most researchers who study eye movements and reading would agree (see Rayner, 1998, for the empirical support for our claims). First, we think it is not at all controversial that there are parafoveal preview effects in reading: When a reader has a preview of the next word in the text, fixations are shorter on the previewed word than when preview is denied. Second, fixation time on a word varies as a function of its frequency and predictability: low frequency words are fixated longer than high frequency words and unpredictable words are fixated longer than predictable words. Third, words are skipped (i.e., not fixated) as a function of their length: Shorter words are skipped more frequently than longer words. These three points serve as points of reference for arguments that we will make. With this as a background, we now turn to each of the five topics that are central to the present volume.

LANDING POSITION EFFECTS

Dunn-Rankin (1978) and Rayner (1979) first demonstrated that where the eyes land in a word is rather systematic. Dunn-Rankin examined words in isolation and Rayner examined words in context. Rayner (1979) found that the eyes tend to land about halfway between the beginning and the middle of an English word (see Deutsch & Rayner, 1999, for similar evidence with Hebrew). Rayner referred to this as the preferred viewing location, and the finding has been replicated many times. Stimulated by research by O'Regan and colleagues (see O'Regan, Lévy-Schoen, Pynte, & Brugaillére, 1984), many European researchers have been fascinated with these landing position effects. Initially, it was believed that landing position effects were due primarily to word length. More recently, Hyönä (1995) found that in comparison to an orthographically regular cluster at the beginning of a word, an orthographically irregular cluster causes the eyes to land further towards the beginning of the word (see also Beauvillain & Doré, 1998; Beauvillain, Doré, & Baudouin, 1996).

The articles by Radach et al. and by White and Liversedge serve to further confirm that orthographic information from the beginning of upcoming words influences where readers initially fixate in those words. Radach et al. also found that lexical and/or morphological processing does not have a modulating effect on where the eyes initially land (so the effect is a low-level one). They further demonstrated that orthographic landing site effects are graded since they used a three-level variation of orthographic regularity and found that the differences between low versus medium regularity words (at the beginning of the word) and between medium and high regularity words were of the same order of magnitude. White and Liversedge likewise show that orthographic familiarity, but not informativeness, of word initial letter sequences influence where words are first fixated.

It is interesting to note that the size of the landing position effect that Radach et al. and White and Liversedge report was quite small (on the order of one-third or one-quarter of a letter). Frankly, our observation is that many researchers outside of the area of eye movement research question how interesting effects this small might be. Our own view is that if the effect is reliable, it is of interest. When researchers first started reporting 25 ms effects on fixation time, such an effect seemed infinitesimally small. Yet, if you add up 25 ms across 100 fixations, the effect soon adds up to a larger effect. It is likewise the case with landing site effects.

WORD SKIPPING EFFECTS

It has long been known that readers do not fixate on each word in the text. Rather, about 30% of the words do not receive a direct fixation. There is good reason to believe that although these words are not fixated, they are still processed (Rayner, 1998). That is, skipping a word is not equivalent to not processing a word. What exactly is the reason that certain words are skipped? Obviously, word length is a factor in that short words are skipped more frequently than longer words (Rayner & McConkie, 1976). But, given that word length is controlled, why are some words skipped more frequently than others? Word predictability influences skipping rates and high frequency words tend to be skipped more than low frequency words particularly when the eyes are near the beginning of the upcoming word (Rayner, Sereno, & Raney, 1996). Brysbaert and Vitu (1998) argued quite strongly that word length was far more important than either predictability or frequency with respect to word skipping.

In their article, Drieghe et al. continue with this proposition and begin by arguing quite strongly that word length is the most important factor in skipping. No one can reasonably argue that word length is not important, but there is also clear evidence that predictability and frequency are both also important (Rayner, 1998). Actually, the experiment that they report in this volume leads them to ultimately conclude that while there is a strong word length effect on skipping behaviour, there is also an effect of contextual constraint. Again, there are simply too many extant studies in which, with word length controlled, an effect of contextual constraint emerges for a purely length account to hold up.

PARAFOVEAL-ON-FOVEAL EFFECTS

Perhaps the most contentious issue in recent research on eye movements in reading is that of parafoveal-on-foveal effects. As clearly explained in the five articles dealing with this issue, parafoveal-on-foveal effects refer to the possibility that characteristics of the word to the right of fixation impacts on the processing of the currently fixated word. Undoubtedly, the reason for why so much work has focused on this issue is because it has been assumed that the demonstration of such effects would be highly damaging to serial attention models of eye movement control such as the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2003; Rayner, Reichle, & Pollatsek, 1998b; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle et al., in press).

At this point, it appears that orthographic properties of the word to the right of fixation can have an effect on the processing of the currently fixated word (and Starr & Inhoff provide further support for this assertion). However, the most recent version of the E-Z Reader model (Reichle et al., in press) can account for this effect. What is more contentious is the extent to which the meaning of the word to the right of fixation can influence the processing of the currently fixated word. Rayner, White, Kambe, Miller, and Liversedge (2003c) have reviewed this question in detail and argued that there is reason for scepticism on this issue. We will not repeat all of the arguments made by Rayner et al. in detail here. However, we will briefly note their main points. First, there is no evidence that preview effects are due to semantic processing. Although readers clearly process a word more efficiently when they have had a preview of that word before fixating it, the basis of the preview effect is not due to semantic priming. Second, when the meaning of the word to the right of fixation is obtained on the current fixation, the parafoveal word is typically skipped on the ensuing saccade. Third, there is evidence that the frequency of the word to the right of fixation does not influence the duration of the current fixation (Carpenter & Just, 1983; Henderson & Ferreira, 1993; Rayner, Fischer, & Pollatsek, 1998a; though see Kennedy, 1998, 2000, for evidence of frequency effects in a reading-like task). Fourth, most demonstrations of parafoveal-on-foveal effects come from experiments in which subjects are not reading, but rather engaged in tasks that may or may not approximate reading. The main concern here is that some of the tasks that have been used seem more like visual search or pattern matching tasks, and there is clear evidence that when subjects engage in such tasks well documented effects in reading (such as the word frequency effect wherein low frequency words are fixated longer than high frequency words) completely disappear (Rayner & Fischer, 1996; Rayner & Raney, 1996). Fifth, even if we grant that parafoveal-onfoveal effects have been demonstrated, they seem to be somewhat fragile and elusive. An important result reported by Murray and Rowan (1998) has proved difficult to replicate (even by the lab that originally demonstrated it, see Murray, in press) and sometimes the effect is one of facilitation and other times it is one of interference. These inconsistencies are good cause for scepticism (though see

Kennedy, Pynte, & Ducrot, 2002 for an attempt to explain the inconsistencies). Furthermore, there are a number of clear failures to find parafoveal-on-foveal effects (see Rayner et al., 2003c).

The articles reported in the present volume, in our opinion, continue the trend for inconsistency with respect to the nature of the effect. Whereas Kennedy et al., Pynte et al., and Vitu et al. provide support for confirmation of parafovealon-foveal effects (and some in the context of reading, rather than reading-liketasks), White and Liversedge find no evidence consistent with the notion. Perhaps the most telling paper is that of Hyönä and Bertram. Across five experiments, they found inconsistent effects. In one experiment (in which the frequency of the first constituent of a compound word was varied), they found that gaze durations were slightly longer when the first constituent of the word to the right of fixation was low frequency. Yet, in a second experiment with exactly the same manipulation, they found the opposite effect. There were also no effects on gaze duration in the experiments in which whole word frequency was manipulated, but they did find an effect on the final fixation before moving to the target word. Likewise, the other measures that were examined (skipping, probability of a refixation) yielded variable and inconsistent effects. As Rayner, Pollatsek, and Reichle (in press) noted, beauty is in the eye of the beholder; Hyönä and Bertram argue that parafoveal-on-foveal effects can occur under some circumstances. Yet, like Rayner et al. (in press) we also believe that a sceptic could argue that such effects are unreliable. In our view, further work is needed to determine more precisely the extent to which parafoveal-on-foveal lexical and semantic effects are reliable. Furthermore, we very much believe that attempts to demonstrate such effects need to be done in the context of reading.

EYE MOVEMENT CONTROL

Underwood and Radach (1998) recently argued that the E-Z Reader model (Rayner et al., 1998; Reichle et al., 1998, in press) is the standard against which alternative models will have to be evaluated. According to the E-Z Reader model, cognitive processes associated with processing a fixated word serve as the engine to drive eye movements through the text. Such a view, which is consistent with many other models of eye movement control (see Reichle et al., in press), contrasts with other theoretical views that attribute a much larger role to low-level visuomotor processes such as the Strategy-Tactics model of O'Regan (1990, 1992). According to this latter model, low-level oculomotor processes are more important than cognitive processes in influencing eye movements. Thus, for example, where readers fixate in a word was argued to be the primary determinant of how long readers look at words. This notion was based largely on research on words presented in isolation (O'Regan et al., 1984) demonstrating that there was a processing cost associated with being fixated in a noneffective location in a word (i.e., being fixated away from the centre of the word). Indeed, it was found that the processing cost amounted to 20 ms per letter that the eyes deviated from the optimal viewing location near the centre of the word.

However, two types of data proved problematic for this view. First, Vitu, O'Regan, and Mittau (1990) found that the processing cost disappeared (or was greatly attenuated) during reading. Second, Rayner et al. (1996) found that independent of where readers fixated in a word, single fixation durations on the word (where readers made only one fixation on the target word) yielded a frequency effect (wherein low frequency words were fixated longer than high frequency words; see also Vitu, McConkie, Kerr, & O'Regan, 2001, for the same finding[1]). Our view is that there is simply too much data indicating that cognitive processes strongly influence the processing of a fixated word (see Rayner, Liversedge, White, & Vergilino-Perez, 2003b, for one recent example) to sustain the position that low level oculomotor activities are more important than cognitive processes in determining when the eyes move.

The claims made in the article by Yang and McConkie, however, are quite reminiscent of arguments made by O'Regan and colleagues (though their second claim goes beyond what O'Regan and colleagues would claim). First, Yang and McConkie argue that cognitive processes serve primarily to inhibit a saccade from being made (as opposed to the view inherent in the E-Z Reader model, and many other models, that cognitive processes serve to trigger an eye movement) and that cognitive processes only affect rather long fixations (this latter point is very similar to claims made by O'Regan and colleagues). Second, they also claim, based on the data obtained from their textreplacement paradigm, that words do not play a critical role in generating saccades (as opposed to the view inherent in E-Z Reader, and many other models, that words are involved in selecting a saccade target). We will not go into great detail with respect to our objections to the claims made by Yang and McConkie. A detailed critique of their paradigm is provided by Rayner, Pollatsek, and Reichle (in press). All we will say here is that we strongly suspect that their results are influenced by unusual strategic processes. Furthermore, while we think that their claims regarding (1) cognitive processes and inhibition of saccades and (2) eye movement control not being word-based are interesting (and deserve further scrutiny), we also think that they are at variance with a great deal of other data.

EYE MOVEMENTS AND WORD PROCESSING/ IDENTIFICATION

For some time now, it has been apparent that eye movements and word identification are intimately related in reading. In the E-Z Reader model, word frequency and word predictability are used as an index of how easy or difficult a word is to process and the ease of processing influences when the eyes move. The articles by Andrews et al., Kliegl et al., and Williams and Morris further continue research aimed at better explicating the relationship between eye movements and word identification.

Kliegl et al. provide further evidence (see also Calvo & Meseguer, 2002) that word length, word frequency, and word predictability all influence fixation times on words. As they point out, the data can be used to serve as a further benchmark for computational models of eye movement control in reading (as the data of Schilling, Rayner, & Chumbley, 1998, have been used). Andrews et al. provide data regarding fixation times on compound words. By varying the frequency of the first and second morphemes, they were able to extend work by Hyönä and Pollatsek (1998; Pollatsek, Hyönä, & Bertram, 2000) on Finnish to English. Hyönä and Pollatsek found that the frequency of both the first and second morpheme in bimorphemic Finnish compound words affected fixation times on compound words. Like Hyönä and Pollatsek, Andrews et al. found effects due to the frequency of both morphemes, and some indication of earlier effects due to first morpheme frequency. This result can be compared and contrasted with results reported by Juhasz, Starr, Inhoff, and Placke (2003), who found stronger effects of second morpheme frequency with English compound words. Clearly, further research is needed on this issue. Finally, Williams and Morris follow up on interesting earlier work by Chaffin, Morris, and Seely (2001) on the processing of unusual and novel words. Like Chaffin et al., Williams and Morris examined how contextual information aids the reader to process novel words. However, they also report the very interesting finding that, with frequency held constant, subjective familiarity of a word led to differences in fixation time on a target word. We find this quite interesting since in a study using regression analysis techniques, we (Juhasz & Rayner, in press) also found that subjective familiarity ratings influence fixation time. Indeed, we also found that age-of-acquisition (see also Juhasz & Rayner, 2001) and concreteness exerted effects (in addition to frequency). We suspect that an important goal for future research should be to further determine to what extent different lexical word properties influence fixation time on a word during reading.

OLD QUESTIONS AND NEW DIRECTIONS

The articles presented in this volume provide a representative overview of current research activities related to eye movements in reading. As such, they provide further data on interesting questions that reading researchers have been asking for the past few years. What is missing from the volume, however, in terms of current research interests, are studies dealing with parsing and discourse

[1] Vitu et al. would perhaps object to our characterising their results as being consistent with Rayner et al. (1996). To be sure, they did find a frequency effect that was independent of landing position (like Rayner et al.). However, they also found that fixations tended to be longer when the reader fixated at the middle of the word than when the reader fixated on the ends of the words. That is, they reported an inverted V-shaped function. This contrasts to the relatively flat function that Rayner et al. reported (actually, Vitu et al., 1990, also found a relatively flat function).

processing. Although there are now many studies, dating back to Frazier and Rayner (1982), examining syntactic ambiguity and parsing strategies that utilise eye movement data, there have been far fewer studies examining discourse representation and processing using eye movements (see Rayner, 1998, for a review of prior research on this issue). We think that one promising line of future research would be to more effectively utilise eye movement data to examine discourse processing.

Related to the issue of parsing and discourse processing, we think that another agenda (for researchers utilising eye movements to study these processes) should be to work out the timing constraints between higher order processing activities and eye movements. Thus, many such studies involve manipulations involving variables such as plausibility or reference relations. Exactly at what point do such variables have an effect? For example, some studies dealing with ambiguity have found immediate effects when a point of disambiguation is reached in a sentence (Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983), whereas in other studies (Pickering & Frisson, 2001) the effect is delayed until the eyes have moved past the word that provides disambiguating information. Likewise, some studies dealing with issues related to discourse processing and plausibility effects find immediate effects (O'Brien, Shank, Myers, & Rayner, 1988) and others find delayed effects (Garrod & Terras, 2000). Generally, researchers doing this type of research (in which eye movements are used as a tool to study language processing) are content enough to simply find effects. However, we would suggest that the time is ripe to determine more precisely what type of higher order effects are immediate and what types yield delayed effects.

At the other end of the spectrum, it is the case that more research is needed to determine exactly how low level effects influence eye movements in reading. Radach and Kennedy (in this issue) suggest that research on eye movements in reading should be tied to more basic research on oculomotor research. On the one hand, we can see the virtue of such an endeavour. Indeed, we (Rayner, Juhasz, Ashby, & Clifton, 2003a) have recently examined inhibition of return effects in reading, and, much to our surprise, found evidence that such an effect occurs in reading: That is, fixations preceding a saccade to a word that was fixated on the immediately preceding fixation are longer than those that cover about the same distance but that are to a word that was not just fixated. On the other hand, Rayner and Liversedge (in press) have documented a number of examples where research on basic oculomotor processes did not generalise to reading (as well as a number that did). Radach and Kennedy also suggest that research on eye movements should be tied more directly to work on basic word recognition processes. Again, we see the virtue of doing so though one can legitimately question the extent to which work on word recognition generalises to reading; one problem with most research on basic word recognition activities is that it assumes that processing of a word begins with the initial fixation on the word, but as a considerable amount of research has demonstrated, readers begin processing a word prior to fixating on it (i.e., there is clear benefit from having a preview of a word prior to fixating it).

Another issue that we suspect needs much more attention is that of regressions. As is well known, readers typically move their eyes backwards in the text on about 10–20% of their saccades, depending on task demands and materials. Although there are lots of assumptions about the nature of regressions, there is in reality very little in the way of hard data (see Kennedy, Brooks, Flynn, & Prophet, 2003; Meseguer, Carreiras, & Clifton, 2002; Vitu & McConkie, 2000, for some notable exceptions). The reason for this is obvious: experimenters can't easily find manipulations that will systematically lead readers to regress. Nevertheless, we do suspect that research activities geared to a better understanding of regressions would be quite useful.

Finally, we would like to point out that the vast majority of research on eye movements and reading has been done in the context of skilled reading. Thus, there are very few studies examining the eye movements of children (see Rayner, 1986, for an exception) and older readers (see the paper by Kliegl et al. in the present issue for an exception). Likewise, there really isn't a great deal of research on differences due to reading skill (see Jared, Levy, & Rayner, 1999). Research on some of the basic findings that have been obtained with skilled readers needs to be carried out with young children and older adults, as well as with readers who are less skilled. Furthermore, there are very little data examining the relationship between oral reading and eye movements. Given that beginning readers spend a lot of time reading aloud, it is the case that more data are needed on this issue (and the recent technological advances that have resulted in fairly accurate eye-tracking devices that do not require a fixed head should make such research more feasible).

In closing, we would like to emphasise again that the articles in this volume represent important contributions to our understanding of eye movements and information processing in reading. Such research is clearly alive and well in Europe. We suspect that it will be very interesting to see what the next 10 years of research on the topic reveals about reading.

PrEview proof published online September 2003

REFERENCES

Beauvillain, C., & Doré, K. (1998). Orthographic codes are used in integrating information from the parafovea by the saccadic computation system. *Vision Research, 38,* 115–123

Beauvillain, C., Doré, K., & Baudouin, V. (1996). The "center of gravity" of words: Evidence for an effect of the word-initial letters, *Vision Research, 36,* 589–603.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–147). Oxford, UK: Elsevier Science.

Calvo, M.G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical and contextual factors. *Spanish Journal of Psychology, 5,* 66–77.

Carpenter, P.A., & Just, M.A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York: Academic Press.

Chaffin, R., Morris, R.K., & Seely, R.E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 225–235.

Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes, 14,* 393–421.

Dunn-Rankin, P. (1978). The visual characteristics of words. *Scientific American, 238,* 122–130.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178– 210.

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

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47,* 201–221.

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

Hyönä, J., & Pollatsek, A. (1998). The role of component morphemes on eye fixations when reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Jared, D., Levy, B.A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General, 128,* 219–264.

Juhasz, B.J., & Rayner, K. (2001, August). *Age of acquisition and word frequency effects during eye fixations in reading.* Paper presented at the 11 th European conference on Eye Movements, Turku, Finland.

Juhasz, B.J., & Rayner, K. (in press). Investigating the effects of a set of inter-correlated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

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

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–180). Oxford, UK: Elsevier Science.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429–456.

Kennedy, A., Brooks, R., Flynn, L., & Prophet, C. (2003). The reader's spatial code. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 193–212). Amsterdam: Elsevier Science.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55A,* 1307– 1337.

Liversedge, S.P., & Findlay, J.M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4,* 6–14.

McConkie, G.W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of garden path sentences. *Memory and Cognition, 30,* 551–561.

Murray, W.S. (in press). The eye movement engine. *Behavioral and Brain Sciences.*

Murray, W.S., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research, 27,* 1–22.

O'Brien, E.J., Shank, D.M., Myers, J.L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 410–420.

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

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

O'Regan, J.K., Lévy-Schoen, A., Pynte, J., & Brugaillére, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 250–257.

Pickering, M.J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 556–573.

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

Pollatsek, A., Reichle, E.D., & Rayner, K. (2003). Modeling eye movements in reading: Extensions of the E-Z Reader model. In J.Hyönä, R.Radach, & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of oculomotor research* (pp. 361–390). Amsterdam: Elsevier Science.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65– 81.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660.

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

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211–236.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J.M.Findlay, R.Walker, & R.W.Kentridge (Eds.), *Eye movement research: Mechanisms, processes and application* (pp. 3–22). Amsterdam: North Holland.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–22.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 358–374.

Rayner, K., & Fischer, M.H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics, 58,* 734–747.

Rayner, K., Fischer, M.H., & Pollatsek, A. (1998a). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.

Rayner, K., Juhasz, B.J., Ashby, J., & Clifton, C. (2003a). Inhibition of saccade return in reading. *Vision Research, 43,* 1027–1034.

Rayner, K., & Liversedge, S.P. (in press). Visual and linguistic processes during eye fixations in reading. In J.M.Henderson & F.Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world.* New York: Psychology Press.

Rayner, K., Liversedge, S.P., White, S.J., & Vergilino-Perez, D. (2003b). Reading disappearing text: Cognitive control of eye movements. *Psychological Science, 14,* 385–388.

Rayner, K., & McConkie, G.W. (1976). What guides a reader's eye movements? *Vision Research, 16,* 829–837.

Rayner, K., Pollatsek, A., & Reichle, E.D. (in press). Eye movements in reading: Models and data. *Behavioral and Brain Sciences.*

Rayner, K., & Raney, G.E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin and Review, 3,* 245–248.

Rayner, K., Reichle, E.D., & Pollatsek, A. (1998b). Eye movement control in reading: An overview and model. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 243–268). Oxford, UK: Elsevier Science.

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

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003c). On the processing of meaning from parafoveal vision during eye fixations in reading. In J.Hyönä, R.Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier Science.

Reichle, E. D, Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

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

Schilling, H.E. H., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26,* 1270–1281.

Starr, M.S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Science, 5,* 156–163.

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

Underwood, G., & Radach, R. (1998). Eye guidance and visual information processing: Reading, visual search, picture perception and driving. In G.Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 1–29). Oxford, UK: Elsevier Science.

Vitu, F., & McConkie, G.W. (2000). Regressive saccades and word perception in adult reading. In A.Kennedy, R.Radach, D.Heller, & J.Pynte (Eds.), *Reading as a perceptual process* (pp. 301– 326). Oxford, UK: Elsevier Science.

Vitu, F., McConkie, G.W., Kerr, P., & O'Regan, J.K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41,* 3513–3533.

Vitu, F., O'Regan, J.K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception and Psychophysics, 47,* 583–600.

Subject index