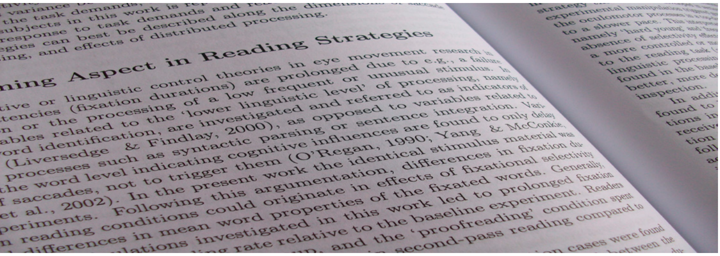
Universität Potsdam



Christiane Wotschack

Eye Movements in Reading Strategies

How Reading Strategies Modulate Effects of Distributed Processing and Oculomotor Control

Spektrum Patholinguistik – Schriften | 1

Spektrum Patholinguistik – Schriften | 1

Spektrum Patholinguistik – Schriften | 1

Christiane Wotschack

Eye Movements in Reading Strategies

How Reading Strategies Modulate Effects of Distributed Processing and Oculomotor Control

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Universitätsverlag Potsdam 2009
http://info.ub.uni-potsdam.de/verlag.htm

Universitätsverlag Potsdam, Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 4623 / Fax: 3474
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Spektrum Patholinguistik – Schriften** wird herausgegeben vom Verband für Patholinguistik e. V.

Umschlagfotos: Christiane Wotschack, Hannelore Gensel

zugl. Potsdam, Univ., Diss., 2009

ISSN (print) 1869-3822
ISSN (online) 1869-3830

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2009/3684/
URN urn:nbn:de:kobv:517-opus-36846
[http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-36846]

Zugleich gedruckt erschienen im Universitätsverlag Potsdam
ISBN 978-3-86956-021-2

for my family

Doctoral Thesis submitted to the Faculty of Human Sciences at the University of Potsdam in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

1st reviewer: Prof. Dr. Reinhold Kliegl
2nd reviewer: Prof. Dr. Ria De Bleser

Day of submission: March 5, 2009
Day of oral defense: September 11, 2009

Acknowledgements

Tough the following dissertation is an individual work, I would like to acknowledge a few people for helping me during my doctoral work. This dissertation was conducted at Prof. Reinhold Kliegl's cognitive psychology unit at the University of Potsdam and I am very grateful for all the support I have received whilst researching and writing up this dissertation.

First and foremost, I would like to thank my two supervisors Prof. Reinhold Kliegl and Prof. Ria De Bleser for giving me the opportunity to work in such a stimulating research environment. Many thanks to Reinhold Kliegl for his generous time and commitment to science. His infectious enthusiasm and unlimited zeal have been major driving forces through my graduate time at Potsdam University. Throughout my doctoral work, he continually stimulated my analytical thinking, guided my development of research skills, taught me many ways of data analysis, and greatly assisted me with scientific writing. He not only laid out the theoretical and methodological foundations of my research; he was also a constructive critic in the best sense. Many thanks to Prof. Ria De Bleser for early sparking my interest in cognitive science, especially in reading research. She encouraged me to develop independent thinking and I wish to thank her for her support and for providing very helpful comments on the first drafts.

I am very grateful for the experience of being a part of the cognitive psychology group at Potsdam University. I want to thank in general all fellows and members of the faculty for creating such a friendly and stimulating context that widened my scientific understanding of (cognitive) science in many aspects. I am extremely grateful for the assistance, generosity, and advice I received from Dr. Jochen Laubrock who assisted me with programming experiments and who read through my draft copies and gave very constructive feedback and inspiring comments. I would like to thank Prof. Ralf Engbert for his open ear and advice to any question I had throughout the years. I've also had the good fortune to have Michael Dambacher as a research colleague and office mate who listened to my ideas as well as my complaining. Working with him has been very productive, rewarding, and a lot of fun. I wish to thank Konstantin Mergenthaler for introducing and helping me with LaTeX and for being a good friend through all the years. The best and worst moments of my doctoral dissertation journey have been shared with many other people from the research group who I want to thank for their support and fruitful discussions: Hans Trukenbrod, Sarah Risse, Olaf Dimigen, Antje Nuthmann, Anja Gendt, Eike Richter, Martin Rolfs, Murilo Babtista, Daniel Schad, Robin Hörnig, and Kathrin Göthe.

I am grateful to the heads of international graduate school 'Computational Neuroscience of Behavioral and Cognitive Dynamics' at the University of Potsdam who supported me financially in the first two years of the dissertation. Thanks to the Institute of German and Dutch Language and Literature at the Department of Philosophy and Humanities at the Freie Universität Berlin for providing the opportunity to continue my research. I am especially grateful to Prof. Gisela Klann-Delius for giving me the time and space to finish my dissertation at the University of Potsdam as well as for her excellent supervision whenever needed.

I want to thank everyone who was involved in the data collection, especially Petra Grüttner and all the student assistants. Many thanks to all the young and old participants in the reading experiments.

I could never have reached the heights or explored the depths without the help, guidance, and efforts of a lot of people outside of Campus. Therefore, I wish to thank my friends and my family for their support and faith in my work. I especially want to thank my husband Philip for his much needed and well placed support and love throughout the years.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Meaning
ANOVA	Analysis of Variance
b	coefficient of fixed effect (in LMM)
cnd	condition
e.g.	exempli gratia (= for example)
F	F-value in analysis of variance
Fig.	Figure
freq.	Word frequency
i.e.	id est (= that is)
IOVP-effect	Inverted optimal viewing position effect
LMM	linear mixed model
M	Mean
min	minute
ms	milliseconds
MF	Multiple fixation
MFC	Multiple fixation case
MSe/ MSE	Mean square error
N	Number of
OVP	Optimal viewing position
p	p-value
PC	Principle component
PCA	Principle component analysis
PVP	Preferred viewing position
PSC	Potsdam Sentence Corpus
regr.	regression
rel.fix.pos.	relative fixation position
sacc.ampl.	saccade amplitude
SD	Standard deviation
SE	Standard error
SF	Single fixation
SFD	Single fixation duration
SFC	Single fixation case
t	t-value in linear mixed model
Tab.	Table
wpm	words per minute

1. Introduction

When we read, we generally read for pleasure or to obtain a specific information from a text. Reading comprehension is the goal in both scenarios. In the context of natural viewing and scene perception, Yarbus (1967) already came to the conclusion that eye movements are not driven by the intrinsic salience of objects, but by their relevance to the task. For example, he demonstrated that eye movement patterns during the examination of a painting differed dependent on the task the subjects were given. In analogy to Yarbus' work, it is plausible to expect that eye movements on the identical reading material may differ depending on the reading intention. At the same time, differences in processing and processing difficulties during sentence reading may originate in variations in text difficulty, in variations in the reader's skills, or in variations in the reading task. Almost all studies in reading research related to psycholinguistic questions focused on determining the relevant factors either related to the reading material or related to the reader's cognitive profiles (Rayner, 1998). There has also been an impressive development in the field of reading research and the improvement of models of eye movement control in "normal" reading, but not much attention has been paid to the variability within skilled reading.

In cognitive science, theories of reading comprehension are assumed to be general, like other cognitive theories. A theory or model of reading is strong, if it is plausible with respect to the psycholinguistic reality of assumed processes and components and if it can account for numerous empirical findings in reading research. Since the early 20th century, the technique of monitoring eye movements during reading has become more and more popular in investigating the mechanisms that control the reading behavior (Huey, 1908). Compared to other performance measures, for example, total sentence reading times or reading comprehension, the advantage of recording eye movements is that this method provides online-measures of the reading process and reflects a moment-to-moment control during reading (Just & Carpenter, 1980). Several indices of processing can be collected simultaneously with high temporal and spatial resolution that allow the examination of variations within the reading process. In addition, by presenting the whole sentence or text on a screen the reader is free to examine any part of the material, which resembles a more natural reading situation than for example, rapid serial visual presentation (RSVP) paradigms. Eye recording systems may vary with respect to their comfort for participants, but they are all useful tools to recover the natural reading process.

The present work merges these aspects into a single framework, that is, it considers at the same time variations at the level of the reading material and differences at the level of the readers that may influence eye guidance in reading. The goal of this work is to investigates for the first time how a reading strategy defined by a given reading goal, and independent from individual differences, systematically modulates effects of oculomotor control and reader-and word-level effects under the assumption of distributed processing in continuous reading. Linear mixed models provide an excellent framework for the analysis of the complex interactions of word-level and reader-level variables on fixation duration measures within the framework of distributed processing.

After a short introduction into the terminology of eye movements in reading (section 2.1), low-level, word- and reader-level variables and their critical role in explaining eye movement behavior during reading will be provided (sections 2.2, 2.3, and 2.4). The concept of distributed processing within the perceptual span is introduced in section 2.3.1. Then, the terminology of 'reading strategy' is discussed and defined (section 2.4.1), before research methods are described (chapter 3), and research question of this work are outlined in more detail in the remaining chapters 4, 5, 6, and 7. Main results are summarized and discussed in chapter 8.

2. Theoretical Background

Eye movements during reading can be interpreted with respect to their spatial parameters, *where* to move the eyes, and their timing parameters, *when* to move the eyes. It has been demonstrated that the when- and where-decisions in eye movement control are to a fairly large extent computed independently (Findlay & Walker, 1999). Next to testing psycholinguistic theories of language processing, one of the final goals of eye movements research in reading is the development of a model that can explain the relevant factors that control eye movement guidance in reading. Those models have been developed in parallel to models of sentence comprehension, but both lines of research are not yet integrated. Models of eye movement control differ in their degree to which higher-level cognitive processes, for example word identification, affect the moment-to-moment decision about when to move the eyes (Reichle, 2006). The early theoretical models of eye movement control could be classified into two types of models: *oculomotor control* models and *cognitive control* models (Rayner, Sereno, & Raney, 1996; Starr & Rayner, 2001; Reichle, 2006).

According to the theory of oculomotor-control, eye movements during reading depend on basic factors such as saccade programming and the related perceptuomotor variability (e.g., O'Regan & Lévy-Schoen, 1987; O'Regan, 1990, 1992; McConkie, Kerr, Reddix, & Zola, 1988). Also low-level visual factors of the material, such as word length, but no higher-order factors, i.e. lexical variables, play a role in those models that emphasize non-cognitive factors. In the framework of this theory spatial parameters of eye movement behavior (the *where*-dimension) are in the focus of analyses. Researches of the opponent theory of cognitive or linguistic control mainly investigate reading time measures (the *when*-dimension). Eye movements are considered to be sensitive to the cognitive processes that act on the material inspected by the reader (e.g., Just & Carpenter, 1980; Morrison, 1984; O'Regan, 1979). A completion of some cognitive event causes the eye to move on and most models are designed to be driven by the lexical access process. Furthermore, variations in eye fixation durations are thought to be indicative of the current difficulties in processing that the reader experiences. Current models of eye movement control integrate both lower-level oculomotor and higher-level cognitive aspects of processing that potentially guide eye movements in reading (e.g., SWIFT, Engbert, Nuthmann, Richter, & Kliegl, 2005; EZ-Reader, Reichle, Pollatsek, Fisher, & Rayner, 1998), but so far no model considers effects beyond the word-level, such as syntactic parsing or reading intentions or reading strategies (see Reichle, Warren, & McConnell, 2009, for a first attempt).

2.1. Basic Eye Movements and Processing Measures in Reading

The general purpose of eye movements during reading is to bring the text into the region of central vision. The region with highest acuity (the fovea) extends out to 2°, which corresponds to 6 - 8 characters depending on font size. The parafoveal region, where acuity of vision is reduced, extends out to 5° (or 15 - 20 characters). Thus, visual processing is best when the input is located in the foveal region. Readers are also able to take up information from the parafoveal region and it has been demonstrated that the perceptual span is asymmetric to the right of fixation in latin alphabetical orthographies such as English, German, or French. In a skilled reader the perceptual span extends about 3 - 4 characters to the left and about 14 - 15 characters to the right of fixation (McConkie & Rayner, 1975; Rayner, 1975).

The continuous eye movements we make during reading are called *saccades*. Saccades are rapid eye movements and last about 30 milliseconds (ms). During *fixations* between saccades our eyes remain relatively still. The duration of a reading fixation is 200 - 300 ms. We do not obtain visual information from static input during a saccade. This phenomenon has been called 'saccadic suppression' (Matin, 1974). Thus, we are effectively 'blind' during these rapid eye movements and are able to perceive information only during fixations. If a word gets more than one fixation, the word is *refixated*. Jumping back with the eyes to previously passed words or parts of the text is called a *regression*. 10 - 15% of the saccades are regressive eye movements (Rayner, 1998). If a word receives no fixation, the word is *skipped*. The basic saccadic eye movements in reading are illustrated in Figure 2.1.

Typically, the unit of analysis for reading time measures is the word. Dependent on the theoretical question different measures of processing time are used in reading research. Important is the differentiation between *first-pass* and *second-pass* reading. First-pass reading consists

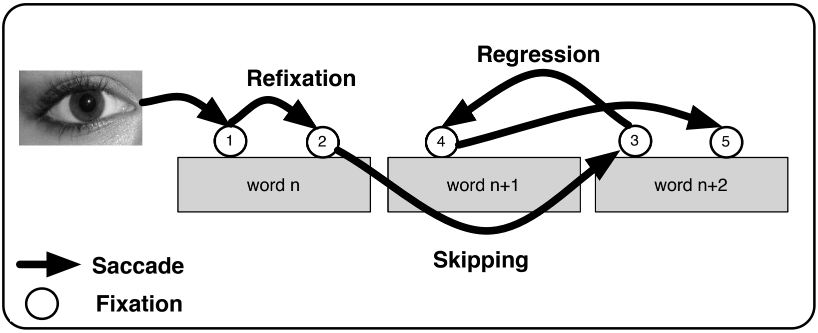


Figure 2.1.: Types of saccadic eye movements: The circles 1 - 5 represent fixations in a sequential order; fixations 1, 2, and 3 are first-pass fixations, fixations 4 and 5 are second-pass fixations; the sum of fixations 1 and 2 is the first-pass gaze duration on word n (as well as its total reading time); the sum of fixations 3 and 5 is the total reading time on word n+2; fixation 3 is a regression origin fixation, fixation 4 is a single fixation regression goal; word n+1 has been skipped in first-pass reading.

of all initial forward (here from left to right) fixations in a given region, e.g. a sentence. The first attempt of reading a sentence is associated with highly automated and efficient processing in reading, at least in a skilled reader. In contrast, second-pass reading includes all fixations after a regressive eye movement on those parts of the text that the eye had already passed during the first pass. The second-pass ends with the fixation of a new word that has not been encountered in the first pass, thus, the first-pass reading continues at that point. Regressions are usually associated with processing difficulties (e.g., Frazier & Rayner, 1982). The end of a non-sentence terminating first-pass segment is determined by a *regression origin fixation*, whereas the beginning of a second-pass is marked as the *regression goal fixation* (see Figure 2.1). In this work, second-pass reading is distinguished from *rereading*. The second pass includes regressive fixations before the end of a sentence has been reached (in other words: jump-backs during the first inspection of the material). Regressive fixations that are made after the last word of a sentence has been fixated are defined as rereading fixations.

Single fixation duration is the dependent time variable if a word is fixated exactly once. *First fixation duration* describes the duration of the first of several fixations on a word. *Second fixation duration* is the duration of the second of multiple fixations on word. *Gaze duration* is the sum of all first-pass fixations on a word, whereas *total reading time* includes all fixations of first- and second-pass reading (cf. Rayner, 1998). Various other, mostly accumulative fixation duration measures are used in studies investigating text processing (cf. Hyönä, Lorch, & Rinck, 2003), but these are not relevant for the studies presented in this work.

2.2. Low-Level Effects on Eye Movement Measures

The *where*-dimension of eye movements in reading is determined by the *saccade amplitudes* and - with respect to word units - by the saccadic *landing position* or the *fixation location*. Mean saccade amplitude in continuous reading is about 2°, corresponding to ca. 8 letters (Rayner, 1998). Mapping the pattern of varying saccade lengths on words in a sentence, the different saccade amplitudes result in single fixation cases, refixation cases, word skippings, and regressions. It has been demonstrated that fixation duration (the *when*-dimension) depends on the size of the previous and the following saccade (e.g., Kliegl, Nuthmann, & Engbert, 2006): The longer the saccade, the longer the previous or following fixation duration. The saccade latency, the time the oculomotor systems needs to programm a saccade, is at least 150 ms (Engbert, Longtin, & Kliegl, 2002). It is open to debate, if the increase in fixation duration reflects the programming time for the next saccade or the processing time of the visual input, but both processes are assumed to work in parallel. It is generally assumed, that the word is the functional target location in reading and thus, word boundaries and word length are mainly determining saccade targeting. A question related to the word as a saccade target is where the reader fixates within the word.

In single word recognition, the optimal viewing position (OVP) is at the center of the word (O'Regan & Lévy-Schoen, 1987). For two reasons this location is considered optimal: First, recognition time at this fixation location is minimal, resulting in a u-shaped curve of reaction times in decision tasks. Second, refixation probability increases, the more letters the fixation location deviates from the OVP. In continuous reading, the fixation location tends to be slightly more to the left of the word center and is called the *preferred viewing location* (PVL; Rayner, 1979; McConkie et al., 1988; McConkie, Kerr, Reddix, Zola, & Jacobs, 1989). Refixation probability is again minimal at the PVL. Fixation duration varies systematically with fixation position, and a somewhat surprising finding is the *inverted optimal viewing position* (IOVP) effect for fixation duration in continuous reading (Vitu, McConkie, Kerr, & O'Regan, 2001). Vitu and colleagues (2001) were the first to demonstrate that in continuous reading fixation durations are longest around the word center and decrease at the ends of the word, the inverse of the effect found in isolated word recognition. One account for the IOVP effect has been provided by Antje Nuthmann and colleagues (Nuthmann, Engbert, & Kliegl, 2005, 2007), who argue that due to saccadic error, we find more mislocated fixations at the ends of words. In the beginning of a word, fixations might be mislocated due to saccadic undershoot, in the end of a word due to overshoot of the word center. Since the oculomotor system is sensitive to saccadic error and mislocations, a corrective saccade programm is immediately started at these locations, entailing that fixation durations at the beginning and end of a word are shorter than fixations located around its optimal viewing position.

In sum, reading saccades usually target a position half way between the beginning and center of the next word, and fixation duration is influenced by the fixation position within the word as well as by processes of saccade targeting. Next to those low-level factors, fixation duration is systematically influenced by higher-level factors, as described in the following section.

2.3. Effects of Reading Material on Eye Movement Measures

In psycholinguistic research, it is widely accepted that the processing of written material in a skilled reader involves several distinctive processes: The visual input needs to be decoded into its orthographic units, the mental lexicon needs to be accessed, and the word needs to be integrated into the sentence's context. In support of these ideas, word variables that affect different processing levels are found to be predictive of eye fixations in sentence reading.

A main part of the variation in fixation durations seems to depend on the processes of word recognition. In general, it can be summarized that the easier a word, the shorter the fixation duration. The difficulty of a word depends on various aspects of a lexem, such as word length, word class, word frequency, or word predictability, etc. Numerous studies demonstrated that low-level variables such as word length modulate the time and number of fixations: As number of letters in a word increases, gaze duration increases (Just & Carpenter, 1980; Kliegl, Grabner, Rolfs, & Engbert, 2004; Kliegl et al., 2006). At the lexical level one of the strongest predictor for fixation duration is word frequency. Many studies have demonstrated that readers fixate longer on low-frequency words than on high-frequency words (e.g., Rayner & Duffy, 1986; Just & Carpenter, 1980). A different factor influencing the timing of eye movements is the age of acquisition. For words that were acquired early in life, shorter single fixations, first fixations, and gaze durations have been found (Juhasz, 2005).

In fact, the correlation of fixational behaviour and lexical variables is somewhat more complex than described above, because in natural reading words are embedded in sentences. It has been found that effects of lexical variables of the fixated word interact with the properties of the neighbouring words (e.g., Henderson & Ferreira, 1990; Kennedy & Pynte, 2005). The data and rationale behind this observation is explained in following section 2.3.1.

Another important factor - that is associated with the word, but only relevant in a sentence context - is the degree of contextual constraints for a given word. It has been found that words that are highly predictable from the preceding context are fixated shorter and skipped more often than contextually unconstrained words (Ehrlich & Rayner, 1981; Balota, Pollatsek, & Rayner, 1985; Rayner & Well, 1996; Kliegl et al., 2004, 2006). A grammatical factor, analyzed on the word unit, is the lexical status of the fixated word. Carpenter and Just (1983) found that content words are fixated about 85% of the time, whereas function words are fixated about 35% of the time. The relatively high skipping rate for function words, such as determiners and prepositions, is related to the confounded variables of lexical status, namely word length and word frequency. Function words tend to be short and of high frequency and therefore, they are prime candidates for skipping.

In addition to word recognition processes, the syntactic and semantic integration of a word into the sentence context have been shown to play a role in the processing times during reading. For example, Frazier and Rayner (1982) have first demonstrated that eye movements differ systematically on structurally ambiguous sentences compared to unambiguous sentences (for a review on studies using *garden path sentences* see Clifton, Staub, & Rayner, 2007). Liversedge and colleagues demonstrated how thematic roles influence the processing of structurally ambiguous sentences (Liversedge, Pickering, Branigan, & Gompel, 1998; Liversedge, Pickering, Clayes, & Branigan, 2003). The eye movement behavior is also determined by the textual demands of the reading material. As text becomes conceptually more difficult, readers make shorter saccades, prolonged fixations, and more regressions (Heller, 1985; Rayner & Pollatsek, 1989; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). In the experiments of the present work, isolated sentence reading was conducted so that textual or discourse related processes will not further be addressed. Though sentences are used as reading materials, syntactic variables are not considered in the analysis. The important information is that processing demands defined by the material, specifically word and contextual word properties, influence eye movement behavior in reading. All evidence listed here supports the 'linguistic control hypothesis' (O'Regan, 1979) of eye movements in reading. This theory assumes that lexical, syntactic, and semantic constraints are used during reading and thus, all of those variables may influence fixational behavior, i.e. fixation probabilities and fixation durations.

2.3.1. Theories of Distributed Processing

Immediacy effects of word length, frequency, and predictability on fixation durations, as reported above, provide evidence for the *eye-mind-assumption* claimed by Just and Carpenter (1980). This hypothesis states that the eye remains fixated on a word as long as the word is being processed. Thus, processing demands by, for example, lexical frequency are directly related to gaze durations. At the same time, the theory proposes that words that are not fixated are not processed. Thus, according to the authors there is no eye-mind-span. The clearest evidence against Just and Carpenter's proposal is the case of skipped words. These words were not fixated during reading, but are processed anyhow, as verifiable by means of comprehension tests.

In contrast to the eye-mind-hypothesis, theories of distributed processing propose that there can be a lag between visual and cognitive processing (an eye-mind span). Distributed processing means that fixation durations reflect processing demands of the fixated word as well as processing of the word to the left (*lag effect*) and to the right (*successor effect*) of fixation. This account rests on evidence of skipping, spillover and parafoveal-on-foveal effects. Skipping provides good evidence for words that have been processed though they have not been fixated. As mentioned earlier, function words are skipped more often than content words (O'Regan, 1979; Carpenter & Just, 1983) but they are still processed and comprehended. Fisher and Shebilske (1985) demonstrated that readers have great difficulties understanding a text if all the words, that a different group of readers had skipped during reading, were deleted from the text. Thus, the skipped words had been processed during the fixation right before skipping or in the subsequent fixation after skipping. In line with the preprocessing account of the skipped word, Ehrlich and Rayner (1981) report that highly predictable content words are skipped more often than content words less constrained by the context.

If processing of a word is not completed before the eyes move on to the next word, effects of word variables can *spill over* to the subsequent fixation duration (Rayner & Duffy, 1986; Kliegl et al., 2006). For example, Kliegl and colleagues (2006) found a strong lag frequency effect of word n - 1 on single fixation and gaze durations on the fixated word n: The less frequent the previous word was, the higher the fixation duration on word n.

Parafoveal-on-foveal effects are a subset of successor effects and mean that properties of the upcoming word n+1, located in the parafoveal region, can influence fixation durations on the fixated word in the foveal region. It implies that the word to the right of fixation has been processed up to a certain degree before the eyes move on to that word or skip it. Several studies have been able to demonstrate that fixation duration was related to the lexical properties of the word to the right of fixation (Kennedy, 2000; Kennedy, Pynte, & Ducrot, 2002; Kennedy & Pynte, 2005; Kliegl et al., 2006; Pynte & Kennedy, 2006). For example, in a corpus analysis Kliegl et al. (2006) found effects of frequency and predictability of the upcoming word on single fixation durations. Further, results indicate that there is a dynamic relation between the extent of preprocessing of the word to the right of fixation and the properties of the fixated word. Kennedy and colleagues demonstrated that only if word n is short, does word frequency of word n+1 influence gaze duration. In other words, there is only a parafoveal-on-foveal word frequency effect, if word n+1 falls into the perceptual span which extends up to 14 characters to the right of fixation (McConkie & Rayner, 1975; Rayner, 1979). The 'perceptual span theory' (Kennedy & Pynte, 2005) further implies that fixations on long words are influenced by sublexical properties of word n+1, such as initial trigramm frequency.

Kliegl et al. (2006) proposed a memory retrieval hypothesis to explain successor effects of word *predictability*. Fixation duration on word n increases with increasing predictability of word n+1. Since this effect was more pronounced if word n was long and thus, word n+1 tended to fall outside the perceptual span, the authors argued that this effect cannot directly be related to the size of the perceptual span but that it is rather due to lexical retrieval from memory, triggered by the given context and the degree of predictability. Highly predictable words (n+1) are predicted and retrieved from the lexicon and this top-down process is reflected in higher fixation durations on word n. In a further analysis, Kliegl (2007) report that predictability successor effects are dependent on the lexical status of word n and n+1. Parafoveal-on-foveal effects were only found if either word n or word n+1 was a function word, thus if one was a highly predictable word.

From a slightly different perspective, Henderson and Ferreira (1990) provide evidence for a dynamical theory of parafoveal preprocessing. They investigated the *preview benefit* of words n+1 in dependence of word n. Preview benefit means a decrease in fixation duration due to prior preprocessing of the word in the parafovea (Inhoff, 1989; Inhoff & Rayner, 1986). Therefore, the dependent variable was fixation duration on word n+1, not on word n. Henderson and Ferreira found that there was a greater preview benefit for word n+1, if word n was easy to process, e.g. if it is a word low in lexical frequency. The authors argue for the 'foveal load hypothesis', implying that the extent of preprocessing of the word to the right of fixation strongly depends on the processing difficulty of the foveal word (see also Schroyens, Vitu, Brysbaert, & d'Idewalle, 1999).

In sum, there is evidence for a dynamically modulated processing span during continuous reading that depends on word variables of the fixated word and as well as on properties of the neighboring words. Therefore, a comprehensive analysis that attempts to explain variance in fixation durations in continuous reading must not only contain word-level parameters of word n, but also word-level parameters of word n-1 and word n+1.

2.4. Effects of Reading Strategy on Eye Movement Measures

In the previous section, suggestive evidence was provided that processing demands induced by the reading material, i.e. via word-level parameters such as length, frequency, and predictability, have a systematic effect on eye movement behavior during reading. In addition, there are several factors at the reader's side that may impact the ease of processing the written input. Turning to the role of reader-level effects on eye movement behavior in reading, the relation between individual differences, reading strategy, reading style, and reader-level parameters, needs to be discussed. The following sections try to disentangle the concepts of reading strategy, reading skill, and individual differences, and end with a coherent definition of 'reading strategy' for this work.

2.4.1. Strategies, Skills, and Parameter Specification

The goal in the daily reading context is generally to obtain a specific information from a text or simply to read for pleasure. Reading comprehension is the ultimate goal in both scenarios (Goodman, 1985). In the context of natural viewing and scene perception, Yarbus (1967) came to the conclusion that eye movements are driven by their relevance to the task, not by the intrinsic salience of objects. In analogy to Yarbus' work, it is plausible to expect that eye movements on the identical reading material may differ depending on the kind of information the reader wants to extract from it. Though the dimensions in reading differ from the dimensions set up in scene viewing, especially because the left to right sequential order of words is inevitable, the idea of *selectivity* can still be applied. As gaze movement in action is highly controlled by task-relevance through top-down processing (see Land, 2007, for an overview), the movements of the eyes on a line of text are possibly modulated by the reader's goal or reading intention, and thus, different reading strategies may be in the repertoire of a skilled reader (Tinker, 1958; Heller, 1985).

A theory of eye-movement control that clearly defined 'strategy' in reading is the 'Strategy-Tactics'-Model by O'Regan (O'Regan & Lévy-Schoen, 1987; O'Regan, 1990). The idea of this model is that saccade targeting and saccade latency during reading are mainly the results of a basic oculomotor strategy that may be modulated by linguistic processing, but only with delay. In this framework, a between-word reading *strategy* is differentiated from within-word (rescue) *tactics*. The 'between-word strategy' a reader adopts based on the available low-level visual information, such as word length, is to target saccades at the general optimal viewing position (OVP) in a word-by-word fashion. If instead the ends of the word are reached by the general scanning routine, e.g. due to oculomotor error, a refixation is made to the other end of the fixated word, a behavior called 'within-word tactics'. This is in line with definitions within other disciplines where strategy as being often practically rehearsed is differentiated from tactics that are defined as immediate actions (e.g., Evered, 1983). Both strategy and tactics are supposed to be automatically applied in a skilled reader. The oculomotor reading strategy as proposed by O'Regan can further be modulated by lexical processing at later stages, implying that the execution of the next saccade can be delayed by e.g., word identification problems. In sum, the authors of the strategy-tactics-theory assume that this oculomotor strategy is the essential eye movement strategy a reader applies to achieve the ultimate goal of reading comprehension in the most efficient way.

Based on the assumptions of this oculomotor control theory of eye movements in reading, the question arises how the reading strategy is adapted when different task demands or reading goals are identified by the reader. Naturally, two aspects of eye movement control come into consideration, namely, the saccade latency (fixation duration) and the saccade size and the related fixation position, as introduced as the *when*- and *where*-dimensions of eye movements in reading. The when-decision is determined by the timing of the described general scanning strategy, that is supposed to be autonomous (O'Regan, 1990; Yang & McConkie, 2001). Supposedly, it can further be modulated by the extent of linguistic processing a reader wishes to perform that would inhibit the subsequent saccade execution and therefore increase fixation duration. The where-decision is realized in the saccade amplitude, that is mainly a result of the careful word-by-word scanning routine and of the within-word tactics in cases when the OVP is missed (O'Regan, 1990). Thus, parameters of saccade timing and saccade targeting need to be set in accordance to the specified reading goal.

The concept of *parameter specification* has also been used in theories of selective attention in motor control (Neumann, 1987, 1989). The idea of parameter specification according to Neumann refers to the abstractness of the skill, in the present case the skill of reading. The author defines a *skill* as acquired control structures that are stored in long term memory. Reading is a skill acquired over a period of several years and necessary control structures are stored in the long term memory, specifically the structures for eye movement control in reading. The effector, the organ or part of the body that can be controlled by the skill is in silent reading always the eye (or better, both eyes in conjunction). Skills have the following two characteristics: First, they are abstract and specifiy a class of movements, e.g. the skill of reading specifies all potential eye movements that can be made during reading. Second, skills are nested, e.g. reading aloud, lexical search, or proofreading are subskills of the general reading skill. Skills are used to obtain action goals. Different reading subskills are applied to obtain a specific action goal, for example the extraction of specific information from a text, entertainment, or the correction of a draft.

According to Neumann (1987, 1989), skills are abstract. With respect to reading it means that the higher-level skill of reading includes the skill of a controlled movement of the eyes, but the parameters of eye movements, such as saccade timing or saccade targeting, need to be specified for each single action of eye movements, i.e. each individual execution of a saccade. Neumann further argues that most of the parameter specification can be solved 'locally' because higher-level skills such as reading include subskills for picking up information that is suitable to specify parameters. If this mode of parameter specification is sufficient to specify all required parameters, the action is called *automatic*.

Unlike hearing or speaking in which linearity of the signal is given, in reading, the visually presented sentence or text as the input provides all information about setting the right parameters for fixating the words. This parallelisms of information forms a problem because the pieces of the input information are simultaneously available and specify the same parameters in mutually exclusive ways, that is, each word of the text could be defined as the next saccade target. Neumann calls this specific situation, where parameter specification does not suffice, *overspecification*. Thus, the parameter of saccade timing and targeting cannot be picked up only by information from the text; it rather needs to be specified elsewhere or there must be a process of object selection. Parameter specification demands that each parameter is given exactly one value at a time, because e.g., the eye movement cannot go into different directions at the same time. The problem of overspecification can be solved by object selection via directing attention. From the left-to-right nature of reading alphabetic script, we know that the next saccade target is specified 'locally': The next one or two words to the right of fixation that fall into the perceptual span (see section 2.1) are considered as saccadic goals (though models of eye movement control may differ in defining saccade targets). The restriction of the space to be processed turns the parameter specification in saccade targeting into a more local process, and thus, the definition of saccade targeting as a fast and automatic process via direct, local parameter specification can be applied, at least in a skilled reader.

Usually, the 'automaticity in reading' refers to higher-level automatic processes, such as letter identification, word identification, acoustic recoding, semantic access etc. These processes are assumed to occur quickly, effortlessly, often beneath conscious awareness, and produce little interference with other tasks. Automatic processes are differentiated from control processes, such as text modeling, that are slower and make high demands on attention and working memory (see Walczyk, 2000, for a review). Most models that try to explain the automaticity in reading processes begin their explanations beyond the level of eye movements (e.g., LaBerge & Samuels, 1985; Perfetti, 1994; Logan, 1997). According to O'Regan the execution of eye movements in a word-by-word fashion is also considered automatic. Since in this work eye movements during reading are recorded to infer specific processing aspects in reading grounded on the alternation of saccades and fixations, the automaticity of eye movements is considered as part of the automatic processes in reading.

In his 'levels-of-control' concept, Neumann argues against the distinction of controlled and automatic processes. According to Neumann (1989) skills, even though they may contain automatic components in the sense of fast, effortless, and unconscious processes, are still a result of some conscious planning or intentionality. In other words, even if the reader might not be aware of the single saccades he executes during reading, he still needs to consciously direct attention to the reading task and to the current reading intention. Thus, eye movements in reading are not a mere passive consequence of the visual stimulus input, but these automatic movements are components of a controlled, voluntary action. In this sense, automaticity is one mode of controlled action. The question of interest is if differences in routinized reading strategies of readers defined by different reading intentions are observable in the highly automatic process of saccade timing and targeting.

A Definition of Reading Strategy

In reading, the borders between the concepts of 'strategy' and 'skill' are fuzzy. According the above mentioned distinction of controlled and automatic processes, eye movements in the beginning reader during reading for comprehension resemble controlled, consciously applied action plans more than they resemble an automatically applied skill. The opposite is found in a normally skilled reader, where the main characteristic of the eye movement are the speed and automaticity, not the conscious efforts related to the reading goal. The subskills defined according to Neumann (1987, e.g., reading for pleasure, reading for comprehension, or proofreading) are equivalent to potential goals in reading. The reading goal is closely tied to the reading strategy applied. In other words, subskills in reading are defined by the reading intention, and the parameter specification for the automatic components within subskills is expected to change in accordance to the defined reading goal. Based on these explications, the following definitions of reading strategy, reading intention, and strategy effects are set: *Reading strategy* is defined as a selective parameter specification in the system of eye movement control in reading that is necessary to achieve a specified reading goal. The set of potential parameters remains underspecified at this point, but any parameter related to saccade programming will be relevant here. *Reading intention* or the reading goal is defined by the reading task and/ or the task demands identified by the reader. *Strategy effects* are differences found between experimental groups that differ with respect to reading instruction or task demands. These strategy effects are differentiated from effects due to individual differences, that is, they are measured in addition to individual variations between subjects. The difference between effects of variation between readers and of reading strategies within readers are explained in more detail in the following section. In sum, a change in reading strategy due to a specific reading goal requires a selective parameter specification e.g., in saccade timing, in saccade targeting, or in the size of the perceptual span.

2.4.2. Reading Strategy and Individual Differences

Screening the literature related to the psychology of reading, the term 'strategy' occurrs in various meanings and contexts, but it is not consistently used. For example, a more 'careful' reading strategy has been observed when difficult text was given to subjects (Rayner & Pollatsek, 1989, for a review). In this sense, different reading strategies are used to make a problem easier to understand or to solve, here the problem of understanding the given text. Eye movement parameters are adapted to fulfil the requirements which are defined by the difficulty of the reading material. In contrast, the term 'reading strategy' was used to describe the observed reading behavior of a subject attributed to the characteristics of the specific individual, e.g. the 'risky reading strategy' of old subjects in contrast to younger readers (e.g., Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006). This type of strategy is not consciously applied but is rather a results of the subject's cognitive features and reading skills. In various other studies (see Rayner, 1998), the term 'individual difference' is used to describe the variance between subjects in their reading behavior and as well in their reading skills, ranging from more or less skilled 'normal' readers to dyslectic readers.

In his 'strategy-tactics model' (O'Regan & Lévy-Schoen, 1987; O'Regan, 1990), O'Regan also differentiates a 'risky' reading strategy from a more 'careful' scanning routine. According to the model's risky reading strategy, readers move their eyes in order to fixate on the word's optimal viewing position (OVP), which is located slightly left of the word center. In a more careful mode, readers make more non-optimal fixations within words and consequently, make many more refixations. The author mentions that a given reader could induce different reading styles by varying the refixation criterion which implies that different strategies are in the repertoire of a skilled reader. Thus, the usage of the term strategy comes resembles the definition of reading strategy stated above.

An early categorisation of reader types based on their eye movement behaviour can be found in the study by Olson and colleagues (Olson, Kliegl, Davidson, & Foltz, 1985). Within the group of readers with dyslexia, the authors described two types of readers: *Plodders* with a very sequential reading style make shorter saccades and fewer regressions and word skippings; *explorers* make relatively long saccades with more word skippings and more regressions. Variance between normally skilled readers has also been reported (Rayner, 1998, p. 393) and a few attempts to describe types of skilled readers with respect to global text processing strategies have been published (Hyönä, Lorch, & Kaakinen, 2002; Hyönä & Nurminen, 2006). Hyönä and colleagues (Hyönä et al., 2002) described individual differences between readers based on the two dimensions reading speed and look-back behavior. The authors identified four different types of reading strategies when reading expository text: 40% of the readers were categorized as *fast linear readers* with short fixation durations and almost no regressions. In contrast, *slow linear readers*, about 25% of the group, make more forward fixations and reinspect each sentence before moving on. Two more types of readers were defined with respect to the processing of the text's topic structure: *Nonselective reviewer*, less than 7% of the readers, showed many regressions to previously read sentences, so called look backs. The reading behavior of *topic structure processors*, about 20% of the participants, was characterized by extra attention (i.e. fixations) spend on the headings of the text. This group had also the largest working memory capacity, measured by reading span, and provided the best text summaries. The authors stress in their discussion, that regarding the reading task, that is summarizing the text after reading, the group of the topic structure processors are the most efficient readers, but more than 80% of competent readers in college do not use this topic-related reading strategy. Results for the three reader types of fast and slow linear readers and topic structure processors have been replicated in Hyönä and Nurminen (2006).

Hyönä and colleagues identified different reading styles of skilled readers as different 'reading strategies'. Given the present definition of *reading strategy* (see page 15), these patterns have to be considered as individual differences, not as reading strategies, because the reading goal and the task demands were identical for all readers. Individual differences during the reading process or individual strategies of reader types have also been labeled 'flexibility', mainly with respect to the flexibility in reading rate (Rayner & Pollatsek, 1989). Considering the difficulty of the text in the same analysis, Walker (1938) reported in an early study that good readers are more adaptable to varying types of materials and requirements of comprehension compared to poor readers. In a recent study comparing the eye movements of readers of two skill-levels, defined by vocabulary size and reading comprehension, Ashby and colleagues found longer fixation durations on low frequency words for average skilled readers than than for high-skilled readers (Ashby, Rayner, & Clifton, 2005). They further demonstrated that average skilled readers make more regressions and reread text more often than highly skilled readers. The authors argue that observed differences in fixation durations and in reading rate are attributable to differences in word recognition time and differences in rereading patterns. The question of what makes someone a 'good' or 'more skilled' reader remains open, but it is the goal of reading research to identify measurable predictors for reading performance. Good candidates that explain individual differences, which were already mentioned, are for example age or vocabulary size. These and a few other factors that are found to be predictive for individual differences in reading are presented in section 2.4.4.

In this work, the term reading strategy is defined as a selective behavior of eye movements due to parameter specification that skilled readers apply in order to achieve a specified reading goal. Hence, different reading strategies lie in the variability within the skilled reader. Next to differences in effects of oculomotor control associated with reading strategy the interaction of reading strategy with word-level effects (see chapter 2.3) are of special interest, and how these are reflected in eye movement measures in addition to effects of individual differences. To this end, effects in eye movements in reading due to task instruction or task demands already demonstrated in the literature are reviewed shortly, before predictors of individual differences are reported.

2.4.3. Effects Due to Reading Instruction

Variations in task demands require different reading strategies. The manipulation of task demands during reading can be achieved by varying reading instructions. Importantly, the reading material needs to be identical in instruction-comparative studies to evaluate the experimental group effects of strategy. Only a few studies focussed on this comparison.

Just, Carpenter, and Masson (1982) monitored eye movements in normal reading, skimming, and speed reading. Reading speed differed significantly between groups: Normal reading rate averaged out 250 words per minute (wpm), skimmers read about 600 wpm, and trained speed readers reached a reading rate of 700 wpm. Differences in reading speed were further expressed in differences in fixation durations and skipping. Normal readers made longer fixations and fixated about two-third of the material, whereas speed readers and skimmers fixated only one-third of the words in a passage with shorter fixation durations. Despite the advantages in reading times in skimming and speed reading, fast readers were disadvantaged in answering detailed comprehension questions. Thus, reading speed was increased at the costs of reading comprehension (Masson, 1985).

One of the few studies investigating the interaction of reader-level (instruction) and word-level (frequency) factors was conducted by Rayner and Raney (1996). They demonstrated that the word frequency effect on fixation duration found in reading for comprehension disappears when a lexical search task was applied on the same text material. The authors claim that in the search task the orthographic match of the target and the input occurred before lexical access took place, and the result of matching triggered the eyes to move on. Therefore, no frequency effect - usually associated with lexical access - was found in first fixation and gaze durations in the visual search. Rayner and Raney (1996) further underline this interpretation with the observation that subject's average fixation durations were shorter in the search than in the reading condition. If analyses had been performed in the framework of theories of distributed processing (cf. section 2.3.1), delayed effects of fequency (lag effects) might have been found, because lexical access is considered as an automatic process, indicated for example by word frequency effects.

In an earlier study evaluation word reading times, Aaronson and Ferres (1984) observed different reading strategies when subjects read for verbatim recall than when they read for comprehension, giving true-false statements. In the 'recall'-strategy, subjects focussed on structure elements such as phrase boundaries and imbedded sentences compared to the 'comprehension'-mode, when readers focussed on semantic encoding. The results suggest that readers adapt their reading behavior in response to the given task demands.

Effects of reading instruction on eye movement patterns at the text-level have also been investigated. Differences in reading behavior have been found comparing the readers' different perspectives on expository text (Kaakinen, Hyönä, & Keenan, 2002). The authors reported that the eye movement patterns changed significantly, depending on the perspective, i.e. the kind of information the reader was instructed to extract from the text. Subjects showed longer fixation durations in first-pass as well as in second-pass reading on those parts of the text that contained perspective-relevant information compared with fixation durations on irrelevant parts of the text. In addition, there was a positive correlation between the time spent on relevant information and the recall performance of this information. In a recent study, Kaakinen and Hyönä (2008) demonstrated the same perspective-driven effects for narrative text comprehension, that is, readers spend more time reading perspective-relevant than perspective-irrelevant information and have a better memory of perspective-relevant than perspective-irrelevant text information after reading. Moreover, the authors showed that prior knowledge modulates the magnitude of the perspective effect (Kaakinen & Hyönä, 2008). Perspective-relevant information overlapping with prior knowledge is faster encoded online and better recalled offline than perspective-relevant information that did not overlap with readers' prior knowledge. Both studies clearly demonstrate how differences in reading goals modulate the eye movement patterns and hence the selectivity of words in text reading. The question arises whether this can also be found on sentence level.

A problem affecting almost all experiments that give precise instructions to their participants is the degree to which subjects fulfill the instructions. Recently, more interest has been found in the role of the reader's global attention during reading, a parameter hard to control in reading experiments. Even if instructed to read comprehensively, the experimenter cannot control nor avoid that the reader's mind wanders away while performing in (a maybe easy) reading experimental setup recording eye movements. So-called *mindless reading* or *zoning out* has hardly been investigated, simply because it can not be instructed and therefore, little is known about how the eye movement pattern look like if one reads without comprehension. Schooler and colleagues tried to experimentally investigate how often readers zone out while reading (Schooler, Reichle, & Halpern, 2004). During text reading, participants had to report via key press whenever they caught themselves not attending to the text. Additionally, subjects were intermittently probed regarding whether they had been zooing out in that particular moment. The authors found that readers do not attend the text in up to 23% of the time.

As a simulation of mindless reading, subjects were asked to 'read' texts of z-strings in order to avoid lexical and semantic effects on eye movement measures and to have a control condition for purely oculomotoric eye guidance while reading (Nuthmann et al., 2007; Vitu, O'Regan, Inhoff, & Topolski, 1995; Rayner & Fischer, 1996). In this experiment, all letters in a word are replaced by Zs, so that word boundaries and case were preserved. Eye movement patterns impressively resembled mindful reading behavior, indicating that subjects have a notion of how they move their eyes during reading, supporting the occulomotor component involved in fixational behavior. More importantly, in experiments investigating reading for comprehension or other reading instructions, eye movements of mindful and mindless reading cannot be differentiated. Therefore, researchers might consider that trials in which subjects zoned out are included in their reading data. It is open to debate how these kinds of trials influence the findings (or non-findings) of experimental effects. But see Kliegl, Masson, and Richter (in press) for a recent analysis on the influence of individual differences, presumably related to task engagement, on the size of experimental group effects.

2.4.4. Effects Due to Individual Differences

Individuals differ widely in their reading abilities. Finding the fundamental sources of individual differences has been an ongoing research interest. Individual differences between readers have been primarily investigated by educational psychologists in the context of developmental reading disorders and the attempt to find causally related factors for developmental dyslexia (for a review see e.g., Perfetti, 1994).

Influential early eye movement research on individual differences in reading has been conducted in the first half of the 20th century. Authors such as W.S. Morse, F.A. Ballantine, W.R. Dixon, and L.C. Gilbert among others investigated the individual reading styles dependent on education level, grade, reading subject, reading material, etc. (see Tinker, 1958, for a review). The general resume of this early research was that deficient eye movements are a symptom not the cause of poor reading and that these questions could also be addressed with less advanced techniques (Tinker, 1946). In neurolinguistics, where the importance of treating patients naturally leads to a focus on individual cases, there is also a longer tradition of single case studies (Caramazza & McCloskey, 1988; Shallice, 1979).

Since the cognitive turn in psychology, less research has been done on individual differences in normally skilled adult readers. This is partially due to a basic psycholinguistic assumption of general, universal processes involved in skilled reading that are acquired during the process of learning to read. Those general processes, common to all readers, have been in the main focus for a long research period that tries to test psycholinguistic, cognitive, and perceptual theories of reading and text comprehension. It is assumed that linguistic skills such as word identification, lexical processing, and syntactic parsing become modular with increasing reading practice (Stanovich, 1990; Caplan & Waters, 1999; Walczyk, 2000), basically meaning that these processes reach the level of 'automaticity'. In addition to language-specific components, reading for comprehension requires also general cognitive processing, such as applying knowledge and making inferences, which may be more prone to individual differences related to variations in processing resources. As mentioned in section 2.4.2, eye movements during reading reveal variance between subjects in fixation duration, in skipping and regression probability, etc., all factors that in sum determine the individual's reading speed. This general speed has been used to assess the level of the subject's reading skill. The 'random' differences in eye movement measures, subsumed under the label reading skill or reading style, were correlated with differences in psychometric variables, such as age, vocabulary size, working memory, or reading span.

Age and Vocabulary

Several studies reported relatively small effects of age in reading fixations (Humphrey, Kemper, & Radel, 2004; Kemper, Crow, &

Kemtes, 2004; Kemper, McDowd, & Kramer, 2006; Kliegl et al., 2004; Laubrock, Kliegl, & Engbert, 2006; Smiler, Gagne, & Stine-Morrow, 2003; Stine-Morrow, Loveless, & Soederberg, 1996) suggesting that reading is a highly automated skill acquired over age and that it is primarily influenced by changes in perceptual accuracy, not by changes in language processing abilities. It is a general finding that there are usually very little if any age-related changes in language comprehension abilities. Generally, the eye movements of old readers are sensitive to the same word variables as those of younger adults, but fixation durations are somewhat prolonged (e.g., Kliegl et al., 2004; Laubrock et al., 2006). Other age differences reported are that old readers tend to show more word skippings and more regressive saccades (Rayner et al., 2006; Laubrock et al., 2006), a behavior labeled as 'risky' strategy by Rayner and colleagues. Reading comprehension is found to decline with age, at least if comprehension requires understanding of complex syntactic structures. The problem arising here is put forward by Caplan and Waters (1999) who stress differences in methods and the precaution that needs to be taken when interpreting the results: In online measures during reading, few age effects are found, indicating a reading related language-specific competence preserved with age. If off-line measures are evaluated, such as answering comprehension questions and reasoning, age differences are tremendous, indicating that working memory limitations rather than a language-specific deficit in older readers are responsible for age effects.

A psychometric variable that is positively correlated with age is vocabulary score (Verhaegen, 2003). The receptive or passive vocabulary size in adults is usually tested by means of standardized questionaires, for example in subtests of Wechslers test of intelligence (Wechsler, 1964). It is open to debate if vocabulary size is a representative of the mental lexicon. But based on the fact that vocabulary knowledge has a unique impact on reading comprehension, in addition to the speed of lexical decoding (e.g., Braze, Tabor, Shankweiler, & Mencl, 2007), vocabulary size is a reasonable predictor for eye fixation patterns in reading. It has been demonstrated that vocabulary size and verbal ability correlate with eye movement behavior (Everatt & Underwood, 1994; Ashby et al., 2005).

Working Memory and Reading Span

Daneman and Carpenter (1980) were the first to introduce a measure of reading span. Subjects had to read a set of a varying number of sentences and to recall the last word in each sentence. The reading span was defined as the maximum number of sentences the subject could read with perfect recall of the final words (Daneman & Carpenter, 1980). Results showed that in college students reading span strongly correlated with reading comprehension. The authors claimed that reading span reflects working memory capacity, which itself is a crucial source for individual differences in language comprehension. Working memory efficiency during reading, measured via reading span, digit span, and recall performance, has early been found to be related to reading comprehension (Dixon, LeFevre, & Twilley, 1988). Calvo (2001) demonstrated that gaze duration across sentences was significantly shorter for high-span readers than low-span readers. Even stronger were the differences between reading span levels in fixation times that reflect late text integration processes, such as drawing inferences. A prominent theory emphasizing the role of limited processing resources in reading comprehension is the 'capacity limitation model' of Just and Carpenter (1992). They extended the initial observation of Daneman and Carpenter (1980) and described a model of individual differences in reading comprehension. There is clear age-related decline in language-related tasks when they tax the efficiency of executive control processes (Caplan & Waters, 1999; Conway & Engle, 1996).

In an eye movement study, Kennison and Clifton (1995) grouped subjects according to reading span levels (Daneman & Carpenter, 1980) and compared effects of parafoveal processing. Generally, low-span readers had significantly longer total reading times per sentence, they made many more forward fixations and regressive eye movements, and had longer gaze durations compared to high-span readers. Working memory capacity did not influence the parafoveal preview benefit, that is, high-span readers did not process more parafoveal information than did low-span readers. At the same time, high-span readers showed a higher skipping rate than low-span readers, indicating a higher parafoveal preprocessing rate.

2.5. Summary

Studies using eye movement recordings during reading have made a major contribution to the understanding of online processes involved in reading. It has been demonstrated that fixation probabilities and durations strongly depend on low-level factors, such saccade amplitude and landing position (cf. section 2.2), as well as on word properties, such as word length, lexical frequency, or contextual predictability (cf. section 2.3). In addition, theories of distributed processing argue that not only the properties of the fixated word influence fixation durations, but also those of the neighbouring words (cf. section 2.3.1). There is considerable variance in the reading behavior of skilled readers that is mostly ignored when research on effects of word-level parameters on reading times are investigated. Individual differences found in fixational behavior correlate with psychometric variables such as age, reading span, or general cognitive ability. In the literature, individual differences were often labeled 'reading strategies' in order to group readers into reader types. Here, reading strategy is defined as a reading behavior resulting from parameter specification within the system of the reading skill. Parameter specification occurs in response to task instruction or task demands that define the reading intention or reading goal (cf. section 2.4).

The present work tries to shed light on the dynamics of reader-level and word-level parameters on the reading behavior in sentence reading, where reader-level factors inlcude reading strategy and individual differences. In addition to variables located at the word-level, the individual as a unit and its associated variance with respect to age, vocabulary size, general reading speed, skipping behavior, and selectivity of fixated words, etc., will be included in the analyses of effects of distributed processing on reading times. Of main interest is the influence of reading strategies on oculomotor and word-level effects. Those strategy effects are experimentally induced and therefore, group differences in effect sizes indicate strategy effects in addition to effects of individual differences. Strategy effects are predominantly tested with linear mixed models (see section 3.4.3).

In contrast to studies that vary text difficulty and investigate the readers' adaption to these demands, the attempt is made to induce different reading strategies on the identical material. Because sentences, not texts, are used as reading material, the space (only a few words) and time (only a few seconds) to find strategy effects is very limited. Therefore, more local effects of reading strategies are tried to identify and describe and the magnitude of effect sizes is expected to fall into ranges of milliseconds. It has not yet been investigated how different task demands or different task instructions influence oculomotor effects and word- and reader-level effects in the context of distributed processing in reading. In this work, reading strategies are directly or indirectly induced, as presented in the next three chapters.

3. Methods for Experiments and Data Analysis

Eye movement data of a baseline experiment and three different experiments, alltogether seven different samples of participants, were analyzed in several group comparisons. Crucial for the study of strategy effects, of effects of individual differences, and of their interaction with oculomotor and word-level effects on reading behavior is the use of the identical reading material across experiments. The sentence material used in all reading experiments is the Potsdam Sentence Corpus (PSC), constructed by E. Grabner. After a description of the composition of the corpus, the baseline experiment is described in detail. Differences in design and procedures between experiments 1, 2, and 3 and the baseline experiment will be characterized in the relevant chapters that address a specific research question (see chapters 4, 5, and 6).

3.1. The Potsdam Sentence Corpus (PSC)

The PSC was used as reading material for all seven experimental groups. The corpus comprises 144 German single sentences (1,138 words) which were designed to represent a large variety of grammatical structure. Each sentence was constructed around one or two target words for which length and frequency are uncorrelated (see Kliegl et al., 2004, for further details on target words). The lengths of sentences range from 5 to 11 words (Mean = 7.9, SD = 1.4). All 144 Sentences are listed in Appendix A. Means of word length, word frequency, and word predictability are specified excluding the first and last word of each sentence, because fixations on those words were never used in the analyses provided in this work. Therefore, eye movement data on 850 words (comprising 550 unique words) of the PSC were included in the analyses.

3.1.1. Word Length and Printed Word Frequency

Excluding the first and last word of each sentence, the number of words for lengths 2 to 13 and more letters are: 50, 217, 117, 119, 99, 83, 60, 52, 12, 18, 9, and 8 (the 13 and more letters category contains 6 words of length 14 to 20). Mean word length is 5.4 letters (SD = 2.6). Frequency norms for the PSC are taken from a German word corpus based on 123 million words (DWDS; Geyken, 2007; Kliegl, Geyken, Hanneforth, & Würzner, 2006). Word frequencies of the 850 words range from 0 to 31,972 per million. The mean log frequency (incremented by 1) is 2.3 (SD = 1.3).

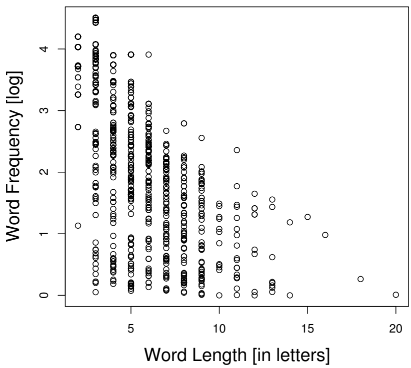


Figure 3.1.: Correlation of length and log frequency of the 850 selected words of the Potsdam Sentence Corpus (PSC).

As illustrated in Figure 3.1, log word frequency and word length are strongly correlated (r = - 0.66). This is a typical result for text statistics due to the distribution of syntactic word classes in natural texts (e.g., Goodman, 1985). Function words, such as determiners, conjunctions, and prepositions are very high frequent words in a language and usually consist of only a few letters. As expected, this correlation is also found in the PSC: Content words have a mean length of 6.4 (SD = 2.6) letters and a mean log frequency of 1.5 (SD = 0.9), whereas function words are on average shorter (mean= 3.5 letters, SD= 1.3) and more frequent (mean log frequency = 3.6, SD = 0.7) than content words. Word length and log frequency are also strongly correlated within the classes of content (r = -0.44) and function words (r = - 0.63). On one side, considering the different frequency ranges of content and function words, one might argue for a differentiation of word frequency dependent on the lexical status (content vs. function word). On the other side, lexical status and word frequency correlate strongly in the selected 850 words of the PSC (r = 0.75), which suggests that variance related to lexical processing of content vs. function words can be captured solely with word frequency (but see Kliegl, 2007, for analysis with word frequency nested within lexical status).

3.1.2. Word Predictability

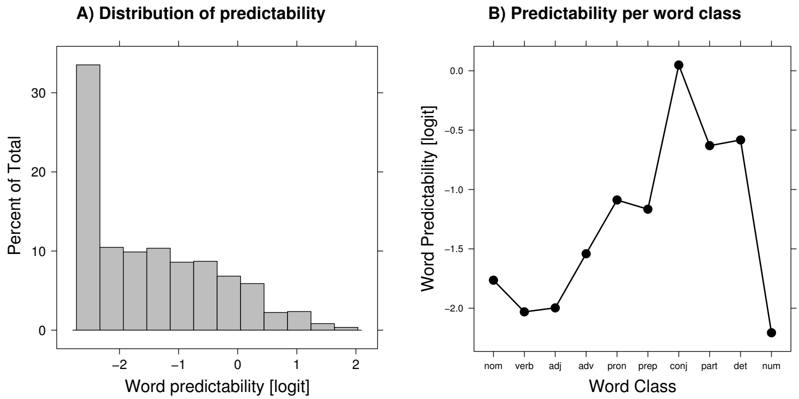


Figure 3.2.: A) Distribution of word predictability of the 850 selected words of the PSC; B) Mean logit predictability in dependency of word class; adj= adjectives, adv= adverbs, pron = pronouns, prep = prepositions, conj = conjunctions, part = particles, det = determiners, num = numerals.

Norms for word predictability in the PSC were collected in an independent study with 272 native speakers of German. Participants were high school students, university students, and older adults, ranging in age from 17 to 80 years. Participants performed an incremental cloze task (Taylor, 1953), and 83 complete predictability protocols for each word of the PSC were collected. Predictability is defined as the probability of predicting a word when knowing the preceding part of the sentence. Excluding the first and last word of each sentence, word predictabilities in the PSC range from 0 to 0.98 with a mean (SD) predictability of 0.16 (0.25). In order to stretch the tails of the distribution, the predictabilities were submitted to a logit transformation[1]. Further details on the norming study are provided in Kliegl et al. (2004). More than 30% of the 850 words are almost impossible to predict from the context. The most predictable words belong to the closed class, such as conjunctions, particles, determiners, pronouns, and prepositions (see Figure 3.2 A and B for an illustration).

3.2. The Baseline Experiment: Reading for Comprehension

In the original reading experiment, eye movements of different age samples were recorded while reading the PSC for comprehension. Data of three of the nine samples (see Kliegl et al., 2006) serve as an age-machted control baseline to evalute the experimental manipulations. Note that this is a post-hoc between subjects design. Question material, apparatus, and procedure of the baseline experiment are described in detail in the following. Subject information and differences in design and procedure compared to the other three experiments are explicated in the chapters 4, 5, and 6 that address the theoretical questions experimentally.

3.2.1. Comprehension Questions for the PSC

In the original reading condition, easy multiple-choice comprehension questions were asked after 27% of the sentence trials. Questions used identical wording of the preceding sentence and three alternative choices were provided. These questions were similar to single word probe tasks and therefore correct responses were possible solely by visual word recognition of the answering options. For example, after the sentence

> *Martins gebrochener Zeh schwoll rasch an.*
> (Martin's broken toe swelled quickly.)

the following comprehension question was asked:

> *Was schwoll an?*
> (What did swell?)

[1] A logit is the logarithm of the odds, here definded as .5*ln(pred/(1- pred). Predictabilities of zero were replaced with 1/(2*83) (= - 2.55) and predictability of one with (2*83 - 1)/(2*83) (= 2.55). For the 850 words of the PSC the mean (SD) logit predictability is -1.48 (1.1). For example, for a word with predictability .50, the odds of guessing are one (1:1 = 1) and the log odds of guessing are zero (ln(1) = 0). Thus, words with predictability larger than .50 yield positive logits, and those with predictabilities smaller than .50 yield negative logits.

and the choices given were the following:

Fuss *Ferse* *Zeh*
(foot heel toe)

Further examples are listed in Appendix A. 32% of the easy questions were subject questions, 28% object questions, 18% verbal questions, and 22% consisted of questions for time or location. The original reading experiment contained questions for 133 of the 144 sentences.

3.2.2. Apparatus

Single sentences of the PSC were presented on the center-line of a 21-in. EYE-Q 650 Monitor (832 pixels x 632 pixels resolution; frame rate 75 Hz; font: regular, New Courier, 12 point) controlled by custom C programs (by S. Kern) on an Apple Power Macintosh G3 computer. Participants were seated at a 60cm-distance in front of the monitor with the head positioned on a chin rest. One letter subtended 0.38° of visual angle. Eye movements of two samples were recorded with an EyeLink II system, manufactured by SR Research Ltd (Osgoode, Ontario, Canada). This eye tracker is an infrared video-based tracking system. There are two cameras, one for each eye, mounted on a headband. Images were sampled at a rate of 500 Hz, and gaze location and pupil size were recorded. The sample 'original old' (cf. section 4.1) was recorded with an EyeLink I system with a 250-Hz-sampling rate. All recordings and calibrations were binocular.

3.2.3. Procedure

Before the reading experiment recording eye movements was conducted, participants' visual acuity was assessed with a standard optical chart (Landolt rings; 5 m distance, e.g. Bach, 1996, 2007) and eye dominance was tested. Vocabulary size was tested with form B of the *Mehrfach-Wortschatz-Test* (MWT-B; Lehrl, 1977). As an index of processing speed, the Digit-Symbol-Test of the *HAWIE* (Wechsler, 1964) was administered. Following these tests, participants were seated in front of the monitor and were calibrated with a standard nine-point grid for both eyes. They were instructed to read the sentence for comprehension and to fixate on 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 center 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 center of the first word was positioned at the location of the fixation spot. Therefore, each sentence-initial word was read from a word-specific optimal viewing position (e.g. O'Regan & Lévy-Schoen, 1987). Sentences were shown until participants looked to the lower right corner of the screen. Then in 27% of the trials, the sentence was replaced by an easy three-alternative multiple-choice question that the participant answered via a mouse click. In the remaining 73% of the trials, no question was shown. After every 15 sentences, a complete recalibration with the nine-point grid was performed. 10 training trials preceded 144 experimental trials. For their participation, subjects either received course credit or were paid 5-7 €/ hour.

3.3. Raw Data Analysis

In all experiments, eye movement records from reading the 144 sentences of the PSC were screened for blinks and loss of measurement. The raw data delivered by the EyeLink system, consisting of time stamps and the horizontal and vertical positions of the left and right eye on the screen (in pixels), were reduced to a fixation format after detecting saccades as rapid binocular eye movements. For the detection of saccades a velocity-based detection algorithm was used, originally developed for the detection and analyses of microsaccades (Engbert & Kliegl, 2003; Engbert & Mergenthaler, 2006). The basic procedure of the algorithm is, that changes in eye velocities are used to distinguish saccades from fixations. Reading saccades were detected binocularly, where binocularity was defined as a temporal overlap of the right and left eye position data. Sentence trials with two or less fixations were excluded. Fixations were assigned to letters, which provide the basis for analyses of the specific fixation positions in reading research. This data treatment was done in MATLAB (The MathWorks, Inc., Natick, MA). For statistical analyses, fixations were selected according to the criteria defining several dependent fixation duration measures.

3.3.1. Data Selection and Computation of Dependent Variables

Even though reading saccades, and thus fixations, were detected as binocular events, all statistical analyses were based only on right eye data. The total number of all detected fixations (1) was submitted to further selection criteria, and different measures were computed (cf. Table 3.1).

Eye movement recordings at the beginnings of sentences can be affected by artefacts due to the initial drift correction and sentence appearance. Recordings at the end of sentences can be influenced by artefacts due to the subject's search for the dot to let the sentence disappear from the screen. Therefore, fixations meeting the following criteria are excluded from all other analyses: (2) first or last fixations in a sentence and fixations on the first or last word of a sentence trial. After applying citerion (2), the remaining fixations are called valid fixations (3), because according to definition, they are all within-sentence fixations and constitute the data pool for various fixation duration measures in *first-pass reading* (4), *second-pass reading* (5), or *rereading* (6) (cf. section 2.1 for a definition of first and second-pass). Rereading fixations are all fixations that are made after the last word of the sentence has been reached once.

Fixations were further distinguished along the sequential order of fixational movement. The following fixation types were detected in all reading passes: (a) *single fixations*, that are preceeded and followed by a forward saccade; (b) single fixations that constitute the end of a first-pass and are labeled *regression origins* (cf. Figure 2.1); (c) *multiple fixation cases*, i.e. cases of two and more fixations on a word, that are borderd by a forward saccades; (d) *multiple fixation regression origins* are multiple fixation cases that mark the end of a pass and are regression origins. In second-pass reading and rereading, three additional fixation types need to be distinguished: (e) single fixations, that are the target of a regression, are preceeded by a backward saccade and followed by a forward saccade; they are labeled *single fixation regression goals*; (f) multiple fixation cases that are the target of a regression, are marked as *multiple fixation regression goals*; and (g) fixations that are the goal of a regressive saccade and a regression origin at the same time.

First-pass single-fixation duration and gaze duration were used as dependent measures in the linear mixed modeling (cf. section 3.4.3). Gaze duration was defined as the sum of fixation durations of all first-pass multiple fixations cases (c). Thus, gaze duration includes fixations different from first-pass single fixations and therefore provides new information to the analyses. Fixations that mark the end of the first-pass (b, d) were excluded from single fixation and gaze duration.

Table 3.1.: Overview of data selection and eye movement variables; note that 1) = 2) + 3); 4) = a)+ b)+ ...+ d); 5) = a.1)+ ...+ g.1); 6) = a.2) + ...+ g.2.

1) TOTAL NUMBER OF FIXATIONS		
2) First/last word; first/last fixation		
3) Number of valid fixations		
4) FIRST-PASS	5) SECOND-PASS	6) REREADING
a) Single fixations (SF)	a.1) SF	a.2) SF
b) SF regression origins	b.1) SF regr. origins	b.2) SF regr. origins
c) Multiple fixations (MF)	c.1) MF	c.2) MF
d) MF regression origins	d.1) MF regr. origins	d.2) MF regr. origins
-	e.1) SF regr. goals	e.2) SF regr. goals
-	f.1) MF regr. goals	f.2) MF regr. goals
-	g.1) SF regr. goals	g.2) SF regr. goals
-	& origins	& origins

3.4. Statistical Data Analysis

Reader-level and fixation-level analyses were distinguished. At the reader level, single-fixation, first-fixation, and gaze duration as well as total word reading time, regression and skipping probability, and the mean number and location of fixations within words as a function of experiment are evaluated. At the fixation level, the impact of reader-level predictors as well as word-level characteristics on first-pass reading measures, namely single-fixation durations and gaze duration, are examined. As a first, traditional inspection of the effects of experimental manipulation, word-based summary statistics averaged across subjects (F1 ANOVA) are computed on the basis of the total number of all detected fixations of the right eye. Analyses were performed in the R system for statistical computing (version 2.7.0; R Development Core Team, 2008). In the following, issues of experimental designs and corpus analyses, and problems of statistical analysis are shortly explicated. Then, data fitting using linear mixed models is explained in more detail.

3.4.1. The Quasi-Experimental Design

The scientific method most amenable for causal explanations is the 'true experimental design'. By definition, experimenters in a true experimental design have control over samples, i.e. participants are randomly assigned to experimental conditions, and experimental factors are systematically manipulated, preferably in an orthogonal design (e.g., Gribbons & Herman, 1997). In contrast to experimental factors, predictors that cannot be fully controlled or purposively manipulated by the experimenter are called quasi-experimental factors. Any experimental setup that includes at least one quasi-experimental factor is a quasi-experimental design.

With respect to research in continuous reading, a true experimental setup includes sentences in which only controlled target words of theoretical interest are the unit of analysis (e.g. Rayner & Fischer, 1996). The target words in the PSC would match this criterion because they are uncorrelated in length and frequency. However, in the present work, all words of the PSC are considered for the so-called 'corpus analyses', acknowledging that several word characteristics might correlate (cf. section 3.1). Experimental condition is a true experimental factor if subjects are randomly assigned to reading conditions. In this work, the 'original' samples were tested before the other three experimental variations were administered. Therefore, reading condition in comparisons with the baseline experiment is a quasi-experimental factor. At the same time, other between subject factors, for example 'age', are quasi-experimental factors, because a random assignment of participants to age groups is not possible. Rather, the characteristics inherent to the subjects define their group belonging. Independent variables of word properties, such as frequency, length, or predictability are also quasi-experimental factors because the word itself defines its category belonging. Even in experiments, in which word frequency is systematically manipulated across words, the selected frequency range probably does not cover the full range of natural language and thus, results of this experimental factor are in fact quasi-experimental. The status of a quasi-experimental design has to be kept in mind, if the interpretation and generalisability of the results is to be discussed, because quasi-experimental factors always yield only correlational evidence (Kliegl, 2007). The advantage of a corpus analysis for reading research is the consideration of the complete text or sentence structure in the analyses. When focussing on effects of eye movement behavior on critical target words, these words are usually content words of a limited frequency range. Thus, these experimental studies are somewhat remote from normal reading and conclusions can only be drawn for the specific sentence structure or word combinations tested (Radach & Kennedy, 2004). In contrast, an analysis taking into account the whole range of words occurring in a sentence or text is able to detect more complex dynamics of eye movement control in reading, for example with respect to eye movement patterns due to the alternation of content and function words. In addition, problems of statistical power are reduced because a larger number of data points of a trial is evaluated. For a critical discussion on corpus analyses see Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007) and Kliegl (2007) for a different perspective.

3.4.2. Repeated Measures and "Randomness"

All experiments analyzed in this work are repeated measures designs. Eye movements from n subjects for 1138 words or 144 sentences are recorded. For each word (or sentence), we have n repeated measures (one from each subject). At the same time, we have 1138 repeated measures for each subject. Sentences, words in the sentences, and subjects are sampled randomly from populations of sentences, words, and participants. These identifiers are not repeatable and thus are called *random factors*. Subjects, words, and sentences are treated as random variables because the interest of this research is not in the experimental effects present in only those individuals who participated in the study or only those sentences used in the experiment, but rather in effects present in all speakers across the language studied (for a discussion on this problem see Clark, 1973). The "random" sample of participants read a "random" sample of 144 sentences. Thus, these factors are crossed (Baayen, Davidson, & Bates, 2008). "Random" is put in quotes because, as for almost all studies, the representativeness of participants and sentences is not exactly known. This argument also refers back to the point mentioned in the previous section 3.4.1 why a predictor is called quasi-experimental. For example, with respect to the old adults, who participated in this study, there is presumably a bias towards high degrees of mental fitness and achievement motivation whereas young adults in this work are representative of high-school students. The sentences of the PSC can be considered as drawn from a population of easy prose. Ignoring the first and last words of the sentences, there is a total of 850 words, but only 550 unique words. These 550 words are specified as a third random factor and are partially crossed with sentences, that is, no word occurs in all sentences but some words occur in more than one sentence and sometimes even more than once in a single sentence (e.g., determiners, prepositions).

A factor is repeatable, if the set of possible levels of that factor is fixed, and if each of these levels can be repeated (Baayen, 2008). These predictors are called *fixed effects*. Repeatable factors of interest in this work include for example age, vocabulary size, or reading condition of the subjects as well as lexical covariates such as word frequency, word predictability, or word length. For several decades, researchers using the eye movement technique averaged their data across subjects (ignoring the randomness of participants), whereas research focussed on word-level effects ignored the variety to items. In contrast, in psycholinguistic research typically subjects and words are treated as random effects. Experimental and quasi-experimental effects are usually assessed with F1 and F2 ANOVAs. Here, the attempt is made to consider both sources of variance in a single framework.

Due to the repeated measures design, there is an inherent structure in the present data. Fixations, the underlying dependent measure, are clustered within subjects, words, and sentences. For example, subjects (higher level) are sampled first, and fixations (lower level) have been sampled within subjects. Therefore, a *multilevel* or *hierarchical* model is needed for the analysis, that simultaneously takes into account the randomness of subjects and items and the fixed effects, so that the model generalizes to the populations of subjects and items. This can be done with *linear mixed models* (Pinheiro & Bates, 2000), because they incorporate fixed and random effects at several levels. The models used for this work will be explained in the next section (see Baayen et al. (2008), Kliegl et al. (in press), Oberauer and Kliegl (2006), for applications of these techniques to test cross-level interactions).

3.4.3. Linear Mixed Models

The *lmer* program of the *lme4* package (Bates, 2007) is used for estimating fixed and random coefficients of linear mixed models (LMM). The packages and programs are supplied in the R system for statistical computing (version 2.7.0; R Development Core Team, 2008). The *lmer* program allows the specification of random and fixed effects, and coefficients are estimated simultaneously. There are simulation studies showing that lme models are much less susceptible to problems of sphericity, design effects, and missing data than F1 and F2 mixed-model ANOVAs or repeated-measures multiple-regression analyses over subjects and items, respectively (Baayen, 2008; Gelman & Hill, 2007; Quené & van den Bergh, 2004). Missing data and thus, imbalances in design are hardly avoidable for eye movements in reading because of, for example, word skipping or loss of measurement due to blinks or technical flaws.

In linear mixed models, random effects are modeled as random variables with a mean of zero and an unknown variance. The standard deviations associated with random effects are parameters that are estimated, just as the coefficients for the fixed effects are parameters that are estimated (Baayen, 2008). According to Gelman and Hill (2007, p. 2), all regression parameters are "random" in the sense that they are considered random outcomes of a process identified with a model predicting them. In the authors' framework, fixed effects are a special case of random effects, in which the higher level (e.g. subject) variance is set to 0. Therefore, they avoid the term "random" and "fixed" and refer to varying or constant coefficients (intercepts and/or slopes) with respect to subjects, words, and sentences (Gelman & Hill, 2007, p. 245). This terminology is partially adopted in the following sections.

Each additional coefficient (varying or fixed) makes the model more complex. Thus, one has to control for the increase in explanatory power in order to justify model complexity. Model comparison is evaluated by measures of goodness of fit. Models were fit by maximum likelihood (ML) for model comparisons with differing fixed effects structure and by restricted maximum likelihood (REML) for model comparisons with identical fixed effects structure and differing number random-effect. For assessment of (differences in) goodness of fit, the *lmer* program provides the following model selection criteria: the Akaike Information Criterion (AIC), which decreases with goodness of fit; the Bayesian Information Criterion (BIC), which also decreases with goodness of fit; and the log likelihood (logLik), that increases with goodness of fit. That is, if comparing two models for a data set, one prefers the model with the smaller AIC or BIC (Pinheiro & Bates, 2000). The AIC and the BIC consider both the fit and the complexity of the models, and the BIC includes a penalty term correcting for the number of estimated parameters and number of observations (see Johnson & Omland, 2004, for a detailed description of model selection approaches). In the case of model comparisons, the likelihood ratio test provides the χ^2 statistic and its associated p-value, and AIC and BIC serve as information against overfitting. Additionally, for an estimation of explained variance for unique predictors, especially for random effects, the R^2 is provided in the results. Models were built up incrementally, following the theoretical interest, i.e., first, random effects were added to the model, then reader-level fixed effects, followed by word-level fixed effects.

As in classical reaction time paradigms, fixation durations are normally distributed but there is a longer tail for long latencies (O'Regan, 1990). Therefore, for statistical analyses with *lmer*, all fixation durations are logarithmically transformed. Continuous predictors are centered around the individual subject mean. In other words, continuous word level predictors include only the residual variance; this means, that the observed effects can clearly be attributed to word properties. For the illustration of results, real-time durations are plotted in the graphics. R-packages mainly used for plotting are *reshape* (Wickham, 2007), *ggplot2* (Wickham, 2008), and *lattice* (Sarkar, 2008).

Linear mixed models specified in this work yield four different types of effects. First, there are estimates of variances relating to the random effects of readers, sentences, and words. Second, they yield estimates of reader-level fixed effects, corresponding to the effects of the experimental and quasi-experimental (e.g. age) manipulations. Third, lme yields estimates of fixed effects at the word level, corresponding to unstandardized regression coefficients. Fourth, there are interactions within and between reader-level and word-level fixed effects.

Random Effects (Varying Intercepts)

In a first step of modeling, readers (subject ID), sentences (sentence ID), and words (word ID) are included as random factors in the model. The program assumes that each reader, sentence, or word can be characterized by a mean fixation duration and that these values are sampled from three normal distributions centered around the overall mean fixation duration. The mean of these random intercepts is zero, so that the grand mean is estimated as the fixed-effect intercept. The variances of these mean fixation durations across readers, sentences, and words are estimated as three random-effects model parameters. Adding subject ID as a random effect in the analyses is one part of the attempt to account for individual differences in addition to factors defined as reading strategies.

Reader-Level Predictors (Constant Coefficients)

Reader-level predictors are the vocabulary score (Lehrl, 1977), the digit-symbol score (Wechsler, 1964) (both centered around 0), and the trial number. These effects are added to the model in a second step. A number of other characteristics of readers, possibly indicative of parameter specification of reading strategy as well as of general individual difference (e.g., the percentage of fixated function words, or skipping probability), also belong in this predictor category. To account for individual differences in general reading style without increasing model complexity, the dimension of these individual profiles need to be reduced. To this end, a principle component analysis (PCA; e.g., Jolliffe, 2002) was run over ten relevant predictors at the subject level. For single fixation duration, these include each subject's 1) mean incoming saccade amplitude, 2) mean outgoing saccade amplitude, 3) mean skipping probability of the previous word, 4) mean skipping probability of the next word, 5) mean frequency, 6) mean length (using the reciprocal value 1/length), and 7) mean content/function word ratio for the fixated word, and 8) mean frequency, 9) mean length (reciprocal value), and 10) mean content/function word ratio for the previous word n-1. For first-pass gaze duration, lag word-level predictors were excluded from the PCA. The scores of the first and second main component are added as reader-level fixed effects into the *lmer*-model.

The main interest, however, lies in the estimation of the impact of experimental manipulation. Therefore, a dummy-coded fixed effect of experimental condition and its interactions with other, basically word-level predictors (cross-level interactions) are included to the model.

Word-Level Predictors (Constant Coefficients)

For the selection of word-level predictors, the set of fixed effects and interactions terms was chosen following the results of the regression model based on 222 readers in Kliegl et al. (2006). This set of predictors includes log word frequency (linear, quadratic, and cubic trend), logit word predictability, word length (using the reciprocal value 1/length), and the linear and quadratic components of relative fixation position (defined as letter-position/word-length scaled to zero, representing the center of the word). According to the idea of distributed processing (cf. section 2.3.1), it is expected that not only the characteristics of the fixated word n but also those of its neighbors (word n-1 and word n+1) influence the target fixation duration. Hence, additional predictors are log word frequency, logit predictability, and word length (reciprocal value) of the previous word n-1 as well as log word frequency, logit predictability, and word length (reciprocal value) of the upcoming word n+1. Finally, the predictor set also included incoming and outgoing saccade amplitude. The report of interactions among these variables will depend on the statistical power for detecting higher-order interactions in the specific experiment. Generally, in a preliminary analysis, interaction terms that were not significant were eliminated. The complete set of predictors comprising a final, selected model is listed in the equivalent appendix.

4. Strategy Effects Due to Task Demands

In the following chapter, the first theoretical question is adressed, namely how task demands in reading for comprehension influence the eye movement behavior during reading. Reading for comprehension is the most natural reading task. In reading experiments, comprehension questions are routinely asked after a certain percentage of trials in order to make sure that participants attentively read the material and do not simply scan across the lines. Comprehension question vary in their difficulty, focussing on various aspects of the content's mental representation. Questions are easier to comprehend when their form matches the form of the original information (Kess, 1992). The question arises if subjects attend to the task throughout the whole reading experiment or if they possibly perform trials of 'mindless' reading, especially when easy comprehension questions are asked. One feasible test would be the investigation of trial effects across the reading experiment. But for example, differences in time spent on a specific task is not necessarily related to attention, but also to practice. The question can also be addressed in an experimental manipulation.

In this study, the attempt is made to manipulate the depth of reading comprehension and the attention to the task by altering the frequency and difficulty of comprehension questions that are routinely asked in sentence reading paradigms. The goal in Experiment 1 is to induce a 'mindful' reading strategy by using difficult questions with reduced verbatim overlap between sentence and question material. Reading with the original 'easy' questions for the PSC (the baseline experiments, cf. section 3.2) is contrasted with reading for comprehension with frequent difficult questions (Experiment 1, cf. section 4.2). Since the reading instruction given to the subjects is identical to the baseline experiment, the manipulation of reading strategy in experiment 1 is indirectly induced via the design of the questions.

Because the sentence material in both experiments is identical (PSC), on the one hand, one could argue that the highly automatized first-pass reading of skilled readers should not depend on the kind of comprehension questions that are asked *after* reading the sentence, and thus, no experimental effects are expected. On the other hand, the cognitive load and attentional demands induced by very frequent and difficult questions, might very well lead to differences in fixational behavior, even in first-pass reading. This might be especially the case in subjects with reduced working memory capacity, for example in older readers. Whereas age effects in reading performance have mainly been found in offline tasks taxing working memory resources, it is resonable to expect differences even in online processing due to memory loads in the offline task. Therefore, the manipulation of comprehension demands is tested both with young and old readers, and it is expected that old readers are more sensitive to the question manipulation than young readers.

I assume that the manipulation of question difficulty and frequency will affect the depth of sentence processing, reflected in the online eye movement behavior. When effects of 'mindfulness' or depths of comprehension are to be evaluated, word predictability might be a lexical variable of special interest. As word predictability is an index of how well a word can be 'foreseen' out of its current sentence context and of how easily the word can be integrated into the processed context, differences in effects of word predictability on fixation durations might be a measure of how deliberate the readers process the sentence. It can be argued that, to be most efficient in reading for comprehension, the reader takes the active role of processing the incoming visual input as well as making (unconscious) predictions based on the current input about what information might come next. If this is true, the effects of word predictability, especially those of the upcoming word n+1, are expected to be stronger when subjects perform deep comprehension during reading than when they perform normal, presumably more superficial reading. In other words, as a possible index of depth of comprehension, word predictability effects are expected to be stronger if mindful reading is induced.

In the following sections, participants of this group comparison and deviating materials and procedures of Experiment 1 are described in detail. In addition to the investigation of young readers' eye movement behavior in the two reading conditions, two groups of older readers are tested in this study. Then, results of raw data analyses and data selection are provided for each experimental group. The impact of the experimental manipulation will be evaluated by considering gobal analyses of word-based statistics across subjects. Here, measures of interest are fixational probabilities, various fixation duration measures and fixation position. Of theoretical interest are also differences between experimental groups in response accuaracies and response latencies for the comprehension questions. The hardest test for the experimental manipulation is the evaluation of differences between reading conditions with respect to word-level and reader-level predictors on fixation level. These effects will be provided in the LMM-results, and focus on cross-level interactions of reading condition and word-level parameters.

4.1. Participants

Table 4.1.: Group statistics on age, vocabulary size, digit-symbol-task, and reading span for the samples 'original young', 'original old', 'hard young', and 'hard old'.

Group	N	M	SD	Range
Age				
Original young	24	17.6	0.58	16 - 18
Original old	32	70.6	3.97	65 - 84
Hard young	30	18.5	0.86	17 - 20
Hard old	23	68.0	3.27	65 - 76
Vocabulary Test				
Original young	24	29.5	2.79	26 - 34
Original old	32	33.1	1.22	30 - 36
Hard young	30	30.7	2.79	25 - 39
Hard old	23	33.0	1.42	30 - 36
Digit-Symbol Test				
Original young	24	61.6	10.34	42 - 86
Original old	32	49.5	10.15	36 - 69
Hard young	30	61.3	9.14	45 - 80
Hard old	23	46.3	8.14	33 - 60
Reading Span				
Original young	24	-	-	-
Original old	32	-	-	-
Hard young	30	0.76	0.12	0.48 - 0.97
Hard old	23	0.44	0.13	0.19 - 0.68

Data of four samples are compared in this study (cf. Table 4.1). A group of 24 high school students (in the following labeled 'original young') and a group of 32 older readers ('original old') read the PSC with easy questions (baseline experiments, cf. section 3.2; see Tables C.1 and C.2 in Appendix C for a complete list of subjects' information)[2]. An age matched group of 30 high school students ('hard young') and 25 old readers ('hard old') read the PSC with frequent, difficult questions (cf. section 4.2). Data of two participants from the 'hard old' group were excluded from analysis because subjects provided less than 150 fixations to the whole data pool. All participants were native speakers of German. They all had normal or corrected to normal vision.

Comparisons between same-age groups revealed no significant differences in age, in scores on Lehrl's (1977) multiple-choice measure of vocabulary or in Wechsler's (1964) Digit-Symbol-Test (all $p > 0.05$). A typical age effect found is the higher vocabulary score ($F(1,105) = 46.23$, $MSe = 4.73$, $p < 0.001$) and a lower score in the digit symbol test ($F(1,105) = 52.79$, $MSe = 90.8$, $p < 0.001$) for older readers in comparison to the young readers. A reading-span test as reported in Oberauer, Süß, Wilhelm, and Wittmann (2003) was administered for the hard groups. Reading span was reliably reduced with age ($F(1,51) = 93.03$, $MSe = 0.01$ $p < 0.001$).

4.2. Experiment 1: Reading with Frequent Difficult Questions

The main differences of experiment 1 in comparison to the baseline experiment are described in this section. These are differences in the difficulty of the comprehension questions for each sentence trial and the frequency of the occurence of comprehension questions.

4.2.1. Hard Comprehension Questions for the PSC

In the difficult reading condition, a three alternative multiple-choice comprehension question was asked after each sentence of the PSC. Examples of difficult questions for the 144 sentences of the PSC are listed in Appendix A. The combination of questions and the alternative choices were designed to reduce the verbatim overlap with the original sentence in order to make a purely visual solution of the question impossible (e.g., by a simple word form recognition). The content of all questions aimed at checking a complete propositional representation of the sentence, also testing inferences. Subjects had to represent a fully integrated model of the sentence to answer the subsequent

[2]Data of these two samples were included in the analyses in Kliegl et al. (2006) and (Kliegl, 2007), labeled sample 4 and sample 9 respectively.

question. For example, given the identical sentence used in section 3.2.1

> *Martins gebrochener Zeh schwoll rasch an.*
> (Martin's broken toe swelled quickly.),

the hard question was:

> *Was passierte mit Martins Zeh?*
> (What happened to Martin's toe?)

and the choices provided were:

> *wurde blau* *wurde steif* *wurde dick*
> (became blue became stiff became thick)

Six different questions types were used that occurred in equal proportions: (1) subject-questions, (2) object-questions, questions asking for (3) verbal information, (4) time and location information, and (5) other adverbial information given in the sentence. Verification questions (6) maintained a statement about the sentence the subject had to respond to with 'yes/no/maybe'- type of answer. Note that questions categorized as a subject-question did not necessarily ask for the grammatical subject given in the sentence but contained a wh-question word, asking for any information given in the sentence. In the following, I will refer to the manipulation of question difficulty and frequency as 'question' or the 'question-effect'.

4.2.2. Apparatus and Procedure

The apparatus was identical to the baseline experiment (cf. section 3.2.2). Eye movements of the samples 'hard young' and 'hard old' were recorded with a 500-Hz-sampling rate.

The basic procedure was identical to the baseline experiment (cf. section 3.2.3). In addition to estimating vocabulary size (Lehrl, 1977) and the index of attention (Wechsler, 1964), the reading span was examined using an adaption of the test by Daneman and Carpenter (Oberauer et al., 2003). A test of visual working memory using abstract figures was administered both for the group of 'hard young' and 'hard old'. In 100% of the trials, the sentence was replaced by a hard three-alternative multiple-choice question that the participant answered via a mouse click.

4.3. Results of Saccade Detection and Fixation Selection

According to the selection criteria defined in section 3.3.1, Table 4.2 and Table 4.3 list the results of the fixation split up for the experimental samples of the baseline experiment and the experiment using hard comprehension questions. In Table 4.2, the top row lists the total number of detected saccades for the right eye (1). The second row provides the number of fixations on first and last words of the sentences and the number of first and last fixations of a trial (2). A breakdown of numbers and percentages of different reading passes relative to valid within-sentence fixations (3) is provided in the bottom part of the Table. Note that in Tables 4.2 and 4.3, the total number and percentages of fixation types per experimental group is reported, but ANOVA statistics are based on subjects' mean percentage per fixation type[3].

Table 4.2.: Distribution of fixation types of fixations for the samples 'original young', 'original old', 'hard young', and 'hard old'; data are from right eye; rows 2) and 3) sum up to 1), rows 4), 5), and 6) sum up to 3).

VARIABLE		ORIG. YOUNG	ORIG. OLD	HARD YOUNG	HARD OLD
1) N of fixations (total)		25,137	36,050	35,368	25,529
2) First/ last word;	*N*	7,276	11,508	9,754	7,860
First/ last fixation	%	29	32	28	31
3) N of valid fixations	*N*	17,861	24,542	25,614	17,669
	%	71	68	72	69
4) First-pass	*N*	15,743	19,085	17,758	11,380
	%	88	78	69	64
5) Second-pass	*N*	1,590	4,264	3,741	3,741
	%	9	17	15	21
6) Rereading	*N*	528	1,193	4,115	2,548
	%	3	5	16	15

In all groups, first and last fixations and fixations on first and last words (2) make about 30% of the total number of fixations, but differences in number of valid fixations (3) between samples were reliable.

[3] Since not all subjects of the experimental samples provided fixations to each fixation type, the degrees of freedom may vary between ANOVA statistics for different fixation types.

Table 4.3.: Number and proportion of fixation types in first-pass reading, second-pass reading, and rereading for the samples 'original young', 'original old', 'hard young', and 'hard old'; data are from right eye; note that the number of fixations types per reading pass sum up to the number of fixations per pass listed in Table 4.2.

VARIABLE		ORIG. YOUNG	ORIG. OLD	HARD YOUNG	HARD OLD
Fixation types in 1st-pass reading					
a) Single fixations (SF)	N	10,828	12,106	10,910	6,445
	%	69	63	61	57
b) SF regression origin	N	875	2,587	1,678	1,741
	%	5.6	14	10	15
c) Multiple fixations (MF)[4]	N	3,976	4,189	4,948	3,008
	%	25	22	28	26
d) MF regression origin	N	64	203	222	186
	%	0.4	1	1	2
Fixation types in 2nd-pass reading					
a.1) Single fixations (SF)	N	500	1,137	1,173	1,215
	%	31	27	31	33
b.1) SF regression origin	N	31	125	108	160
	%	2	3	3	4
c.1) Multiple fixations (MF)	N	230	575	638	694
	%	14	13	17	19
e.1) SF regression goals	N	725	1,985	1,401	1,180
	%	46	47	37	31
f.1) MF regression goals	N	41	178	169	163
	%	3	4	5	4
g.1) SF regr. goals & origin	N	63	264	252	329
	%	4	6	7	9
Fixation types in rereading					
a.2) Single fixations (SF)	N	110	164	1,722	727
	%	21	14	42	28
b.2) SF regression origin	N	13	32	126	146
	%	2	3	3	6
c.2) Multiple fixations (MF)	N	55	132	636	394
	%	10	11	16	15
e.2) SF regression goals	N	217	521	755	582
	%	41	43	18	23
f.2) MF regression goals	N	40	106	224	169
	%	8	9	5	7
g.2) SF regr. goals & origin	N	93	238	652	530
	%	18	20	16	21

Old readers produced proportionally fewer valid fixations than young readers (F(1, 105) = 38.9, MSe = 8.32, $p < 0.001$). Readers in the 'hard' samples produced more valid fixations than readers in the 'original' samples (F(1, 105) = 6.4, MSe = 8.32, $p < 0.05$). Effects of age and reading condition are also found in the distributions of fixation types provided in the bottom part of Table 4.2. Old readers made proportionally fewer first-pass fixations (4) (F(1, 105) = 4.2, MSe = 134.6, $p < 0.05$) and more second-pass fixations (5) (F(1, 105) = 15, MSe = 57.2, $p < 0.001$). In comparison to the age-matched samples 'original young' and 'original old', the 'hard' groups showed a reduced proportion of of first-pass fixations (F(1, 105) = 44.2, MSe = 134.6, $p < 0.001$), but made more second-pass (F(1, 105) = 13.7, MSe = 57.2, $p < 0.001$) as well as rereading (6) fixations (F(1, 103) = 36.8, MSe = 62.1, $p < 0.001$).

The numbers and percentages of different fixation types for each reading pass are listed in Table 4.3. Significant main effects of age and reading condition are found in first-pass and rereading fixation types; second-pass fixation types revealed effects of condition.

First-pass Reading. The effects of age are reliable in the proportion of single fixation regression origins (b) and of multiple fixation cases (c). Old readers produced proportionally more single fixations that are regression origins (F(1, 105) = 25.9, MSe = 52.1, $p < 0.001$), but fewer multiple fixation cases (F(1, 105) = 4.3, MSe = 80.9, $p < 0.05$) than young readers.

Effects of reading condition are found in first-pass single fixation cases (a), in single fixation regression origins (b), as well as in multiple fixation origins (d). The 'hard' samples had a smaller proportion of single fixations (F(1, 105) = 9.2, MSe = 162, $p < 0.01$), but a reliably larger proportion of single fixation cases that were regression origins (F(1, 105) = 5.8, MSe = 52, $p < 0.05$). Readers in the 'hard' condition made proportionally more regressions out of multiple fixation cases than readers in the 'original' reading condition (F(1, 97) = 11.2, MSe = 0.96, $p < 0.01$).

Second-pass Reading. No effects of age are found between samples in second-pass fixation types. Condition effects were reliable in the proportion of single fixation cases (a.1) and of multiple fixations (c.1). In both cases, readers in the 'hard' condition produced proportionally more fixations of both types (SF: F(1, 105) = 6.3, MSe = 83, $p < 0.05$; MF: F(1, 99) = 6.5, MSe = 80, $p < 0.05$). Furthermore, in comparsion

[4]Multiple fixation cases sum up to 1,787 valid first-pass gaze durations for the sample 'original young', to 1,837 gaze durations for the sample 'original old', to 2,161 gaze durations for 'hard young', and to 1,270 gaze durations for the sample 'hard old'.

to the 'original' samples, the 'hard' samples had a smaller proportion of single fixations that were the goal of a regressive eye movement (e.1) ($F(1, 104) = 45$, $MSe = 151.9$, $p < 0.001$), but a larger proportion of single fixation regression goals that were the origin of a regressive movement at the same time (g.1) ($F(1, 97) = 6.9$, $MSe = 12.9$, $p < 0.05$).

Rereading. Reliable age effects in rereading fixation types are found in the number of single fixation cases (a.2) and the number of single fixation regression goals, that are also the origin of a regressive eye movement (g.2). Old readers had a smaller proportion of single fixation cases ($F(1, 88) = 12.8$, $MSe = 196$, $p < 0.001$), and a larger proportion of single fixation regression goals and origins ($F(1, 94) = 7.9$, $MSe = 144$, $p < 0.01$).

Significant effects of reading condition are found in the proportion of single fixation cases and of single and mulitple fixation goals. Readers in the 'hard' samples made proportionally more single fixations in rereading than the 'original' samples ($F(1, 88) = 15.5$, $MSe = 196$, $p < 0.001$). Compared to the baseline samples, the 'hard' groups produced fewer single fixation regression goals ($F(1, 102) = 25.2$, $MSe = 340$, $p < 0.001$) and fewer multiple fixation regression goals ($F(1, 88) = 9.2$, $MSe = 34$, $p < 0.01$).

In sum, old readers made significantly fewer 1st-pass fixations, but more 2nd-pass fixations than young readers. The groups in the 'hard' conditions produced generally more fixations than the 'original' groups. They made significantly more 2nd-pass and rereading fixations than the 'original' samples.

4.4. Mean Probabilities, Positions, and Durations of Fixations

In Table 4.4 results of word-based summary statistics averaged across subjects are listed for the four groups 'original young', 'original old', 'hard young', and 'hard old'. Means are based on all fixations per sample (cf. top row of Table 4.2). ANOVAs revealed several main effects of age and of reading condition. Old readers made significantly more skippings ($F(1, 105) = 25.2$, $MSe = 0.01$, $p < 0.001$), had a lower probability of fixating a word once ($F(1, 105) = 29$, $MSe = 0.01$, $p < 0.001$) or twice ($F(1, 105) = 4.0$, $MSe = 0.002$, $p < 0.05$), and made significantly more regressions ($F(1, 105) = 7.2$, $MSe = 0.01$, $p < 0.01$). The age differences between skipping and regression probability are linked: In comparison to young readers, old readers made significantly more regressions back to words that they had previously skipped (conditional probability: $F(1, 105) = 24.7$, $MSe = 0.001$, $p < 0.001$). Compared to the young groups, the first fixation in the old group was significantly located further to the right within a word ($F(1, 105) = 50$, $MSe = 0.01$, $p < 0.001$), whereas the second fixation on a word was located further to the left ($F(1, 105) = 38$, $MSe = 0.01$, $p < 0.001$).

Effects of experimental condition primarily appeared in differences in fixation durations. In comparison to the original groups, the hard group showed prolonged single fixation durations ($F(1, 105) = 6.3$, $MSe = 1118$, $p < 0.05$), 2nd fixation durations ($F(1, 105) = 10.4$, $MSe = 878$, $p < 0.01$), gaze durations ($F(1, 105) = 11.2$, $MSe = 1800$, $p < 0.01$), as well as total reading times ($F(1, 105) = 49.7$, $MSe = 4342$, $p < 0.001$). Furthermore, regression probability was significantly increased in the hard condition ($F(1, 105) = 27.8$, $MSe = 0.01$, $p < 0.001$) and single fixations were located further to the left within words ($F(1, 105) = 4.1$, $MSe = 0.001$, $p < 0.05$). The difference in skipping probability between 'original old' and 'hard old' was significant ($F(1,53) = 4.2$, $MSe = 0.006$, $p < 0.05$). The only interaction of age and reading condition that reached significance is the probability of fixating a word three times or more: Old readers in the hard condition made more multiple fixation cases than the sample 'original old', whereas the young readers show no differences between conditions ($F(1, 105) = 5.4$, $MSe = 0.0002$, $p < 0.05$). The overall reading speed decreased more than 50 words per minute in the young readers, and more than 80 wpm in the old readers in the hard condition.

Table 4.4.: Word-based summary statistics averaged across subjects for the samples 'original young', 'original old', 'hard young', and 'hard old'.

VARIABLE		ORIG. YOUNG	ORIG. OLD	HARD YOUNG	HARD OLD
N of readers		24	32	30	23
N of fixations/ sentence	*M*	8.7	8.6	11	12
	SD	1.5	1.9	2.5	3.1
N of sentences	*M*	122	131	108	97
	SD	23	12	24	27
Fixation probabilities					
Skipping	*M*	.16	.25	.16	.21
	SD	.07	.09	.07	.06
Single fixation	*M*	.68	.59	.64	.59
	SD	.06	.08	.06	.06
Double fixation	*M*	.10	.08	.11	.10
	SD	.04	.04	.03	.04
Three-plus fixation	*M*	.02	.01	.02	.02
	SD	.01	.01	.01	.02
Regression (Regr.)	*M*	.07	.14	.18	.25
	SD	.04	.10	.08	.17
Regr. goal after skipping	*M*	.02	.05	.03	.05
(cond. probability)	*SD*	.01	.04	.02	.03
Relative fixation position					
Pos. single fixation	*M*	.53	.51	.52	.49
	SD	.04	.03	.04	.05
Pos. 1st fixation	*M*	.25	.40	.25	.39
	SD	.06	.15	.07	.12
Pos. 2nd fixation	*M*	.66	.53	.64	.54
	SD	.06	.12	.07	.11
Fixation duration (ms)					
Single fixation	*M*	231	224	242	245
	SD	31	31	36	37
1st of multiple	*M*	211	218	216	233
	SD	23	29	29	37
2nd of multiple	*M*	190	184	203	208
	SD	26	32	29	31
Gaze duration	*M*	261	250	277	289
	SD	40	37	42	51
Total reading time	*M*	281	279	358	383
	SD	47	52	76	84
Reading rate (wpm)		242	256	186	175

4.5. Response Accuracy and Latency

For the readers of the samples 'original young', 'original old', 'hard young', and 'hard old', response accuracy and response latency for the three-alternative multiple choice questions were analyzed. The young and old groups in the original reading condition, that is, with relatively easy mulitple-choice comprehension questions in about 30% of the trials, showed a high response accuracy of above 96% correct answers (see Table 4.5). Both groups in the hard reading condition performed poorer in answering the relatively harder comprehension questions, but accuracy was still above 90% correct. ANOVA analysis of response accuracy over all four groups showed a main effect of reading condition ($F(1, 105) = 41.14$, $MSe = 0.0008$, $p < 0.001$) and a significant interaction of age and reading condition ($F(1,105) = 11.43$, $MSe = 0.0008$, $p < 0.01$).

Response latency was defined as the time interval from the onset of question presentation on the screen until the recording of the button press for response. Response latencies for answering the comprehension questions were log-transformed and corrected for the length of the stimulus (string length), i.e. the length of the question plus the length of the answering options in letters. Therefore, the response latency values provided in this section are the residuals of a linear regression of string length on the logarithmic response latency (in ms). This explains, why positive values indicate relatively 'long' latencies, whereas negative values represent relatively 'short' response latencies. As can be seen in Table 4.5, old readers of both reading condition responded significantly slower than the young readers ($F(1, 105) = 78.82$, $MSe = 0.0332$, $p < 0.001$)

4.5.1. Spillover Effects in the Hard Reading Condition

For the hard groups, accuracy profiles and profiles of mean residual log. latency for response per question type are illustrated in Figure 4.1. Old adults were less accurate ($F(1, 51) = 14.76$, $MSe = 0.00687$, $p < 0.001$) and slower ($F(1, 51) = 47.59$, $MSe = 0.1734$, $p < 0.001$) than young readers in answering the hard questions. More interestingly, verification questions seemed to be most troublesome for both age groups, reflected by significantly lower accuracies ($F(1, 51)= 85.66$, $MSe = 0.00664$, $p < 0.001$) and larger response latencies ($F(1, 51) = 128.34$, $MSe = 0.006$, $p < 0.001$) for verification questions in contrast to all other question types. The interaction with age was significant for

Table 4.5.: Response accuracy (in % correct) and residual log. latency (in ms) to comprehension questions for the samples 'orig. young', 'orig. old', 'hard young', and 'hard old'.

GROUP	RESPONSE ACCURACY (% correct)		
	Mean	SD	Range
Original young	96.9	2.5	91 - 100
Original old	97.3	2.5	89 - 100
Hard young	95.2	2.7	90 -100
Hard old	91.7	4.0	85 - 98
	(Res.log.) RESPONSE LATENCY (in ms)		
	Mean	SD	Range
Original young	-0.183	0.168	-0.50 - 0.09
Original old	0.131	0.210	-0.26 - 0.56
Hard young	-0.129	0.155	-0.49 - 0.11
Hard old	0.193	0.188	-0.05 - 0.71

accuracy ($F(1, 51) = 21.63$, $p < 0.001$) and log. residual latency ($F(1, 51) = 29.98$, $p < 0.001$), indicating a proportional age effect.

Verification questions were designed to test the integrated, complete meaning of a sentence. Thus, this type of question makes large demands on working memory because a full-integrated model needs to be represented for chosing the correct answer. If verification was the hardest question type, the query arises whether the obvious struggle over this question is reflected in the eye movement data of the subsequent trial. Since our experimental manipulation occurred after the sentence was read and after eye movements were recorded, the existence of spillover effects of question type on reading times of the following trial was tested in a post-hoc analysis. Two conditions were contrasted: The first condition includes sentence trials that followed a verification question; the second condition includes all remaining sentence trials following the other five question types. Based on the number of valid fixations (cf. row 3 in Table 4.2), the mean total reading time per sentence for each group and condition was evaluated. Numerically, for the sample 'hard young' the total sentence reading time was 40 ms longer after verification questions than after other question types. In the sample 'hard old' the difference between conditions in total sentence reading time was in the order of 130 ms. Since number of words and word lengths differed between sentences, total reading

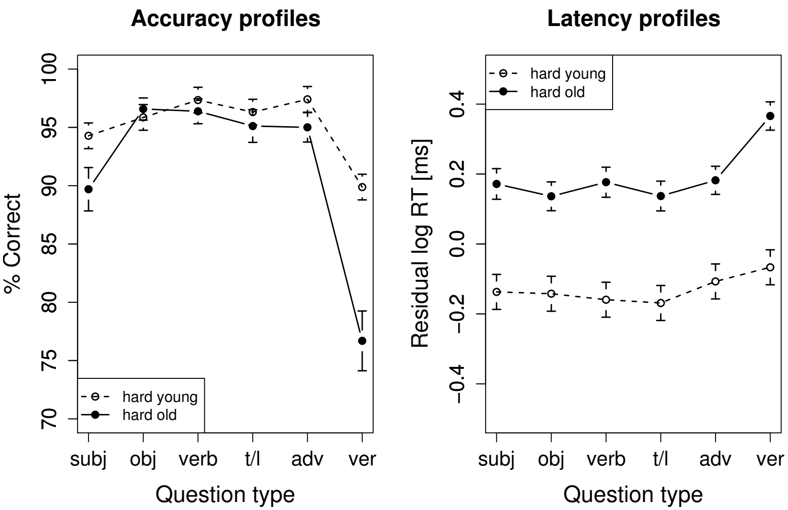


Figure 4.1.: Accuracy and latency profiles for the samples 'hard young' and 'hard old'; residual log. latencies are plotted; question types: subj = subject, obj = object, verb = verbal, t/l = time/ location, adv = adverbial, ver = verification.

time was log-transformed and corrected for the length of sentence in letters. ANOVA results show a significant interaction of age and question type ($F(1, 51) = 5.6$, $MSe = 0.006$, $p < 0.03$). Whereas the young readers did not show a spillover effect after verification questions ($F(1, 51) = 3.2$, $MSe = 0.006$, $p > 0.05$), the old readers were significantly slower in reading the sentence subsequent to a verification question.

4.6. Linear Mixed Modeling: Effects of Reader and Word Variables

For the statistical analysis on fixation level, linear mixed models were used (Gelman & Hill, 2007; Baayen, 2008; Pinheiro & Bates, 2000). In order to differentiate between variance attributed to the randomness of subject and item sampling, variance attributed to the reader-level, and variance attributed to word-level predictors, all models were built up step by step. First, random effects of readers (subject ID) and material (word ID, sentence ID) were included. In a second step, fixed effects related to the subjects profile and experimental condition were included (vocabulary score, digit-symbol score, first and second PC of individual differences, trial, reading condition). In a last step, all predictors related to the word-level and their interaction with reading condition were tested for significance. Different models were built for each dependent time variable. Main effects of interest are reported in the following sections, the complete final model for each measure is listed in the equivalent appendix. As a form of cross-validation and test of generalizability, all linear mixed models were run seperately for the groups of young and old readers.

4.6.1. Selectivity Effects Related to Reading Strategy

In section 4.4 results of mean fixation durations, probabilities, and fixation locations, aggregated over subjects with the word as the unit of analysis, were presentend. Results showed that reading speed was significantly reduced in the hard samples compared with the original reading groups. Furthermore, old readers in the hard condition skipped fewer words than old readers in the original reading condition. Before the influence of word properties and reader-level factor at the level of fixations is analyzed, effects of word selectivity between reading conditions are investigated. The idea behind the selectivity approach is, that different fixation patterns between groups due to different reading strategies may lead to varying sets of fixated words within each group. Even though all experimental groups read the Potsdam Sentence Corpus, groups may differ in the number of fixated words, and consequently differences in frequency, predictability, and length ranges may occurr. This can influence the interpretation of the results of the *lmer*-analyses on various fixation durations measures provided in the following sections.

In this section, group differences in the potential word-level and oculomotor predictors, that are used in the *lmer*-analyses, are tested by ANOVA. To test for group differences, the values are first aggregated for each subject. The word-level predictors are word frequency, word predictability, word length, and the proportion of fixated function words with respect to the total number of fixated words. Thus, the larger the function word percentage, the more function words are fixated; the smaller the function word percentage, the more content words are fixated. Because LMM are tested in the framework of the distributed processing account, all four word-level predictors are provided for the fixated word n, word n-1, and word n+1. Oculomotor variables of interest are mean incoming and outgoing saccade amplitudes, and mean relative fixation position on the word. Mean statistics of fixated word properties and related oculomotor variables for the data sets of first-pass single fixation cases (SFC) and multiple fixation cases (MFC) for each experimental group are provided in Tables 4.6 and 4.7.

Group Differences in Selectivity in First-Pass Single Fixations

In the data set of first-pass single fixation durations, no differences in word-level or oculomotor variables were found between the samples 'original young' and 'hard young' (all $p > 0.05$). Comparing the samples 'hard old' and 'original old' revealed various differences in the subset of first-pass SFCs. Mean lag word frequency was significantly lower in the sample 'hard old' ($F(1,53) = 4.4$, $MSe = 0.02$, $p < 0.05$), and current word frequency was significantly higher compared with 'original old' ($F(1,53) = 15.9$, $MSe = 0.03$, $p < 0.001$). Mean predictability of the fixated words and mean predictability of the words to the right of fixation were lower in the group 'original old' than in 'hard old' ($F(1,53) = 16.2$, $MSe = 0.01$, $p < 0.001$; $F(1,53) = 4.6$, $MSe = 0.01$, $p < 0.05$). Mean length of word n-1 were longer in the sample 'hard old' ($F(1,53) = 5.8$, $MSe = 0.0001$, $p < 0.03$), whereas mean length of the fixated word was longer in the group 'original old' ($F(1,53) = 15.8$, $MSe = 0.0002$, $p < 0.001$). Furthermore, the proportion of fixated function words was much higher in the sample 'hard old' compared to the sample 'original old' ($F(1,53) = 15.8$, $MSe = 0.004$, $p < 0.001$). The average percentage of function words to the left of single fixations was lower in the group 'hard old' than in 'original old' ($F(1,53) = 8.5$, $MSe = 0.003$, $p < 0.01$). Saccade amplitudes in the data set of 'hard old' were shorter than in the sample 'original old', but only the difference in mean outgoing saccade length reached the level of significance ($F(1,53) = 6.6$, $MSe = 1.4$, $p < 0.03$). Mean fixation position of single fixations was marginally shifted further to the left in the sample 'hard old' compared to the 'original old' ($F(1,53) = 3.97$, $MSe = 0.002$, $p = 0.05$).

In sum, no selectivity differences were found in single fixation cases between the groups of young readers. Old readers in the 'hard' group made on average shorter saccades and fixated significantly more short, high frequent, highly predictable and function words than readers in the 'original' group. Words to the right of SFCs were lower predictable and words to the left of SFCs were longer, lower in frequency, and more often content words in the 'hard' sample than in the 'original' sample.

Interestingly, the selectivity differences in SFCs between the old reader groups entail that the selectivity pattern of the 'hard' old sample is more similar to the selectivity pattern of young readers. 'Original young', 'hard young' and 'hard old' reveal comparable proportions of fixated function words as well as comparable mean length, frequency, and predictability of SFCs.

Group Differences in Selectivity in Multiple Fixation Cases

Means of word-level predictors for the data set of first-pass multiple fixation cases are listed in the second columns in Tables 4.6 and 4.7. Predictors of oculomotor control are not considered for gaze durations because saccade amplitudes are only related to the first or last fixation of multiple fixations on the word. The word statistics for frequency and predictability in the multiple fixation data sets of 'original young' and 'hard young' are comparable. Note that mean frequency and predictability of the fixated word n are very low in comparison to the corpus value (PSC), indicating that mostly low frequent and low predictable words are the targets for multiple fixations. At the same time the function word proportion of words with multiple fixations is only at 11 - 13%. The mean length of the fixated words that received multiple fixations differed between the samples 'original young' and 'hard young'. Words in the group 'original young' are slightly longer than fixated words in the group 'hard young' ($F(1,52) = 4.5$, $MSe = 0.0002$, $p < 0.05$). In sum, in both samples target words for multiple fixations are mostly content words, words low in frequency and predictability, and longer in length.

Comparing the multiple fixation data sets of the two samples 'original old' and 'hard old' revealed no differences between group in word-level predictors. Similar to the younger readers, mostly longer content words of lower frequency and lower predictability are the targets for multiple fixation cases.

Table 4.6.: Means and standard deviations of word properties of fixated words and of oculomotor variables for the samples 'original young' and 'hard young' for first-pass single fixation cases (SFC) and multiple fixation cases (MFC).

		ORIGINAL YOUNG		HARD YOUNG	
VARIABLE		SFC	MFC	SFC	MFC
Frequency (log/ mio)					
word n	*M*	2.24	1.26	2.26	1.35
	SD	0.20	0.20	0.16	0.21
word n - 1	*M*	2.28	2.69	2.26	2.69
	SD	0.08	0.27	0.09	0.28
word n + 1	*M*	2.36	2.20	2.34	2.22
	SD	0.06	0.18	0.07	0.17
(PSC: *M* = 2.3, *SD* = 1.3)					
Predictability (logit)					
word n	*M*	-1.58	-1.98	-1.55	-1.98
	SD	0.12	0.08	0.1	0.13
word n - 1	*M*	-1.65	-1.49	-1.66	-1.48
	SD	0.05	0.14	0.05	0.16
word n + 1	*M*	-1.20	-1.24	-1.19	-1.23
	SD	0.05	0.10	0.06	0.25
(PSC: *M* = -1.48, *SD* = 1.1)					
Length (n of letters)					
word n	*M*	4.5	7.5	4.5	7.1
	SD	0.4	0.7	0.3	0.8
word n - 1	*M*	4.2	3.9	4.3	3.8
	SD	0.1	0.3	0.1	0.3
word n + 1	*M*	4.3	4.7	4.4	4.6
	SD	0.1	0.4	0.1	0.3
(PSC: *M* = 5.4 , *SD* = 2.6)					
Function word proportion					
word n	*M*	0.32	0.11	0.33	0.13
	SD	0.06	0.06	0.04	0.05
word n - 1	*M*	0.40	0.57	0.39	0.56
	SD	0.03	0.1	0.04	0.11
word n + 1	*M*	0.39	0.32	0.38	0.34
	SD	0.02	0.07	0.03	0.08
(PSC: *M* = 0.37, *SD* = 0.48)					
Oculomotor variables					
incoming sacc.ampl.	*M*	6.9	-	7.0	-
	SD	0.9	-	0.9	-
outgoing sacc.ampl.	*M*	7.1	-	7.1	-
	SD	0.9	-	0.9	-
rel. fix. position	*M*	0.42	-	0.42	-
	SD	0.04	-	0.03	-

Table 4.7.: Means and standard deviations of word properties of fixated words and of oculomotor variables for the samples 'original old' and 'hard old' for first-pass single fixation cases (SFC) and multiple fixation cases (MFC).

		ORIGINAL OLD		HARD OLD	
VARIABLE		SFC	MFC	SFC	MFC
Frequency (log/ mio)					
word n	*M*	2.03	1.41	2.23	1.43
	SD	0.22	0.22	0.13	0.32
word n - 1	*M*	2.31	2.62	2.23	2.53
	SD	0.15	0.29	0.11	0.36
word n + 1	*M*	2.39	2.21	2.39	2.21
	SD	0.08	0.22	0.07	0.20
(PSC: *M* = 2.3, *SD* = 1.3)					
Predictability (logit)					
word n	*M*	-1.65	-1.96	-1.52	-1.95
	SD	0.12	0.09	0.1	0.17
word n - 1	*M*	-1.63	-1.51	-1.63	-1.53
	SD	0.06	0.19	0.07	0.17
word n + 1	*M*	-1.17	-1.30	-1.10	-1.22
	SD	0.09	0.20	0.12	0.24
(PSC: *M* = -1.48, *SD* = 1.1)					
Length (n of letters)					
word n	*M*	4.9	6.8	4.6	6.7
	SD	0.4	0.7	0.3	0.9
word n - 1	*M*	4.3	3.9	4.4	4.1
	SD	0.2	0.3	0.2	0.5
word n + 1	*M*	4.3	4.6	4.3	4.7
	SD	0.1	0.3	0.1	0.3
(PSC: *M* = 5.4 , *SD* = 2.6)					
Function word proportion					
word n	*M*	0.27	0.15	0.33	0.15
	SD	0.07	0.09	0.04	0.11
word n - 1	*M*	0.41	0.54	0.37	0.51
	SD	0.06	0.11	0.04	0.16
word n + 1	*M*	0.40	0.35	0.40	0.36
	SD	0.03	0.09	0.03	0.08
(PSC: *M* = 0.37, *SD* = 0.48)					
Oculomotor variables					
incoming sacc.ampl.	*M*	8.4	-	7.7	-
	SD	1.3	-	1.1	-
outgoing sacc.ampl.	*M*	8.4	-	7.6	-
	SD	1.4	-	0.9	-
rel. fix. position	*M*	0.41	-	0.39	-
	SD	0.04	-	0.04	-

4.6.2. Effects on First-Pass Single Fixation Duration in Young Readers

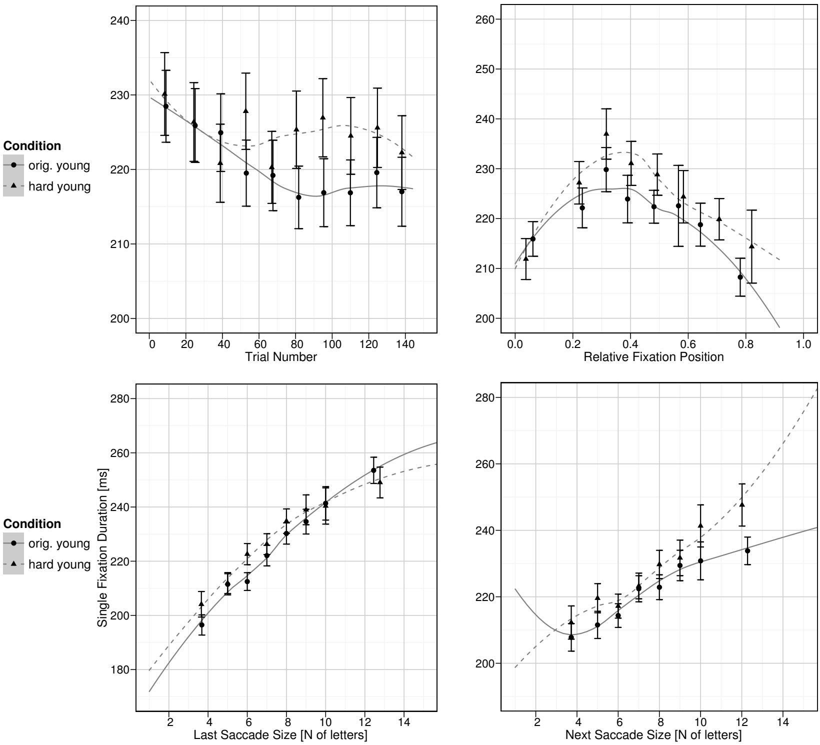


Figure 4.2.: Effects of trial number, relative fixation position, and of incoming and outgoing saccade amplitude on single fixation duration for the samples 'original young' and 'hard young'.

The largest part of first-pass fixations in the groups of the young readers are single fixation cases (cf. Table 4.3). Adding the individual (subject ID) as the only random effect to the model explained 15% of the variance in the data. Including word ID (14%) and sentence ID (1%) as additional random effects improved the goodness of fit of the model significantly ($\chi^2(2) = 2400$, $p < 0.001$). Thus, the random sampling of subjects, words, and sentences contributed 30% of the variance in single fixation durations in young readers. Adding reader-level fixed effects to the model significantly improved the model fit ($\chi^2(5) = 44$, $p < 0.001$). There is a reliable interaction of trial number and reading condition (cf. left panel in top row in Fig. 4.2): In the group 'original young', SFD decreased significantly with increasing trial number (b = -4.053*10^{-04}, SE = 6.971*10^{-05}, t = -5.81); in the group 'hard young', this effect was strongly reduced (b = 2.946*10^{-04}, SE = 9.804*10^{-05}, t = 3.00). Effects of the first and second PC are presented in chapter 7. Word-level predictors and their interactions with reading condition increased the amount of explained variance to 35%, but more importantly, the final model (listed in Table D.2 in Appendix D) was superior to the model without word-level effects ($\chi^2(28) = 1630$, $p < 0.001$). Differences between reading conditions in the effects of oculomotor control and of word-level predictors of the immediately fixated word, and the lag and successor word are presented seperately.

Effects of Oculomotor Control on Single Fixation Duration

Effects of saccade amplitudes and relative fixation position are illustrated in Figure 4.2. The strongest predictor on single fixation duration is the size of the incoming saccade amplitude, that correlates positively with SFD (b = 3.286*10^{-02}, SE = 1.379*10^{-03}, t = 23.83). Outgoing saccade amplitude shows the same pattern with increasing SFDs with increasing saccade length (b = 1.318*10^{-02}, SE = 1.553*10^{-03}, t = 8.49). The influence of incoming saccade amplitude is significantly weaker in the sample 'hard young' compared to 'original young' (b = -6.967*10^{-03}, SE = 1.747*10^{-03}, t = -3.99). On the other hand, the effect of outgoing saccade amplitude is more pronounced in the group 'hard young' (b = 1.332*10^{-02}, SE = 2.184*10^{-03}, t = 6.10), indicated by a steeper slope (cf. fig 4.2). The effect of landing position on SFD shows the expected IOVP-curve for both samples with longer SFDs around the center of word (see right panel in top row in Fig. 4.2). In the group 'original young', the peak of the curve is slighty shifted to the right compared with the group 'hard young' (b = 7.418*10^{-02}, SE = 1.878*10^{-02}, t = 3.95).

Immediacy Effects on Single Fixation Duration

The influence of the frequency, predictability, and length of the fixated word on single fixation duration is illustrated in Figure 4.4. In both samples, word frequency shows a negative correlation with single fixation duration: The more frequent a word, the shorter the fixation duration on the word. Since the shape of the frequency curve shows a cubic trend in Figure 4.4 (see left panel, top row), the three orthogonal polynomials of word frequency are added to the model. Next to the linear (b = -5.016, SE = 1.350, t = -3.71) and the quadratic trend (b = -2.805, SE = 1.136, t = -2.47), the cubic term of word frequency is a very strong predictor (b = -5.317, SE = 6.398*10^{-01}, t = -8.31). Word length is positively correlated with SFD: Fixation duration increases with increasing number of letters in a word (see right panel, top row in Fig. 4.4). The interaction of word length and frequency differs significantly between reading conditions (b = -1.642*10^{-01}, SE = 4.677*10^{-02}, t = -3.51) and is illustrated in Figure 4.5. In the group 'original young', there is only a pronounced frequency effect on long words. In the group 'hard young', high frequent words that are short or long are fixated shorter than low frequency words of equal lengths. In both samples, high predictable words are fixated shorter than low predictable words (b = -1.979*10^{-02}, SE = 3.566*10^{-03}, t = -5.55). The predictability effect does not differ between groups (cf. bottom left panel in Fig. 4.4).

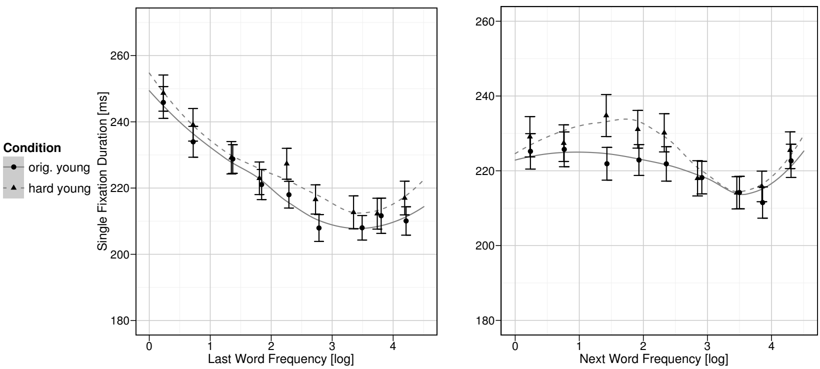


Figure 4.3.: Effects of frequency of word n-1 and word n+1 on single fixation duration for the samples 'original young' and 'hard young'.

Lag and Successor Effects on Single Fixation Duration

For both samples, the frequency of word n-1 is negatively correlated with SFD on word n (b = -3.664*10^{-02}, SE = 3.887*10^{-03}, t = -9.43) (cf. left panel in Figure 4.3). The lag length effect shows the opposite pattern with longer SFD, the longer the word to the left of fixation (b = 3.296*10^{-01}, SE = 4.296*10^{-02}, t = 7.67). Increasing frequency of the upcoming word n+1 reduces SFD on the fixated word in both groups (b = -2.135*10^{-02}, SE = 3.841*10^{-03}, t = -5.56) (cf. right panel in Figure

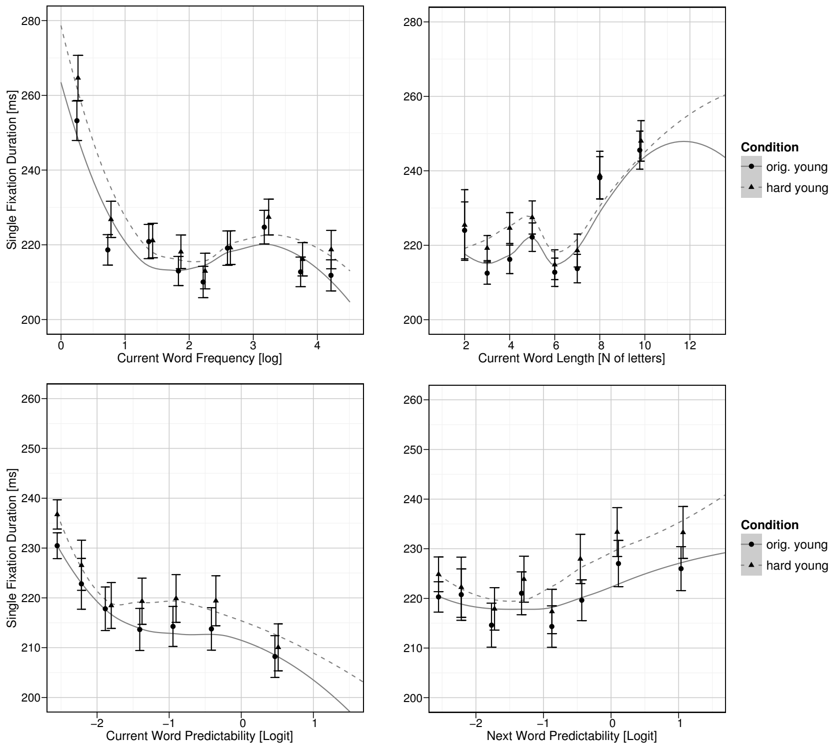
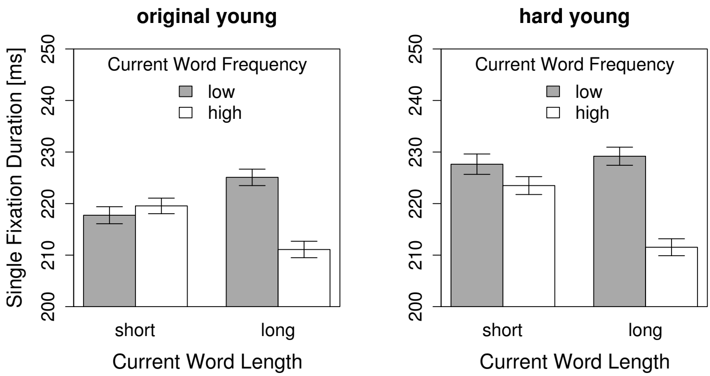


Figure 4.4.: Effects of frequency, length, and predictability of word n, and of predictability of word n+1 on single fixation duration for the samples 'original young' and 'hard young'.

4.3). The only difference in successor effects between the two samples is found for the predictability of the upcoming word. The effect of increasing SFD on word n with increasing predictability of word n+1 ($b = 1.168*10^{-02}$, $SE = 3.540*10^{-03}$, $t = 3.30$) is reliably stronger in the sample 'hard young' ($b = 8.428*10^{-03}$, $SE = 3.351*10^{-03}$, $t = 2.52$). As illustrated in Figure 4.4 (right panel, bottom row), this is mainly due to words of higher predictability classes.

Figure 4.5.: Interaction of length and frequency of the fixated word on single fixation duration for the samples 'original young' and 'hard young'; categories of frequency and length are created by a median split.



4.6.3. Effects on First-Pass Single Fixation Duration in Old Readers

In the sample 'original old', 63% of first-pass fixations were single fixations. In the sample 'hard old' single fixations made up 57% of all first-pass fixations (cf. Table 4.3). Subject ID explained 17% of the variance in SFD, word ID 10%, and sentence ID 2% of the variance. In sum, random factors of participants, words, and sentences contributed 29% of the variance in the data. Adding reader-level fixed effects of trial number, vocabulary size, first and second PC, and reading condition increased the model fit significantly ($\chi^2(5) = 32$, $p < 0.001$). The final model is listed in Table D.3 in Appendix D; the influence of the first and second PC are discussed in chapter 7. There was a reliable interaction of reading condition and trial number, as illustrated in the top left panel in Figure 4.6. SFD decreased with increasing trial number in the group 'original old' (b = -2.958*10^{-04}, SE = 6.496*10^{-05}, t = -4.55); this effect was significantly reduced in the sample 'hard old' (b = 3.935*10^{-04}, SE = 1.113*10^{-04}, t = 3.54). Including word-level fixed effects to the model strongly increased the model fit ($\chi^2(26) = 1390$, $p < 0.001$).

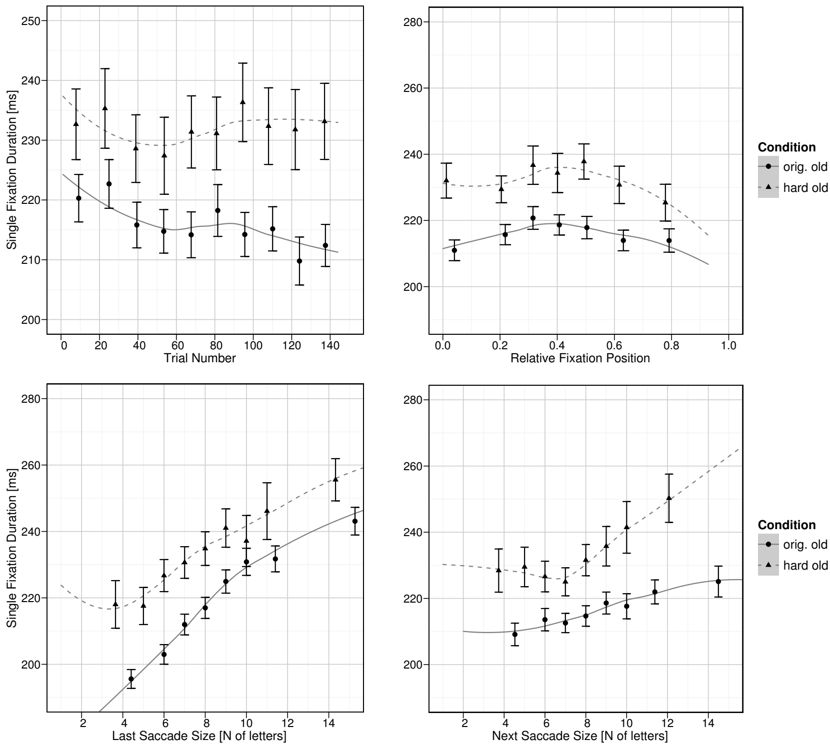


Figure 4.6.: Effects of trial number, relative fixation position, and of incoming and outgoing saccade amplitude on single fixation duration for the samples 'original old' and 'hard old'.

Effects of Oculomotor Control on Single Fixation Duration

The relation between relative fixation position on the word and SFD showed the IOVP-effect, with longest fixation durations at fixations located slightly left to the word center (linear trend of rel. fix. pos.: b = -2.667*10^{-02}, SE = 1.186*10^{-02}, t = -2.25; quadratic trend of rel. fix. pos.: b = -1.679*10^{-01}, SE = 3.644*10^{-02}, t = -4.61). The samples did not differ in the effect of landing position on SFD (cf. right panel, top row in Fig. 4.6). The effects of saccade amplitudes on SFD are illustrated in the bottom row of Figure 4.6. The incoming saccade amplitude is the strongest predictor for SFD, with longer fixation durations, the longer the saccade (b = 2.722*10^{-02}, SE = 9.937*10^{-04}, t = 27.40). This effect was significantly weaker in the sample 'hard old' (b = -5.996*10^{-03}, SE = 1.589*10^{-03}, t = -3.77). Groups differed in the opposite direction concerning the influence of outgoing saccade amplitude. The sample 'original old' produced longer fixation duration with increasing saccade length (b = 8.937*10^{-03}, SE = 1.231*10^{-03}, t = 7.26). This effect was more pronounced in the sample 'hard old' (b = 7.864*10^{-03}, SE = 2.017*10^{-03}, t = 3.90).

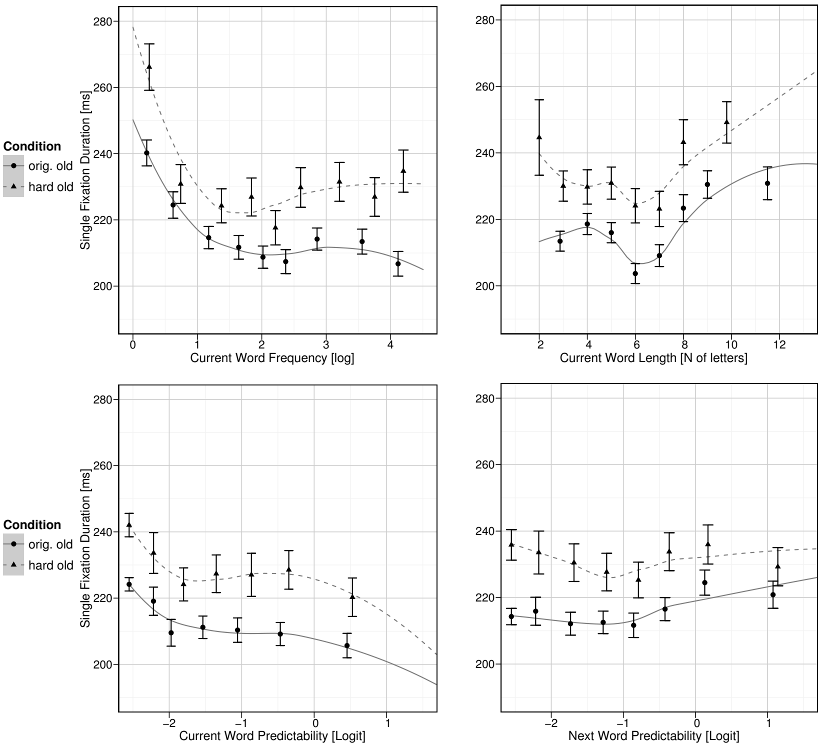


Figure 4.7.: Effects of frequency, length, and predictability of word n, and of predictability of word n+1 on single fixation duration for the samples 'original old' and 'hard old'.

Immediacy Effects on Single Fixation Duration

Word predictability was negatively correlated with SFD (b = -2.011*10^{-02}, SE = 3.631*10^{-03}, t = -5.54), and word length was positively correlated with fixation duration. Samples did not differ with respect to effects of word predictability and word length (cf. bottom, left panel and top, right panel in Fig. 4.7). The linear and cubic trend of the polynomial for word frequency were reliable in the sample 'original old' (linear: b = -5.055, SE = 1.058, t = -4.78; quadratic: b = 1.469, SE = 0.097, t = 1.61; cubic: b = -1.861, SE = 5.506*10^{-01}, t = -3.38). Overall, SFD decreases with increasing word frequency, but there is a slight increase in SFD again on high frequent words. The sample 'hard old' differed signficantly in the linear and cubic trend of word frequency from the sample 'original old' (linear: b = 2.626, SE = 6.482*10^{-01}, t = 4.05; quadratic: b = 5.285*10^{-01}, SE = 6.577*10^{-01}, t = 0.80; cubic: b = -1.512, SE = 6.485*10^{-01}, t = -2.33). As illustrated in the top left panel in Figure 4.7, the readers in 'hard old' show a steep negative slope on SFD for low frequency words, but increasing SFD for higher frequency words. Thus, the general trend of decreasing SFD with increasing word frequency is reduced in the sample 'hard old'.

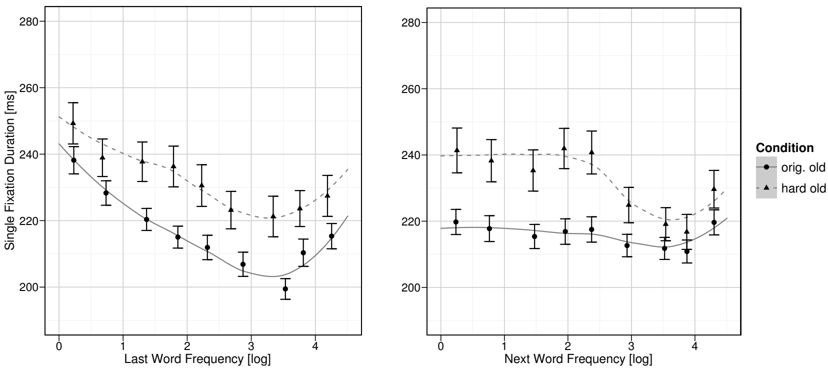


Figure 4.8.: Effects of frequency of word n-1 and word n+1 on single fixation duration for the samples 'original old' and 'hard old'.

Lag and Successor Effects on Single Fixation Duration

Differences between reading groups in lag and successor effects for word-level predictors are found for word frequency. The sample 'original old' shows reliably reduced SFD on word n, if word n-1 was high frequent (b = -3.539*10^{-02}, SE = 4.154*10^{-03}, t = -8.52). The lag word frequency effect was weaker in the sample 'hard old' (b = 8.438*10^{-03}, SE = 3.588*10^{-03}, t = 2.35) (cf. left panel in Figure 4.8). The successor word frequency effect is illustrated in the right panel in Figure 4.8.

Whereas the sample 'original old' shows a significant negative correlation between SFD and upcoming word frequency (b = -1.690*10^{-02}, SE = 4.006*10^{-03}, t = -4.22), the successor word frequency effect in the sample 'hard old' was reliably stronger (b = -1.342*10^{-02}, SE = 3.494*10^{-03}, t = -3.84). This is mainly due to the high SFD for low frequent words n+1. Word predictability of the upcoming word n+1 was positively correlated with SFD (b = 6.548*10^{-03}, SE = 3.101*10^{-03}, t = 2.11). No significant differences were found between reading conditions (cf. bottom, right panel in Figure 4.7).

4.6.4. Effects on First-Pass Gaze Duration

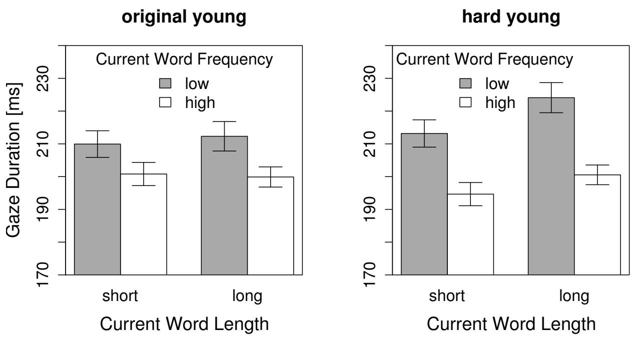
Next to single fixation cases, multiple fixation cases are the second largest part of first-pass fixations. Again, results of LMMs are presented seperately for young and old reader groups. In the analysis of gaze duration, no factors of oculomotor control were included in the LMM, because the saccade amplitudes can only be assigned to one of multiple fixations and the landing position would be an averaged position across all fixations of the gaze. These values are not meaningful for an accumulated variable, such as gaze duration.

Effects on First-Pass Gaze Duration in Young Readers

The random factor of subject ID explained 15% of the variance in gaze duration, word ID explained additional 32%, and sentence ID only 0.3% of the variance. Adding word ID and sentence ID improved the model fit significantly ($\chi^2(2) = 912$, $p < 0.001$). In sum, the random sampling of subjects, words, and sentences explained 47% of the variance in gaze duration in young readers. Including reader-level fixed effect, such as vocabulary score, trial number, condition, and 1st and 2nd PC, into the LMM increased the model fit reliably ($\chi^2(5) = 37$, $p < 0.001$). Further significant improvement of the data fit was achieved by adding word-level fixed effects ($\chi^2(7) = 121$, $p < 0.001$). The final model is listed in the Table D.1 in Appendix D. Results of the LMM did not reveal any differences between reading conditions. The intercept between conditions did not differ (b = 0.0232783, SE = 0.0348903, t = 0.67). Trial number (b = -0.0004057, SE = 0.0001088, t = -3.73), the 1st PC (b = 0.0329241, SE = 0.0088280, t = 3.73), and marginally vocabulary score (b = -0.0124083, SE = 0.0060463, t = -2.05) are significant reader-level predictors. In both samples, gaze duration decreased with increasing word frequency (b = -0.0637593, SE = 0.0098872, t = -

6.45), and with increasing word predictability (b = -0.0283618, SE = 0.0086406, t = -3.28). The word length effect on gaze duration, using the reciprocal value of word length, was reliable (b = -0.6277035, SE = 0.1438315, t = -4.36). The interaction of word length and word frequency on gaze duration is significant (b = 0.6720729, SE = 0.0835288, t = 8.05), showing a more pronounced frequency effect on long words in both reading conditions (cf. Figure 4.9).

Figure 4.9.: Interaction of length and frequency of the fixated word on gaze duration for the samples 'original young' and 'hard young'; categories of frequency and length are created by a median split.



Effects on First-Pass Gaze Duration in Old Readers

The random effects of subject ID (23%), word ID (24%), and sentence ID (1%) contributed 48% of the variance in first-pass gaze duration for the samples 'original old' and 'hard old'. Except for the factor of reading condition (b = 0.116457, SE = 0.042298, t = 2.75), no further reader-level fixed effect is significant and therefore, they are excluded from the model without a decline in model fit ($\chi^2(4) = 7.8$, $p > 0.05$). The final model including word-level fixed effects produces a significantly higher data fit ($\chi^2(8) = 142$, $p < 0.001$) and is listed in Table D.4 in Appendix D.

The effect of word predictability on gaze duration for the old groups is not reliable (b = -0.012633, SE = 0.008128, t = -1.55). The interaction of word frequency and word length is significant (b = 0.357388, SE = 0.078577, t = 4.55), showing a stronger frequency effect on long words. The only significant difference between reading conditions is found in

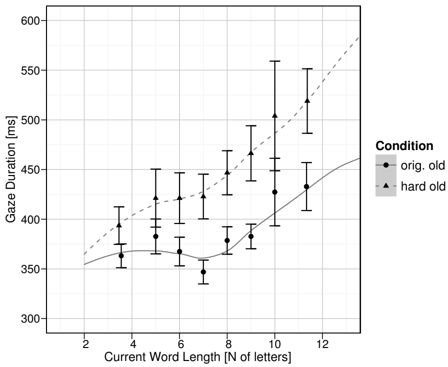


Figure 4.10.: Word length effect on first-pass gaze duration for the samples 'original old' and 'hard old'.

the size of the word length effect on gaze duration which is stronger in the sample 'hard old' (b = -0.467159, SE = 0.146423, t = -3.19). The word-length effect for both samples is illustrated in Figure 4.10.

4.7. Summary of Results and Discussion

Strategy effects due to taks demands on eye movements during reading isolated sentences have been investigated in different age groups using global summary statistics and linear mixed modeling. Specifically, it has been tested in young and old readers, to what extent the frequency and difficulty of comprehension questions influence the reading behavior on the identical sentence material. The eye movement behavior of the samples 'hard young' and 'hard old', who received a difficult comprehension question after each trial, was compared to the samples 'original young' and 'original old', who received easy comprehension question in about 27% of the trials. Results of the lme model revealed robust word level effects of frequency, length, predictability, landing position, and saccade amplitude that are in line with previous reports (Kliegl et al., 2004, 2006; Kliegl, 2007; Nuthmann et al., 2005, 2007; Rayner, 1998). This is important for the interpretation of the cross-level interactions for two reasons: First, the present reading data of both reading conditions are consistent with previously established influences of word variables on reading times. Second, higher order cross-level interactions involving question exist in addition to these word-level effects, or in other words, these interactions do not overlap with variance explained by word-level effects but explain additional variance.

There are three main results of relevance for the analysis and interpretation of eye movements in reading. First, as expected, the difficulty of comprehension questions influences reading strategies in young as well as in old readers. Second, a critical measure linking reader-level and word-level effects is the selectivity of fixated words. Third, given a dynamic modulation of the reading behavior, first-pass reading cannot be construed as a simple reflection of automatic fast sentence processing but rather depends strongly on reading intention and processing depth.

4.7.1. Modulation of Reading Strategy in Response to Difficult Questions

It has clearly been demonstrated in this study that the manipulation of "question" impacts significantly on eye movements in old as well as in young readers. More difficult and frequent questions induce slower reading, reflected in increases in almost all fixation-duration measures and decreases in reading rates. This implies a quantitative change in reading strategy. Subjects were dramatically slowed by the difficulty of the questions, even though reading material and reading instruction were exactly the same in both conditions. Both 'hard' groups produced many more multiple fixations cases in first-pass reading as well as in second-pass reading. A different reading behavior in the 'hard' reading condition is also represented by the large increase in number of second-pass and rereading fixations, related to the increased number of regressive eye movements. This indicates a fixation strategy in the 'hard' samples that puts more weight on refixating words and rereading whole parts of the sentences. A similar change in reading strategy to a more 'careful' reading strategy has been observed when difficult text was given to subjects (for a review see Rayner & Pollatsek, 1989). In the present study, stimulus material as well as reading instruction was identical and the change in reading strategy was indirectly induced by the demands of the comprehension questions. This result brings up the methodological issue of how one controls for the reader's intention and/or attention in reading experiments (Radach & Kennedy, 2004). Differences in the design of reading studies across laboratories may give rise to differences in mean fixation durations, e.g. caused by subtle differences in reading material, question material, or reading purpose.

I argue that using difficult comprehension questions changed the reading behavior into the direction of a more 'mindful' reading. The difficulty of the reading task was increased in the hard condition in two ways. First, a comprehension question had to be answered after *each* sentence of the reading experiment, that forced the participants to keep the attention on the reading task. Second, the wording of the questions was mostly paraphrased with respect to the sentence material. Thus, linguistic processing had to reach the semantic level because questions could not be solved by simple visual word form recognition of the answering options in the multiple choice question. Longer response latency and lower response accuracy in both hard groups, especially in the sample 'hard old', prove the higher task demands associated with more difficult and frequent questions. The higher task demand resulted in prolonged fixation durations and in an increase of number of fixations per sentence in both old and young readers. The idea of differences in the amount of attention and processing time allocated to the sentences is supported by varying effects of trial number on SFD between conditions. Whereas SFD decreased with increasing trial number in the 'original' groups, the size of the trial effect was reliably reduced in the 'hard' groups. Interestingly, old readers in the hard condition skipped significantly fewer words than the old readers in the 'original' group. In line with this, the selectivity pattern in SFC of the 'hard old' sample resembled more the pattern of young readers than their age-matched control, 'original old'. Presumably, with easy questions after less than 30% of all trials, old adults may have 'zoned out', as this can happen up to 23% of the time while reading (Schooler et al., 2004). It can be argued that the higher proportion of skipping (and subsequent regressions) reported for old compared to young adults (Kliegl et al., 2004; Laubrock et al., 2006; Rayner et al., 2006) may be less an inclination towards greater risk taking, as argued by Rayner and colleagues, but simply a higher propensity to lapses of attention.

The effects of distributed processing in the analysis of first-pass single-fixation durations also support the interpretation of a change in reading strategy towards a focus on the sentence semantics for the 'hard' groups. As hypothesized, in young readers the effect of predictability of word n+1 on single fixation durations on word n was larger in the 'hard' condition than in the 'original' condition. The influence of predictability of the upcoming word with longer fixation durations on word n for highly predictable words n+1 compared to low predictable words n+1 has been interpreted as an effect of memory retrieval (Kliegl et al., 2006). The more predictable the context for a given word, the more likely is the preprocessing and prediction of word n+1 during the fixation of word n. Thus, a delay (or even cancellation) of a saccade to word n+1 due to the prediction of word n+1 may lead to an increase of the fixation duration on word n. For young adults, the relevance of sentence context for comprehension becomes more explicit in the hard condition. Importantly, the difference in the successor word predictability effect cannot be explained by varying predictability ranges between the two samples for single fixation cases, because means of word properties and of oculomotor variables did not differ between 'original young' and 'hard young'. As predicted, effect sizes of upcoming word predictability seem to be an index of the mindfulness or depth of processing during reading.

A differences in the successor word predictability effect could not be found for the groups of old readers. The influence of upcoming word predictability is significant for both groups, with increasing SFD with increasing predictability of word n+1. This result is in line with the memory retrieval hypothesis (Kliegl et al., 2006). Comparable effects for both experimental conditions implies that for old readers there was no demonstrable extra focus on semantic and syntactic preprocessing in the hard condition compared to the original reading condition.

Generally, SFDs of the sample 'hard old' were located at words of higher predictability than of the sample 'original old'. Also, the predictability range of the upcoming words with respect to single fixation cases was larger in 'hard old' group than in the 'original old' group. Instead of finding a stronger preprocessing effect out of context, the experimental manipulation led to a very different fixational behavior for single fixation cases in the old group, resulting in various differences in mean word properties and oculomotor variables between the samples 'original old' and 'hard old'. The sample 'hard old' fixated significantly more function words than the sample 'original old'. This observation fits the differences in frequency, predictability, and length ranges of single fixation cases between the two reading conditions. Readers of the 'hard old' group fixated on average shorter words, words of higher frequency, and words of higher predictability, all characteristics of function words. This fixation pattern is in line with the reduced skipping rate of the sample 'hard old' compared to the sample 'original old'. Since function words are the prime candidates for skipping, a reduced skipping rates lead to a higher proportion of fixated function words in the data set. Interestingly, the sample 'hard old' showed a stronger successor word frequency effect than the 'original' sample, especially for words n+1 of low frequency that lead to higher SFD on word n. I argue that the difference in the parafoveal-on-foveal effect of word frequency is strongly related to the fixational selectivity of function words. It cannot be related to differences between conditions in the frequency ranges of words n+1 because these did not exist[5]. An increased prevalence of fixated short and highly frequent function words in the difficult condition leads to strong evidence for preprocessing of word n+1. This increased preprocessing effect is compatible with distributed processing in the perceptual span and, possibly, its modulation by the difficulty of the fixated word, as predicted by the foveal-load hypothesis (Henderson & Ferreira, 1990; Kennedy & Pynte, 2005). Since function words are easier to process, the upcoming word n+1 can be preprocessed to a greater extent. If more fixations are located on function words, as found in the sample 'hard old' compared to the sample 'original old', a stronger parafoveal-on-foveal word frequency effect can be observed.

Other cross-level interactions are found in immediacy word frequency effects. In young readers, the interaction of word length and frequency on SFD differed reliably between conditions. Groups did not differ in mean word frequency or length of the fixated words. The sample 'hard young' showed a word frequency effect with shorter SFD for high frequent words on both short and long words. The sample 'original young' showed only on long words a frequency effect on SFD. The differences in effect sizes is probably related to increased single-fixation durations in response to question difficulty, providing more time for word frequency to come into play, even on short words. In other words, readers in the 'original young' group appear to leave short words with their eyes, before lexical access was achieved and thus, no frequency effect can be observed on short words (see Rayner & Raney, 1996, for a similar argumentation).

In old readers, the linear immediacy word frequency effect is basically reduced in the 'hard' group compared to the 'original' group, that fixated on average words of lower frequency. The cubic trend of word frequency on SFD was more pronounced in the sample 'hard old', that showed increasing SFD for very high frequent words. Again,

[5] Note that the weaker lag word frequency effect of the sample 'hard old' compared to the sample 'original old' can clearly be attributed to the higher mean frequency of words n-1 in the 'original' group. Consistent with a higher frequency of word n-1 is the shorter mean length and larger proportion of function words for words to the left of fixation in the difficult reading condition (cf. Table 4.7).

this can be interpreted as a preprocessing effect, as proposed by the foveal-load hypothesis: The sample 'hard old' fixated more function words, and thus fixated words of higher mean frequency. Therefore, more preprocessing of word n+1 could be done while fixating easy words n, reflected in longer SFD on high frequent words in the 'hard' sample compared to the 'original' sample.

In addition to the general strategy effects of a more careful and slower reading in response to higher task demands, two somewhat different results related to distributed processing can be summed up for young and old readers. On the one hand, if task demands did not influence fixational selectivity, as was the case in young readers, the expected difference in the successor predictability effect can be found. Mindful readers show a stronger effect of upcoming word predictability, that is in line with the memory retrieval hypothesis (Kliegl et al., 2006). Furthermore, due to prolonged fixation durations in the hard condition, frequency effects can be found even on short words. On the other hand, if the experimental manipulation leads to differences in fixational selectivity between groups and therefore to differences in the properties of the fixated words, as was the case in old readers, differences of successor effects are found that fit the perceptual span hypothesis (Henderson & Ferreira, 1990; Kennedy & Pynte, 2005). Mindful readers at higher age make more fixations and the more function word are fixated, the stronger the parafoveal-on-foveal word frequency effect. Presumably, the extended preprocessing of words n+1 in the hard condition is also reflected in increasing SFDs on very high frequent, that is, easy words n.

4.7.2. Resources in Sentence Comprehension and Age Effects

Visual word recognition in a skilled reader is considered to be an automatic and obligatory process (Fodor, 1983; Rayner & Pollatsek, 1989; LaBerge & Samuels, 1985). Equally, eye movements during parsing and understanding a sentence are supposed to reflect the efficient and highly automated process of reading. In the framework of sentence comprehension, Caplan and Waters (1999) call the processes of lexical access, construction of syntactic representations, assigning thematic roles, focus and other aspect of semantics "interpretive processes". Interpretive processes are obligatory, highly specialized, and independent of working-memory capacity and associated effects of aging. In contrast, "post-interpretive processes" such as using the sentence's meaning for other tasks, reasoning, and planning are supposed to be mediated through verbal working memory and are therefore sensitive to age because the efficiency of executive control processes is found to decline with age. In the present study, the online recorded eye movements reflect interpretive processing of sentence comprehension, but the experimental manipulation took place during offline post-interpretive processing. According to Caplan and Waters (1999) one may not expect that question manipulation affects reading behavior in first-pass reading because the task of reading relatively easy isolated sentences and the task of answering questions tap different processing resources. In fact, not only did the intrinsic manipulation of comprehension demands lead to changes in the proportion of second-pass and re-reading fixations in both young and old readers, but it also had a significant effect on fixation probabilities and fixation duration measures in first-pass reading. Readers in the hard reading condition made significantly more multiple fixation cases and showed prolonged mean single fixation duration and gaze duration in first-pass reading compared to the participants in the original reading condition. This implies that the task demands during post-interpretive processing induced a different reading strategy during first-pass interpretive processing.

The cognitive challenge induced by the difficulty and frequency of the comprehension questions had an impact on subsequent eye-movement dynamics during reading of simple sentences. Evidence for this direct impact is provided in the spill-over effect of verification questions, the most difficult type of comprehension questions based on response accuracy and latency measures. Subjects' reading time on the subsequent sentence after a verification questions was longer than after other, easier question types. As this effect was more prominent in the old readers, who showed lower working memory scores than young readers, it appears that interpretive processes are not encapsulated; they depend on post-interpretative top-down control. This conclusion raises problems for Caplans and Waters' (1999) modularity view, which posits that first-pass reading strategies that are sensitive to age and executive-control processes express themselves in supposedly age-invariant interpretive processes. Therefore, the results may be taken as evidence that there are resources used by offline processing tasks (e.g., answering questions) that are shared with online processing tasks during sentence reading (e.g., Just & Carpenter, 1992). I argue that in the difficult condition, reading intention was changed by tasks demands, inducing old readers to allocate more effort (i.e. time) to accomplish sentence comprehension. A similar effect has been discussed in the context of self-regulatory processes in language understanding in old adults (Smiler et al., 2003). The mechanism behind the idea of the modulation of local eye movements in accordance with the reading intention fits the concept of parameter specification in action control as suggested by Neumann (Neumann, 1987, 1989). Even though the control processes of saccade execution are to a large extent automatic, the parameter specification for each single saccade is still a component of a controlled, voluntary action. Old adults seem to be more challenged by the difficult questions than young adults, possibly related to their reduced working memory capacity, and thus, the reading behavior is immediately adapted to the larger effort that is needed to reach the reading goal.

In sum, the results demonstrate that changes in reading strategy, induced by task demands during post-interpretive processing, are predictive of eye movement patterns during isolated sentence reading. Task demands change eye-movement measures at the level of the reader (e.g., skipping and regression probability, fixational selectivity of function words, average fixation duration) and effects of word properties (e.g., frequency, predictability) on single fixation durations. Furthermore, cross-level interactions demonstrated that the parameter specification at the reader level, as a results of the reader's intention and effort on sentence comprehension, entailed a reading strategy that focused on a detailed analysis of the sentence material, that in the end gave rise to differences in parafoveal-on-foveal effects. On-line processes in reading are susceptible to reading intention and the reader's age, possibly tied in with age-related efficiency of executive control processes, and individual differences in reading style. In this context, linear mixed models allow us to analyze subject-related, sentence-related, and word-related factors that influence eye movements in reading in a unified framework.

The described effects of 'question' are mainly interpreted as results of the difficulty of the questions. Questions in the hard reading condition used paraphrased wording and therefore made answering more challenging because it demanded deep semantic processing. In principle, the differences in fixational behavior between conditions could also originate solely in the frequency of occurrence of comprehension questions. Comprehension questions in the hard reading condition, though more demanding, were asked after each sentence. This clearly forced subjects to focus their attention on the instructed task, as evidenced by the weaker trial effects on fixation duration in the hard conditions compared to the original condition. But answering a question after each single trial might also disrupt the reading process that is already somewhat unnatural in the conditions of the experimental setup[6]. At this point, it is not possible to clearly distinguish between those two explanations for the effects of question provided here, though the spill-over effects of verification questions on the reading time of the subsequent sentence favor the explanation of question difficulty. Therefore, in the next study reading for comprehension was tested when only the frequency of questions was manipulated, not their difficulty.

[6]The unnaturalness of the reading condition refers only to the general experimental settings when using eye tracking technique, such as the use of a chin rest, the mounting of a cameras on the participant's head, the reading of sentences displayed on the monitor, etc. With respect to the eye movement recording in these studies, the paradigm of free viewing on the screen is much more 'natural' than for example reading in RSVP paradigms, where no regressive eye movements are impossible, or moving window paradigms, where no skipping is possible (Just, Carpenter, & Woolley, 1982; Haberlandt, 1994).

5. Strategy Effects Due to Question Frequency

This chapter adresses a question arising in the interpretation of the experimental effects found in the comparison of the reading behavior of the first four experimental groups (cf. section 4.7). It needs to be tested in a control study, if not the difficulty, but solely the frequency of comprehension questions during sentence reading manipulates the eye movement patterns systematically. The experimental manipulation is attained by using different percentages of comprehension questions that are asked after the sentences. In the 'original' sample, simple comprehension questions, that can purely be solved by word form recognition without deeper sentence understanding, are asked in about one third of the trials (cf. section 3.2). In the sample 'frequent' the identical comprehension questions are asked after each single trial (see section 5.2). The reading condition of the 'frequent' group serves as a control condition for the experimental manipulation in the hard groups (cf. section 4.2), where *difficult* questions were asked after each sentence.

The chapter is organized along the lines of the previous chapter 4. First, the participants of this group comparison and the experimental details of Experiment 2 are shortly described. Then the results of saccade detection, global subject-based summary statistics, and LMMs for first-pass single fixation duration and gaze duration, focussing on effects of reading strategy, are provided.

5.1. Participants

A group of 33 college students (in the following labeled 'original') performed the baseline experiment (cf. section 3.2)[7]. 30 college students read the PSC for comprehension and had to answer an easy comprehension question after each sentence. This sample is called 'frequent'. All participants were native speakers of German. They all had normal or corrected to normal vision. Psychometric data of both samples are listed in Table 5.1; a complete list of subject information is provided in Tables C.3 and C.6 in Appendix C. In comparison to the 'original' group, the sample 'frequent' was marginally older ($F(1,61) = 5.8$, MSe

= 10.68, $p < 0.05$). The 'frequent' group had a slighty lower vocabulary size (Lehrl, 1977) ($F(1,60) = 5.2$, $MSe = 4.02$, $p < 0.05$) and scored lower in Wechsler's (Wechsler, 1964) Digit-Symbol-Test ($F(1,60) = 5.9$, $MSe = 89.1$, $p < 0.05$) than the 'original' group.

Table 5.1.: Group statistics on age, vocabulary size, and digit-symbol task for the samples 'original' and 'frequent'; a detailed list of subject information is listed in Tables C.3 and C.6 in Appendix C.

GROUP	N	M	SD	RANGE
Age				
Original	33	21.9	2.18	19 - 28
Frequent	30	23.9	4.15	19 - 36
Vocabulary test				
Original	33	32.8	0.95	31 - 35
Frequent	30	31.7	2.8	26 - 35
Digit-Symbol test				
Original	33	67.9	8.05	53 - 85
Frequent	30	62	11	43 - 89

5.2. Experiment 2: Reading with Frequent Easy Questions

The important difference of Experiment 2 in comparison to the baseline experiment is the frequency of the occurence of comprehension questions.

5.2.1. Easy Comprehension Questions for the PSC

The 133 questions used in the original experiment (cf. section 3.2) served as question material for this experiment. Eleven additional questions of the same format were designed, filling up the set to one easy question for each of the 144 sentences of the PSC. The easy questions are listed in the Appendix A. All questions used the exact verbatim wording of the given sentence. Thus, a correct response is possible by simple word form recognition without further understanding

[7] Data of this sample were included in the analyses in Kliegl et al. (2006) and (Kliegl, 2007), labeled sample 1.

of the sentence (see example in section 3.2.1). Therefore, the comprehension questions used for the samples 'original' and 'frequent' make low demands on cognitive processing and working memory.

5.2.2. Apparatus and Procedure

The basic apparatus and procedure with respect to sentence presentation, subject setup, and calibration were identical to the baseline experiment (see sections 3.2.2 and 3.2.3)[8]. Monitor resolution for sentence prensentation in the 'frequent' sample was 1024 pixels x 768 pixels. Therefore, letters subtended 0.41° of visual angle. Eye movements were recorded with a 500-Hz-sampling rate. For the 'frequent' sample, in 100% of the trials the sentence was replaced by a three-alternative multiple-choice comprehension question that participants answered via a mouse click. The task was initiated with five training trials. The sample 'original' received in 27% of the trials a three-alternative multiple-choice question.

Table 5.2.: Distribution of fixation types for the samples 'original' and 'frequent'; data are from right eye; rows 2) and 3) sum up to 1), rows 4), 5), and 6) sum up to 3).

VARIABLE		ORIGINAL	FREQUENT
1 N of fixations (total)		37,175	34,071
2) First/ last word;	*N*	11,238	9,689
First/ last fixation	%	30	28
3) N of valid fixations	*N*	26,035	24,382
	%	70	72
4) First-pass	*N*	22,812	21,327
	%	88	88
5) Second-pass	*N*	2,750	2,464
	%	10	10
6) Rereading	*N*	473	591
	%	2	2

[8]Experiment 2 was written with MATLAB (The MathWorks, Inc., Natick, MA), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) as well as the Eyelink Toolbok extensions (Cornelissen, Peters, & Palmer, 2002).

Table 5.3.: Number and proportion of fixation types in first-pass reading, second-pass reading, and rereading for the samples 'original' and 'frequent'; data are from right eye; note that the number of fixations types per reading pass sum up to the number of fixations per pass listed in Table 5.2.

VARIABLE		ORIGINAL	FREQUENT
Fixation types in first-pass reading			
a) Single fixations (SF)	*N*	15,350	14,288
	%	67	67
b) SF regression origin	*N*	1,651	1,356
	%	7	6
c) Multiple fixations (MF)[9]	*N*	5,647	5,555
	%	25	26
d) MF regression origin	*N*	164	128
	%	1	1
Fixation types in second-pass reading			
a.1) Single fixations (SF)	*N*	746	715
	%	27	29
b.1) SF regression origin	*N*	37	35
	%	1	1
c.1) Multiple fixations (MF)	*N*	320	328
	%	12	13
e.1) SF regression goals	*N*	1,424	1,194
	%	52	49
f.1) MF regression goals	*N*	83	85
	%	3	4
g.1) SF regr. goals & origin	*N*	140	107
	%	5	4
Fixation types in rereading			
a.2) Single fixations (SF)	*N*	70	119
	%	15	20
b.2) SF regression origin	*N*	8	9
	%	2	2
c.2) Multiple fixations (MF)	*N*	31	67
	%	6	11
e.2) SF regression goals	*N*	254	239
	%	54	41
f.2) MF regression goals	*N*	28	48
	%	6	8
g.2) SF regr. goals & origin	*N*	82	109
	%	17	18

5.3. Results of Saccade Detection and Fixation Selection

According to the selection criteria defined in section 3.3.1, Table 5.2 provides the results of the fixation split up for the experimental groups labeled 'original' and 'frequent'. The top row lists the total number of detected saccades for the right eye. The second row gives the number of fixations for criterion (2), indexing fixations on first and last words of the sentences and first and last fixations of the trials. A breakdown of numbers and percentages of different reading passes relative to valid within-sentence fixations (3) is provided in the bottom part of the Table. There are no significant differences between reading conditions in the distribution of fixations types listed in Table 5.2 (all $p > 0.05$). The relative proportion of valid fixation (3), first-pass fixations (4), second-pass fixations (5), and rereading fixations (6) is identical in both samples.

The number of fixations per types for each reading pass is listed in Table 5.3. In first-pass and second-pass reading, no reliable differences between the samples 'frequent' and 'original' in the percentage of fixations per type can be found.

5.4. Mean Probabilities, Positions, and Durations of Fixations

Word-based summary statistics averaged across subjects for the two samples 'original' and 'frequent' are listed in Table 5.5. Means are based on all fixations (cf. Table 5.2). The mean number of fixations per sentence is comparable in both groups. The number of sentences per subject included in the ANOVA analyses is slightly smaller in the 'frequent' sample than in the 'original' sample.

While holding the reading instruction constant ('reading for comprehension'), the manipulation of the frequency of comprehension questions has a significant effect on several fixation duration measures, but not on fixation locations or fixation probabilities (all $p > 0.05$). Single fixation duration ($F(1, 61) = 4.15$, $MSe = 850$, $p < 0.05$), the duration of first ($F(1, 61) = 4.2$, $MSe = 802$, $p < 0.05$) and second fixation ($F(1, 61) = 5.2$, $MSe = 695$, $p < 0.05$), and total reading time (F(1,

[9] Multiple fixation cases sum up to 2,546 valid first-pass gaze durations for the sample 'original' and to 2,461 gaze durations for the sample 'frequent'.

61) = 4.77, MSe = 2117, $p < 0.05$) are prolonged in the 'frequent' reading condition compared to the corresponding durations in the 'original' reading group. The difference in gaze duration shows a trend into the same direction ($F(1, 61) = 3.54$, MSe = 1461, $p = 0.06$). Consistent with prolonged fixation durations, the reading rate of the 'frequent' group is reduced in comparison to the 'original' sample.

In sum, the frequency of occurrence of the comprehension question influences fixation durations and reading speed, but there is no evidence for a systematic influence on fixation probabilities. Using linear mixed modeling including reader-level and word-level predictors, it will be investigated in section 5.6, if longer fixation durations on the identical sentence material may lead to differences in local effects of word properties and oculomotor control.

5.5. Response Accuracy

Response accuracy for the samples 'original' and 'frequent' is listed in Table 5.4. For the sample 'original', accuracy represents the mean percentage of correct answers to the easy mulitple-choice comprehension questions, that occurred in 27% of the trials. The sample 'frequent' recieved the identical easy mulitple-choice comprehension questions in 100% of the trials. Overall performance for responses for the PSC was very high with more than 96% correct. Mean accuracy of the sample 'original' did not differ from the performance of the sample 'frequent' ($F(1,61) = 0.89$, $p > 0.05$).

Table 5.4.: Response accuracy for comprehension questions of the samples 'original' and 'frequent'.

GROUP	RESPONSE ACCURACY (in % correct)		
	Mean	SD	Range
Original	97.6	2.4	91 - 100
Frequent	98.1	1.2	95 - 100

Table 5.5.: Word-based summary statistics averaged across subjects for the samples 'original' and 'frequent'.

VARIABLE		ORIGINAL	FREQUENT
N of readers		33	30
N of fixations/ sentence	*M*	8.2	8.6
	SD	1.3	1.4
N of sentences	*M*	138	133
	SD	8	10
Fixation probabilities			
Skipping	*M*	.20	.17
	SD	.08	.07
Single fixation	*M*	.65	.67
	SD	.06	.05
Double fixation	*M*	.09	.10
	SD	.04	.05
Three-plus fixation	*M*	.01	.01
	SD	.01	.01
Regression	*M*	.07	.07
	SD	.04	.05
Regression goal after skipping	*M*	.03	.02
(cond. probability)	*SD*	.02	.01
Relative fixation position			
Pos. single fixation	*M*	0.5	0.5
	SD	.04	.04
Pos. 1st fixation	*M*	0.3	0.3
	SD	.12	.09
Pos. 2nd fixation	*M*	0.6	0.7
	SD	.11	.07
Fixation duration (ms)			
Single fixation	*M*	210	225
	SD	31	27
1st of multiple	*M*	198	213
	SD	30	27
2nd of multiple	*M*	168	183
	SD	26	26
Gaze duration	*M*	235	253
	SD	39	37
Total reading time	*M*	249	274
	SD	45	47
Reading rate (wpm)		283	251

Table 5.6.: Means and standard deviations of word properties of fixated words and of oculomotor variables for the samples 'original' and 'frequent' for first-pass single fixation cases (SFC) and multiple fixation cases (MFC).

		ORIGINAL		FREQUENT	
VARIABLE		SFC	MFC	SFC	MFC
Frequency (log/ mio)					
word n	*M*	2.14	1.30	2.22	1.36
	SD	0.23	0.26	0.17	0.23
word n - 1	*M*	2.29	2.66	2.28	2.64
	SD	0.10	0.27	0.11	0.27
word n + 1	*M*	2.37	2.16	2.35	2.15
	SD	0.06	0.22	0.06	0.19
(PSC: *M* = 2.3, *SD* = 1.3)					
Predictability (logit)					
word n	*M*	-1.62	-2.02	-1.58	-1.98
	SD	0.12	0.11	0.11	0.14
word n - 1	*M*	-1.66	-1.53	-1.65	-1.54
	SD	0.05	0.16	0.05	0.13
word n + 1	*M*	-1.20	-1.25	-1.20	-1.32
	SD	0.07	0.21	0.05	0.22
(PSC: *M* = -1.48, *SD* = 1.1)					
Length (n of letters)					
word n	*M*	4.7	7.3	4.6	7.2
	SD	0.4	1.0	0.3	0.9
word n - 1	*M*	4.3	3.9	4.3	3.9
	SD	0.1	0.3	0.2	0.3
word n + 1	*M*	4.3	4.7	4.4	4.7
	SD	0.1	0.3	0.08	0.3
(PSC: *M* = 5.4 , *SD* = 2.6)					
Function word proportion					
word n	*M*	0.30	0.13	0.32	0.14
	SD	0.07	0.07	0.06	0.08
word n - 1	*M*	0.40	0.56	0.40	0.55
	SD	0.04	0.09	0.04	0.10
word n + 1	*M*	0.39	0.32	0.38	0.33
	SD	0.02	0.08	0.02	0.07
(PSC: *M* = 0.37, *SD* = 0.48)					
Oculomotor variables					
incoming sacc.ampl.	*M*	7.5	-	7.1	-
	SD	1.3	-	0.9	-
outgoing sacc.ampl.	*M*	7.7	-	7.2	-
	SD	1.3	-	0.9	-
rel. fix. position	*M*	0.43	-	0.43	-
	SD	0.04	-	0.03	-

5.6. Linear Mixed Modeling: Effects of Reader and Word Variables

Linear mixed models were built up step by step as described in section 4.6. After including random effects of reader (subject ID) and items (word ID and sentence ID), reader-level fixed effects were added to the model (vocabulary score, digit-symbol score, first and second PC of individual differences, trial, and reading condition). Finally, word-level fixed effects and their interaction with reading condition were added to the model. These cross-level interactions are of main interest in the analyses to test for effects of reading strategy. Two separate models were fit for first-pass single fixation duration and gaze duration.

5.6.1. Selectivity Effects Related to Reading Strategy

As outlined in section 4.6.1, differences in reading strategy might lead to varying fixation patterns between reading conditions. For a better interpretation of the LMM results, it needs to be tested if groups differ in the ranges of word properties and oculomotor variables with respect to the fixated word and the selected fixation duration measure. Values are aggregated for each subject before entering the ANOVA. Word-level predictors are word frequency, word predictability, word length, and the proportion of fixated function words for the fixated word (n), and for the word left (word n-1) and right of fixation (word n+1). Oculomotor variables include mean incoming and outgoing saccade amplitude, and mean relative fixation position on the word. These variables are only relevant for single fixation cases, not for multiple fixation cases. Mean statistics of fixated word properties and related oculomotor variables for the data sets of first-pass single fixation cases (SFC) and multiple fixation cases (MFC) for the samples 'original' and 'frequent' are listed in Table 5.6.

In both groups, 67% of all first-pass fixations were single fixation cases. For SFC, no differences in word-level predictors or oculomotor variables are found between the samples 'original' and 'frequent'. Multiple fixation cases made up about 25% of all first-pass fixations. ANOVA revealed no differences between conditions in word variables of fixated words with multiple fixations (all $p > 0.05$). Compared to the corpus means, target words for multiple fixations tended to be long content words of low predictability and low frequency.

5.6.2. Effects on First-Pass Single Fixation Duration

The final model of fitting single fixation duration of the two samples 'original' and 'frequent' is listed in Table D.5 in Appendix D. The random sampling of subjects (17%), words (11%), and sentences (1%) explained a total of 30% of the variance in SFD. Adding reader-level fixed effects to the model improved the model fit significantly ($\chi^2(8)$ = 66.6, p < 0.001). The strongest reader-level predictor for SFD was trial number (b = -3.342*10^{-04}, SE = 4.099*10^{-05}, t = -8.15). Both reading groups showed decreasing SFD with increasing trial number, as illustrated in the top left panel in Figure 5.1. Other reliable reader-level predictors were the first and second PC (pc1: b = 1.489*10^{-02}, SE = 5.256*10^{-03}, t = 2.83; pc2: b = -4.283*10^{-02}, SE = 1.180*10^{-02}, t = -3.63), that is further discussed in chapter 7, as well as reading condition (b = 1.034*10^{-01}, SE = 3.011*10^{-02}, t = 3.43). The sample 'frequent' has a significantly larger intercept, that is, longer SFDs than the sample 'original'. Adding word-level fixed effects to the model increased the model fit significantly ($\chi^2(24)$ = 1858, p < 0.001). Relevant effects of oculomotor control and of word variables are presented in the following.

Effects of Oculomotor Control on Single Fixation Duration

Effects of saccade amplitude and relative fixation position for the two reading groups are illustrated in Figure 5.1. The strongest predictor for single fixation duration is the size of the incoming saccade amplitude (b = 2.461*10^{-02}, SE = 1.019*10^{-03}, t = 24.14). The longer the saccade amplitude to word n, the longer is the single fixation duration on word n (cf. bottom left panel in Fig. 5.1). This effect is more pronounced in the sample 'frequent' than in the sample 'original' (b = 3.849*10^{-03}, SE = 1.404*10^{-03}, t = 2.74). SFD increases reliably with increasing outgoing saccade amplitude (b = 1.088*10^{-02}, SE = 1.162*10^{-03}, t = 9.36). Again, this effect is stronger in the sample 'frequent' than in the original group (b = 7.568*10^{-03}, SE = 1.596*10^{-03}, t = 4.74), as illustrated in the lower right panel in Figure 5.1. An IOVP-effect in single fixation duration is found in both reading conditions (linear trend of relative fixation position: b = -1.370*10^{-01}, SE = 1.319*10^{-02}, t = -10.38; quadratic trend of rel.fix.pos.: b = -4.670*10^{-01}, SE = 4.164*10^{-02}, t = -11.22). SFDs are longest slightly left to the center of the word. In the sample 'frequent', the peak of the curve is shifted somewhat to the right compared to the sample 'original' (linear trend: b = 4.320*10^{-02},

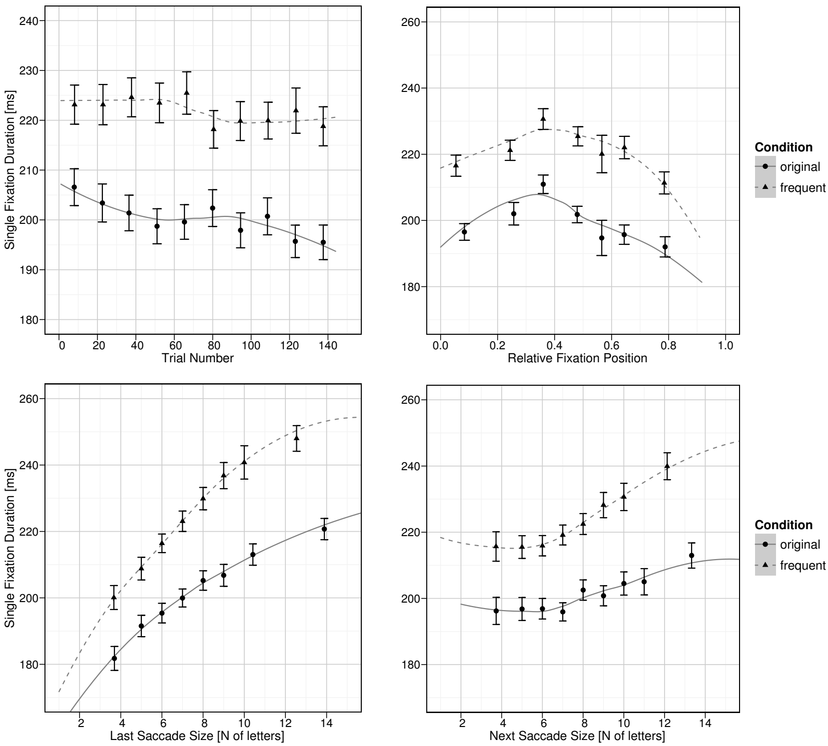


Figure 5.1.: Effects of trial number, relative fixation position, incoming and outgoing saccade amplitude on single fixation duration for the samples 'original' and 'frequent'.

SE = $1.884*10^{-02}$, t = 2.29).

Immediacy Effects on Single Fixation Duration

In both reading conditions, the influence of word frequency, word length, and word predictability on SFD is reliable. SFD decreases with increasing word frequency and for very high frequent word, SFD is found to increase again (cf. top left panel in Figure 5.2; linear trend: b = -3.864, SE = 1.319, t = -2.93; quadratic trend: b = -1.544, SE = 1.079, t = -1.43; cubic trend: b = -4.880, SE = $5.467*10^{-01}$, t = -8.93). The more predictable word n, the shorter is the SFD on word n (cf. bottom left panel in Figure 5.2; b = $-2.113*10^{-02}$, SE = $3.101*10^{-03}$, t = -6.81). Word

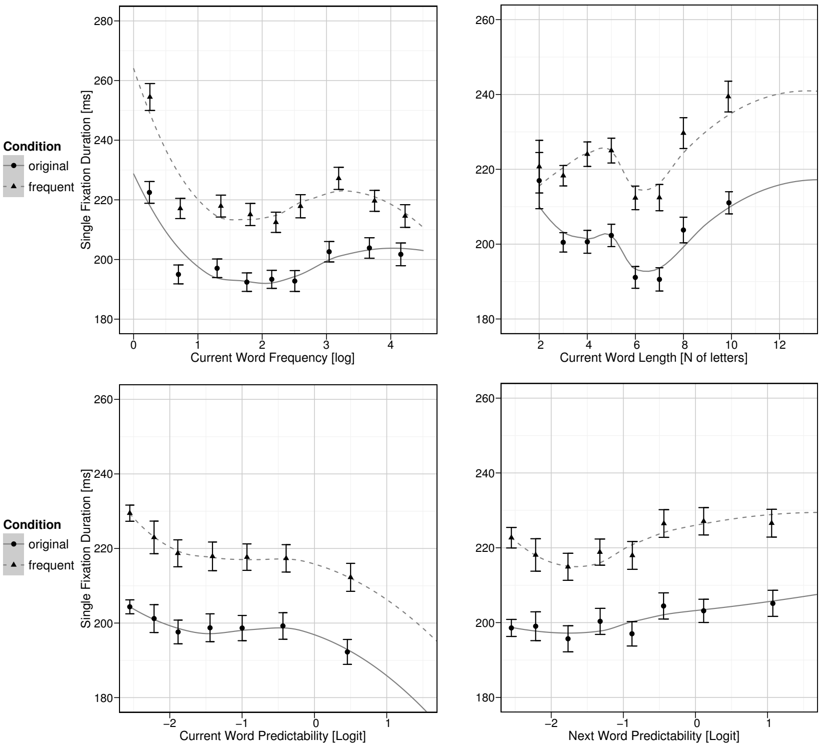


Figure 5.2.: Effects of frequency, length, predictability of word n, and of predictability of word n+1 on single fixation duration for the samples 'original' and 'frequent'.

length is a reliable predictor for SFD (b = 1.899*10^{-01}, SE = 7.006*10^{-02}, t = 2.71; cf. top right panel in Figure 5.2). No differences in word-level immediacy effects were found between reading conditions.

Lag and Successor Effects on Single Fixation Duration

Differences in word-level lag and successor effects have not been found between reading conditions. Lag word frequency is negatively correlated with SFD on word n (b = -3.092*10^{-02}, SE = 3.416*10^{-03}, t = -9.05). The more frequent word n-1, the shorter the SFD on word n (cf. left panel in Figure 5.3). The same pattern is found for upcoming word frequency: The more frequent word n+1, the shorter the SFD on word n (b = -1.908*10^{-02}, SE = 3.392*10^{-03}, t =-5.63; see right panel in Figure 5.3). The successor word predictability effect was also significant (b = 9.921*10^{-03}, SE = 2.777*10^{-03}, t = 3.57). High predictable words n+1 are correlated with the longer SFD on word n (cf. bottom right panel in Figure 5.2).

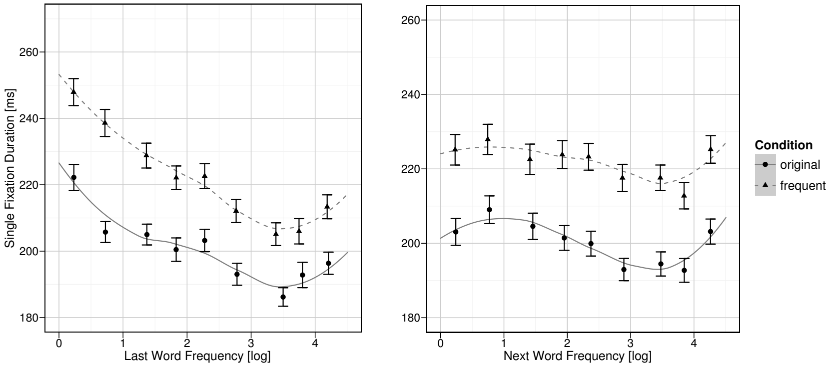


Figure 5.3.: Effect of frequency of word n-1 and of word n+1 on single fixation duration for the samples 'original' and 'frequent'.

5.6.3. Effects on First-Pass Gaze Duration

The random factor of subject ID explained 18% of the variance in gaze duration, word ID explained additional 27%, and sentence ID extra 0.5% of the variance in the data. Adding word ID and sentence ID to the model increased the model fit significantly ($\chi^2(2) = 1046$, $p < 0.001$). Model fit was further improved by including reader-level fixed effects ($\chi^2(5) = 40$, $p < 0.001$). Additional improvement in the fit of gaze duration was achieved by adding word-level fixated effects into the LMM ($\chi^2(8) = 126$, $p < 0.001$). The final model for the samples 'original' and 'frequent' fitting first-pass gaze duration is listed in Table D.6 in Appendix D.

At the reader level, the first PC influenced gaze duration significantly (b = 0.0263919 , SE = 0.0080858, t = 3.26). The interaction of trial number and reading condition on gaze duration was the only significant interaction with condition. Gaze duration decreased with increasing trial number (b = -0.0004627, SE = 0.0001259, t= -3.68); this effect was reduced in the 'frequent' sample (b = 0.0003999, SE = 0.0001810, t = 2.21).

No differences between reading conditions could be found for word-level predictors on gaze duration. Gaze duration decreased with increasing word frequency (b = -0.0449364, SE = 0.0087095, t = -5.16) and with increasing word predictability (b = -0.0207809, SE = 0.0071711, t = -2.90). The reciprocal of word length was negatively correlated with gaze duration (b = -0.5866613, SE = 0.1292970 , t = -4.54), that is, gaze duration was longer, the longer word n.

5.7. Summary of Results and Discussion

The influence of the frequency of comprehension questions on eye movements in isolated sentence reading was tested in an experimental manipulation. The 'frequent' sample received multiple-choice comprehension questions in 100% of the trials, the sample 'original' in 27% of the sentences. Both experimental groups revealed robust effects of word-level predictors and oculomotor control variables on fixation duration that are in line with the general findings of chapter 4 and that are consistent with previous research reports (e.g., Kliegl et al., 2006; Rayner, 1998). Single fixation duration is shorter, the more frequent, the more predictable, or the shorter the fixated word is. Lag word frequency and successor word frequency are negatively correlated with single fixation duration, whereas successor word predictability is positively correlated with SFD. Interestingly, the experimental manipulation of question frequency only had an impact on fixation durations and effects of oculomotor control on fixation duration, but not on word-level effects on fixation duration.

5.7.1. Modulation of Reading Strategy in Response to the Frequency of Questions

The fixation pattern across sentences was similar in the two reading conditions. Groups did not differ in the number of first-pass, second-pass, and rereading fixations. The proportion of single fixation cases, multiple fixation cases, word skippings, and regressions was identical in both reading conditions. A major difference between conditions was found in reading rate. Compared to the 'original' sample, the 'frequent' group showed significantly prolonged single fixation durations, first and second fixation durations, and total word reading times along with a reduced reading speed. The difference between groups in fixation duration measures was also evident in significantly varying intercepts in the LMMs. Fitting the largest part of first-pass fixations, namely single fixation durations, revealed an identical pattern of significant word-level predictors for the two experimental groups. Even though in the 'frequent' condition, the eyes stayed longer on the words, no difference in effects sizes of word variables on fixation duration could be observed. During sentence reading, both groups showed a similar sensitivity to word frequency, word predictability, and word length of the fixated words and its neighboring words.

A second main difference between reading conditions was found in the effect size of oculomotor predictors on single fixation duration. The length of incoming and outgoing saccade amplitude had a stronger effect on SFD in the 'frequent' sample than in the 'original' sample, even though groups did not significantly differ in mean saccade lengths. In the positive correlation of saccade amplitude and fixation duration, the sample 'frequent' showed a much steeper slope, that is, single fixation duration was even longer before and after long saccades. In addition, a minor difference in the IOVP-effect was found between the experimental groups. In the 'frequent' sample, the peak of the IOVP-curve was shifted slightly further to the right, towards the center of the word. Again, groups did not differ in mean relative fixation position of single fixation cases.

Interestingly, even though reading rate was much lower in the 'frequent' sample compared to the 'original' sample, the fixational selectivity was not affected by the frequency of comprehension questions. Readers in the 'frequent' group fixated words for a longer time, but on average did not fixate different words. A different selectivity pattern could have been expected with a reduced reading speed, because a smaller reading rate is usually associated with shorter saccade amplitudes and a lower skipping rate, as found for example in less skilled readers or reading beginners (Rayner, 1998; Underwood, Hubbard, & Wilkinson, 1990). Word skipping probability and incoming and outgoing saccade amplitudes for single fixation cases in first-pass reading were numerically smaller in the 'frequent' sample compared to the 'original' sample, but these differences did not reach the level of significance. Thus, it seems that the reading strategy induced by frequent, easy questions was basically a slower reading with slightly shorter saccade amplitudes. The stronger effects of saccade amplitudes on SFD in the 'frequent' condition can be explained as a results of the reduced reading speed: The execution of long saccades, that occurred somewhat less often in the 'frequent' than in the 'original' condition, needed longer saccade latency times. In sum, the experimental manipulation of the frequency of occurrence of easy comprehension questions had a major impact on reading speed and its related oculomotor effects. Importantly, the results demonstrate that a reduction of reading rate is not necessarily associated with a significant decrease of saccade amplitudes and skipping rates, providing further evidence for an independence of the *when*- and *where*-dimension in saccade programming (Findlay & Walker, 1999).

5.7.2. Reading Speed and Attention to the Task

Despite the fact that the sample 'frequent' was slightly older than the group 'original', vocabulary scores and the index of processing speed were lower in the 'frequent' group than in the 'original' group. One could argue, that prolonged fixation durations in the 'frequent' group are related to a lower reading skill, indicated by vocabulary knowledge and the general processing speed, and that they are not a results of the experimental manipulation. However, this possibility is not consistent with the absence of differences in word frequency effects, that have been found in the study of eye movements of highly skilled and average readers (Ashby et al., 2005). If reading skill would differ significantly between the two reading groups, and thus, processes of lexical access would vary in their efficiency, than the word frequency effect on fixation duration would have been expected to be more pronounced in the 'frequent' group. Instead, the two reading groups did not differ in their effect sizes of word frequency and this indicates a comparable word processing in both groups, that is, lexical access presumably occurred in both reading groups at the same point in time. Thus, for the interpretation of the results the group differences in psychometric data can be neglected because it can be assumed that these minor differences (advantage in age, disadvantage in vocabulary size and processing speed of the sample 'frequent') cancel each other. With respect to this claim it must be explained why the readers' eyes of the 'frequent' sample stayed on the words, even after lexical access was achieved.

It appears that lexical access processes did not always control when the eyes moved. Rather, it seems that the main characteristic of the reading strategy applied when frequent comprehension questions are asked is a general slowing in reading speed, though mostly independent of a change in saccade targeting. Again, one could argue that the increase in fixation duration in the 'frequent' sample can be attributed to a higher attention on word processing and the integration of words. Contrary to this claim, no evidence could be found for differences in lexical processing between the two reading groups. Moreover, longer fixation durations in the 'frequent' sample cannot be a result of variation in local lexical processing, because lexical difficulty is identical in both reading conditions. In addition, accuracy rates on the comprehension questions were very high in both reading conditions. As argued in chapter 4, the comprehension questions of the baseline experiment are simple because they can be solved by pure visual word form recognition. This is true even if those easy questions are asked after each sentence. Therefore, it is implausible to argue for an increased high-level linguistic processing in the 'frequent' sample compared to the 'original' sample.

Instead, it appears that a different amount of attention at a different level of processing is allocated in the two reading conditions, namely in the general attention to the reading task. The frequent interruption of the reading process by asking questions after each trial might have distracted subjects in the 'frequent' condition from showing a smoother and faster fixational behavior. Subjects were reminded of the reading task after every sentence and this presumably influenced the parameter setting for the eye movement control each trail anew. The prominent decrease in reading speed in the sample 'frequent' compared to the 'original' sample can clearly be attributed to the manipulation of question frequency and the entailed increase in attention to the reading task. Saccade timing is influenced by some top-down mechanism triggered by reading intention. Importantly, the results of the sample 'frequent' confirm the interpretation that the effects of selectivity and word-processing in the 'hard' samples (cf. section 4.7) can be attributed to the higher task demands due to the difficulty of the comprehension questions. Frequent questions lead to a reading strategy with increased fixation times, but fixational selectivity and word-level effects on fixation duration measures are not affected.

6. Strategy Effects Due to Task Instruction

In the present chapter, the influence of task instruction on the reading of the identical sentence material is investigated. In contrast to Experiment 1, the goal in the third experiment is to induce a more 'mindless' reading strategy with respect to reading comprehension. A lexical search task (as used for example in Rayner & Raney, 1996) does not seem to be the adequate method to induce superficial reading because it does not guarantee that readers attend to each word form in the text. In lexical search, readers may simply move on with their eyes, triggered by the recognition of the first three letters of a fixated word and the identified match or mismatch with the target word. Instead, a reading strategy that needs word form recognition for each item, but not necessarily semantic processing is needed. Proofreading seems to be the perfect reading condition for this test, especially if misspellings affect only word form, not syntactic or semantic congruency. In almost all studies using proofreading - and especially proofreading of homophones - the role of phonology in silent reading has been investigated (see Jared, Levy, & Rayner, 1999, for an example and review). Here, proofreading is used to induce a specific reading strategy via instruction.

In the following study, eye movement data of two samples with different reading instructions but identical sentence material are compared. Reading for comprehension as a mindful reading strategy is contrasted with a more superficial reading strategy, namely proofreading. As demonstrated in the preceeding chapter 5, the frequency of the occurrence of comprehension questions has only a significant impact on fixation duration measures, not on fixational selectivity or word-level effects on fixation duration. It is argued that due to the constant interruption of the reading process by frequent questions in the 'frequent' sample the saccade timer is set to a slower pace, resulting in the reduction of the reading speed. Since the frequency of questions did not affect word processing mechanisms, the sample 'frequent' serves as a good control group for the proofreading condition, in which the instruction differed, but questions were also asked after each trial.

Skilled readers are able to detect word form errors in text or sentence reading (Daneman & Stainton, 1993; Pilotti & Chodorow, 2009; Pilotti, Chodorow, & Thornton, 2005; Levy & Begin, 1984; Levy, 1983). Proofreaders differ in the accuracy of error detection in familiar vs. unfamiliar text (Pilotti & Chodorow, 2009; Levy & Begin, 1984). To avoid effects of rereading (Raney & Rayner, 1995), subjects in the proofreading condition were unfamiliar with the PSC. In a proofreading task that focus on the accuracy of the word form subjects are expected to attend strongly to the orthography of words and less to the word and sentence meanings, especially if no comprehension questions are asked (Haber & Schindler, 1981; Pilotti et al., 2005; Rayner & Pollatsek, 1989). Thus, a proofreading strategy is expected to minimize reading comprehension and to resemble a more 'mindless' reading strategy. In contrast to deep semantic processing during reading, a superficial reading strategy such as proofreading might influence the effects of distributed processing in a different direction. Word predictability effects of the upcoming word turned out to be indexical for depth of processing (cf. section 4.7). Because it is expected that during proofreading the reader puts less weight on the processing of the word's and sentence's meaning, that is, minimal processing of the sentence context is needed, weaker predictability effects on fixation duration are expected in the proofreading condition. In addition to reduced preprocessing effects, stronger immediacy effects of word variables might be expected because the focus in proofreading is very local, that is, on the fixated word. In contrast, while proofreading the reader is expected to put more weight on visual accuracy, here on the accuracy of word spellings. Visual acuity is highest at the fovea and word recognition times are minimal at fixation locations at the word center, the so-called optimal viewing position (OVP; O'Regan & Lévy-Schoen, 1987). In other words, fixation locations at the word center are optimal for saccade targeting and foveal word processing. Therefore, one would expect to find a more precise saccade targeting to the word center in the proofreading condition. In sum, with respect to processing depth proofreading is expected to induce a more superficial reading strategy than reading for comprehension. With respect to oculomotor control, proofreading is expected to induce a more precise word targeting strategy than reading for comprehension.

After a detailed description of the subject samples and the proofreading experiment, results of the analyses on eye movement patterns comparing the proofreading experiment with the sample 'frequent' are provided. Results are presented analogous to the previous result chapters 4 and 5. In addition to providing analyses of fixational distributions, word-based summary statistics averaged across subjects, and results of lmer-modeling comparing the samples ‘proofreading’ and ‘frequent’, results of the eye movements on erroneous sentences in the proofreading experiment are presented. Here, the critical questions are if and how the eye movement behavior reflects the detection of a spelling mistake and what effects are causally related to the error detection.

6.1. Participants

Data of two age-matched samples of 30 college students were compared in this study (cf. Table 6.1). The group instructed to proofread the material is labeled ‘proofreading’, whereas the group that read for comprehension is labeled ‘frequent’. In contrast to the previous group comparisons, where groups were tested at different points in time (post-hoc between-subjects design), subjects were randomly assigned to the ‘proofreading’ or the ‘frequent’ groups. Therefore, reading condition is a true experimental factor. Note that the ‘frequent’ group is the identical sample described in chapter 5 and that results of this group are repeated in this chapter in comparison to the data of the ‘proofreading’ sample. All subjects were native speakers of German and they all had normal or corrected to normal vision. Groups did not differ in age, on Lehrl’s (Lehrl, 1977) vocabulary score, or in Wechsler’s (Wechsler, 1964) Digit-Symbol-Test (all $p < 0.05$).

Table 6.1.: Group statistics on age, vocabulary size, and digit-symbol task for the samples ‘proofreading’ and ‘frequent’; a complete list of subject information is listed in Tables C.7, and C.6 in Appendix C.

GROUP	N	M	SD	RANGE
Age				
Frequent	30	23.9	4.15	19 - 36
Proofreading	30	22.8	3.0	20 - 32
Vocabulary test				
Frequent	30	31.7	2.8	26 - 35
Proofreading	30	31.6	1.64	29 - 35
Digit-Symbol test				
Frequent	30	62	11	43 - 89
Proofreading	30	63.2	8.1	48 - 82

6.2. Experiment 3: Proofreading the PSC

To ensure a data comparison on the identical sentence material with different reading instructions, the 144 sentences of the PSC had to be errorfree. Therefore, additional sentence material that contained word form errors were used in the proofreading condition. Group-comparative results of the data analysis are based only on the 144 correctly spelled sentences of the PSC.

6.2.1. Additional Erroneous Sentence Material

For the proofreading condition 56 additional sentences were used in which one misspelling per sentence occurred. The erroneous sentence material is listed in Appendix B. 64% of the misspellings occurred in content words, 36% in function words. Misspellings consisted of letter elisions (*schleichen* → *schlechen*), letter additions (*die* → *diie*), letter permutations (*Preis* → *Presi*), and letter substitutions (*Gegend* → *Kegend*). The four error types occurred in equal proportions and always resulted in a pronouncable pseudoword. The occurrence of word form errors was restricted from the second up to the second last word in the sentences.

6.2.2. Apparatus and Procedure

The monitor resolution used in the proofreading experiment was 1024 x 768 pixels. Letters subtended 0.41° of visual angle. Eye movements were recorded with a 500-Hz-sampling rate. Sentence presentation and calibration were identical to the baseline experiment (cf. sections 3.2.2 and 3.2.3)[10]. In the proofreading condition, each sentence was replaced by a two-choice accuracy-question ("Was each word in the sentence spelled correctly?" yes/no). Participants answered via a mouse click. The proofreading task was initiated with five training trials and subjects received feedback about the number of misspellings in the five training trials. Due to the 56 filler sentences, subjects in the proofreading condition read a total of 200 sentences, whereas subjects in the comprehensive reading condition (sample 'frequent') read a total of 144 sentences. The sample 'frequent' received an easy comprehension question after each trial (cf. section 5.2).

[10]Experiment 3 was written with MATLAB (The MathWorks, Inc., Natick, MA), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) as well as the Eyelink Toolbok extensions (Cornelissen et al., 2002).

6.3. Results of Saccade Detection and Fixation Selection

For group-comparative data analyses, only eye movements on the 144 sentences of the PSC were evaluated. Hence, fixations on the identical and correctly spelled sentence material provides the data base in both reading conditions; solely reading instruction differed between the samples 'frequent' and 'proofreading'.

Table 6.2.: Distribution of fixation types for the samples 'frequent' and 'proofreading'; data are from right eye; rows 2) and 3) sum up to 1), rows 4), 5), and 6) sum up to 3); data of the proofreading sample are devided into fixations on the PSC and fixations on the erroneous filler sentences.

VARIABLE		FREQUENT	PROOFREADING	
			PSC	ERR. SENT.
1) N of fixations (total)		34,071	44,816	15,152
2) First/ last word;	*N*	9,689	12,183	6,869
First/ last fixation	%	28	27	45
3) N of valid fixations	*N*	24,382	32,633	8,283
	%	72	73	55
4) First-pass	*N*	21,327	25,584	6,910
	%	88	79	84
5) Second-pass	*N*	2,464	4,004	1,178
	%	10	12	14
6) Rereading	*N*	591	3,045	195
	%	2	9	2

The distribution of types of detected fixations for the samples 'frequent' and 'proofreading' are listed in Table 6.2. For the 'proofreading' sample, number of fixations types are listed for eye movements on the PSC (central column) as well as for eye movements on the filler sentences (right column). The total number of detected saccades for the right eye are given in the first row (1). The second row lists the number of fixations for criterion (2), marking fixations on first and last words of the sentences and first and last fixations of the trials. In the bottom part of the Table, a breakdown of numbers and percentages of different reading passes relative to valid within-sentence fixations (3) is provided. Note that in Tables 6.2 and 6.3, the total number and percentages of fixation types per experimental group is reported, but ANOVA statistics are based on subjects' mean percentage per fixation type[11].

Similar to the results of the previously reported experimental groups, about 70% of the fixations in proofreading are valid within sentences fixations in the data set of the 144 sentences of the PSC. The sample 'frequent' and the sample 'proofreading' do not differ in the percentage of valid fixations detected ($p > 0.05$). In contrast, eye movement data on the erroneous filler sentences in proofreading result in only 55% valid fixations, a significantly smaller proportion in comparison to the data of the 'frequent' sample ($F(1, 58) = 129.5$, $MSe = 32.2$, $p < 0.001$) as well as to the data of 'proofreading' the PSC ($F(1, 58) = 155.6$, $MSe = 30.9$, $p < 0.001$). Remember that fixations of criterion (2) are a) the first and last fixation of a trial and b) fixations on the first and last word of a trial. Usually, these fixations overlap, because if reading starts at the first word and ends at the last word of a trial, the first and last fixation of this trial are located at the first, respectively last word of the sentence. In reading the erroneous sentences, it is likely that subjects ended reading the sentences if they detected the error. Thus, the last fixation in a trial does not always coincides with fixations on the last word. Indeed, a posthoc analysis shows that on average in 15% (range: 4 - 34%) of the erroneous trials subject did not produce first-pass fixations after reading the misspelled word. This supports the assumption, that the last fixation in a filler sentences was often located before the last word of the trial. The fact that the first reading pass of erroneous trials was often 'early interrupted' might also influence word skipping probability on these trials.

The sample 'proofreading' produced significantly fewer first-pass fixations than the 'frequent' sample ($F(1, 58) = 6.8$, $MSe = 102.4$, $p < 0.05$). The proportion of second-pass fixations did not differ between samples ($p > 0.05$), but the 'proofreading' group showed proportionally more rereading fixations than the 'frequent' sample ($F(1, 57) = 8.2$, $MSe = 45.7$, $p < 0.01$). A different picture arises from the erroneous trials of the proofreading sample. In comparison to the 'frequent' group, again fewer first-pass fixations are found in 'proofreading' the filler sentences ($F(1, 58) = 4.4$, $MSe = 49.3$, $p < 0.05$), but more second-pass fixations ($F(1, 58) = 6.0$, $MSe = 34.3$, $p < 0.05$) and no differences between the proportions of rereading fixations ($p > 0.05$). Again, it is

[11] Since not all subjects of the experimental samples provided fixations to each fixation type listed, the degrees of freedom may vary between ANOVA statistics for different fixation types.

Table 6.3.: Number and proportion of fixation types in first-pass reading, second-pass reading, and rereading for the samples 'frequent' and 'proofreading'; data are from right eye; note that the number of fixations types per reading pass sum up to the number of fixations per pass listed in Table 6.2.

VARIABLE		FREQUENT	PROOFREADING	
			PSC	ERR. SENT.
Fixation types in first-pass reading				
a) Single fixations (SF)	*N*	14,288	13,525	3,397
	%	67	53	49
b) SF regression origin	*N*	1,356	1,627	690
	%	6	6	10
c) Multiple fixations (MF)[12]	*N*	5,555	10,133	2,727
	%	26	40	40
d) MF regression origin	*N*	128	299	96
	%	1	1	1
Fixation types in second-pass reading				
a.1) Single fixations (SF)	*N*	715	1,120	337
	%	29	28	29
b.1) SF regression origin	*N*	35	100	34
	%	1	2	3
c.1) Multiple fixations (MF)	*N*	328	1,058	308
	%	13	27	26
e.1) SF regression goals	*N*	1,194	1,338	381
	%	49	33	32
f.1) MF regression goals	*N*	85	190	62
	%	4	5	5
g.1) SF regr. goals & origin	*N*	107	198	56
	%	4	5	5
Fixation types in rereading				
a.2) Single fixations (SF)	*N*	119	682	39
	%	20	22	20
b.2) SF regression origin	*N*	9	133	12
	%	2	5	6
c.2) Multiple fixations (MF)	*N*	67	859	35
	%	11	28	18
e.2) SF regression goals	*N*	239	539	53
	%	41	18	27
f.2) MF regression goals	*N*	48	338	25
	%	8	11	13
g.2) SF regr. goals & origin	*N*	109	494	31
	%	18	16	16

likely that subjects did not finish reading the sentence in all erroneous trials. That way, only a few rereading fixations can be expected. The difference in proportion of rereading fixation between 'proofreading' erroneous and correct trials is significant ($F(1, 58) = 6.7$, $MSe = 46.4$, $p < 0.05$).

According to the definitions in section 3.3.1, Table 6.2 provides the number of fixation types per reading pass for the sample 'frequent' (in the left column) and the sample 'proofreading' on the PSC (central column) and on the filler sentences (right column). Significant main effects of reading condition are found in first-pass, second-pass, and rereading fixation types. The distribution of fixation types on the erroneous sentences differed from the data on the correct sentences as well as from the data of reading for comprehension.

First-pass Reading. In comparison to the 'frequent' group, the 'proofreading' group produced proportionally fewer single fixations (a) ($F(1, 58) = 22.8$, $MSe = 124.3$, $p < 0.001$), but reliably more multiple fixation cases (c) ($F(1, 58) = 23.2$, $MSe = 115.2$, $p < 0.001$), indicating a reading strategy that emphasizes refixations. The proportion of multiple fixation regression origins (d) was larger in 'proofreading' (1.1) than in the 'frequent' sample (0.7) ($F(1, 58) = 6.7$, $MSe = 0.35$, $p < 0.05$).

The eye movement behavior on the erroneous trials in comparison to the 'frequent' sample, also reveals significantly more multiple fixation cases ($F(1, 58) = 24.7$, $MSe = 114$, $p < 0.001$) at the cost of 'pure' single fixations, bordered by forward saccades ($F(1, 58) = 40.7$, $MSe = 123.3$, $p < 0.001$). The proportion of regression origins in single fixation cases (b) and multiple fixation cases (d) was significantly larger in 'proofreading' the error trials than in reading for comprehension ($F(1, 58) = 10.4$, $MSe = 20.1$, $p < 0.01$; $F(1, 51) = 13.3$, $MSe = 0.7$, $p < 0.001$). 'Proofreading' on erroneous trials compared to correct trials revealed significantly more single fixations that are the origin of a regression ($F(1, 58) = 15.6$, $MSe = 14.4$, $p < 0.001$). Thus, the existence of a misspelled word in a trial seems to entail an increase in regressive eye movements. In section 6.7 the targeting of regressions in the erroneous filler trials is further analyzed.

Second-pass Reading. The 'proofreading' sample produced proportionally more multiple fixations on a word (c.1) compared to the 'frequent' sample ($F(1, 58) = 12.2$, $MSe = 83.7$, $p < 0.001$). Consistently, in proofreading a regressive movement was less frequently finished af-

[12] Multiple fixation cases sum up to 2,461 valid first-pass gaze durations for the sample 'frequent' and to 4,193 gaze durations for the sample 'proofreading'.

ter one fixation (e.1) in comparison to reading for comprehension ($F(1, 58) = 16.2$, $MSe = 133.9$, $p < 0.001$).

A similar patterns is found for the proportion of fixation types of 'proofreading' the incorrect trials in comparison to the 'frequent' sample. While proofreading erroneous sentences, subjects produce more multiple fixation cases ($F(1, 54) = 9.5$, $MSe = 101.9$, $p < 0.01$), fewer single fixations as regression goals ($F(1, 58) = 14.4$, $MSe = 178.8$, $p < 0.001$), but more multiple fixation regression goals (f.1) ($F(1, 44) = 8.6$, $MSe = 13.4$, $p < 0.01$). The number of single fixations in second-pass that are the origin of a regressive eye movement was reliably larger in 'proofreading' incorrect sentences than in reading for comprehension ($F(1, 27) = 8.2$, $MSe = 14.2$, $p < 0.01$).

The distribution of second-pass fixation types in proofreading the correct sentences is comparable to the one on incorrect filler sentences. A difference in these two conditions is found in the proportion of single fixations that are regression origins ($F(1, 39) = 13.3$, $MSe = 10.5$, $p < 0.001$).

Rereading. The distribution of fixation types in rereading is similar to the the pattern found in second-pass reading. Proofreaders reading correct sentences produced proportionally more multiple fixation cases (c.2) in comparison to subjects that read for comprehension ($F(1, 42) = 8.0$, $MSe = 68.1$, $p < 0.01$). In turn, they produced reliably fewer single fixations that were regression goals (e.2) ($F(1, 57) = 16.1$, $MSe = 528.3$, $p < 0.001$).

The differences between proofreaders reading incorrect sentence and the sample 'frequent' reveal a similar picture. Proofreading incorrect trials produced proportionally more multiple fixation cases ($F(1, 31) = 10.7$, $MSe = 84.8$, $p < 0.01$), their proportion of multiple fixation regression goals was increased ($F(1, 33) = 5.2$, $MSe = 318.5$, $p < 0.05$), and their proportion of single fixation regression origins was larger ($F(1, 10) = 5.8$, $MSe = 77$, $p < 0.05$) than in reading for comprehension.

Note that the proportion of rereading fixations in the group 'frequent' and 'proofreading' on erroneous trials made up only 2% of the total number of valid fixations (see Table 6.2). Therefore, these differences between reading conditions are of minor importance.

In sum, the experimental manipulation of reading instruction has an impact on the distribution of fixation types. In comparison to reading for comprehension, subjects in a proofreading mode more frequently fixate words twice or more often during the first inspection of a sentence. In proofreading, more regressions can be observed if an error is present in the sentence. During the second inspection of a sentence section, proofreaders again produce many more multiple fixation cases than comprehensive readers. After reaching the end of a sentence, proofreaders compared to the comprehension readers reread a sentence more often if no error is present in the trial. In this case, the proportion of multiple fixation cases is again much higher than in reading for comprehension. In the next section, word-based statistics averaged across subjects concerning fixation probabilities, fixation locations, and fixation durations will broaden this picture.

6.4. Mean Probabilities, Positions, and Durations of Fixations

Word-based summary statistics averaged across subjects for the samples 'frequent' and 'proofreading' are listed in Table 6.4. Means are based on all fixations per sample (cf. top row of Table 6.2). For both reading for comprehension and proofreading the PSC, 92% of the sentences that the subjects read, were included in the analyses. Proofreaders made on average three more fixations per sentence than comprehension readers on the identical sentence material.

Considering the eye movement data on the 144 sentences of the PSC, ANOVA results in reliable differences between the samples 'frequent' and 'proofreading' with respect to fixation probabilities and fixation durations, but not with respect to fixation locations. In proofreading, first-pass skipping probability was reduced compared to the skipping rate in reading for comprehension ($F(1, 58) = 5.9$, $MSe = 0.004$, $p < 0.05$). Whereas the probability of fixating a word once during first-pass reading was significantly lower in proofreading ($F(1, 58) = 14.5$, $MSe = 0.003$, $p < 0.001$), proofreaders showed a much higher probability of fixating a word twice ($F(1, 58) = 12.9$, $MSe = 0.002$, $p < 0.001$) or three times and more often ($F(1, 58) = 29.6$, $MSe = 0.0005$, $p < 0.001$) compared to the subjects that read for comprehension. Regression probability was much higher in proofreading than in the 'frequent' sample ($F(1, 58) = 9.1$, $MSe = 0.006$, $p < 0.01$). Note that this word related regression probability also includes regressions made out of second-pass reading or rereading. The difference between the regression probability for proofreading the PSC listed in Table 6.4 (13 %) and the regression probability of the 'frequent' sample (7%) originates from differences in the number of second-pass regression origin fixations and rereading regression origin fixations (cf. Table 6.3 rows b.1 and c.1).

Table 6.4.: Word-based summary statistics averaged across subjects for the samples 'frequent' and 'proofreading'; proofreading data are split up for the 144 sentences of the PSC and the 56 erroneous filler sentences.

VARIABLE		FREQUENT	PROOFREADING	
			PSC	ERR. SENT.
N of readers		30	30	30
N of fixations/ sentence	*M*	8.6	11.3	9.7
	SD	1.4	3.1	1.4
N of sentences	*M*	133	133	52
	SD	10	14	5
Fixation probability				
Skipping	*M*	.17	.13	.26
	SD	.07	.05	.07
Single fixation	*M*	.67	.62	.53
	SD	.05	.06	.08
Double fixation	*M*	.10	.15	.11
	SD	.05	.05	.04
Three-plus fixation	*M*	.01	.04	.04
	SD	.01	.03	.02
Regression	*M*	.07	.13	.11
	SD	.05	.10	.04
Regr. goal after skipping	*M*	.02	.01	.01
(cond. probability)	*SD*	.01	.01	.01
Relative fixation position				
Position single fixation	*M*	0.5	0.5	0.5
	SD	.04	.06	.05
Position 1st fixation	*M*	0.3	0.2	0.3
	SD	.09	.03	.05
Position 2nd fixation	*M*	0.7	0.6	0.6
	SD	.07	.05	.05
Fixation duration (ms)				
single fixation	*M*	225	235	239
	SD	27	32	36
1st of multiple	*M*	213	224	235
	SD	27	34	35
2nd of multiple	*M*	183	212	230
	SD	26	34	40
Gaze duration	*M*	253	302	307
	SD	37	59	63
Total reading time	*M*	274	373	362
	SD	47	133	84
Reading rate (wpm)		251	175	216

The groups did not differ in single fixation duration or first fixation duration (all $p > 0.05$). Means of second fixation duration ($F(1, 58) = 13.6$, $MSe = 913$, $p < 0.001$), gaze duration ($F(1, 58) = 14.9$, $MSe = 2401$, $p < 0.001$), as well as of total reading time ($F(1, 58) = 14.7$, $MSe = 10005$, $p < 0.001$) were significantly longer in the 'proofreading' sample than in the 'frequent' sample. No differences between samples were found in means of relative fixations position of single fixations, first fixations, and second of multiple fixations (all $p > 0.05$). The overall reading rate in proofreading was reduced by 76 words per minute (wpm).

Comparison of the means in reading for comprehension and proofreading on erroneous sentences reveals a slightly different picture. Skipping probability in proofreading incorrect sentences is significantly higher than in the sample 'frequent' ($F(1, 58) = 26.3$, $MSe = 0.005$, $p < 0.001$). As mentioned above, the relatively high first-pass skipping rate in proofreading on erroneous sentence is supposely due to an early stopping of sentence reading after detecting the error (cf. section 6.3). Similar to the fixational behavior found in proofreading correct sentences, the probability of fixating a word once is reduced ($F(1, 58) = 71.2$, $MSe = 0.004$, $p < 0.001$), and the probability of fixating a word three times or more is reliably increased ($F(1, 58) = 22.7$, $MSe = 0.0003$, $p < 0.001$) in proofreading incorrectly spelled sentences compared to the sample 'frequent'. The mean probability of double fixations does not differ between reading for comprehension and proofreading incorrect sentences ($p > 0.05$). The higher regression probability in proofreading incorrect sentences compared to the 'frequent' sample is significant ($F(1, 58) = 8.6$, $MSe = 0.002$, $p < 0.01$). Differences in single fixation durations are not reliable ($p > 0.05$), but first ($F(1, 58) = 7.4$, $MSe = 980$, $p < 0.01$), but second fixation duration ($F(1, 58) = 29.3$, $MSe = 1147$, $p < 0.001$), gaze duration ($F(1, 58) = 16.7$, $MSe = 2677$, $p < 0.001$), and total reading time ($F(1, 58) = 24.8$, $MSe = 4672$, $p < 0.001$) are significantly longer in proofreading erroneous sentences in comparison to the sample 'frequent'. Again, no differences were found in relative fixation location of single and multiple fixation cases (all $p > 0.05$). Reading speed was much slower in proofreading incorrect sentences than in reading for comprehension.

The main differences in proofreading between the eye movement data on correct sentences and on the erroneous material are found in fixation probabilities. First-pass skipping rates are much higher on erroneous sentences than on correct sentences ($F(1, 58) = 66.7$, $MSe = 0.004$, $p < 0.001$). Probabilites of fixating a word once ($F(1, 58) = 25.5$, $MSe = 0.005$, $p < 0.001$) or twice ($F(1, 58) = 7.7$, $MSe = 0.002$, $p < 0.01$)

are significantly lower on erroneous sentences. Regression probability and the number of multiple (three plus) fixations cases do not differ between sentence types. No differences between eye movement data on correct and incorrect sentences are found in fixation durations; only the second of multiple fixation cases is marginally longer in erroneous sentences ($F(1, 58) = 3.8$, $MSe = 1362$, $p = 0.055$). Consistent with the higher word skipping rate on erroneous trials, the reading rate was about 40 wpm faster in proofreading incorrect sentences than in proofreading correctly spelled trials. Except for the relative fixation position of the first of multiple fixation on a word, which is located further to the right in erroneous sentences ($F(1, 58) = 12.1$, $MSe = 0.002$, $p < 0.001$), no differences in mean fixation locations are found between the subdata in proofreading. Remember that the sentence material used in proofreading erroneous trials differed from the one used in reading for comprehension and proofreading the PSC. Therefore, the comparability of eye movements of the 'proofreading' sample on incorrect trials to the other two samples is reduced.

6.5. Response Accuracy

Response accuracy for the samples 'frequent' and 'proofreading' is listed in Table 6.5. The sample 'frequent' recieved easy mulitple-choice comprehension questions in 100% of the trials. The accuracy for the proofreading-sample is split up into a) percentage of correct acceptances of the spelling of the 144 sentences of the PSC, and b) percentage of correct error detections in the 56 filler sentences that included a spelling error. Error detection accuracy for erroneous sentences is further broken down into accuracy for the four error types 1) letter addition, 2) letter elision, 3) letter permutation, and 4) letter substitution. Since the PSC was created using the German writing rules based on Duden (2000) and the proofreading experiment was conducted after the German spelling reformation in August 2006 (Wermke, Kunkel-Razum, & Scholze-Stubenrecht, 2006), sentences with wrong spelling according to the new German spelling rules were excluded from accuracy statistics for 'proofreading' the PSC.

Overall performance of responses on the PSC was very high with more than 96% correct. ANOVA analyses revealed a significant differences in accuracy between the samples 'frequent' and 'proofreading' ($F(1,58) = 7.93$, $MSe = 0.0004$, $p < 0.01$). Within the experimental group 'proofreading', subjects were reliably more accurate in accept-

Table 6.5.: Response accuracy for comprehension questions of the sample 'frequent', and for spelling check of the sample 'proofreading'; responses in proofreading are broken down by the correct sentences of the PSC and the erroneous filler sentences; response accuracy on erroneous sentences is broken down by error type: 1) letter addition, 2) letter elision, 3) letter permutation, and 4) letter substitution.

GROUP		RESPONSE ACCURACY (% correct)		
		Mean	SD	Range
Frequent		98.1	1.2	95 - 100
Proofreading	a. PSC	96.6	2.6	88 - 100
	b. erroneous sent.	90.2	6.2	70 - 96
Error types in b.	1. addition	91.4	0.07	79 - 100
	2. elision	87.9	0.11	57 - 100
	3. permutation	92.6	0.07	69 - 100
	4. substitution	89.1	0.08	60 - 100

ing a correctly spelled sentence than detecting spelling mistakes in the erroneous filler sentences ($F(1, 58) = 19.6$, $MSe = 0.002$, $p < 0.001$). Examining the accuracy of error detection for the four error types, letter permutation (3) seems to be the easiest error to perceive, whereas letter elisions (2) are harder to trace; but the difference between the accuracy for these two error types is only marginal ($F(1, 58) = 3.6$, $MSe = 0.009$, $p < 0.06$). Independent of error type, error detection accuracy does not differ between errors in content (91 % correct) or function words (88% correct) ($F(1, 58) = 2.0$, $MSe = 0.006$, $p > 0.05$).

Table 6.6.: Means and standard deviations of word properties of fixated words and of oculomotor variables for the samples 'frequent' and 'proofreading' for first-pass single fixation cases (SFC) and multiple fixation cases (MFC).

		FREQUENT		PROOFREADING	
VARIABLE		SFC	MFC	SFC	MFC
Frequency (log/ mio)					
word n	*M*	2.22	1.36	2.40	1.35
	SD	0.17	0.23	0.19	0.19
word n - 1	*M*	2.28	2.64	2.23	2.70
	SD	0.11	0.27	0.06	0.21
word n + 1	*M*	2.35	2.15	2.36	2.16
	SD	0.06	0.19	0.05	0.13
(PSC: *M* = 2.3, *SD* = 1.3)					
Predictability (logit)					
word n	*M*	-1.58	-1.98	-1.47	-2.00
	SD	0.11	0.14	0.12	0.10
word n - 1	*M*	-1.65	-1.54	-1.66	-1.51
	SD	0.05	0.13	0.03	0.11
word n + 1	*M*	-1.20	-1.32	-1.17	-1.29
	SD	0.05	0.22	0.05	0.10
(PSC: *M* = -1.48, *SD* = 1.1)					
Length (n of letters)					
word n	*M*	4.6	7.2	4.3	7.0
	SD	0.3	0.9	0.3	0.6
word n - 1	*M*	4.3	3.9	4.3	3.8
	SD	0.2	0.3	0.1	0.2
word n + 1	*M*	4.4	4.7	4.4	4.6
	SD	0.1	0.3	0.1	0.2
(PSC: *M* = 5.4 , *SD* = 2.6)					
Function word proportion					
word n	*M*	0.32	0.14	0.37	0.13
	SD	0.06	0.08	0.06	0.05
word n - 1	*M*	0.40	0.55	0.37	0.57
	SD	0.04	0.10	0.03	0.08
word n + 1	*M*	0.38	0.33	0.38	0.33
	SD	0.02	0.07	0.02	0.05
(PSC: *M* = 0.37, *SD* = 0.48)					
Oculomotor variables					
incoming sacc.ampl.	*M*	7.1	-	6.4	-
	SD	0.9	-	0.7	-
outgoing sacc.ampl.	*M*	7.2	-	6.4	-
	SD	0.9	-	0.8	-
rel. fix. position	*M*	0.43	-	0.51	-
	SD	0.03	-	0.07	-

6.6. Linear Mixed Modeling: Effects of Reader and Word Variables

Linear mixed models (Pinheiro & Bates, 2000; Gelman & Hill, 2007) were used to evaluate differences between the samples 'frequent' and 'proofreading' in reader-level and word-level effects on first-pass single fixation duration and gaze duration. Models were built in the same way as described in sections 4.6 and 5.6. The random sampling of subjects, words, and sentences, as well as reader-level effects, such reading condition, vocabulary size etc., and word-level predictors, such as word frequency, word predictability, etc., were considered simultaneously during modeling. Of main interest are cross-level interactions with reading condition. First-pass single fixation duration and gaze duration were fitted separately.

6.6.1. Selectivity Effects Related to Reading Strategy

The results described in the word-based summary statistics across subjects in section 6.4 demonstrate that there is a difference between the samples 'frequent' and 'proofreading' in total number of fixations per sentence and also in fixation probabilities, e.g. in first-pass skipping rate. This indicates clearly that reading instruction affected the selectivity of fixated words. Therefore, effects related to reader- and word-level predictors in models fitting first-pass single fixation duration and gaze duration need to be interpreted with simultaneous consideration of the composition of the 'sub-corpora' of each reading condition. Means of word-level predictors and variables of oculomotor control for first-pass single and multiple fixation cases are listed in Table 6.6.

In proofreading, 53% of first-pass fixations are single fixation cases. In the 'frequent' sample, single fixations constitute 67% of all first-pass fixations (cf. Table 6.3). In the data set of first-pass single fixation cases, ANOVA testing revealed several differences in word-level and oculomotor variables between reading conditions. At the word-level, fixated words n were of significantly higher frequency ($F(1, 58) = 15.4$, $MSe = 0.03$, $p < 0.001$), of higher predictability ($F(1, 58) = 13.4$, $MSe = 0.01$, $p < 0.001$), and shorter ($F(1, 58) = 12.7$, $MSe = 0.0003$, $p < 0.001$) in the 'proofreading' sample than in the 'frequent' sample. Consistent with this result, the proportion of fixated function words was reliably larger in 'proofreading' than in the 'frequent' group ($F(1, 58) = 12$, $MSe = 0.003$, $p < 0.001$). The words to the left of the single fixation differed also in frequency and the proportion of function words. In proofreading, words n-1 were significantly less often a function word ($F(1, 58) = 11.3$, $MSe = 0.001$, $p < 0.001$) and marginally lower in frequency ($F(1, 58) = 3.9$, $MSe = 0.007$, $p = 0.05$) than in reading for comprehension.

Main effects of reading condition were also found in oculomotor variables. Single fixations in 'proofreading' were bounded by shorter incoming ($F(1, 58) = 10.7$, $MSe = 0.67$, $p < 0.01$) and shorter outgoing saccade amplitudes ($F(1, 58) = 12.9$, $MSe = 0.74$, $p < 0.001$) than in the 'frequent' sample. Mean relative fixation position within the word was significantly located further to the right in the 'proofreading' sample compared to the 'frequent' group ($F(1, 58) = 39.9$, $MSe = 0.002$, $p < 0.001$). Single fixations in proofreading were located exactly at the word center, whereas readers in the 'frequent' group fixated slightly left of the word center.

The means of word properties of multiple fixation cases did not differ between groups. Compared to the corpus means (PSC), target words for multiple fixations have the following characteristics in both reading groups: They are most often content words, that are long, low in frequency, and low in predictability.

6.6.2. Effects on First-Pass Single Fixation Duration

Adding the individual (subject ID) as the only random effect to the model explained 11% of the variance in the data. Including word ID and sentence ID as additional random effects improved the goodness of fit of the model significantly ($\chi^2(2) = 3338$, $p < 0.001$) and explained additional 16% of the variance. Thus, the random sampling of subjects, words, and sentences contributed 27% of the variance of the data. The final model is listed in Table D.7 in Appendix D.

Including the reader-level fixed effects (condition, trial number, digit-symbol score, vocabulary score, 1st and 2nd PC) into the model resulted again in a significant improvement of goodness of fit ($\chi^2(5) = 50$, $p < 0.001$). Reading condition as a simple term did not have a significant effect on SFD. Subject's vocabulary score or index of processing speed (digit-symbol score) did not affect SFD and were therefore excluded from the model. In both groups, SFD was reduced with increasing trial number, that is, towards the end of the experiment ($b = -2.560 * 10^{-04}$, $SE = 3.737 * 10^{-05}$, $t = -6.85$). This trial effect is illustrated in the top left panel in Figure 6.1. Due to the 56 filler sentences in the 'proofreading' condition, trial numbers ranged from 1 to 200, whereas trial number in the 'frequent' sample were equivalent to the number of sentences of the PSC (n = 144). The first principle component is a significant predictor for SFD (b = 1.946*10^{-02}, SE = 5.413*10^{-03}, t = 3.60) (cf. chapter 7 for further analysis of PC-effects). Adding word-level predictors to the model improved the model fit significantly (χ^2 (29) = 2247 , p < 0.001). Effects of oculomotor control and word-level predictors are presented separately in the next two sections. The focus of the analysis lies on cross-level interactions of reading condition and fixation-level fixed effects.

Effects of Oculomotor Control on Single Fixation Duration

Effects of saccade amplitudes and relative fixation position on single fixation duration are illustrated for both reading groups in Figure 6.1. The strongest predictor in the model for SFD was the size of incoming saccade amplitude (b = 2.752*10^{-02}, SE = 9.490*10^{-04}, t = 29.00). The longer the previous saccade, the longer the fixation duration on word n (see bottom left panel in Figure 6.1). Groups do not differ in the effect sizes of incoming saccade amplitude. Outgoing saccade amplitude correlates positively with SFD (b = 1.825*10^{-02}, SE = 1.461*10^{-03}, t = 12.49), that is, SFD is longer, the larger the next saccade size. This effect is significantly stronger in proofreading compared to the 'frequent' sample (b = 8.197*10^{-03}, SE = 2.076*10^{-03}, t = 3.95), as illustrated in the right bottom panel in Figure 6.1.

The relation between relative fixation position and single fixation duration shows the IOVP-effect with longer durations at fixation locations around the center of the word (linear trend: b = -8.632*10^{-02}, SE = 1.451*10^{-02}, t = -5.95; quadratic trend: b = -3.155*10^{-01}, SE = 4.595*10^{-02}, t = -6.87). The peak of the curve is significantly shifted closer to word center in the 'proofreading' sample compared to the 'frequent' sample (b = 1.244*10^{-01}, SE = 1.755*10^{-02}, t = 7.09), and the curvature is increased as well (b = 1.507*10^{-01}, SE = 5.510*10^{-02}, t = 2.73; cf. top right panel in Fig. 6.1).

Immediacy Effects on Single Fixation Duration

The influence of word frequency, word length, and word predictability of the fixated word on single fixation duration is illustrated in Figure 6.2. Word predictability is negatively correlated with SFD (b = -1.877*10^{-02}, SE = 3.981*10^{-03}, t = -4.72). The word predictability effect is significantly stronger in the 'proofreading' sample than in the 'frequent' sample (b = -1.175*10^{-02}, SE = 4.155*10^{-03}, t = -2.83; cf. bottom

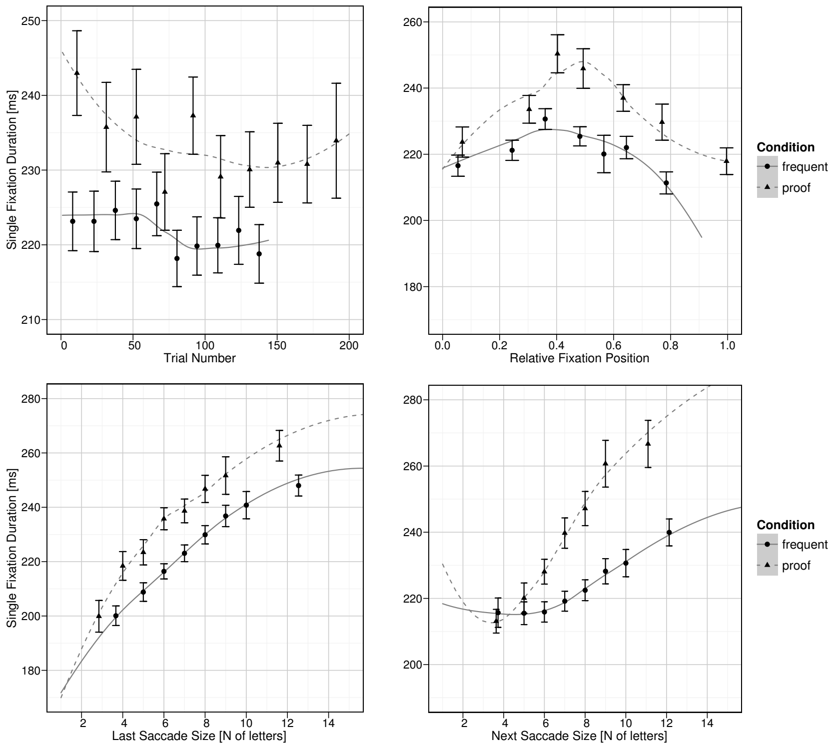


Figure 6.1.: Effects of trial number, relative fixation position, incoming and outgoing saccade amplitude on single fixation duration for the samples 'frequent' and 'proofreading'.

left panel in Fig. 6.2). Since the shape of the frequency curve for SFD shows a cubic trend, especially for the sample 'frequent', word frequency was fitted in the model using its three orthogonal polynomials. The cubic trend of word frequency is a reliable predictor for SFD in the 'frequent' sample (linear: b = -2.753, SE = 1.554, t = -1.77; quadratic: b = 2.030, SE = 1.294, t = 1.57; cubic: b = -3.703, SE = 7.129*10^{-01}, t = -5.19). In the 'proofreading' condition, the linear (b = -4.834, SE = 9.128*10^{-01}, t = -5.30) and the cubic trend (b = -2.434, SE = 6.576*10^{-01}, t = -3.70) of word frequency differ significantly from the 'frequent' sample. These differences in effect sizes mean that the relation of word frequency and SFD is much more linear in the proofreading sample, with decreasing SFD with increasing word frequency (cf. top left panel in Fig. 6.2).

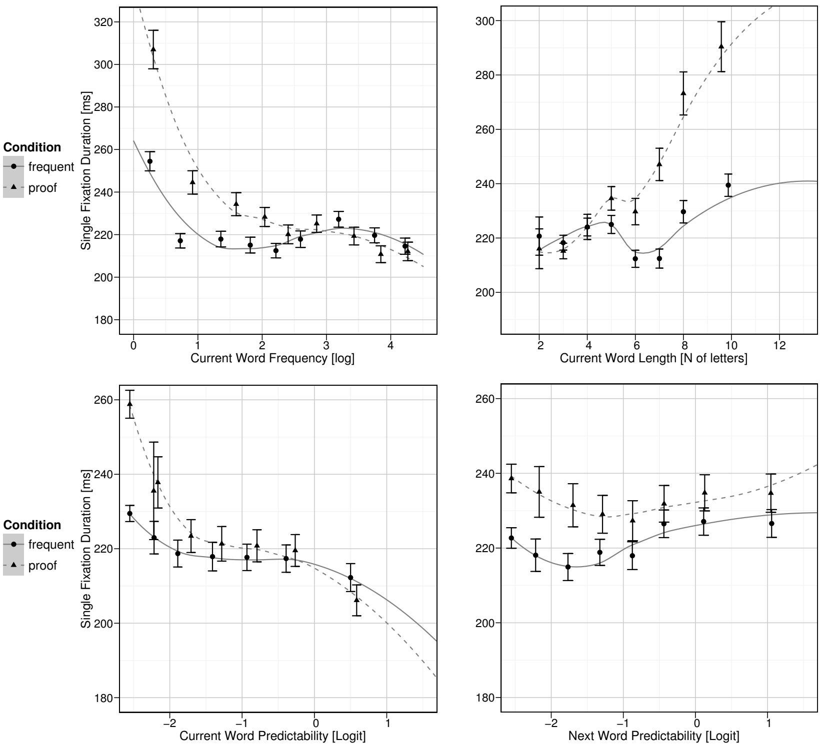


Figure 6.2.: Effects of frequency, length, and predictability of word n, and of predictability of word n+1 on single fixation duration for the samples 'frequent' and 'proofreading'.

The word length effect on SFD is reliably stronger in the 'proofreading' condition than in reading for comprehension (b = -2.580*10^{-01}, SE = 6.011*10^{-02}, t = -4.29, cf. top right panel in Fig. 6.2), but the word length effect as well as the word frequency effect are qualified by a higher order interaction of word length and frequency.

The interaction of word length and frequency on SFD is illustrated in Figure 6.3. Whereas comprehensive readers ('frequent' sample) only show a frequency effect on long words (b = 2.224*10^{-01}, SE = 6.410*10^{-02}, t = 3.47), the frequency effect in 'proofreading' is also significant for short words (b = 2.413*10^{-01}, SE = 4.386*10^{-02}, t = 5.50).

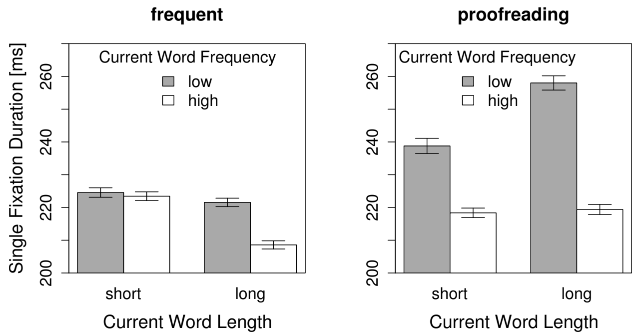


Figure 6.3.: Interaction of length and frequency of the fixated word on single fixation duration for the samples ‘frequent’ and ‘proofreading’; categories of length and frequency are created by a median split.

Lag and Successor Effects on Single Fixation Duration

Lag word frequency was negatively correlated with SFD (b = -3.917*10^{-02}, SE = 4.021*10^{-03}, t = -9.74), that is, single fixation duration on word n was shorter, the more frequent word n-1 was (see left panel in Figure 6.4). This effect was significantly weaker in the ‘proofreading’ sample than in the ‘frequent’ sample (b = 1.289*10^{-02}, SE = 2.942*10^{-03}, t = 4.38). The successor word frequency effect showed the same pattern with decreasing SFD with increasing frequency of word n+1 (b = -1.591*10^{-02}, SE = 3.734*10^{-03}, t = -4.26). No difference in the successor word frequency effect could be found between reading conditions (cf. right panel in Figure 6.4). Influences of the upcoming word differed between conditions with respect to word predictability (cf. bottom right panel in Fig. 6.2). Whereas in the ‘frequent’ sample predictability of word n+1 is positively correlated with SFD on word n (b = 6.773*10^{-03}, SE = 3.376*10^{-03}, t = 2.01), the effect of successor word predictability is weaker in the ‘proofreading’ sample (b = -6.609*10^{-03}, SE = 3.170*10^{-03}, t = -2.09). Considering the t-values of both conditions (+2, -2), there is effectively no effect of upcoming word predictability for the ‘proofreading’ sample.

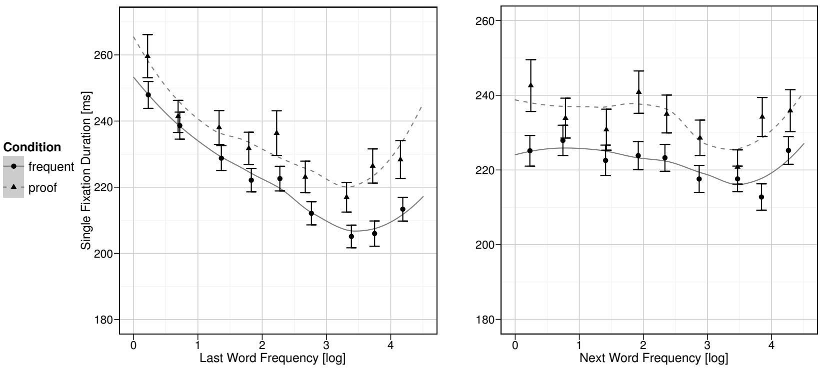


Figure 6.4.: Effects of frequency of word n-1 and word n+1 on single fixation duration for the samples ‘frequent’ and ‘proofreading’.

6.6.3. Effects on First-Pass Gaze Duration

First-pass gaze duration encompasses all cases of two or more fixations on a word. Predictors related to oculomotor control were not included in the model, because means of fixation position and saccade amplitudes of all fixations of the gaze are meaningless in the analysis. The final model fitting first-pass gaze duration can be found in Table D.8 in Appendix D.

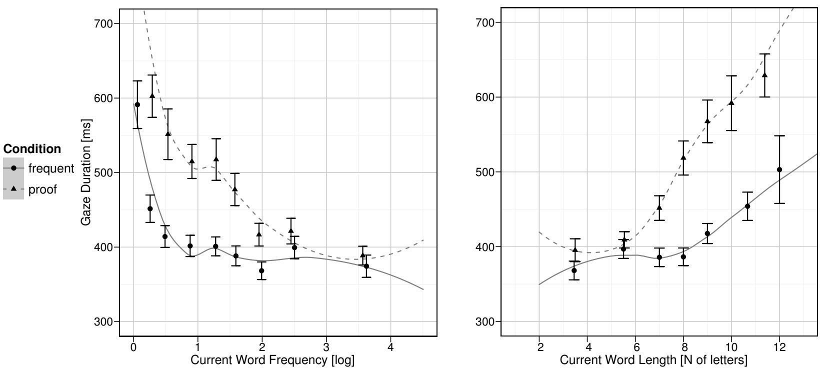


Figure 6.5.: Effects of frequency and length of word n on gaze duration for the samples ‘frequent’ and ‘proofreading’.

The random sampling of subjects, words, and sentences explained 45% of the variance in first-pass gaze duration. Including reader-level fixed effects into the model improved the model fit reliably ($\chi^2(6) = 76$, $p < 0.001$). Significant reader-level predictors for gaze duration were the the first PC (b = 3.081*10^{-02}, SE = 8.991*10^{-03}, t = 3.43) and reading condition. The sample 'frequent' had reliably shorter gaze durations than the sample 'proofreading' (b = 1.239*10^{-01}, SE = 4.035*10^{-02}, t = 3.07). The effect of trial number on gaze duration interacted with reading condition. Proofreaders showed a stronger decrease in gaze duration with increasing trial number than did the sample 'frequent' (b = -5.128*10^{-04}, SE = 1.804*10^{-04}, t = -2.84).

Adding word-level fixed effects to the model significantly improved the model fit ($\chi^2(13) = 522$, $p < 0.001$). In the 'frequent' sample, gaze duration decreased with increasing word frequency (b = -6.085*10^{-02}, SE = 1.186*10^{-02}, t = -5.13; cf. left panel in Figure 6.5) and with decreasing word length (b = -5.512*10^{-01}, SE = 1.866*10^{-01}, t = -2.95; cf. right panel in Figure 6.5). These effects were qualified by a higher order interaction of word length and word frequency which differed significantly between conditions, as illustrated in Figure 6.6. In the sample 'frequent', the frequency effect is much more prominent on long words than on short words (b = 3.253*10^{-01}, SE = 1.010*10^{-01}, t = 3.22). In the 'proofreading' sample, a clear reduction of gaze duration for high frequent words compared to low frequent words can be found on long as well as on short words.

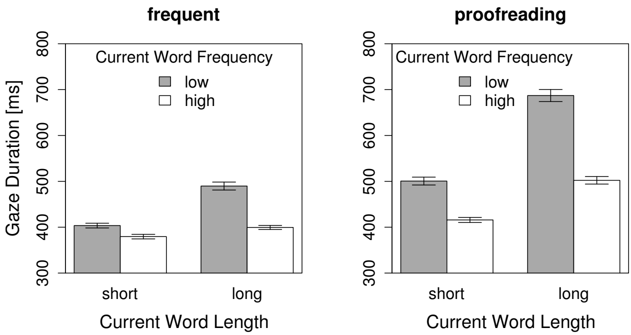


Figure 6.6.: Interaction of length and frequency of the fixated word on gaze duration for the samples 'frequent' and 'proofreading'; categories of length and frequency are created by a median split.

6.7. Eye Movement Behavior on Erroneous Sentences

In this section, eye movement data on the erroneous filler sentences in the proofreading experiment are further examined. Of main interest is the subject's behaviour with respect to the error occurrence. In each of the 56 sentences, one word form error occured, either a letter addition, a letter elision, a letter permutation, or a letter substitution (cf. section 6.2.1).

6.7.1. Single Fixation Durations in First-Pass Reading

The first question of interest is if there are differences in first-pass fixation durations on the correctly spelled words before the error word occured and fixation durations on the erroneous words. In the framework of distributed processing it could be expected that fixation durations before the erroneous word are affected by the erroneous word via parafoveal-on-foveal preprocessing. Single fixations are the largest part of all first-pass fixations in proofreading the filler sentences (cf. Table 6.3) and are therefore analyzed in the following.

As illustrated in Figure 6.7, there is no indication of parafoveal preprocessing of the error word (n+1) on word n in either trials that were correctly answered (see panel A), that is, when the error was detected, or in trials with an incorrect response (panel C). In both cases, single fixation durations do not differ depending on the accuracy of the upcoming word to the right. In contrast, there is a strong foveal error-effect with dramatically increased single fixation durations on erroneous words in comparison to single fixation durations on correct words. This effect is only found in trials that were correctly answered, that is, when the reader reported the detected error (panel B), but not in trials in which the subject missed the error (panel D).

The dependency of single fixation duration on response accuracy can also be observed from a regression-perspective. In Figure 6.8 A, the negative regression effect for correctly answered trials is illustrated. It seems that if the error is detected during the first inspection of the word (and reported), single fixation duration is strongly increased and the error word will not be the target of a later regression (left bar 'no regression goal' in Figure 6.8 A). In contrast, if the misspelling is not reported (see Figure 6.8 B), first-pass single fixation duration is comparable to the fixation time of correctly spelled words (cf. Figure 6.7). Furthermore, no difference in fixation duration can be found depended on the cases if the erroneous word will be the tar-

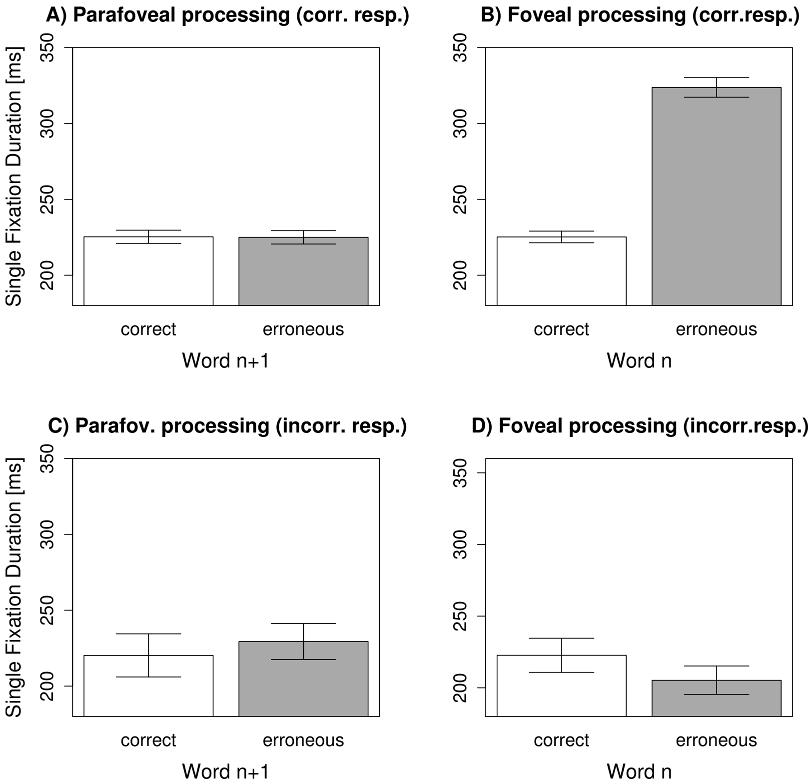


Figure 6.7.: Parafoveal-on-foveal and foveal typo-effect on single fixation duration; A) single fixation durations on word n are plotted in dependency of the status of the upcoming word n+1 (correct or erroneous) for correctly answered trials; B) single fixation durations are plotted in dependency of the status of the fixated word n (correct or erroneous) for correctly answered trials; C) single fixation durations on word n are plotted in dependency of the status of the upcoming word n+1 (correct or erroneous) for incorrectly answered trials; D) single fixation durations are plotted in dependency of the status of the fixated word n (correct or erroneous) for incorrectly answered trials.

get of a regressive eye movement later on or if it is not a regression goal (cf. Figure 6.8 B). In correctly answered trials, the error detection is also evident in the slight increase in SFD if the erroneous word is the target of a regression (cf. right panel in Figure 6.8 A) compared to the mean first-pass SFD of regression targets in incorrectly answered

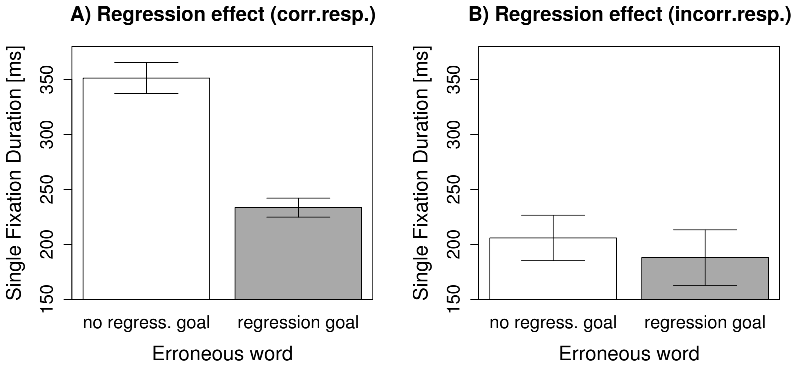


Figure 6.8.: Regression effect on first-pass single fixation duration on erroneous words for A) correctly answered trials and B) for incorrectly answered trials.

trials (cf. right panel in Figure 6.8 B). It can not be answered if this slight increase in fixation duration already indicates the detection of the error and maybe the programmed next saccade could simply not be cancelled, or if the error is detected during the second inspection. At least it can be summarized that a slightly increased single fixation duration on the error word during the first inspection *plus* a regression to that word seems to represent an alternative procedure of error detection, next to the procedure of immediate detection of the spelling error during the first reading pass. Therefore, a further examination of the regressive behavior in reading incorrect sentences is done to illuminate the eye-mind span of error detection in proofreading.

6.7.2. Regressions in Erroneous Sentences

The number of regressions originating in first-pass reading is about 4% higher in proofreading an erroneous sentence compared to proofreading a correct sentence (cf. line b and d in Table 6.3). It can be reasoned that the occurence of an error causes the increase in number of regressions in erroneous sentences. Considering the 11% of regressions originating in first-pass reading in erroneous sentences, two questions arise directly. 1. When do subjects notice the error, or in other words, how long is the eye-mind-span in these cases? 2. Where do regressions target at?

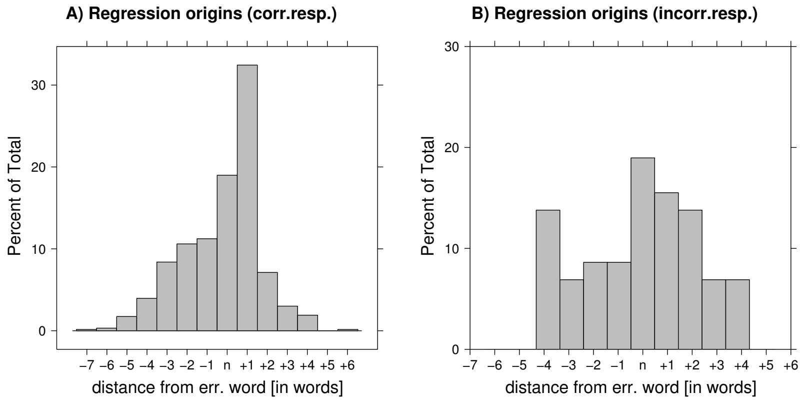


Figure 6.9.: Distances from the erroneous word of regression origins in first-pass reading; A) percentage of distances of regression origins [in words] centered around the error word [n] for correctly answered trials; B) percentage of distances of regression origins [in words] centered around the error word [n] for incorrectly answered trials; 'n' marks the location of the erroneous word in the sentence.

In Figure 6.9 A), the distribution of first-pass regression origins (n = 632) relative to the erroneous word in correctly answered trials is visualized. The erroneous word is considered as the reference word (position 'n'). Most of the regressions, about 33% of all regressions in correctly answered trials, start at word n+1, and about 12% of first-pass regressions start from word n+2 or later. The eye-mind lag of one word corresponds to a lag of 297 ms. Interestingly, the second largest part of regressions, namely 19%, originate directly on the erroneous word. In trials in which the spelling error was not detected by the reader, the systematics of the regressive behavior is less clear (cf. Figure 6.9 B). Since there were only 58 regression origins identified in incorrectly answered trials, the result is not further evaluated.

For correctly answered trials, the distribution of regression goals for all regressions, made after the erroneous word n has been passed, is illustrated in in Figure 6.10 A. Almost all regressive eye movements (84 %) go directly back to the target word, only 13% of regressions undershoot the target word, and ca. 2% of regressions overshoot the erroneous word. Considering the 19% of regressions that start on the error word (cf. Figure 6.9 A), 87% of those regressions go back to the word directly left to the erroneous word (see Figure 6.10 B).

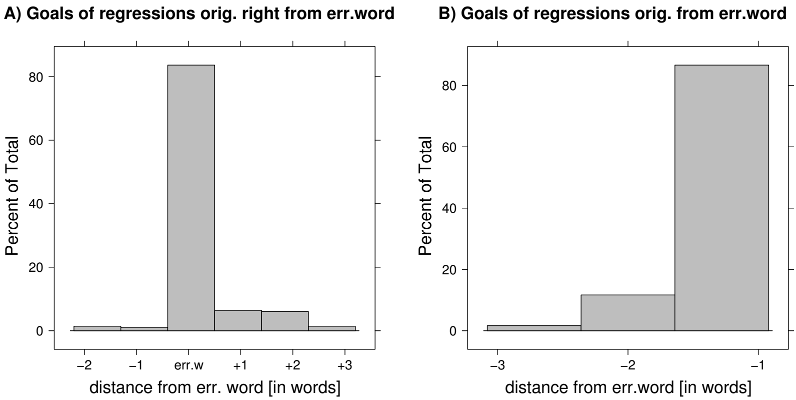


Figure 6.10.: Regression goals in correctly answered erroneous trials; A) distances of regression goals from the erroneous word for regressions starting to the right of the error word [in words] (= distances 1-4 in in Figure 6.9 A); B) distances of regression goals from the erroneous word for regressions originating from the error word [in words].

6.8. Summary of Results and Discussion

In the study presented here, the influence of task instruction on the eye movement behavior during reading isolated sentences was investigated. The 'frequent' group was instructed to read for comprehension and received an easy comprehension question after each trial. The age-matched 'proofreading' sample was instructed to proofread the identical sentence material and was asked for word spelling accuracy after each sentence. The experimental manipulation had a clear effect on the reading behavior, reflected in differences in reading speed, in fixational selectivity, as well as in word-level and oculomotor effects on fixation duration. In comparison to reading for comprehension, in the proofreading mode subjects produced fewer first-pass single fixation cases but many more multiple fixation cases in first-pass reading as well as in second-pass reading. The increase in number of multiple fixation cases is associated with longer gaze durations; the increase in number of second-pass fixations with longer total word reading times. Overall reading speed was much slower in proofreading than in reading for comprehension. Word skipping probability was reduced in the 'proofreading' sample and the proportion of regressions was increased. This eye movement pattern clearly indicates a more serial and 'careful' reading strategy during proofreading that seeks to fixate all words in a sentence.

6.8.1. Word Selectivity and Precision of Word Targeting

Associated with the reduced first-pass skipping rate the selectivity of fixated words was significantly affected in the proofreading condition. Regarding single fixation cases, proofreaders fixated more function words than comprehensive readers. Consistently, fixated words in the proofreading condition were shorter, higher in frequency, and higher in predictability than in the 'frequent' sample. Further in line with a lower word skipping proportion, saccade amplitudes of proofreaders were found to be shorter than those of the comprehensive readers, indicating the 'careful' reading mode in proofreading as proposed by the careful word-by-word reading strategy by O'Regan (O'Regan & Lévy-Schoen, 1987; O'Regan, 1990). The seriality or word-by-word fashion in proofreading is further evident in the CW/FW-ratio of 0.37 which is identical to the Corpus (PSC) mean. In addition to reduced saccade lengths, mean fixation position of single fixations differed between conditions. As expected, proofreaders fixated words exactly at the word center, the optimal viewing position (OVP), whereas readers during reading for comprehension fixated words slightly left of the word center, at the preferred viewing position (PVP) in continuous reading. In reading for comprehension, the major goal is to grasp the words' and sentences' meaning that can best be achieved by fixating the most informative part of the words which is often the first half of a given word (e.g., Hyönä, Niemi, & Underwood, 1989; Pynte, Kennedy, & Murray, 1991). Because word identification (and a comparison to the internal visual word representation) is the primary goal in the proofreading mode and word recognition is optimal in foveal region (O'Regan & Lévy-Schoen, 1987), it was hypothesized that in proofreading saccades should target the word center more often. This claim was clearly supported by the results.

Further differences between conditions were found in the effect sizes of oculomotor variables on single fixation duration. Since proofreaders' mean fixation location was at the word center, the IOVP-curve is shifted to the right in comparison to the 'frequent' sample. That is, single fixation duration in proofreading was maximal at the word's central relative fixation position. In linear mixed modeling, the influence of the outgoing saccade amplitude on single fixation duration was strongly increased in the proofreading condition. This effect can also be attributed to the supposed attempt during proofreading to target saccades precisely at the OVP. It is known that accuracy in saccade targeting increases with longer preceding fixation durations (O'Regan & Lévy-Schoen, 1987). Thus, when longer saccades were executed in proofreading, even longer saccade latencies or fixation durations were needed to achieve a high targeting accuracy. In sum, oculomotor control processes in proofreaders are characterized by shorter saccades and fewer word skippings than in reading for comprehension, and by fixations locations at the optimal viewing position. The hypothesis that proofreading is a more precise reading strategy with respect to targeting the OVP is supported.

6.8.2. The Extent of Distributed Processing in Proofreading

In the framework of the distributed processing account it was expected to find less evidence for preprocessing effects in proofreading than in reading for comprehension. The rationale behind this hypothesis was the assumption that a reading strategy that focuses on the accuracy of word forms will reveal local, immediate effects of word variables, but preprocessing based on the sentence context will be minimized. In line with this assumption, proofreaders fixated more words in first-pass reading than comprehensive readers, entailing that more words are in the foveal region of highest visual acuity. Thus, proofreading is a more serial reading strategy than reading for comprehension. As mentioned above, the reduced skipping rate in proofreading resulted in a different composition of the set of words that received single fixations. Single fixation cases in proofreading were on average more often function words, more frequent, more predictable, and shorter than those fixated by the 'frequent' sample. This is important for the evaluation of effects of distributed processing.

The claim of reduced preprocessing in proofreading, especially of the sentence content, was supported by differences in successor word predictability effects compared to reading for comprehension. The 'frequent' sample showed increasing single fixation durations with increasing predictability of word n+1, interpreted as an effect of memory retrieval during the fixation of word n (Kliegl et al., 2006). In contrast, successor word predictability was found to have almost no effect on fixation duration in the proofreading condition, indicating the absence of preprocessing based on the sentence context. Furthermore, no differences were found in the successor word frequency effect between conditions that could have been expected according to the perceptual span and foveal load hypothesis (Henderson & Ferreira, 1990;

Kennedy & Pynte, 2005; Kliegl et al., 2006). Even though subjects in the proofreading condition fixated many more function words, such that the eyes were located on easy words, the word frequency effect of the upcoming word n+1 on single fixation duration was not increased, contrasting the results in the sample 'hard old' (see chapter 4). Despite the fact that the fixation was located on a short and high frequent word, proofreaders appear to hold attention on the currently fixated word, contradicting the foveal load and perceptual span hypothesis claimed for reading for comprehension and demonstrated in Chapter 4.

The observation of an intensified focus on the fixated word in the proofreading condition is further supported by the strong linear trend of the negative correlation of word frequency and single fixation duration. The 'frequent' sample shows the expected cubic trend in the word frequency curve, with decreasing fixation duration with increasing frequency, and increasing fixation duration on very high frequent word, previously interpreted as an effect of preprocessing word n+1 (see section 4.7.1). In contrast, the 'proofreading' sample shows a much more linear relation of decreasing single fixation duration with increasing word frequency, indicating a pure immediacy effect of word frequency. Generally, the word frequency effect on single fixation duration in proofreading was stronger than in reading for comprehension. The illustration of the frequency effect in proofreading suggests that especially for low frequency words more time was needed for the instructed accuracy check.

Word predictability of the fixated word was also found to have a stronger impact on single fixation duration in proofreading than in reading for comprehension. Again, this is in line with the interpretation that on low predictable (as on low frequent) words a prolonged fixation duration is needed to perform a detailed spelling check. Since word frequency and word predictability correlate with word length, it is not surprising that the immediacy word length effect in proofreading is also found to be more pronounced than in reading for comprehension. In fact, the very linear length effect of increasing SFD with increasing number of letters in the proofreading condition almost resembles a letter-by-letter reading strategy of words, though the magnitude of the length effect is not comparable with a effect size in a letter-by-letter reader in acquired dyslexia (e.g., Arguin & Bub, 1993; Patterson & Kay, 1982; Warrington & Shallice, 1980). It is reasonable to expect that in the proofreading condition of the present experiment, in which errors occurred at the letter-level, the evaluation of the word's orthography was performed more in a letter-by-letter fashion than in a holistic mode. But the immediacy word length effects in both reading conditions are qualified by an interaction with word length that needs appropriate consideration.

The interaction of word length and word frequency of the currently fixated word on SFD is more pronounced in the 'proofreading' sample than in the 'frequent' sample. Whereas proofreaders show a sensitivity to word frequency on short as well as on long words, this correlation can only be found on long words in the 'frequent' sample. This can be interpreted as an effect of reading instruction. Proofreaders focus more on the fixated words than comprehensive readers, as supported by stronger word frequency and word length effects, as well as their interaction. During proofreading, low frequent words, even when they are short, make higher demands on the process of spelling check, than high frequent, and therefore easily accessible words. In reading for comprehension, short words are fixated for an equally long period of time independent of their frequency. Additionally, the frequency range in the proofreading sample is presumably larger than in the 'frequent' sample due to the fixational selectivity described above. Interestingly, in the 'frequent' sample, the mean single fixation duration of the category of high frequent long words is shorter than the mean fixation duration on short words (see Fig. 6.3). The correlation of log word frequency and word length of the words in the PSC is -0.66, but there seems to be an exception for words of about six and seven letter length. These words are proportionally high frequent in the PSC and thus, single fixation durations on these words are relatively short and appear to disrupt the linear correlation of word length and fixation duration. This effect can clearly be observed in the length effect for the sample 'frequent' (see Fig. 6.2). Therefore, the somewhat surprising shorter fixation duration on long than on short words of this frequency category can be attributed to the unique composition of words in the PSC.

The interpretation of the pronounced immediacy effects due to a higher attention focus on the fixated word is further supported by the weaker lag word frequency effect in proofreading compared to data in reading for comprehension. The smaller spillover effect of word frequency on single fixation duration in proofreading indicates that word processing is more often finished when the eyes left word n-1, confirming the idea of a very local and serial processing strategy in proofreading.

Statistical modeling on gaze duration reveals similar results as summarized for single fixation duration. Again, the interaction of length and frequency on first-pass gaze duration is more pronounced in proofreading than in reading for comprehension. Note that all differences between reading conditions found in word-level and oculomotor effects on single fixation duration are qualitative and that the intercepts between conditions in the LMM did not differ. In sum, little evidence for distributed processing during fixations could be found in the proofreading condition. Instead, the local focus on fixated words dominates in the proofreading mode. Based on the absence of a successor word predictability effect, proofreading appears to be a more mindless reading strategy with respect to the sentence content.

6.8.3. Strategies Related to the Spelling Error

During proofreading the erroneous sentences, subjects detected the spelling error in 90% of the trials. Letter elisions were less likely perceived than the other three error types of letter substitution, letter addition, or letter permutation. Unlike previous research results (Haber & Schindler, 1981; Levy, 1983), the proportion of detected spelling errors in function words was not significantly smaller than the proportion of error detection in content words (though there was a tendency towards this finding). The point in time of the error detection can clearly be observed in the eye movement behavior. The detection of an error in first-pass reading was characterized by a large increase in fixation duration, presumably related to the inhibition of a subsequent saccade. An increase in fixation duration has been found when the eyes fall on a misspelled word in a eye-contingent display paradigm (Underwood & McConkie, 1985). The authors reported an increase of 9 ms in mean fixation durations on words with a letter replacement compared to the fixation duration on correct words. But since in that study subjects read for comprehension, it is plausible that in proofreading, where the detection of a letter error was part of the reading intention, fixation durations are much more prolonged (about 80 ms) when fixating an erroneous word.

There was no evidence for parafoveal-on fovea error processing. This is in line with previous findings, that the region in which letter distinctions can be made is very small and that letter replacements on the words to the right of fixation do not influence the current fixation duration (Underwood & McConkie, 1985). If errors were not detected during the first pass, the eye-mind lag was usually one word, that is, most regressive eye movements on the erroneous sentences originated on word n+1 with respect to the target error word. 80% of all regressions originating after the erroneous word targeted at the error word, indicating that a subsequent word form processing after leaving the erroneous word with the eyes triggered the precise regression. Overall, the error detection is usually performed when the word is located in the foveal region, supporting the notion of a very locally focussed reading strategy in the proofreading mode claimed above. In a few cases the error detection occurred with a delay of one or two fixations.

Proofreading was defined as a subskill of the general reading skill. One could argue that proofreading in the tested college students is indeed not a well trained task, especially in times where error correction is automatically applied when using software with authoring programs. An argument against this constraint is that the proofreading task was more like a detection task and all errors lead to a creation of nonwords. Therefore, the detection task could in fact be accomplished on a word-nonword discrimination without the need to know the correct spelling. It remains to future research to compare the current results with the performance of professional proofreaders or to investigate eye movements during proofreading with errors at the syntactic or semantic level.

Generally, it can be discovered that in the proofreading mode the proposed reading strategy of subjects focussing on the orthography of each single word in the sentence turned into a reading strategy of 'finding the error'. This is evident in the very high first-pass skipping rates in proofreading erroneous sentences compared to the skipping rate in proofreading correctly spelled sentences or to those in reading for comprehension. It seems that proofreaders stopped reading the sentence as soon as a misspelling was detected, in other words, as soon as the reading goal was achieved. In addition, the very large proportion of second-pass and rereading fixations in eye movement behavior on correctly spelled sentences can be interpreted as a search for the error. Subjects presumably noticed after a few trials that maximally one error occurred per sentences, and thus, they tended to double-check the words when no error was found by the end of the first pass. In sum, the eye movement behavior on erroneous sentence material provides further evidence, that proofreading is a serial reading strategy that focus on the immediate processing of the fixated words.

7. Strategy Effects and Individual Differences

In the previous chapters, the influence of reading strategy on local eye movement behavior in sentence reading was investigated and clear evidence for specific strategy effects has been demonstrated. As theoretically motivated in chapter 2 and methodologically explained in section 3.4, effects of reading strategy were tested under the assumption of the co-occurence of individual differences between readers. In other words, it was assumed and demonstrated in the previous chapters that strategy effects in response to a specific reading goal exist *in addition* to variance between individual readers, so-called reader types (cf. section 2.4.2). It is reasonable to assume that even if a reading intention, such as reading for comprehension or proofreading systematically modulates the reading strategy applied, there are still differences between individual readers in how they achieve the same reading goal. Individual differences needed to be considered in this work because different subjects read under various task demands and no subject participated in more than one of the experiments. The between-subjects design was chosen to avoid prior knowledge of the sentences of the PSC and rereading effects (e.g., Raney & Rayner, 1995) in any of the reading conditions tested.

In the following, the differences in eye movement behavior between readers in the present data sets are further inspected and the attempt is made to identify different reading styles between individual subjects. To this end, the results of the principle component (PC) analyses and the role of reader-level fixed effects, namely vocabulary score (Lehrl, 1977), digit-symbol-score (Wechsler, 1964) and the first and second PC, in the various LMMs of chapters 4, 5, and 6 fitting first-pass single fixation duration are investigated.

7.1. Psychometric Variables

An index for vocabulary size was administered for all subjects using form B of Lehrl's multiple choice vocabulary test (MWT-B; Lehrl, 1977). An index for processing speed was administered with the Digit-

Symbol-Test of the *HAWIE* (Wechsler, 1964). Scores for all subjects are listed in Appendix C. The age-matched experimental groups did not differ in vocabulary size or digit-symbol score, except for the samples 'original' and 'frequent' (see section 5.1).

The influence of vocabulary size and processing speed on fixation duration was tested by including both scores as fixed effects in the lmer-models. In the majority of cases, none of the predictors reached the level of significance and were therefore excluded from the models. Only in in the comparison of the samples 'original young' and 'hard young', the influence of vocabulary size was a reliable predictor for single fixation duration and for gaze duration. Fixation duration was shorter, the higher the vocabulary score of the individual (SFD: $b = -1.273*10^{-02}$, $SE = 5.893*10^{-03}$, $t = -2.16$; Gaze: $b = -0.0124083$, $SE = 0.0060463$, $t = -2.05$). This is in line with the finding that vocabulary knowledge is related with the ease of word identification and word processing in reading, especially during the acquisition of reading (see Perfetti, 1994; Stanovich, 1990, for a review). The samples 'original young' and 'hard young' were the youngest samples tested and it might be the case that in highschool students the vocabulary size still plays a greater role during the process of reading than in the older samples of college students or in elderly adults. In sum, the index of vocabulary size and of processing speed do not impact on online-measures in reading such as single fixation duration and gaze duration.

7.2. Principle Component Analysis of Reader-level Variables

To account for individual differences at the reader-level in the LMM without increasing model complexity, principle components (Jolliffe, 2002) were computed based on subjects' mean incoming saccade amplitude, mean outgoing saccade amplitude, mean skipping probability of the previous and the next word, mean log word frequency, mean word length (reciprocal value), and mean content/function word ratio of the fixated word n and the word n-1 to the left of fixation. In all group comparisons, the 1st plus the 2nd PC covered between 92 and 95% of the variance of the data and were therefore selected as fixed effects for the LMM.

7.2.1. Interpreting the 1st Principle Component

The loadings of the 1st principle component for all group comparisons are listed in Table 7.1. It is obvious that all of the ten selected variables load equally strong on the first component, though the loadings of the variables of word n-1 are somewhat smaller and the lag variables load strongly on the second PC (see section 7.2.2). Considering the first three columns in Table 7.1, the oculomotor variables (saccade amplitude and skipping rate) and the lexical variables of word n-1 (frequency, length, and CW/FW-ratio) behave similarly, because they are all negatively correlated with the component. At the same time, they behave differently from the the lexical variables of the fixated word n (frequency, length, and CW/FW-ratio) that are positively correlated with the first PC. In the comparison of the samples 'frequent' and 'proofreading', the signs of the loadings are reversed. In a post-hoc analysis, a PCA was computed separately for the sample 'frequent' and 'proofreading' and results revealed that the signs of the loadings in both groups were identical to the first three group comparisons. Thus, the signs of the weights seem to be arbitrary in the last group comparison. Therefore, negative loadings for oculomotor variables and lexical variables of word n-1, and positive loadings for lexical variables of word n are interpreted. To keep the metric in lmer-modeling constant across group comparisons, the sign of the scores of the first PC in the comparison 'frequent' - 'proofreading' were reversed.

The following relations can be observed: A high 1st PC score is correlated with short amplitudes and a low skipping rate. At the same time, a high 1st PC score is correlated with higher frequent and shorter words n, and a higher proportion of fixated function words. As discussed above, the similar behavior of the variables frequency, length, and CW/FW-ratio of word n can be attributed to the collinearity of those variables, due to the finding that functors are usually short and high frequent words. Hence, subjects with a low PC1 score have a long mean saccade amplitude, have a higher skipping rate, and fixate words that are lower in frequency, longer in length, and that are more often content words. Words to the left are high frequent and short in length for those subjects. Presumably, long saccade amplitudes (and a high skipping rate) co-occurr with lower mean frequency of fixated words because short and high frequent function words are frequently skipped, as possibly indicated by the higher frequency of words n-1. It seems that the influence of the first PC on reader types is best summa-

Table 7.1.: Loadings of the 1st principle component of oculomotor and lexical processing variables for group comparisons 'original young' - 'hard young', 'original old' - 'hard old', 'original'- 'frequent', and 'frequent' - 'proofreading'.

VARIABLES	GROUP COMPARISON			
	ORIG./ HARD Y.	ORIG. HARD O.	ORIG./ FREQU.	FREQU./ PROOF.
Outgoing sacc. ampl.	-0.351	-0.360	-0.335	0.343
Incoming sacc. ampl.	-0.333	-0.317	-0.329	0.339
Skipping word n-1	-0.327	-0.316	-0.334	0.334
Skipping word n+1	-0.344	-0.354	-0.341	0.340
Log frequency word n	0.346	0.360	0.344	-0.343
1/ Length word n	0.346	0.359	0.340	-0.339
CW/FW-ratio word n	0.326	0.346	0.333	-0.342
Log frequency word n-1	-0.248	-0.234	-0.257	0.234
1/ Length word n-1	-0.202	-0.210	-0.234	0.198
CW/FW-ratio word n-1	-0.305	-0.359	-0.295	0.309

rized as the extent of *distributed processing* or the size of the perceptual span, a subject applies during reading. Low scores in the first PC indicate a high amount of distributed processing. If words are processed in a more parallel fashion, more, especially high frequent and easily processable words can be skipped and thus, saccade amplitudes are on average longer. The composition of the words selected for fixation are therefore lower in frequency. In sum, the first principle component can be interpreted as an index of distributed processing.

In the LMMs of chapters 4, 5, and 6, the scores of 1st PC were included as reader-level variables to account for individual differences in addition to effects of experimentally induced reading strategies. The 1st PC impacts positively on single fixation duration in all group comparisons, as illustrated in Figure 7.1. In all models fitting first-pass single fixation duration, expect for the group comparison of 'original old' and 'hard old', the influence of the 1st PC reached the level of significance ('original young' - 'hard young': $b = 1.737*10^{-02}$, $SE = 5.815*10^{-03}$, $t = 2.99$; 'original old' - 'hard old': $b = 1.157*10^{-03}$, $SE = 7.808*10^{-03}$, $t = 0.15$; 'original'- 'frequent': $b = 1.489*10^{-02}$, $SE = 5.256*10^{-03}$, $t = 2.83$; 'frequent' - 'proofreading': $b = 1.946*10^{-02}$, $SE = 5.413*10^{-03}$, $t = 3.60$; cf. Appendix D). A low score of the 1st PC predicts shorter single fixation durations. Thus, readers with a larger extent of distributed processing (indicated by a low 1st PC) are faster readers, as directly linked to long saccade amplitudes, a higher skipping rate, and fixating fewer high frequent and short function words.

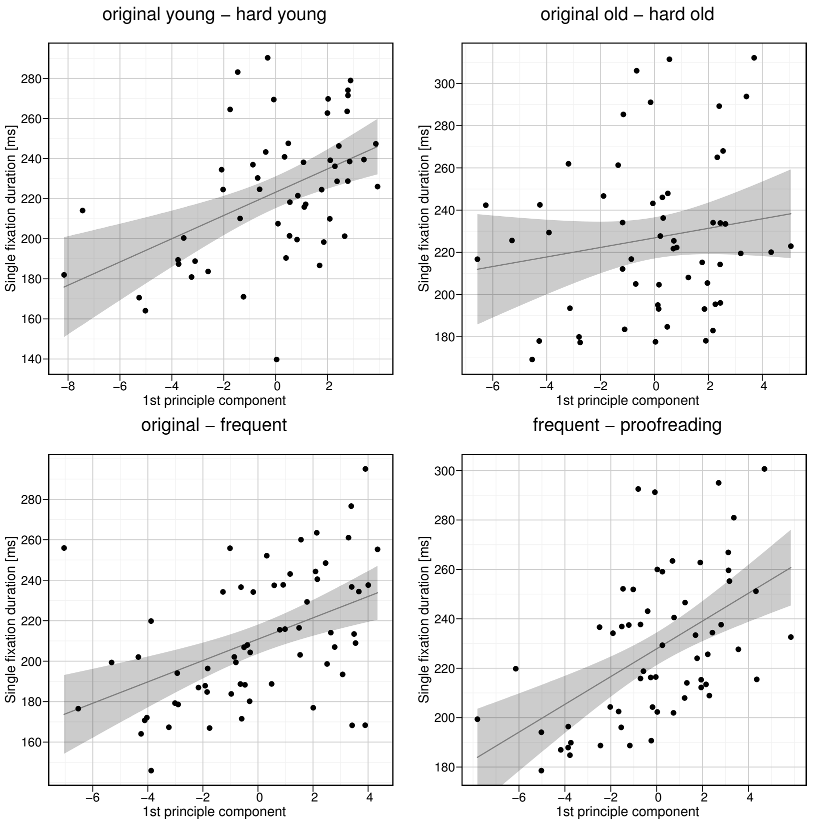


Figure 7.1.: Effects of the 1st PC on SFD for group comparisons ‘original young’ - ‘hard young’, ‘original old’ - ‘hard old’, ‘original’- ‘frequent’, and ‘frequent’ - ‘proofreading’.

7.2.2. Interpreting the 2nd Principle Component

The loadings of the 2nd principle component are listed in Table 7.2. Across all samples, the lag word properties of frequency, length, and content/function word ratio load the highest on the 2nd component.

Table 7.2.: Loadings of the 2nd principle component of oculomotor and lexical processing variables for group comparisons 'original young' - 'hard young', 'original old' - 'hard old', 'original'- 'frequent', and 'frequent' - 'proofreading'.

VARIABLES	GROUP COMPARISON			
	ORIG./ HARD Y.	ORIG./ HARD O.	ORIG./ FREQU.	FREQU./ PROOF.
Outgoing sacc. ampl.	-0.139	-0.131	-0.216	-0.170
Incoming sacc. ampl.	-0.216	-0.304	-0.244	-0.204
Skipping word n-1	-0.166	-0.257	-0.151	-0.116
Skipping word n+1	-0.177	-0.177	-0.153	-0.131
Log frequency word n	0.119	-	0.123	0.131
1/ Length word n	0.128	-	0.158	0.193
CW/FW-ratio word n	0.164	-	0.128	0.121
Log frequency word n-1	0.535	0.498	0.535	0.569
1/ Length word n-1	0.621	0.545	0.583	0.642
CW/FW-ratio word n-1	0.382	0.484	0.406	0.342

Lag word properties are positively correlated with the 2nd principle component. In other words, a high 2nd PC is associated with words n-1 of higher frequency, shorter length, and a higher CW/FW ratio. Frequency, length, and lexical status of word n-1 are again correlated, hence, in the following word frequency is used to capture the property of words n-1. Thus, next to the first dimension of the extent of distributed processing (1st PC), the *selectivity of word n-1* best determines the 2nd dimension of the variance in the data.

Including the scores of the 2nd PC as a reader-level factor in the LMMs of chapters 4, 5, and 6 resulted in reliable pc2-effects on single fixation duration (cf. Appendix D). In all model-fits, though not reliable in the comparison of the samples 'frequent' - 'proofreading', the 2nd PC impacts negatively on single fixation duration, as illustrated in Figure 7.2 ('original young' - 'hard young': b = -3.109*10^{-02}, SE = 1.238*10^{-02}, t = -2.51; 'original old' - 'hard old': b = -3.334*10^{-02}, SE = 1.301*10^{-02}, t = -2.56; 'original'- 'frequent': b = -4.283*10^{-02}, SE = 1.180*10^{-02}, t = -3.63; 'frequent' - 'proofreading': b = -1.721*10^{-02}, SE = 1.075*10^{-02}, t = -1.60). A subject's low score of the 2nd PC is associated with longer single fixation durations. Slow readers have a low 2nd PC score, linked with a composition of words n-1 of larger

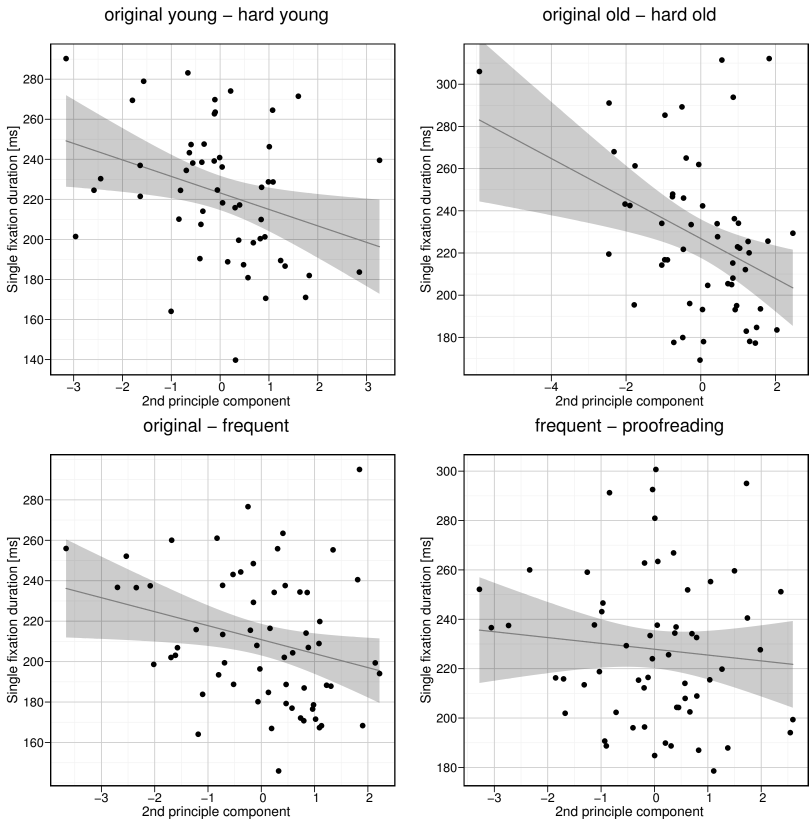


Figure 7.2.: Effects of the 2nd PC on SFD for group comparisons 'original young' - 'hard young', 'original old' - 'hard old', 'original'- 'frequent', and 'frequent' - 'proofreading'.

mean length and lower mean frequency, that are more often content words. This can be interpreted in two ways: First, as a spill-over effect, e.g. low frequent words n-1 entail longer fixation durations on word n due to unfinished processing of word n-1 (Rayner & Duffy, 1986). Second, as a preprocessing effect within the perceptual span, e.g. during fixating easy, that is, high frequent words n-1, word n is preprocessed to a certain extent, entailing reduced mean fixation durations on word n (Henderson & Ferreira, 1990; Schroyens et al., 1999). In sum, independent from the subject's amount of distributed processing the selectivity of words n-1 explains additional variance between reader types. The individual composition of words n-1 has a second major impact on single fixation duration.

7.3. Discussion

Reader-level fixed effects were included in the lmer-models to account for effects of individual differences. Vocabulary size and an index of general processing speed were no reliable predictors for fixation duration. A principle component analysis was computed over the means of several oculomotor and lexical variables per subject to account for individual differences, especially for those in fixational behavior. The detailed analysis of the results of the PCA revealed two aspects that best describe differences between readers. The 1st principle component indicates the *extent of distributed processing* of the individual. A second main factor in individual differences is the *selectivity of words n-1*.

An illustration of the meaning of the orthogonal relation between the 1st and 2nd principle component is provided in Figure 7.3. Along the horizontal dimension, the 1st PC, the extent of distributed processing is represented: Subjects with a low 1st PC perform more parallel processing, indicated by making on average long saccade amplitudes, having a high word skipping rate, and fixating more infrequent words (quadrants A and C). On the other side, we find individuals with shorter mean saccade amplitude and a lower mean skipping rate, that perform a more serial reading and select more frequent words as saccade targets (quadrants B and D).

This behavior is independent from the second dimension, the 2nd PC, namely the selectivity of words to the left of fixation. The averaged composition of words n-1 range from low frequent or long words (low PC2; quadrants C and D) to more high frequent or short words (high PC2; areas A and B). As demonstrated above in Figure 7.2, a low PC2 score is associated with longer mean single fixation duration. Hence, the selection of low frequent words n-1 may entail a spill-over effect or less preprocessing of word n, such that SFD on word n is increased.

Quadrant-A and -C readers are generally fast readers with a high extent of distributed processing, indicated by long saccade amplitudes, high skipping rates, and shorter mean single fixation durations. But these readers may differ from each other in the selectivity of (presumably) skipped words. Readers in quadrant A skip many words and these words are most often short and high frequent function words.

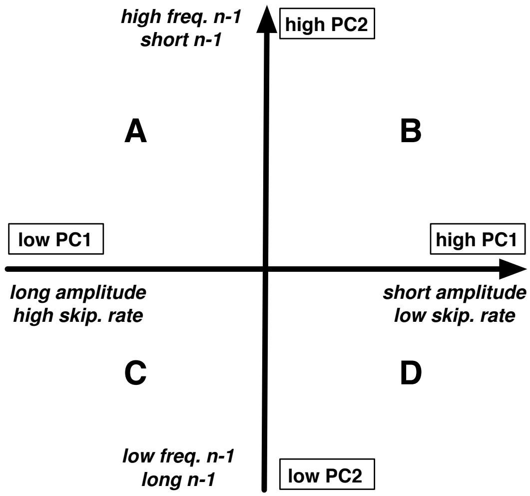


Figure 7.3.: Illustration of the relation between the 1st (PC1) and 2nd principle component (PC2).

These words are easy to process and therefore no spill-over effect can be observed on single fixation duration. Thus, quadrant-A subjects are the fastest readers of the four reader types. They selectively skip high frequent words and fixate lower frequent words. Quadrant-C readers are also generally fast readers, but they presumably skip more often lower frequent and longer content words, entailing that they are slowed down in those cases by a lag effects of those skipped words n-1 (see Haberlandt, Graesser, & Schneider, 1989, for a similar observation in fast readers).

Individuals with generally more serial processing in reading (quadrant-B and -D readers), associated with shorter saccade amplitudes, a lower skipping rate, and longer mean single fixation durations, also differ with respect to their selectivity of words n-1. In quadrant-D subjects, words n-1 are on average lower frequent or longer words that presumably produce lag effects on single fixation duration on word n. Thus, quadrant-D readers are the slowest readers of those four theoretical groups. Quadrant-B readers read sequentially, but words to the left of fixation are more often high frequent and short function words that, due to their ease in processing, usually do not generate spill-over effects.

In sum, the factors that play a main role in characterizing the subject's reading style are the extent of distributed processing and the selectivity of fixated and skipped words. The extent of distributed processing in a reader may be a function of reading experience. Furthermore, the degree of parallel processing is directly related to the size of the perceptual span. Skilled readers with more reading practice have a larger perceptual span, make longer saccade amplitudes, and have a higher skipping rate than poor readers or beginning readers (Rayner, 1986; Everatt & Underwood, 1994; Rayner, 1998). The results further show that readers are differently affected by the lag word properties. Both dimensions demonstrate that the lexical status of the fixated and skipped words play a major role in defining individual differences in reading and that readers are sensitive to the lexical status of the words. Most important for the interpretation of the LMM results in the previous chapters is the fact that the experimentally induced group differences (condition effects) were reliable while individual differences were taken into account. The results of the analysis of the 1st and 2nd PC further unfold the complex dynamics of saccade targeting and the properties of the fixated or skipped words and their impact on fixation duration that have also been observed in the effects of reading strategies on eye movement behavior in reading.

8. General Discussion

The goal of the present work was to investigate the role of reading strategies on the eye movement behavior during sentence reading. Although the development in the field of reading research and eye movement control in reading has been impressive, the variability of skilled reading has not gained much of attention in the last decades of research as compared to the first half of the 20th century (see e.g., Tinker, 1958). Rayner and Pollatsek (1989) argued that results from studies investigating proofreading, skimming, or visual search do not inform us at all about normal reading. They remark that findings from proofreading or visual search are not generalizable to normal reading, because reading comprehension is not the goal. The authors consider the comprehension process involved in reading as the core aspect of reading and this process is obviously not the focus in visual search or proofreading. In contrast to Rayner and Pollatsek (1989), the present work considers proofreading with respect to the associated eye movements as a sub-skill of the general reading skill (cf. Neumann, 1987), as would be oral reading, visual letter search, or reading for comprehension with varying processing depth. Therefore, theories of reading and especially models of reading and oculomotor control must be able to handle data of sub-skills of reading in order to claim generalizability. In this work, a reading strategy is viewed as the result of a parameter specification at the reader's side concerning e.g., saccade timing, saccade targeting, or the size of the perceptual span. Thus, the results presented here need to be taken into account when building a realistic model of eye movement control in reading. A model of eye movement control should be able to account for diverse reading strategies a reader can apply, or for different reading modes a reader can adopt, determined by a given reading goal.

In the present work, reading strategies were directly (via instruction) or indirectly (via task demands) induced using the identical sentence material (the PSC). Reading strategy was differentiated from individual differences or differences in reading styles (see chapter 7). The experimental manipulation of reading strategies by varying task demands (see chapters 4 and 5) and varying task instruction (see chapter 6) clearly affected variables related to the *when*-dimension and *where*-dimension of eye movements as well as indicators of cognitive control and distributed processing during reading. Robust effects across reading conditions were found that are in line with previous research (Rayner, 1998). These robust, mostly word-level effects in all reading conditions provide further evidence that the results reported for the PSC do not depend on having a large amount of data, as argued by Rayner et al. (2007), but they generalize across individual, smaller samples (Kliegl, 2007). Given the definition of reading strategy as a parameter specification in accordance with an identified reading intention, all these effects are causally related to top-down influences.

8.1. Individual Differences in Reading Style

In this work, individual differences in reading styles are considered to be present in addition to a defined reading strategy that is applied in accordance to a specific reading intention, e.g., reading for comprehension. Variance between readers with respect to their eye movement behavior was considered by two procedures: first, subject ID was included as a random factor in the linear mixed modeling; second, reader-level fixed effects were added to the model. These fixed effects included an index of vocabulary size and of processing speed, trial number, and the first and second principle component comprising individual means of oculomotor variables (skipping rates, selectivity of function and content words, and saccade amplitudes) and lexical variables (for example, word frequency or word length) (cf. section 3.4.3 and chapter 7).

Table 8.1 summarizes the proportion of variance explained by the three random effects, namely, subject ID, word ID, and sentence ID, that were considered in the LMM for the different experimental manipulations. For both fits of single fixation duration and gaze duration, the random sampling of subjects and words explained the largest amount of the variance in the fixation data across all experimental comparisons. Variance related to the unique composition of the sentences is much smaller. The variance due to the uniqueness of the words was even larger for fits of gaze than of single fixation duration. This effect can be explained by the finding that refixations are predominantly associated with word processing difficulties, due to e.g., longer word length, lower word frequency, or lower word predictability (e.g., Just & Carpenter, 1980; Rayner & Duffy, 1986; Kliegl et al., 2004). Thus, the unique properties of words play a greater role for the observed (an accumulated) gaze duration than for the observed (shorter) single fixation duration.

Table 8.1.: Summary of unique explained variance of random effects (subject ID, word ID, sentence ID), fixed effects, and total variance explained in LMMs for all group comparisons ('original young' - 'hard young', 'original old' - 'hard old', 'original' - 'frequent', 'frequent' - 'proofreading'); total variance explained applies for the complete, finals models as listed in Appendix D.

GROUP COMPARISON	EXPLAINED VARIANCE (in %)			
	ORIG./ HARD Y.	ORIG./ HARD O.	ORIG./ FREQU.	FREQU./ PROOF.
Single fixation duration				
Subject ID (random)	15	17	17	11
Word ID (random)	14	10	11	15
Sentence ID (random)	1	2	1	1
Fixed effects	5	4	4	4
TOTAL	*35*	*33*	*33*	*31*
Gaze duration				
Subject ID (random)	15	23	18	14
Word ID (random)	32	24	27	31
Sentence ID (random)	.3	1	.5	0.3
Fixed effects	<1	<1	<1	1.5
TOTAL	*48*	*48*	*46*	*47*

The inclusion of reader-level and word-level fixed effects to the LMM explained extra 4-5% of the variance in single fixation duration and ca. 1% of the variance in gaze duration. In comparison to the variance explained by the random sampling of readers and words, the variance attributed to fixed effects of reader- and word-level variables is relatively small. This observation has early been formulated by W.R. Dixon and W.S. Morse (as quoted in Kliegl, 1982) who stated in 1951 that individual differences were much greater than group differences, that is, differences between reading conditions. For the current data, this basically means that the existing differences between readers were not fully captured by the fixed effects such as vocabulary size (MWT-B Lehrl, 1977), the digit-symbol score (HAWIE Wechsler, 1964), or the PCA on subjects' means of lexical and oculomotor variables. The selected reader-level and word-level fixed effects impacted significantly on fixation duration, as demonstrated in sections 4.6, 5.6, and 6.6, though the amount of explained variance is relatively small in comparison to the amount of variance attributed to random effects. Especially at the reader-level the 1st and 2nd PC comprising means of individual oculomotor behavior and selectivity effects reliably influenced fixation duration. As discussed in chapter 7, the reading styles that could be identified by means of the loadings on the 1st and 2nd PC can best be captured by the following two dimensions: the extent of distributed processing and the selectivity of fixated words. Possibly due to differences in the size of the perceptual span, individual readers differ mainly in their mean saccade amplitude, skipping rate, and selectivity of the fixated word, and independently from these aspects they differ with respect to the selectivity of fixated and/ or skipped words to the left of fixation. Irrespective of this variance between subjects, readers are able to adapt their general reading style to the task demands, as demonstrated in chapters 4, 5, and 6. The variability within subjects in this work is represented by the different reading strategies observed in response to task demands and reading instructions. Differences in reading strategies can best be described along the dimensions of saccade timing, saccade targeting, and effects of distributed processing.

8.2. The Timing Aspect in Reading Strategies

According to cognitive or linguistic control theories in eye movement research in reading, saccade latencies (fixation durations) are prolonged due to e.g., a failure in word recognition or the processing of a low frequent or unusual stimulus. In most studies, variables related to the 'lower linguistic level' of processing, namely those related to word identification, are investigated and referred to as indicators of cognitive control (Liversedge & Findlay, 2000), as opposed to variables related to higher linguistic processes such as syntactic parsing or sentence integration. Variables related to the word level indicating cognitive influences are found to only delay the execution of saccades, not to trigger them (O'Regan, 1990; Yang & McConkie, 2001; Engbert et al., 2002). In the present work the identical stimulus material was used in all experiments. Following this argumentation, differences in fixation duration between reading conditions could originate in effects of fixational selectivity and associated differences in mean word properties of the fixated words. Generally, all experimental manipulations investigated in this work led to prolonged fixation durations and a decrease in reading rate relative to the baseline experiment. Readers in the 'hard' samples, the 'frequent' group, and the 'proofreading' condition spent more time on the words in first-pass as well as in second-pass reading compared to the 'original' groups.

Differences in the selectivity of words in first-pass single fixation cases were found between the groups 'hard old' and 'original old' (cf. section 4) and between the 'proofreading' and the 'frequent' sample (cf. section 6). Since single fixation duration was prolonged in the 'hard old' as well as in the 'proofreading' condition relative to the baseline condition, according to the linguistic control theory it might have been expected that these two groups fixated more difficult words, that is, longer words of lower frequency and lower predictability. In fact, the opposite was found: Single fixation cases in these two conditions were more often function words of higher frequency, higher predictability, and shorter in length. Although the 'hard old' sample and the 'proofreading' group fixated on average easier words than their age matched comparison groups, fixation duration was reliably longer (comparison: 'hard old' - 'original old') or equal (comparison: 'proofreading' - 'frequent'). Therefore, the prolongation of fixation durations cannot be a result of local processing difficulties of the visual input during sentence reading. Instead prolonged fixations seem to be result of a cognitive mechanism related to the reading intention, in other words, a results of a general parameter setting triggered by the task demands. The selective parameter specification within the general system of eye movement control in response to the specified reading intention is exactly the definition of reading strategy that has been developed in section 2.4.1. Thus, it can be argued that the experimental manipulation triggered general cognitive control processes to influence the oculomotor processes in a top-down fashion, precisely, to set the saccade timing to a slower pace. This interpretation also holds for the other experimental groups, namely 'hard young' and 'frequent', that showed prolonged fixation durations in the absence of selectivity effects. A reading mode at a lower pace can be considered as a more controlled or more careful reading strategy, with more attention allocated to the reading task (and the reading material). A new saccade is triggered when linguistic processing has proceeded further as compared to the reading strategy found in the baseline experiment. This strategy can be applied in order to achieve a better, more deliberate sentence comprehension or to perform a more precise word inspection.

In addition, this top-down modulation of the general reading speed has been found to be dynamic as demonstrated by the spillover effects of verification questions in the 'hard' samples (see section 4.5.1). Both young and old readers who received difficult comprehension questions after each sentence produced longer fixation durations. This question-effect was even more pronounced on trials immediately following verification questions that - according to the behavioral data of response accuracy and latency - turned out to be the most difficult question type for the subjects. This result favors the interpretation that the parameter setting for the general reading speed is triggered by a cognitive control process in a top down-fashion and parameters are voluntarily adapted in a flexible way (Neumann, 1987, 1989). The fact that the spillover effect of verification questions was more pronounced in old than in young readers has been interpreted with a resource account (cf. section 4.7.2).

8.3. The Targeting Aspect in Reading Strategies

A careful reading strategy has been postulated with respect to saccade target locations within the word (O'Regan & Lévy-Schoen, 1987; O'Regan, 1990). According to their 'Strategy-Tactics-Model' of eye movement control in reading, readers usually aim to target the optimal viewing position (OVP) around the word center in a word-by-word fashion ("careful strategy"). Whenever the landing position deviates greatly from the intended OVP, a refixation ("local tactics") is made to the other half of the word, indicating the carefulness of the reading behavior. In a mode with greater risk taking, readers may lax the refixation criterion and show saccades of lower accuracy with respect to the OVP but no additional refixations.

For single fixation cases, differences in mean landing position were found between the samples 'hard old' and 'original old' as well as between the groups 'proofreading' and 'frequent'. Whereas readers in the 'hard old' group located single fixations slightly, but significantly more to the left of the word center (relative fixation position: 0.39) than readers in the 'original old' group (relative fixation position: 0.41), proofreaders targeted single fixations at the word center (relative fixation position: 0.51), unlike readers in the 'frequent' group (relative fixation position: 0.43). Thus, the carefulness of the 'hard old' group and the 'proofreading' group that was attested by the prolonged fixation durations differs from the carefulness in saccade targeting. Since it was argued in section 4.7 that the 'hard' samples applied a more 'mindful' reading strategy than the 'original' samples, the slight leftward shift of relative fixation position could be associated with an increased focus on informative parts of the word (e.g., Hyönä et al., 1989; Pynte et al., 1991). In contrast, the shift of the relative fixation position to the word center in the proofreading condition is interpreted as a more precise saccade targeting at the OVP to bring words into the region of maximal visual acuity. Targeting the OVP might be the optimal targeting strategy in proofreading, because it might be the most informative location for error detection with a single fixation, since errors could occur at any position between the second and the second last letter in a word. The instruction to proofread the words in the sentences induced a reading strategy that focussed on the serial and immediate word processing which resembles the careful word-by-word scanning routine proposed by O'Regan and Lévy-Schoen (1987). The two authors further argue that:

> "How precisely the eye must be fixated near the word middle for a single fixation to occur is a parameter of global oculomotor strategy which is set in advance as a function of the desired reading speed." (O'Regan & Lévy-Schoen, 1987, p. 381)

According to O'Regan, the 'parameter of global oculomotor strategy' is determined within the continuum from a careful to a risky scanning strategy. As outlined above, I argue that the reading intention, the identified task demands, at a reader's cognitive level specifies the desired reading speed. Thus, the global timing is set by a top-down process. In O'Regans quote, saccade timing and saccade targeting are interrelated. This is found to be true regarding the precision of saccade targeting, namely, that saccadic accuracy (with respect to targeting the OVP) is higher with longer preceding fixation duration (O'Regan & Lévy-Schoen, 1987). But single fixation duration in 'proofreading' did not differ from single fixation duration in the 'frequent' sample. Therefore, the precision of targeting the OVP in single fixation cases in proofreading is not solely the result of prolonged fixation durations. The results of the other experimental samples also demonstrate that the decision about when and where to move the eyes are relatively independent.

In 'proofreading', the high incidence of single fixations that targeted the OVP can be associated with the word-form oriented reading strategy. The demanded carefulness in word identification is optimally applied when fixating at the word center, and the increased proportion of double and multiple fixation cases in the 'proofreading' condition compared to the 'frequent' sample provide further evidence for the careful mode. Words are more often refixated during proofreading than during reading for comprehension, supporting the idea of a serial, and local word processing in proofreading. Along the idea of O'Regan, the proofreading task sets a very strict refixation criterion: Only at a precise fixation location at the OVP a single fixation can occur, as proven by the mean fixation location of single fixation cases located exactly at the word center. A minor deviation in saccade targeting leads to a refixation ("local tactics"), explaining the higher proportion of two-fixation and multiple fixation cases in 'proofreading' compared to the 'frequent' sample. Refixations occur at the cost of reading speed, because they naturally lead to longer gaze durations, as found in the 'proofreading' group compared to the 'frequent' sample.

An additional explanation for the precision of proofreaders to target the OVP can be found in reduced saccade lengths. In single fixation cases, outgoing and more important incoming saccade amplitudes in the 'proofreading' sample were much shorter than in the 'frequent' sample. With shorter incoming saccade amplitudes the saccadic range error (McConkie et al., 1988) is reduced as well, entailing that saccades land more frequently at the OVP and the IOVP is more symmetric, as is the case in proofreading. However, the reduction of saccade lengths in proofreading compared to the mode of reading for comprehension mainly supports the classification of proofreading as a more serial word processing strategy, because along with shorter saccades the skipping rate is systematically reduced and more words of the sentence are inspected foveally.

An important influence determining saccade targeting and the extent of distributed processing is the effective size of the perceptual span. A wider span enables the reader to process words in parallel that affects word targeting and skipping of potentially preprocessed words. A narrow and more symmetrical perceptual span could be assumed for the proofreading sample to explain the reduced skipping rate and the increased proportion of fixated function words. The perceptual span in the 'hard young' and the 'original young' sample should be of comparable size because no changes in fixational selectivity were found. As outlined above, the sample 'hard old' also shows effects of selectivity that in analogy to the proofreading condition could be explained by a narrower perceptual span in comparison to the 'original old' group. A smaller, but more asymmetric span has been obtained for older readers in comparison to younger readers for example in SWIFT-simulations (Laubrock et al., 2006). Here, the difference between 'hard old' and 'original old' readers is possibly a more symmetric, but still narrow perceptual span in the 'hard' reading condition.

8.4. Effects of Cognitive Control and Distributed Processing in Reading Strategies

Possible indicators of of cognitive control that have been investigated in this work are effects of word length, word frequency, and word predictability on fixation measures. The effect sizes have been tested in first-pass single fixation duration and gaze duration. Along the ideas of the distributed processing account (cf. Schroyens et al., 1999; Engbert et al., 2002; Kliegl et al., 2006), the impact of the fixated word (n), of the word left to fixation (n-1), as well as of the word to the right of fixation (n+1) has been simultaneously tested with linear mixed models. It has been demonstrated that the effects of cognitive control indicators in distributed processing differ between reading strategies and their associated top-down control. Importantly, these differences are independent of variance in fixation duration.

While keeping the reading instruction constant (reading for comprehension), the manipulation of the frequency of comprehension questions (see chapter 5) did not show an influence on cognitive control indicators. Despite producing longer fixation durations, the sample 'frequent' did not differ from the sample 'original' in effects of word frequency, word predictability, or word length. As discussed in section 5.7.2 and above in section 8.2, the prolonged fixation durations in the 'frequent' group are interpreted as an increased allocation of general attention to the task in response to the frequent test of comprehension, resulting in a slower pace of saccade executions and thus, in longer saccade latencies.

The manipulation of the frequency *and* difficulty of the comprehension questions clearly impacts on indicators of cognitive control in reading for comprehension (see section 4), with different outcomes in old readers compared to young readers. As expected according the idea of a deep linguistic processing, young readers in the 'hard' group showed an increased effect of upcoming word predictability. In line with the memory retrieval explanation (Kliegl et al., 2006), very high predictable words n+1 were associated with even longer single fixation durations on word n, indicating an increased preprocessing of upcoming words based on the sentence context. Thus, the size of the successor word predictability effect can be interpreted as an index of processing depth. Additionally, the interaction of immediate word length and word frequency in the 'hard young' group was marginally stronger than in the 'original young' group. Young readers receiving frequent and difficult comprehension questions revealed a frequency effect even on short words, with longer fixation durations on low frequent than on high frequent words. This local processing effect might be explained by a strengthened focus on deliberate and more controlled processing in the 'hard' condition, as already associated with the general slowing in reading speed.

Evidence for deliberate reading strategy is clearly found in the old readers who received difficult comprehension questions after each trial. As discussed above, 'hard old' readers read much slower than 'original old' readers interpreted as a consequence of global parameter setting with respect to saccade timing (as a consequence of the increased task demands). Furthermore, changes in saccade targeting led to a reduced skipping rate entailing a different composition of words that received single fixations: The 'hard old' sample fixated many more easily processable function words of high frequency and short length than the 'original old' sample. Interestingly, this fixational selectivity was associated with an increased preprocessing effect of upcoming word frequency. Lexical parafoveal-on-foveal effects can be explained with distributed processing within the perceptual span (Henderson & Ferreira, 1990; Schroyens et al., 1999; Kennedy & Pynte, 2005; Kliegl et al., 2006; Pynte & Kennedy, 2006). The foveal load hypothesis (Henderson & Ferreira, 1990) suggests that during fixations on easy words, that produce a low foveal load, upcoming words in the perceptual span can be processed in parallel. Thus, the occurrence of frequently fixated easy function words may give rise to the increased parafoveal-on-foveal word frequency effect in the 'hard old' condition.

Lag word frequency on the other side had a smaller impact on single fixation duration in the 'hard old' group than in the 'original old' group. This can either be explained by the prolonged fixation durations under which words are further processed before the eyes move on to the subsequent word, or by the reduced size of the perceptual span that has been assumed for this reading group. Furthermore, the immediate word frequency effect in the 'hard old' condition revealed a stronger cubic trend, indicating a preprocessing of word n+1 on high frequent words n. According to Kliegl's (2007) demonstration of the dependency of single fixation duration on lexical status of the fixated word and skipping of the upcoming word for PSC-data, it can be assumed that predominantly the sequence function word (n) (e.g., determiner) followed by a content word (n+1) (e.g., noun), which is skipped, gives rise to the described successor effect, since this is the only pattern that led to skipping costs (i.e. an increase in SFD) on word n.

In sum, there is evidence for a strengthened focus on local processing (supported by the reduced spill over effect of word frequency and the fixational selectivity) and at the same time evidence for an increased focus on distributed processing in reading direction (supported by the increased successor word frequency effect and the increased fixation duration on high frequent words n). Low-level parafoveal-on-foveal effects are very well documented, but the influence of lexical parafoveal-on-foveal effects is controversily discussed (Rayner & Juhasz, 2004). Here, an impact of upcoming word-frequency and word-predictability on fixation duration is clearly demonstrated in both young and old readers.

In proofreading, the effects of cognitive control indicators differ clearly from reading for comprehension. As expected from the idea that in a proofreading mode readers put less weight on the processing of the sentence context, the successor word predictability effect on single fixation duration was weaker in the 'proofreading' sample than in the 'frequent' sample. Thus, the effect size of the upcoming word predictability proved again to be an indicator for linguistic processing depth. The results indicate that proofreaders put more weight on the immediate processing of individual words, as supported by the fixational selectivity in favor of function words discussed in the previous section. More evidence for a focus on local processing is found in different effect sizes of the immediate cognitive control indicators between proofreading and reading for comprehension. Despite equally long single fixation durations, the 'proofreading' group showed a stronger influence of current word predictability and of the interaction of word frequency and word length. Moreover, the word frequency effect on SFD on the fixated word is much more linear in the 'proofreading' sample, implying that minor preprocessing of word n+1 is performed on high frequent words n, in contrast to the 'frequent' sample. The weaker lag word frequency effect in the 'proofreading' sample compared to the 'frequent' sample further supports the notion of the focus on the immediate and more serial word processing during proofreading. The increased immediate effects of word frequency, word length, and word predictability fit to the assumption of a very local and serial word processing in proofreading due to a reduced perceptual span and the different composition of fixated words due to this selectivity of saccade targets. A narrow span would diminish parallel processing to the left as well as to the right of fixation, as has been found in the proofreading condition.

The more pronounced immediacy effects along with reduced lag word frequency effects and differences in word selectivity have been explained with a reduced size of the perceptual span in both the 'proofreading' group and the 'hard old' group that read for comprehension. In contrast to the proofreading group, the 'hard old' group produced stronger parafoveal-on-foveal word frequency effects than the 'original old' sample. The only explanation for this difference is to assume a dynamically modulated perceptual span, as suggested by Henderson and Ferreira (1990). That is, whenever a word of minor lexical difficulty is fixated, e.g. a high frequent word n, the perceptual span is instantly widened to the right so that preprocessing of word n+1 is increased. This preprocessing is measurable in the fixation duration on the 'easy' word n. Hence, in the 'hard old' condition the attention span is generally narrowed in a top-down fashion in response to the task demands, which can explain the selectivity effects, but at the same time the span varies across fixations and can locally be widened if foveal load is low. In contrast, for proofreading the assumption must be made that the size of the perceptual span is not dynamically modulated as a function of lexical difficulty of word n so that preprocessing effects are not increased even though words n are more often easier and shorter words in proofreading than in reading for comprehension. Hence, the size of the perceptual span is not only locally modulated (Henderson & Ferreira, 1990), but must be set to some degree by a top-down process in reponse to the desired task demands and reading intention. Since proofreading requires a very serial word reading strategy, the perceptual span of proofreaders cannot as strongly be locally modulated as is presumably the case in the reading for comprehension.

In sum, the different effects of distributed processing and saccade timing between reading strategies demonstrate that reading intention specifies parameters of the otherwise automatic processes of eye movement control in reading. Even though the reading material was identical in all tested conditions, on which parameters could be specified locally in a bottom-up fashion, the reading mode seems to preselect the range of possible parameters in a top-down fashion in order to achieve the reading goal. This is in line with Neumann's idea of the 'levels-of-control' concept (Neumann, 1989) in which he describes different modes of controlled action that can well be applied to eye movement control in reading.

8.4.1. Fixational Selectivity and the Role of Function Words in Reading

Reading is a selective process and eye movements during reading are selective. It is known that not all words are focussed during reading and that skipped words can be processed during the fixation of other words (Fisher & Shebilske, 1985; Just & Carpenter, 1980). Especially short and highly frequent words are prime candidates for skipping. An additional factor increasing skipping probability is the syntactic category a word belongs to, as demonstrated by the higher proportion of skipped function words compared to content words that were machted in length and frequency (e.g., Drieghe, Pollatsek, Staub, & Rayner, 2008). In accordance with a reduced skipping rate, both the 'hard old' and the 'proofreading' group fixated more function words than their control groups.

As their names indicate, functors such as determiners, conjunctions, pronouns, or prepositions play a minor semantic role in sentence comprehension compared to content words. Considering this idea, the increased proportion of fixated function words stays in contrast to the proposed in-depth reading strategy in the 'hard' reading condition because readers would not benefit in sentence comprehension from fixating functors. But at the same time, function words provide most of the text cohesion in narrative and nonnarrative texts (Goodman, 1985). Even though cohesion is not as relevant in isolated sentence reading, the additional focus on function words in the 'hard old' sample can be attributed to a more deliberate comprehension processing. Whereas old readers usually tend to skip many more words than younger readers (Rayner et al., 2006; Laubrock et al., 2006), it has been demonstrated that this effect is reduced when high comprehension demands are made. In the light of the structural account of reading (Koriat & Greenberg, 1994), the increased amount of fixated function words in the 'hard old' group would reflect a strengthened role of phrase structures. Following Koriat and Greenberg (1994), the coding of the sentence's structure preceedes the coding of meaning and lays the pathway for the following semantic analysis. A similar argument is made by Schmauder, Morris, and Poynor (2000) who stress the different roles of function words in lexical compared to sentence or text processing. Function words have a relational function and are therfore often reread in order to create a sentence integration (Schmauder et al., 2000). Since function words were focused more often in the 'hard old' than in the 'original old' condition, it seems that readers put more effort in the precise interpretation of the sentence, grounded on an exact structural representation of the sentence. Along these ideas, the increased amount of fixated functors in the 'hard' condition would be a reflection of the reader's more detailed structural and semantic analysis, thus a change in processing depth. Therefore, the higher proportion of fixated function words in the 'hard old' sample is not contradictory to the proposed mindful reading strategy assumed in the 'hard' readers[13]. In contrast to the old reader groups, the higher proportion of fixated function words during proofreading compared to reading for comprehension can clearly be attributed to the serial scanning routine in order to spell check most of the words in the sentence, as demanded by the reading instruction.

As described in section 3.1, there is a strong correlation between lexical frequency and word length in the sentences used in the present reading experiments. This is mainly due to a basic characteristic of language that content words, especially uncommon content words, are defined by their coreferences and their imbedding in contexts of high frequent, noncontent words. The problem of collinearity of word properties of the PSC and other corpora with respect to the results has been critically discussed by Rayner and colleagues (2007). Since in this work the eye movement dynamics across a whole sentence was investigated with respect to effects of reading strategies, the selection of only a few target words for the analysis would have restricted the picture, hence the generalisability to normal reading (without restrictions by the experimental setup, cf. footnote on page 80). The artificial control of word frequency can make texts less predictable and therefore less comprehensible to readers (Goodman, 1985), an effect attributed to the status of function words. Furthermore, the limited frequency range of selected target words in the experimental material may give rise to inconsistencies of results between experiments

[13]Since the young readers generally fixate more words, and thus more function words, than old readers (Rayner et al., 2006; Laubrock et al., 2006), the increased mindfulness and focus on sentence comprehension in the 'hard young' readers is predominantly evident in the stronger successor word predictability effect.

and corpus analysis (Kliegl, 2007). Thus, only the corpus based analysis, that naturally includes a complete range of word properties, enables the researcher to unfold the complex dynamics of reading. Effects related to distributed lexical processing, especially those related to the alternation of content and function words and skipping of these words, might be detected only in a corpus based analysis (see Kliegl, 2007, for an analysis with word frequency nested within function and content words). From a methodological point of view, corpus analyses in eye movement research that include different syntactic word classes yields a good picture of the natural reading processes.

8.5. Outlook and Conclusion

In this work, reading strategies were experimentally induced in different reading groups to explore the associated variance in the eye movement behavior in a systematic way. The strategies investigated were reading for comprehension, in the continuum from superficial over normal to deep, mindful processing, and proofreading, as a case of a more mindless reading strategy with respect to the sentence context. The observed reading strategies fall all in the range of reading sub-skills that are inherent to a skilled reader. Therefore, even though reading strategies were tested in a between-subjects design, the results indirectly demonstrate the variability of reading skill within the subject. To get an even more precise picture of the differences between effects of individuality and effects of top-down reading strategy, as a next step the experimental manipulations ought to be tested in a within-subjects design.

Individual differences within a defined reading skill were differentiated from the effects of reading strategies. Of course, when it comes to the investigation of individual differences for example, in reading for comprehension, the aspects of a more mindful or more mindless reading proven in this work are very well translatable into reading styles or types of readers. Evidence has been provided that the differences between readers fall along similar dimensions that have been described in different reading strategies, namely, the extent of distributed processing and the selectivity of fixated words. As demonstrated in this work, not only the reading instruction alters the reading behavior. Also variation in the way the instructed task is tested after reading (here, the difficulty and frequency of comprehension questions) clearly influences reading speed, the selectivity of fixated words, as well as local effects of indicators of cognitive control. A similar finding has recently reported by Radach and colleagues (Radach, Huestegge, & Reilly, 2008) who investigated the two factors 'reading depth' and 'material format' on local eye movement control in reading using a between-subjects design. Therefore, the variability within skilled reading and the methodological concern that the limited control over the subjects' engagement in the given task must be taken more seriously into account in interpreting eye movement data in reading as well as in model building.

Even though a large amount of eye movements in sentence reading can be explained without the integration of a language processing theory, an adequate model of eye movement control in reading will have to include an account for language processing as claimed by Reichle, Rayner, and Pollatsek (2003). The results of the current work further suggest that there is a need to consider the top-down control of eye movements triggered by reading intention or differences in comprehension depth, and to try to simulate the variability of reading strategies in skilled reading (Radach et al., 2008). More precisely, it is plausible to assume a saccade timing independent from saccade targeting and a dynamically modulated size of the perceptual span, and that all components are sensitive to the demands of the reading task. This assumption is compatible with the concept of parameters specification in selective attention theory of motor control (Neumann, 1987, 1989). Reading intention or the reading goal specifies a subset of parameters for eye movement control in reading. Other parameters are locally determined by the given input, that is, the reading material, and those processes represent the automatic components of a controlled, voluntary action of reading with a specified reading goal (cf. Neumann, 1987, 1989).

Current computational models of eye movement control in reading differ widely in their architecture and their underlying assumptions of what triggers a saccade (e.g., Reichle et al., 1998; Reilly & O'Regan, 1998; Engbert et al., 2002). The SWIFT model of eye movement control (Engbert et al., 2002, 2005) might be a good starting point to simulate reading strategies based on the current data because it is psychologically and neurologically plausible with respect to saccade timing and saccade targeting. Saccade programming is guided by separate temporal and spatial pathways. The model is built upon a parallel processing framework, incorporating mechanisms of the perceptual span and spatially distributed processing, all aspects that are relevant for the theoretical embedding of reading strategies into a model of reading. In SWIFT, individual differences are simulated by selecting different values for different 'subjects', for example, the parameters chosen for the random saccade timing. Furthermore, the model has been repeatedly tested on eye movement data from reading studies using the PSC (Engbert et al., 2002, 2005; Richter, Engbert, & Kliegl, 2006; Laubrock et al., 2006; Nuthmann et al., 2007). Therefore, the SWIFT model would serve as a good framework to test the ideas of a (top-down) parameter specification for eye movement control in response to a given reading intention.

Linear mixed models proved as a useful tool for the analysis of the complex eye movement behavior on sentence reading. Random effects attributed to the selection of subjects and material as well as fixed effects attributed to the reader- and word-level, that all potentially influence fixation duration, were estimated in a single sweep. Confounded variables, such as word length and word frequency, were captured by estimating coefficients of these fixed effects simultaneously. The fact that there is more than 15% of variance in the fixation data due to the random selection of readers underlines the need for further research on the source of individual differences in reading, formulated as an urgent issue in reading research (Radach & Kennedy, 2004). Within LMM, this goal translates into the detection of reader-level fixed effects that explain some of the relatively large amount of random variance in readers. Vocabulary size and the index of general processing speed turned out to have only minor influence on first-pass online-measures in reading. The selection of a principle component analysis on several reader-level variables proved useful to capture individual differences in reading and to identify different reader types. Readers differ with respect to the extent of distributed processing and their selectivity of fixated and skipped words, that both turned out to affect fixation duration measures at the reader-level.

The present work is the first to examine effects of reading strategy and individual differences on eye movements during reading within the framework of the distributed processing account using identical sentences material. The results provide evidence that the eye movement behavior is sensitive to specific task demands and reading goals, and that reading intentions are strongly reflected in highly automatic first-pass sentence reading. There is a complex, dynamic interaction of oculomotor and cognitive control processes in reading: The parameters of the oculomotor system set by the intended reading strategy (via instruction or task demands) affect saccade timing, saccade targeting, and presumably the size of the perceptual span. To what extent the perceptual span is adapted according to the intended reading goal could be experimentally tested in a moving window paradigm (McConkie & Rayner, 1975). Evidence has been provided that the entailed selectivity of words, especially the amount of fixated closed-class words, can again influence effects of cognitive control, especially those critically discussed lexical spill-over and preprocessing effects. Effect sizes of upcoming word predictability have proven to be indexical for the mindfulness in reading. Result indicate that the size of the perceptual span is dynamically modulated by foveal load in reading for comprehension, but to a much smaller extent in more serial reading strategies. In sum, the results of this work greatly expand the understanding of the variability within skilled reading, differentiated from individual differences. They provide evidence for voluntary, top-down influences on eye movement control in reading, that - next to the automatic, bottom-up control processes based on the reading material - need to be taken into account for the improvement of models of eye movement control in reading.

References

Aaronson, D. & Ferres, S. (1984). Reading strategies for children and adults: Some empirical evidence. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 189-220.

Arguin, M. & Bub, D. (1993). Single-character processing in a case of pure alexia. *Neuropsychologia*, *31*, 435-458.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: differential effects of frequency and predictability. *Quartely Journal of Experimental Psychology A*, *58*(6), 1065-86.

Baayen, R. (2008). *Practical data analysis for the language sciences with R.* Cambrigde, MA: Cambridge University Press.

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Bach, M. (1996). The "Freiburg Visual Acuity Test" - automatic measurement of visual acuity. *Optometry and Vision Science*, *73*, 49-53.

Bach, M. (2007). The freiburg visual acuity test - variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *245*(7), 965-971.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364-90.

Bates, D. (2007). *lme4: Linear mixed-effects models using S4 classes.* (R package version 0.99875-9.)

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433-436.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226-243.

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory*, *9*(4-6), 365-381.

Caplan, D. & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77-94; discussion 95-126.

Caramazza, A. & McCloskey, M. (1988). The case for single patient studies. *Cognitive Neuropsychology*, *6*, 501-513.

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* (p. 275-307). New York: Academic Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335-359.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (p. 341-371). New York: Elsevier Ltd.

Conway, A. R. & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, *4*(6), 577-590.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, and Computers*, *34*(4), 613-617.

Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.

Daneman, M. & Stainton, M. (1993). The generation effect in reading and proofreading. *Reading and Writing: An Interdisciplinary Journal*, *5*, 297-313.

Dixon, P., LeFevre, J.-A., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, *80*(4), 465-472.

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1552-1560.

Drosdowski, G., Scholze-Stubenrecht, W., & Wermke, M. (Eds.). (2000). *Der Duden, Die deutsche Rechtschreibung* (Bd.1 ed.). Mannheim: Meyers Lexikonverlag.

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

Engbert, R. & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035-45.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*(5), 621-636.

Engbert, R. & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 7192-797.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777-813.

Everatt, J. & Underwood, G. (1994). Individual differences in reading subprocesses: Relationships between reading ability, lexical access, and eye movement control. *Language and Speech*, *37*(3), 283-297.

Evered, R. (1983). So what is strategy? *Long Range Planning*, *16*(3), 57–72.

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

Fisher, D. & Shebilske, W. (1985). There is more that meets the eye than the eyemind assumption. In R. Groner, G. McConkie, & C. Menz (Eds.), *Eye movements and human information processing* (p. 149-158). Amsterdam: North Holland.

Fodor, J. A. (1983). *The modularity of mind.* MIT Press.

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

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/ hierarchical models.* Cambridge: University Press.

Geyken, A. (2007). The DWDS-Corpus: A reference corpus for the German language of the 20th century. In C. Fellbaum (Ed.), *Collocations and idioms: Linguistic, lexicographic, and computational aspects* (p. 23-40). London: Continuum Press.

Goodman, K. S. (1985). Unity in reading. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes of reading* (p. 813-840). Newark: International Reading Association.

Gribbons, B. & Herman, J. (1997). True and quasi-experimental designs. *Practical Assessment, Research, and Evaluation*, *5*(14).

Haber, R. N. & Schindler, R. M. (1981). Error in proofreading: Evidence of syntactic control of letter processing? *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 573-579.

Haberlandt, K. (1994). Methods in reading research. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (p. 1-31). Orlando, FL: Academic Press. Michigan Press.

Haberlandt, K. F., Graesser, A. C., & Schneider, N. J. (1989). Reading strategies of fast and slow readers. *Journal of Experimental Psychology*, *15*(5), 815-823.

Heller, D. (1985). Eye movements in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (p. 139-154). Berlin: Deutscher Verlag der Wissenschaften.

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.

Huey, E. (1908). *The psychology and pedagogy of reading.* New York: Macmillan.

Humphrey, H. E., Kemper, S., & Radel, J. D. (2004). The time course of metonymic language text processing by older and younger adults. *Experimental Aging Research*, *30*(1), 75-94.

Hyönä, J., Lorch, R. F. J., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology*, *94*(1), 44-55.

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

Hyönä, J., Niemi, P., & Underwood, G. (1989). Reading long words embedded in sentences: Informativeness of words halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 142-152.

Hyönä, J. & Nurminen, A. M. (2006). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology*, *97*(1), 31-50.

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*(3), 544-555.

Inhoff, A. W. & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics*, *40*(6), 431-439.

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*(3), 219-264.

Johnson, J. B. & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology and Evolution*, *19*(2), 101-108.

Jolliffe, I. (2002). *Principle component analysis.* New York: Springer.

Juhasz, B. (2005). Age-of-acquisition effects in word and picture processing. *Psychological Bulletin*, *131*, 684-712.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122-149.

Just, M. A., Carpenter, P. A., & Masson, M. E. (1982). *What eye fixations tell us about speed reading and skimming.* (Tech. Rep.). Department of Psychology, Canergie-Mellon University.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228-238.

Kaakinen, J. & Hyönä, J. (2008). Perspective-driven text comprehension. *Applied Cognitive Psychology*, *22*, 319-334.

Kaakinen, J., Hyönä, J., & Keenan, J. (2002). Perspective effects on online text processing. *Discourse Processes*, *33*(2), 159-173.

Kemper, S., Crow, A., & Kemtes, K. (2004). Eye-fixation patterns of high- and low-span young and older adults: Down the garden path and back again. *Psychology and Aging*, *19*(1), 157-170.

Kemper, S., McDowd, J., & Kramer, A. (2006). Eye movements of young and older adults while reading with distraction. *Psychology and Aging*, *21*(1), 32-39.

Kennedy, A. (2000). Parafoveal processing in word recognition. *The Quarterly Journal of Experimental Psychology*, *53A*(2), 429-455.

Kennedy, A. & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153-168.

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

Kennison, S. M. & Clifton, J., 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*(1), 68-81.

Kess, J. (1992). *Psycholinguistics.* Amsterdam: Benjamins.

Kliegl, R. (1982). *On relations between cognition and reading style: Individual differences and developmental trends.* Unpublished Ph.D.Thesis: University of Colorado.

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007). *Journal of Experimental Psychology: General*, *136*(3), 530-537.

Kliegl, R., Geyken, A., Hanneforth, T., & Würzner, K. (2006). Corpus matters: A comparison of German DWDS and CELEX lexical and sublexical frequency norms for the prediction of reading fixations. *Unpublished manuscript*.

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., Masson, M. E., & Richter, E. M. (in press). A linear mixed model analysis of masked repetition priming.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*(1), 12-35.

Koriat, A. & Greenberg, S. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin and Review*, *1*(3), 345-356.

LaBerge, D. & Samuels, S. (1985). Toward a theory of automatic information processing in reading. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes of reading* (p. 689-718). Newark: International Reading Association.

Land, M. (2007). Fixation strategies during active behaviour: A brief history. In R. van Gompel, M. Fischer, W. H. Murray, & R.L. (Eds.), *Eye movements: A window on mind and brain* (p. 75-95). Elsevier Ltd.

Laubrock, J., Kliegl, R., & Engbert, R. (2006). SWIFT explorations of age differences in eye movements during reading. *Neuroscience and Biobehavioral Reviews*, *30*(6), 872-884.

Lehrl, S. (1977). *Mehrfach-Wortschatz-Test Form B [multiple choice vocabulary test].* Erlangen, Germany: Perimed.

Levy, B. A. (1983). Proofreading familiar text: Constraints on visual processing. *Memory and Cognition*, *11*(1), 1-12.

Levy, B. A. & Begin, J. (1984). Proofreading familiar text: Allocating resources to perceptual and conceptual processes. *Memory and Cognition*, *12*(6), 621-632.

Liversedge, S. & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science*, *4*(1), 6-14.

Liversedge, S., Pickering, M., Branigan, H., & Gompel, R. van. (1998). Processing arguments and adjuncts in isolation and context: The case of by-phrase ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 461-475.

Liversedge, S., Pickering, M., Clayes, E., & Branigan, H. (2003). Thematic processing of adjuncts: Evidence from an eye-tracking experiment. *Psychonomic Bulletin and Review*, *10*(3), 667-675.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly*, *13*(2), 123-146.

Masson, M. (1985). Rapid reading processes and skills. *Reading Research: Advances in Theory and Practice*, *4*, 183-230.

Matin, E. (1974). Saccadic suppression: A review. *Psychological Bulletin*, *81*, 899-917.

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

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), 1107-18.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception and Psychophysics*, *46*(3), 245-253.

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.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (p. 361-393). Hillsdale, NJ: Lawrence Erlbaum Associates.

Neumann, O. (1989). On the origins and status of the concept of automatic processing. *Zeitschrift für Psychologie*, *197*, 411-428.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*(17), 2201-17.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling;. *Vision Research*, *47*(7), 990–1002.

Oberauer, K. & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, *55*(4), 601–626.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167-193.

Olson, R., Kliegl, R., Davidson, B., & Foltz, G. (1985). Individual and developmental differences in reading disability. *Reading Research: Advances in Theory and Practice*, *4*, 1-64.

O'Regan, J. K. (1990). Eye movements and reading. *Review of Oculomotor Research*, *4*, 395-453.

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: Vol. 12. The psychology of reading* (p. 363-383). Hillsdale, NJ: Erlbaum.

O'Regan, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics*, *25*(6), 501-509.

O'Regan, 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* (p. 333-354). New York: Springer.

Patterson, K. & Kay, J. (1982). Letter-by-letter reading: Psychological description of a neurological syndrome. *Quarterly Journal of Experimental Psychology*, *34A*, 411-441.

Pelli, D. (1997). The videotoolbox software for visual psychophsysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437-442.

Perfetti, C. A. (1994). Psycholinguistics and reading ability. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (p. 849-894). Orlando, FL: Academic Press. Michigan Press.

Pilotti, M. & Chodorow, M. (2009). Error detection/ correction in collaborative writing. *Reading and Writing*, *22*(3), 245-260.

Pilotti, M., Chodorow, M., & Thornton, K. C. (2005). Effects of familiarity and type of encoding on proofreading. *Reading and Writing*, *18*, 325-341.

Pinheiro, J. & Bates, D. (2000). *Mixed-effects models in S and S-PLUS.* New York: Springer.

Pynte, J. & Kennedy, A. (2006). In influence over eye movements in reading exerted from beyond the level of the word: Evdidence from reading English and French. *Vision Research*, *46*(22), 3686-3801.

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*(2), 458-470.

Quené, H. & van den Bergh, H. (2004). On multi-level modeling of data from repeated mesures designs: A tutorial. *Speech Communication*, *43*(103-121).

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria.

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye movement control in reading. *Psychological Research*, *72*, 675-688.

Radach, R. & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *Journal of Cognitive Psychology*, *16*(1-2), 3-26.

Raney, G. E. & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology*, *49*(2), 151-172.

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*, 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. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 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. & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics*, *58*(5), 734-747.

Rayner, K. & Juhasz, B. J. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology*, *16*(1/2), 340-352.

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

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General*, *136*(3), 520-529.

Rayner, K. & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin and Review*, *3*(2), 245-248.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*(3), 448-65.

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*(3), 21-49.

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), 1188-200.

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

Reichle, E. D. (2006). Computational models of eye-movement control during reading: Theories of the "eye-mind" link. *Cognitive Systems Research*, *7*, 2-3.

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

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

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin and Review*, *16*(1), 1-21.

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

Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research*, *7*, 23-33.

Sarkar, D. (2008). *Lattice: Lattice graphics.* (R package version 0.17-6)

Schmauder, A., 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*(7), 1098-1108.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zooning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (p. 203-226). Cambridge: MIT Press.

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

Shallice, T. (1979). Case-study approach in neuropsychological research. *Journal of Clinical Neuropsychology*, *1*, 183-211.

Smiler, A. P., Gagne, D. D., & Stine-Morrow, E. A. (2003). Aging, memory load, and resource allocation during reading. *Psychology and Aging*, *18*(2), 203-9.

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. *Developmental Review*, *10*, 72-100.

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

Stine-Morrow, E. A., Loveless, M. K., & Soederberg, L. M. (1996). Resource allocation in on-line reading by younger and older adults. *Psychology and Aging*, *11*(3), 475-86.

Taylor, W. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, *33*, 42-48.

Tinker, M. A. (1946). The study of eye movements in reading. *Psychological Bulletin*, *43*(2), 93-120.

Tinker, M. A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin*, *55*(4), 215-231.

Underwood, G., Hubbard, A., & Wilkinson, H. (1990). Eye fixations predict reading comprehension: The relationship between reading skill, reading speed, and visual inspection. *Language and Speech*, *33*(1), 69-81.

Underwood, N. R. & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly*, *20*(2), 153–162.

Verhaegen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, *18*(2), 332-339.

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), 3513-3533.

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), 352-64.

Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly*, *35*(4), 554–566.

Walker, R. Y. (1938). A qualitative study of the eye movements of good readers. *The American Journal of Psychology*, *51*(3), 472-481.

Warrington, E. & Shallice, T. (1980). Word form dyslexia. *Brain*, *103*, 99-112.

Wechsler, D. (1964). *Der Hamburger Wechsler Intelligenztest für Erwachsene (HAWIE).* Bern, Switzerland: Huber.

Wermke, M., Kunkel-Razum, K., & Scholze-Stubenrecht, W. (Eds.). (2006). *Der Duden, Die deutsche Rechtschreibung* (Bd. 1. 24. ed.). Mannheim: Dudenverlag.

Wickham, H. (2007). *Reshape: Flexibly reshape data.* (R package version 0.8.0)

Wickham, H. (2008). *ggplot2: An implementation of the grammar of graphics.* (R package version 0.6)

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

Yarbus, A. (1967). *Eye movements and vision.* New York: Plenum Press.

A. The Potsdam Sentence Corpus (PSC) with Questions and Answering Options

ID	LENGTH	WORD MATERIAL		
1	7	Den Ton gab der Künstler seinem Gehilfen.		
easy question		Was gab der Künstler seinem Gehilfen?		
		Ton	Stein	Figur
hard question		Wem gab der Künstler den Ton?		
		seinem Assistenten	seiner Putzfrau	seinem Kollegen
2	8	Der Hof lag weit außerhalb des eigentlichen Dorfes.		
easy question		Was lag außerhalb des Dorfes?		
		Haus	Hof	Stall
hard question		Wo lag der Hof?		
		im Dorfzentrum	jenseits der Dorfgrenzen	am Stadtrand
3	10	Die Wanderer sahen Rehe auf einer Lichtung im Wald äsen.		
easy question		Was taten die Rehe?		
		fressen	kauen	äsen
hard question		Wen sahen die Wanderer im Wald?		
		Kühe	Vögel	Wild
4	10	Den Kopf hieb man früher nur Mördern und Verrätern ab.		
easy question		Was hieb man Mördern und Verätern ab?		
		Kopf	Hände	Bein
hard question		Was hieb man Mördern und Verätern ab?		
		das Haupt	die Hand	den Fuß
5	8	Vorne am Bug sah man eine prächtige Galionsfigur.		
easy question		Wo war die Galionsfigur?		
		Heck	Bug	Mast
hard question		Wo war die Galionsfigur zu sehen?		
		am Heck	an der Spitze	am Mast

ID	LENGTH	WORD MATERIAL		
6	7	Sogar aus Raps läßt sich Kraftstoff herstellen.		
easy question		Woraus läßt sich Kraftstoff herstellen?		
		Flachs	Hanf	Raps
hard question		Was läßt sich sogar aus Raps produzieren?		
		Körpermilch	Speiseöl	Treibstoff
7	8	Torsten beobachtete gestern eine Maus, die Efeu fraß.		
easy question		Was fraß die Maus?		
		Efeu	Gras	Laub
hard question		Wann beobachtete Torsten die Maus?		
		einen Tag zuvor	am heutigen Morgen	vor zwei Tagen
8	9	Der schüchterne kleine Gnom mied die Nähe der Elfen.		
easy question		Wer mied die Nähe der Elfen?		
		Troll	Zwerg	Gnom
hard question		Wie war der Gnom?		
		vorlaut	zurückhaltend	elfenfreundlich
9	9	Claudia hatte ihr Fahrrad auf der Straße stehen lassen.		
easy question		Wo stand Claudias Fahrrad?		
		Hof	Einfahrt	Straße
hard question		Wo stand Claudias Fahrrad?		
		im Flur	im Keller	draußen
10	11	Wir hätten schon vor einer Stunde wissen sollen, ob ihr kommt.		
easy question		Seit welchem Zeitraum hätten wir wissen sollen, ob ihr kommt?		
		Stunde	Woche	Monat
hard question		Wußten sie vor einer Stunde bescheid?		
		nein	ja	möglich
11	9	Die Eltern konnten ihre Kinder im Garten raufen hören.		
easy question		Wo rauften die Kinder?		
		Terrasse	Straße	Garten
hard question		Wer hörte die Kinder raufen?		
		Oma und Opa	Mutter und Vater	Tante und Onkel
12	9	Er hätte nicht auch noch am Telefon nörgeln sollen.		
easy question		Was tat er am Telefon?		
		meckern	schimpfen	nörgeln
hard question		Was tat er am Telefon?		
		lachen	niesen	meckern
13	11	Wegen ihrer Diät hatte die Gräfin leider keine Auster nehmen dürfen.		
easy question		Was konnte die Gräfin nicht nehmen?		
		Auster	Krebs	Muscheln
hard question		Was steht nicht auf dem Speiseplan der Gräfin?		
		Muscheln	Krabben	Langusten

ID	LENGTH	WORD MATERIAL		
14	9	Die meisten Hamster bleiben bei Tag in ihrem Häuschen.		
easy question		Wer bleibt tagsüber im Häuschen?		
		Meerschweinchen	Hamster	Ratten
hard question		Wann bleiben die Hamster drinnen?		
		bei Nacht	wenn es hell ist	im Winter
15	10	Man sollte nie Geschirr mit einem dreckigen Lappen spülen müssen.		
easy question		Wie sollte der Lappen nicht sein?		
		sauber	schmutzig	dreckig
hard question		Was sollte man nicht mit einem dreckigen Lappen erledigen?		
		den Hausputz	die Gartenarbeit	den Abwasch
16	9	Man kann Spargel dämpfen oder in viel Wasser kochen.		
easy question		Wie kann man Spargel zubereiten?		
		dämpfen	braten	backen
hard question		Was braucht man für die Zubereitung von Spargel?		
		Flüssigkeit	Rauch	Gewürze
17	9	Manchmal sagen Opfer vor Gericht nicht die volle Wahrheit.		
easy question		Wer sagt vor Gericht nicht immer die Wahrheit?		
		Opfer	Zeugen	Täter
hard question		Wer sagt im Prozess nicht immer die Wahrheit?		
		Geschädigte	Zeugen	Täter
18	7	Die meisten Befragten hören Musik zur Entspannung.		
easy question		Was hören viele zur Entspannung?		
		Hörspiel	Radio	Musik
hard question		Was machen viele bei Musik?		
		verdauen	nachdenken	sich erholen
19	7	Kinder essen Quark am liebsten mit Früchten.		
easy question		Was mögen Kinder am liebsten mit Früchten?		
		Quark	Brei	Yoghurt
hard question		Welchen Quark essen Kinder am liebsten?		
		Obstquark	Sahnequark	Magerquark
20	10	Bei Wölfen leben Rudel nicht verwandter Tiere in getrennten Revieren.		
easy question		Wer lebt bei den Wölfen in getrennten Revieren?		
		Gruppen	Rudel	Herden
hard question		Wer lebt bei den Wölfen in unterschiedlichen Gebieten?		
		verwandte Rudel	nicht verwandte Rudel	konkurrierende Rudel

ID	LENGTH	WORD MATERIAL		
21	11	Die Frauen in den Andendörfern weben Stoff noch auf traditionellen Webstühlen.		
easy question		Was weben Frauen in den Anden?		
		Tücher	Teppiche	Stoff
hard question		Auf welche Weise weben die Frauen in den Anden?		
		auf neuartige Weise	auf fortschrittliche Art	auf altherköm-mliche Art
22	11	Die Platzwarte ebnen Stück für Stück den Rasen nach dem Spiel.		
easy question		Was machen die Platzwarte mit dem Rasen?		
		ebnen	sprengen	mähen
hard question		Wo wird der Rasen geglättet?		
		auf dem Platz	im Garten	im Park
23	7	In den Fässern gären Beize und Lauge.		
easy question		Was gärt in den Fässern?		
		Säure	Beize	Wein
hard question		Wofür lagern Beize und Lauge in Fässern?		
		zur Kühlung	zur Zersetzung	zur Aufbewahrung
24	8	Die Förster küren Ahorn zum Baum des Jahres.		
easy question		Wen küren die Förster?		
		Esche	Eiche	Ahorn
hard question		Wann ernennen die Förster bestimmte Bäume?		
		alle zwei Jahre	jeden Monat	alle zwölf Monate
25	6	Wolfgangs Töchter studieren Literatur und Maschinen-bau.		
easy question		Was studiert eine von Wolfgangs Töchtern?		
		Literatur	Linguistik	Germanistik
hard question		Wie viele Kinder hat Wolfgang?		
		mindestens zwei	eins	keins
26	9	In der Klosterschule herrschen Schwester Agathe und Schwester Maria.		
easy question		Was tun die Schwestern in der Klosterschule?		
		regieren	amtieren	herrschen
hard question		Wer hat die Macht in der Klosterschule?		
		die Vorsteherin	zwei Nonnen	der Bischof
27	7	Hier scheinen Klempner am Werk zu sein.		
easy question		Wer ist am Werk?		
		Installateure	Flaschner	Klempner
hard question		Arbeiten hier Klempner?		
		auf keinen Fall	mit Sicherheit	wahrscheinlich

ID	LENGTH	WORD MATERIAL		
28	8	Im Aussehen gleichen Bratsche und Geige sich sehr.		
easy question		Was ähnelt der Geige?		
		Bratsche	Viola	Cello
hard question		Was tun Bratsche und Geige?		
		sich äußerlich ähneln	sich klanglich gleichen	sich gleich handhaben
29	7	Angeblich flunkern Künstler oft bezüglich ihrer Einnahmen.		
easy question		Was tun Künstler bezüglich ihrer Einnahmen oft?		
		mogeln	flunkern	übertreiben
hard question		Was wird von Künstlern falsch angegeben?		
		ihre Wertpapiere	ihre Einkünfte	ihre Kosten
30	6	Manchmal krakeelen Politiker genauso wie Demonstranten.		
easy question		Was tun Politiker ebenso wie Demonstranten?		
		lärmen	grölen	krakeelen
hard question		Wie sind Politiker manchmal?		
		nachdenklich	unbeholfen	laut
31	9	Die Armen plündern Speicher und Vorratskeller der reichen Bauern.		
easy question		Was tun die Armen?		
		plündern	stehlen	rauben
hard question		Wie sind die Plünderer?		
		bedürftig	reich	krank
32	8	Die Richter der Landwirtschaftsschau prämieren Rhabarber und Mangold.		
easy question		Was machen die Richter mit Rhabarber und Mangold?		
		küren	prämieren	wählen
hard question		Was passiert mit Rhabarber und Mangold?		
		werden verkocht	werden gekürt	werden verzehrt
33	9	Schon immer war der Besitz von Land sehr wichtig.		
easy question		Was gilt als wertvoller Besitz?		
		Grundstück	Wald	Land
hard question		Als was gilt Landeigentum?		
		unbedeutend	überschätzt	wertvoll
34	9	Ein berühmter Maler hat sich selbst ein Ohr abgeschnitten.		
easy question		Welches Körperteil hat ein Maler sich abgeschnitten?		
		Ohr	Nase	Daumen
hard question		Wie ist der Maler?		
		bekannt	einsam	geächtet
35	9	Das Pferd ist seinem Reiter auf den Fuß getreten.		
easy question		Worauf trat das Pferd?		
		Hand	Fuß	Zeh
hard question		Womit hat das Pferd den Reiter getroffen?		
		Schweif	Huf	Kopf

ID	LENGTH	WORD MATERIAL		
36	7	Kein einziges Tor fiel im gestrigen Spiel.		
easy question		Was fiel in dem Spiel nicht?		
		Entscheidung	Foul	Tor
hard question		Wie ging das Spiel aus?		
		Niederlage	Sieg	unentschieden
37	9	Der Skandal hat dem Ruf des Politikers deutlich geschadet.		
easy question		Was nahm durch den Skandal Schaden?		
		Ruf	Ansehen	Ehre
hard question		Wie wurde der Ruf des Politikers geschädigt?		
		stark	gering	gar nicht
38	7	Als Kapitalanlage ist Gold nicht zu empfehlen.		
easy question		Was eignet sich nicht als Kapitalanlage?		
		Geld	Gold	Öl
hard question		Eignet sich demnach Gold als Anlage?		
		manchmal	nein	ja
39	9	Markus klettert gern auf den alten Baum im Garten.		
easy question		Worauf klettert Markus gern?		
		Berg	Turm	Baum
hard question		Steht der Baum schon lange dort?		
		nein	vielleicht	ja
40	7	Sarah hat ihrem Opa ein Bild gemalt.		
easy question		Was hat Sarah gemalt?		
		Bild	Zeichnung	Blume
hard question		Wem malte Sarah ein Bild?		
		ihrem Großvater	ihrem Onkel	ihrem Vater
41	7	Medizinisch gesehen ist das Herz ein Hohlmuskel.		
easy question		Welches Organ ist ein Hohlmuskel?		
		Lunge	Herz	Milz
hard question		Wer bezeichnet das Herz als Hohlmuskel?		
		Fleischer	Ärzte	Organspender
42	7	Jede Sprache der Welt besitzt eine Grammatik.		
easy question		Was besitzt jede Sprache?		
		Vokale	Wörter	Grammatik
hard question		Wo besitzen Sprachen Grammatiken?		
		in Europa	auf der Erde	im Westen
43	7	Unsere Küche müßte dringend neu gestrichen werden.		
easy question		Was müßte gestrichen werden?		
		Küche	Flur	Bad
hard question		Wann sollte die Küche renoviert werden?		
		möglichst bald	im nächsten Jahr	nach dem Sommer

ID	LENGTH	WORD MATERIAL		
44	8	Der Politiker reagierte auf keine Frage der Journalisten.		
easy question		Worauf reagierte der Politiker nicht?		
		Vorwurf	Frage	Kritik
hard question		Wie oft antwortete der Politiker?		
		einmal	gar nicht	gelegentlich
45	9	Die Insel ist nur mit dem Flugzeug zu erreichen.		
easy question		Was ist nur per Flug zu erreichen?		
		Dorf	Berg	Insel
hard question		Wo reisen Touristen zu der Insel ab?		
		am Bahnhof	am Yachthafen	am Flughafen
46	8	Es sollte mehr Strom mit Solarenergie erzeugt werden.		
easy question		Was sollte mit Solarenergie erzeugt werden?		
		Strom	Energie	Wärme
hard question		Was sollte mehr gefördert werden?		
		Sonnenenergie	Windenergie	Wasserenergie
47	8	Die monotone Arbeit machte den Angestellten keinen Spaß.		
easy question		Was machte den Angestellten keinen Spaß?		
		Tätigkeit	Arbeit	Job
hard question		Wie war die Tätigkeit?		
		interessant	langweilig	aufregend
48	9	In dem kleinen Zimmer standen viel zu viele Möbel.		
easy question		Wo standen zu viele Möbel?		
		Raum	Haus	Zimmer
hard question		Wie war das Zimmer?		
		zu groß	zu leer	zu voll
49	9	Das Fenster im Flur klemmt seit ein paar Tagen.		
easy question		Was klemmt seit Tagen?		
		Fenster	Tür	Schublade
hard question		Was klemmt seit Tagen?		
		das Korridorfenster	das Küchenfenster	das Badfenster
50	9	Die Sekretärin informierte den Kanzler erst am nächsten Morgen.		
easy question		Wen informierte die Sekretärin?		
		Minister	Kanzler	Rektor
hard question		Wann informierte die Sekretärin den Kanzler?		
		sofort	am nächsten Tag	in der nächsten Stunde
51	9	Vielleicht gibt es bald im Dezember Ostereier zu kaufen.		
easy question		Wann gibt es vielleicht bald Ostereier?		
		November	Januar	Dezember
hard question		Wann gibt es vielleicht bald Ostereier?		
		zum Tag der Einheit	zu Ostern	zu Weihnachten

ID	LENGTH	WORD MATERIAL		
52	10	Das kleine Unternehmen konnte sich die teure Maschine nicht leisten.		
easy question		Was war zu teuer für das kleine Unternehmen?		
		Maschine	Gerät	Apparat
hard question		Wie war die Maschine?		
		nicht billig	günstig	bezahlbar
53	6	Der Franzose gewann gegen den Belgier.		
easy question		Wer besiegte den Belgier?		
		Spanier	Franzose	Holländer
hard question		Wen besiegte der Mann aus Frankreich?		
		den spanischen Gegner	den belgischen Gegner	den niederländischen Gegner
54	9	Jan hat sich zum dritten Mal die Schulter ausgekugelt.		
easy question		Was hat sich Jan ausgekugelt?		
		Hüfte	Daumen	Schulter
hard question		Wie oft hat sich Jan die Schulter ausgekugelt?		
		einmal	zweimal	dreimal
55	10	Der Bischof erschien mit seinem neuen Sekretär auf der Konferenz.		
easy question		Mit wem erschien der Bischof zur Konferenz?		
		Sekretär	Mitarbeiter	Dekan
hard question		Wer wurde kürzlich eingestellt?		
		der Sekretär	der Bischof	die Sekretärin
56	7	Das Schicksal führte die Freunde wieder zusammen.		
easy question		Wodurch wurden die Freunde zusammengeführt?		
		Zufall	Glück	Schicksal
hard question		Welche Menschen wurden wieder vereint?		
		verwandte Menschen	befreundete Menschen	zerstrittene Menschen
57	6	Vor Gericht wurde die Situation nachgestellt.		
easy question		Was wurde vor Gericht nachgestellt?		
		Unfall	Hergang	Situation
hard question		Wann wurde die Situation nachgestellt?		
		in der Beratung	in der Pause	im Prozeß
58	6	Das Wetter im September spielte verrückt.		
easy question		Wann spielte das Wetter verrückt?		
		September	August	November
hard question		Wie war das Wetter im September?		
		sehr wechselhaft	wie gewohnt	gleichbleibend
59	8	Die diesjährige Konferenz der Wissenschaftler dauerte vier Tage.		
easy question		Was dauerte vier Tage?		
		Tagung	Konferenz	Symposium
hard question		Wer hielt eine viertägige Tagung?		
		Wirtschaftler	Forscher	Dozenten

ID	LENGTH	WORD MATERIAL		
60	8	Der Hirte wanderte mehrere Kilometer durch die Wüs-te.		
easy question		Wer wanderte mehrere Kilometer?		
		Schäfer	Hirte	Tierpfleger
hard question		Beschäftigt sich dieser Wanderer beruflich mit Tieren?		
		möglicherweise	auf keinen Fall	bestimmt
61	7	Yvonne trat unglücklicherweise auf eine Tube Kleb-stoff.		
easy question		Worauf trat Yvonne?		
		Tube	Dose	Schachtel
hard question		Wo hängt Yvonne jetzt der Kleber?		
		am Schuh	an der Hand	am Knie
62	5	Manuela reagiert auf Senf allergisch.		
easy question		Worauf ist Manuela allergisch?		
		Wurst	Senf	Meerrettich
hard question		Wie bekommt Manuela Senf?		
		normal	schlecht	gut
63	6	Martins gebrochener Zeh schwoll rasch an.		
easy question		Was schwoll an?		
		Fuß	Ferse	Zeh
hard question		Was passierte mit Martins Zeh?		
		wurde blau	wurde steif	wurde dick
64	11	Der alte Kapitän goß stets ein wenig Rum in seinen Tee.		
easy question		Was goß der Kapitän in den Tee?		
		Rum	Milch	Zitrone
hard question		Wann trank der Kapitän seinen Tee mit Rum?		
		immer	manchmal	nie
65	9	Die schmale Öse ist zu klein für den Faden.		
easy question		Was ist zu klein?		
		Nadel	Öse	Loch
hard question		Wie ist die Öse?		
		zu lang	zu eng	zu weit
66	7	Die Tänzer probten ihre Kür besonders intensiv.		
easy question		Was probten die Tänzer besonders?		
		Tanz	Zugabe	Kür
hard question		Wie trainierten die Tänzer?		
		selten	so wie immer	sehr häufig
67	10	Nach dem Streit schien alles wieder im Lot zu sein.		
easy question		Wonach schien alles wieder im Lot zu sein?		
		Streit	Gefecht	Kampf
hard question		Ist nach dem Streit alles wieder in Ordnung?		
		scheinbar	nein	sicherlich

ID	LENGTH	WORD MATERIAL		
68	8	Der frisch gekochte Brei war noch zu heiß.		
easy question		Was war zu heiß?		
		Brei	Suppe	Püree
hard question		Was wurde zubereitet?		
		warmer Brei	kalter Brei	warmes Muß
69	7	Die Schneiderin steckte die Naht sorgfältig ab.		
easy question		Was steckte die Schneiderin ab?		
		Rock	Saum	Naht
hard question		Wie arbeitete die Schneiderin?		
		hastig	unsicher	gewissenhaft
70	8	Auf dem höchsten Mast hielt der Pirat Wache.		
easy question		Wo hielt der Pirat Wache?		
		Mast	Deck	Bug
hard question		Was tat der Pirat?		
		aufpassen	schlafen	singen
71	8	Claudia kann Salatsaucen mit viel Essig nicht ausstehen.		
easy question		Was mag Claudia nicht in Salatsaucen?		
		Senf	Essig	Öl
hard question		Was mag Claudia nicht zu sauer?		
		Suppe	Dressing	Obstsalat
72	8	Die Torte erwies sich als ein wahrer Leckerbissen.		
easy question		Was war lecker?		
		Kaffee	Kuchen	Torte
hard question		Was schmeckte besonders gut?		
		der Kaffee	die Kekse	der Kuchen
73	7	Sie machten einen Spaziergang am Deich entlang.		
easy question		Wo gingen sie spazieren?		
		Deich	Damm	Graben
hard question		Was taten sie am Deich?		
		gehen	radfahren	joggen
74	7	Ulf hat schon wieder eine Niete gezogen.		
easy question		Was hat Ulf gezogen?		
		Los	Niete	Gewinn
hard question		Hat Ulf diesmal Glück im Spiel?		
		ja	nein	vielleicht
75	7	Der Giebel des alten Hauses drohte einzustürzen.		
easy question		Was stürzt bald ein?		
		Erker	Dach	Giebel
hard question		Was schien unsicher?		
		Teile am Balkon	Teile am Fundament	Teile am Dach

ID	LENGTH	WORD MATERIAL		
76	7	Das Karamellbonbon blieb Julia am Gaumen kleben.		
easy question		Wo klebte das Bonbon?		
		Gaumen	Zähne	Rachen
hard question		Wo klebte das Bonbon?		
		im Mund	am Daumen	am Schuh
77	7	Tamara führte mit der Hebamme mehrere Gespräche.		
easy question		Mit wem sprach Tamara?		
		Ärztin	Hebamme	Schwester
hard question		War die Geburtshelferin gesprächig?		
		nein	ja	vielleicht
78	9	Nach der Trauung wartete eine Kutsche vor der Kirche.		
easy question		Was wartete vor der Kirche?		
		Limousine	Gespann	Kutsche
hard question		Wann sollte das Brautpaar die Kutsche nehmen?		
		vor der Trauung	während des Gottesdienstes	nach der Zeremonie
79	8	Robert ließ sich den Schinken in Scheiben schneiden.		
easy question		Was ließ sich Robert schneiden?		
		Schinken	Speck	Wurst
hard question		Was geschah mit dem Schinken?		
		wurde zerteilt	wurde zerkocht	wurde zerhackt
80	8	Die Kinder hüpften auf der alten Matratze herum.		
easy question		Worauf hüpften die Kinder?		
		Bett	Matratze	Kissen
hard question		Was taten die Kinder?		
		rennen	springen	herumsitzen
81	8	Laura stellte eine Schüssel Kirschen auf den Tisch.		
easy question		Worin waren die Kirschen?		
		Schale	Korb	Schüssel
hard question		Was steht jetzt auf dem Tisch?		
		eine Dose	ein Korb	eine Schale
82	5	Das entscheidende Telefonat verzögerte sich.		
easy question		Was verzögerte sich?		
		Telefonat	Gespräch	Treffen
hard question		Ist das Gespräch wichtig?		
		ja	nein	vielleicht
83	6	Mandy aß die Mandarine sofort auf.		
easy question		Was aß Mandy?		
		Nektarine	Mandarine	Clementine
hard question		Wartete Mandy mit dem Essen der Mandarine?		
		ja	nein	vielleicht

ID	LENGTH	WORD MATERIAL		
84	8	Der Teppich mit dem Schnörkel gefiel Bettina nicht.		
easy question		Was gefiel Bettina an dem Teppich nicht?		
		Muster	Kringel	Schnörkel
hard question		Was gefiel Bettina an dem Teppich nicht?		
		die Farbe	die Größe	die Verzierung
85	7	Der majestätische Gletscher wurde schon oft bestiegen.		
easy question		Wer ist majestätisch?		
		Gletscher	Berg	Eisberg
hard question		Wen besteigen die Leute häufig?		
		den imposanten Gletscher	den kleinen Gletscher	den tauenden Gletscher
86	8	Der Großvater fand in seinem Pantoffel eine Wäscheklammer.		
easy question		Wo war die Wäscheklammer?		
		Hausschuh	Pantoffel	Schlafrock
hard question		Wo war die Wäscheklammer?		
		im Stiefel	im Hausschuh	im Wäschekorb
87	5	Johannes wollte unbedingt Karussell fahren.		
easy question		Womit wollte Johannes fahren?		
		Achterbahn	Eisenbahn	Karussell
hard question		Fährt Johannes gerne Karussell?		
		nein	unwahrscheinlich	wahrscheinlich
88	8	Die Olympiade findet dieses Jahr in Australien statt.		
easy question		Was fand in Australien statt?		
		Olympiade	Paralympics	Meisterschaft
hard question		Was macht Australien mit diesen Wettkämpfen?		
		austragen	boykottieren	unterstützen
89	5	Sonja kam als Einzige pünktlich.		
easy question		Was tat Sonja?		
		ging	kam	rannte
hard question		Was tat Sonja?		
		abreisen	eintreffen	spazieren
90	8	Der alte Mann zog einen Karren zum Marktplatz.		
easy question		Was tat der Mann mit dem Karren?		
		schob	zerrte	zog
hard question		Was ist das Ziel des Mannes?		
		der Kirchplatz	die Markthalle	der Verkaufsplatz
91	6	Keiner wußte, was zu tun war.		
easy question		Wer wusste, was zu tun war?		
		alle	jeder	keiner
hard question		Hatte jemand Erfahrung mit solch einer Situation?		
		nein	ja	vielleicht

ID	LENGTH	WORD MATERIAL		
92	8	Heute morgen saß auf unserer Terrasse ein Frosch.		
easy question		Was tat der Frosch auf unserer Terrasse?		
		hockte	saß	lag
hard question		Was tat der Frosch auf der Terrasse?		
		springen	hocken	schwimmen
93	8	Der Wandersmann bat den Wirt um etwas Wasser.		
easy question		Was tat der Wanderer?		
		fragte	bettelte	bat
hard question		Worum bat der Wanderer?		
		Essen	Unterkunft	Trinken
94	8	Seit gestern geht es dem Patienten deutlich besser.		
easy question		Wem geht es besser?		
		Patienten	Klienten	Pfleger
hard question		Was tut der Kranke?		
		genesen	krank bleiben	Medizin nehmen
95	7	Der Waffenstillstand hält seit fast vier Monaten.		
easy question		Seit wann hält der Waffenstillstand?		
		Tagen	Monaten	Jahren
hard question		Wie ist der Waffenstillstand bisher?		
		instabil	andauernd	in Diskussion
96	6	Freundlich zu sein kann nie schaden.		
easy question		Wie zu sein kann nie schaden?		
		höflich	nett	freundlich
hard question		Was soll man tun?		
		schaden	unfreundlich sein	nett sein
97	8	Die Mutter gibt ihren Kindern jeden Montag Taschengeld.		
easy question		Wem gibt die Mutter Taschengeld?		
		Neffen	Kindern	Enkeln
hard question		Wann erhalten die Kinder ihr Geld?		
		am Wochenbeginn	am Monatsanfang	zu Tagesbeginn
98	7	Jeder Vorschlag gilt, der rechtzeitig eingereicht wird.		
easy question		Was gilt?		
		Antrag	Vorschlag	Idee
hard question		Gilt Fristgerechtes?		
		nein	manchmal	ja
99	9	Ich weiß nicht, ob wir uns noch sehen werden.		
easy question		Was tut die Person nicht?		
		wissen	kennen	ahnen
hard question		Steht ein baldiges Treffen schon fest?		
		vielleicht	ja	nein

ID	LENGTH	WORD MATERIAL		
100	7	Die meisten Kinder gehen gerne zur Schule.		
easy question		Wohin gehen die meisten Kinder gern?		
		Schule	Kindergarten	Sport
hard question		Was tun viele Kinder gerne?		
		Unterricht besuchen	Schulaufgaben machen	Schule schwänzen
101	6	Viele Kinder lesen nur noch selten.		
easy question		Was tun manche Kinder nur noch selten?		
		erzählen	lesen	singen
hard question		Wie oft lesen viele Kinder?		
		häufig	kaum	immer
102	8	Die Journalisten fragen den Bürgermeister nach seiner Meinung.		
easy question		Wonach fragten die Journalisten den Bürgermeister?		
		Ansicht	Meinung	Sichtweise
hard question		Was soll der Bürgermeister im Interview sagen?		
		sein Wissen	seine Ansicht	sein Programm
103	7	Kevin und Marie spielen oft im Garten.		
easy question		Was tun Kevin und Marie im Garten?		
		spielen	toben	raufen
hard question		Wo spielen Kevin und Marie oft?		
		draußen	drinnen	im Wintergarten
104	7	Die Geschworenen glauben dem Beklagten bestimmt alles.		
easy question		Was tun die Geschworenen?		
		zweifeln	glauben	wissen
hard question		Sind die Geschworenen überzeugt?		
		auf keinen Fall	mit hoher Wahrscheinlichkeit	mit Sicherheit
105	10	Am besten stellen wir das Klavier nicht direkt ans Fenster.		
easy question		Wohin sollen sie das Klavier nicht stellen?		
		Heizung	Fenster	Tür
hard question		Was soll nicht am Fenster stehen?		
		das Cello	die Harfe	das Piano
106	7	Meistens wünschen Kinder sich Spielzeug zu Weihnachten.		
easy question		Was wünschen Kinder sich zu Weihnachten?		
		Tiere	Spielzeug	Gameboy
hard question		Wollen Kinder Spielsachen zum Fest geschenkt bekommen?		
		selten	sehr häufig	grundsätzlich

ID	LENGTH	WORD MATERIAL		
107	6	Einige Häftlinge sprechen nicht gern miteinander.		
easy question		Was tun einige Häftlinge nur ungern miteinander?		
		reden	sprechen	arbeiten
hard question		Was tun einige Inhaftierte nur ungern miteinander?		
		arbeiten	spielen	reden
108	7	Nur wenige Menschen brauchen ein Handy wirklich.		
easy question		Was brauchen die wenigsten Menschen?		
		Telefon	Handy	Computer
hard question		Was brauchen wenige Menschen?		
		ein Telefon	ein Mobiltelefon	ein Faxgerät
109	6	Einige der Angestellten arbeiten nur vormittags.		
easy question		Was tun einige der Angestellten nur vormittags?		
		arbeiten	einkaufen	schlafen
hard question		Wie arbeiten ein paar der Angestellten?		
		halbtags	ganztags	am Wochenende
110	9	Die Spieler hoffen, daß sie ihre Gegner schlagen werden.		
easy question		Wer hofft zu gewinnen?		
		Sportler	Spieler	Mannschaft
hard question		Wer soll besiegt werden?		
		Mitspieler	die andere Mannschaft	die Trainer
111	10	Ich bin nicht sicher, ob alle die Prüfung schaffen werden.		
easy question		Was werden einige nicht schaffen?		
		Prüfung	Test	Examen
hard question		Werden alle den Test bestehen?		
		vielleicht	mit Sicherheit	auf keinen Fall
112	10	Die meisten Geschäfte schließen samstags früher als unter der Woche.		
easy question		Wer schließt samstags früher?		
		Läden	Märkte	Geschäfte
hard question		Sind die Öffnungszeiten wochentags länger?		
		nein	selten	oft
113	7	Die Astronauten antworten seit Tagen nicht mehr.		
easy question		Wer antwortet nicht mehr?		
		Raumfahrer	Taucher	Astronauten
hard question		Wen konnte man nicht mehr erreichen?		
		die Basisstation	die Piloten	die Raumfahrer

ID	LENGTH	WORD MATERIAL		
114	8	Die Schüler schreiben ihrer kranken Lehrerin einen Brief.		
easy question		Wer schreibt einen Brief?		
		Klasse	Schüler	Kinder
hard question		Wer soll einen Brief erhalten?		
		der kranke Lehrer	die kranke Lehrerin	die kranke Schülerin
115	6	Die Beschuldigten schweigen zu den Vorwürfen.		
easy question		Wozu schweigen die Beschuldigten?		
		Vorwürfe	Anschuldigungen	Fragen
hard question		Was äußern die Angeklagten zu den Beschuldigungen?		
		gar nichts	Verteidigungen	Ausreden
116	7	Die beiden Mädchen schütteln sich vor lachen.		
easy question		Wer schüttelt sich vor Lachen?		
		Mädchen	Kinder	Frauen
hard question		Haben die zwei Kinder Spaß ?		
		keinesfalls	möglich	offensichtlich
117	8	Dorothea log oft bei Fragen nach ihrem Alter.		
easy question		Was tat Dorothea oft?		
		lügen	mogeln	schummeln
hard question		Steht Dorothea zu ihrem Alter?		
		selten	nie	immer
118	9	Die Großmutter wog die Zutaten beim Backen sehr genau.		
easy question		Wer wog die Zutaten ab?		
		Oma	Großmutter	Tante
hard question		Wie buk die Großmutter?		
		nach freiem Rezept	meßgenau	nach Gefühl
119	7	Die Mutter sagte, Nina übe gerade Klavier.		
easy question		Wer sagte, Nina übe Klavier?		
		Mutter	Oma	Tante
hard question		Wann übt Nina Klavier?		
		jetzt	morgen	nachher
120	8	Der Gehilfe des Gärtners sät Kresse und Radieschen.		
easy question		Wer sät Pflanzen?		
		Helfer	Gehilfe	Lehrling
hard question		Wer sät Pflanzen?		
		der Gärtnerlehrling	der Gärtnergehilfe	der Gärtnerbursche
121	7	Die zwei Nichten öden sich gegenseitig an.		
easy question		Wer ödet sich an?		
		Cousinen	Tanten	Nichten
hard question		Was tun die beiden Mädchen?		
		sich unterhalten	sich interessieren	sich langweilen

ID	LENGTH	WORD MATERIAL		
122	10	Sei so gut und miß bitte die Tiefe des Regals.		
easy question		Was soll am Regal gemessen werden?		
		Breite	Höhe	Tiefe
hard question		Was soll vorgenommen werden?		
		eine Berechnung	eine Schätzung	eine Messung
123	8	Jetzt rate doch mal, wen ich gesehen habe!		
easy question		Was soll getan werden?		
		spekulieren	denken	raten
hard question		War der Befragte bei der Begegnung dabei?		
		bestimmt	vielleicht	keinesfalls
124	9	Sei so lieb und lies mir die Angaben vor!		
easy question		Was soll vorgelesen werden?		
		Anleitung	Angaben	Hinweise
hard question		Was soll getan werden?		
		eine Arbeit	ein Gefallen	eine Pflicht
125	8	Bitte wirf den Ball nicht wieder aufs Dach.		
easy question		Wohin soll der Ball nicht geworfen werden?		
		Dach	Regenrinne	Garage
hard question		Was soll man mit dem Ball nicht tun?		
		schmeißen	rollen	schießen
126	6	Bitte hilf deiner Schwester beim Aufräumen.		
easy question		Wem soll geholfen werden?		
		Schwester	Bruder	Geschwister
hard question		Was soll getan werden?		
		Ordnung machen	saubermachen	renovieren
127	8	Die streikenden Fahrer konnte man kilometerweit hupen hören.		
easy question		Wer streikt?		
		Bauern	Fahrer	Spediteure
hard question		Wie laut waren die Streikenden?		
		leise	sehr laut	normal
128	9	Die Gärtner mähen den Rasen im Park jeden Mittwoch.		
easy question		Wer mäht den Rasen?		
		Förster	Gärtner	Arbeiter
hard question		Was machen die Gärtner wöchentlich mit dem Grün?		
		düngen	kürzen	rechen
129	9	Gute Beziehungen ebnen vielen Unternehmern den Weg zum Erfolg.		
easy question		Wem helfen gute Beziehungen?		
		Geschäftsleuten	Unternehmen	Unternehmern
hard question		Wem helfen gute Beziehungen?		
		Geschäftspartnern	Unternehmen	Unternehmern

ID	LENGTH	WORD MATERIAL		
130	7	Mäuse und Ratten nagen gerne an Stromkabeln.		
easy question		Was tun Mäuse mit Kabeln?		
		knabbern	fressen	nagen
hard question		Woran knabbern Mäuse und Ratten gerne?		
		an Wollfäden	an Gummibändern	an Elektrizitäts-leitern
131	6	Tierärzte impfen keine Kaninchen gegen Tollwut.		
easy question		Wen impfen Tierärzte nicht?		
		Hasen	Kaninchen	Hamster
hard question		Wer impft keine Kaninchen?		
		Förster	Mediziner	Schwestern
132	8	Die Hunde der Wächter bellen beim geringsten Anlaß.		
easy question		Wem gehören die Hunde?		
		Wärter	Wächter	Wachmänner
hard question		Wem gehören die Hunde?		
		den Wärtern	den Wachmännern	den Lieferanten
133	7	Affen kraulen sich oft stundenlang das Fell.		
easy question		Was tun Affen gegenseitig mit ihrem Fell?		
		streicheln	lausen	kraulen
hard question		Was tun Affen oft über Stunden?		
		sich schrubben	sich umarmen	sich sanft kratzen
134	9	Die Forscher stapfen durch den Schnee zurück zum La-ger.		
easy question		Wer stapft zum Lager zurück?		
		Wanderer	Alpinisten	Forscher
hard question		Wer geht zum Lager zurück?		
		Wanderer	Skifahrer	Wissenschaftler
135	9	Viele Babys schielen nach der Geburt eine Weile lang.		
easy question		Wer schielt oft nach der Geburt?		
		Babys	Säuglinge	Kinder
hard question		Was ist nach der Geburt bei Babies oft beeinträchtigt?		
		die Augenstellung	der Geruchssinn	das Gehör
136	9	Die Bäume in den Wäldern speichern sehr viel Wasser.		
easy question		Wer speichert viel Wasser?		
		Bäume	Sträucher	Pflanzen
hard question		Was speichern die Bäume?		
		Flüssigkeit	Nährstoffe	Wärme
137	7	Die meisten Leute schummeln beim Spielen gelegent-lich.		
easy question		Was tun viele Leute?		
		mogeln	schummeln	betrügen
hard question		Mogeln viele Menschen beim Spielen?		
		nein	ab und zu	ja

ID	LENGTH	WORD MATERIAL		
138	10	Nach dem Spiel massieren die Therapeuten den Spielern die Beine.		
easy question		Wonach werden den Spielern die Beine massiert?		
		Match	Training	Spiel
hard question		Wen sehen die Spieler nach dem Match?		
		die Trainer	die Ärzte	die Masseure
139	7	Die beiden Mädchen tuscheln während des Unterrichts.		
easy question		Wann tuscheln die Mädchen?		
		Stunde	Unterricht	Pause
hard question		Wer schwatzt während der Stunde?		
		zwei Schüler	zwei Schülerinnen	drei Schülerinnen
140	8	Die Häuser am Horizont flimmern in der Sonne.		
easy question		Was flimmert am Horizont?		
		Städte	Gebäude	Häuser
hard question		Ist es an dem Ort warm?		
		nein	unwahrscheinlich	wahrscheinlich
141	10	Den ganzen Tag über konnte man die Raben krächzen hören.		
easy question		Wer krächzte?		
		Krähen	Vögel	Raben
hard question		Wann konnte man die Raben hören?		
		am Abend	nachts	von morgens bis abends
142	7	Vor dem Auftritt schminken die Schauspieler sich.		
easy question		Wer schminkt sich?		
		Schauspieler	Künstler	Sänger
hard question		Wie gehen die Schauspieler auf die Bühne?		
		geschminkt	ungeschminkt	frisiert
143	6	Die zwei frechen Jungs heucheln Unschuld.		
easy question		Wer heuchelt Unschuld?		
		Buben	Jungs	Männer
hard question		Was tun die Jungs?		
		lächeln	lügen	provozieren
144	5	Manche Menschen stottern bei Nervosität.		
easy question		Bei welchem Zustand stottern manche Menschen?		
		Nervosität	Anstrengung	Streß
hard question		Wer stottert bei Nervosität?		
		einige Leute	alle Leute	sehr viele Leute

B. Sentences for Proofreading with Word Form Errors

ID	LENGTH	WORD MATERIAL
145	8	Thomas hat die geliehene Kassette noch nichtt zurückgegeben.
146	9	Ich habe mit großem Vergnügen die amüsnaten Filmkritiken gelesen.
147	9	Man sollte versuchen, abwechselnd kalt und warmm zu duschen.
148	7	Seiner Tchter erteilte der Vater eine Rüge.
149	8	Am Montag fahren wir anch Hamburg zum Einkaufen.
150	9	Gerade von einem Therapeiten hätte man es verlangen können.
151	9	Dieses Jahr wird die Erlte mal wieder schlecht ausfallen.
152	10	Spielsucht ist einar der Gründe, warum er ständig pleite ist.
153	7	Als Jugendlicher war er eiu ernster Mensch.
154	8	Ohne Zeichen inerer Anteilnahme hörte er das Urteil.
155	10	Nach seinen Niederlagen steht Gerhard ein Ertfolg gut zu Gesicht.
156	9	Von seinen Eltern hat er einen seeltenen Vornamen bekommen.
157	7	Fast drei Stückchen Kulchen hat er gegessen.
158	7	Wie alle mussten sie sich esrt bewerben.
159	7	Das was mal wieder typisch für Klaus.
160	8	Der Privoant bestand aus Speck, Butter und Erbsen.
161	7	Bei schwerem Strum ist ein Frachter gesunken.
162	10	Polizeikräfte bereiten sich in volller Montur auf die Paraden vor.
163	7	Aufwendig dekorierte Fassahden prägen das Zentrum Londons.

ID	Length	Word material
164	7	Unter den Verletzten war acuh der Bräutigam.
165	8	Meine Schwägerin hate sich schon ihre Mütze aufgesetzt.
166	8	Nicht der Preis ist entscheidnd, sondern die Qualität.
167	8	Offen blieb noch, wan das Enkelkind getauft wird.
168	9	Wir wären ja törcht, wenn wir das nicht täten.
169	10	Alte Männer haben manchmal einen sletsamen Blick auf die Dinge.
170	10	Für ein paar Tage treffen sick Menschen aus aller Welt.
171	12	Der gute alte Mann ist mit dem Plerd auf dem Eis eingebrochen.
172	12	Die Künstlerin musste wegen Trunkenheit am Steuer für zwi Tage ins Gefängnis.
173	7	Große Krisen schlechen sich meist langsam an.
174	9	Sie sorgt für den Gesang, er frü die Musik.
175	10	Bislang war er ein Mannt mit Verständnis für große Sorgen.
176	7	Anders ist es wohl niht zu verstehen.
177	7	Unser Nachbar hat sich die Nsae gebrochen.
178	8	Die Idee hatten ein Fotograhf und eine Galeristin.
179	9	Seit Jahren wird über eine Saniirung des Gebäudes diskutiert.
180	7	Sie wohnen imer noch bei ihren Eltern.
181	8	Gegen Mittag warn zwei Drittel der Kandidaten ausgeschieden.
182	7	Offizielle Reden hat er selten holten müssen.
183	7	Ihr Bruder und sie könen Reifen wechseln.
184	8	Nach wachsendem Druck hat siach Kurt offiziell entschuldigt.
185	8	Die Mediziner verloren den Wettkampf met der Zeit.
186	8	Sie sind sich gegenseitig nur Projektion umd Hoffnung.
187	9	Vielleicht haben sich diie Tänzer ihren Traum selbst ruiniert.
192	8	Eure Berpe sind nicht so schön wie unsere.
193	7	Heute sind alle Bouern draußen und ernten.
194	8	Es muss einmal klipp und kla gesagt werden.
195	8	Gerüchte verwies die Managerin isn Reich der Fantasie.

ID	LENGTH	WORD MATERIAL
196	9	Die japanische Raumfahrt erleabte am Wochenende einen weiteren Absturz.
197	9	In der einst vergifteten Kegend soll ein Naturpark entstehen.
198	9	Der Prinz kann Alkhol und schönen Frauen nicht widerstehen.
199	8	Die Leute trieben die Mülchkühe auf die Weide.
200	11	Neben einem Geldpreis gehört zum ersten Presi die Verfilmung der Idee.
188	9	Die Erinnerung daran triebt mich schaudernd ins kalte Badezimmer.
189	9	Reichlich Pfnude auf Po und Rippen sind ausdrücklich erwünscht.
190	9	Ganz schön ist der russzische Walzer in prachtvollen Kostümen.
191	10	Mittlerweile sucht die Seuche acht andere Staahten der Region heim.

C. Psychometric Data for all Participants

Table C.1.: Psychometric data for subjects in 'original young'.

ID	Age in years	Vocabulary score (Lehrl, 1977)	Digit-Symbol-Test (Wechsler, 1964)
1	17	30	56
2	17	32	86
3	18	34	58
4	17	33	64
5	16	26	48
6	18	33	42
7	18	29	61
8	18	32	71
9	18	29	54
10	18	30	63
11	17	29	66
12	18	32	47
13	18	26	66
14	18	28	69
15	18	26	72
16	18	33	61
17	17	27	61
18	18	30	61
19	17	33	73
20	18	27	51
21	18	26	66
22	18	26	65
23	17	31	45
24	18	26	73

Table C.2.: Psychometric data for subjects in 'original old'.

ID	Age in years	Vocabulary score (Lehrl, 1977)	Digit-Symbol-Test (Wechsler, 1964)
1	66	33	39
2	70	32	58
3	74	35	44
4	73	33	47
5	74	34	61
6	72	32	63
7	74	32	50
8	73	34	66
9	71	36	60
10	66	31	50
11	74	33	55
12	71	35	43
13	75	33	43
14	71	33	38
15	70	34	52
16	67	34	45
17	66	33	65
18	65	34	50
19	74	33	36
20	68	34	48
21	74	31	42
22	69	32	44
23	72	34	44
24	72	33	41
25	71	33	69
26	66	32	36
27	70	33	60
28	66	33	38
29	84	33	48
30	71	33	69
31	66	33	45
32	66	30	36

Table C.3.: Psychometric data for subjects in 'original'.

ID	Age in years	Vocabulary score (Lehrl, 1977)	Digit-Symbol-Test (Wechsler, 1964)
1	20	35	78
2	19	33	65
3	28	34	71
4	21	33	62
5	24	32	77
6	19	32	72
7	20	33	67
8	21	33	55
9	24	32	64
10	22	32	85
11	20	33	60
12	21	34	83
13	27	33	53
14	22	32	63
15	20	34	64
16	24	34	66
17	21	32	68
18	22	32	59
19	20	34	68
20	26	33	76
21	21	33	77
22	22	32	64
23	21	33	71
24	22	32	76
25	22	32	78
26	24	31	65
27	22	32	61
28	20	33	58
29	21	33	67
30	24	34	78
31	20	33	67
32	23	34	64
33	21	31	58

Table C.4.: Psychometric data for subjects in 'hard young'.

ID	Age in years	Vocabulary (Lehrl, 1977)	Digit-Symbol (Wechsler, 1964)	reading span	fig. memory
1	19	32	53	0.9	0.88
2	19	28	78	0.97	0.87
3	18	31	57	0.73	0.81
4	19	30	68	0.73	0.83
5	19	32	46	0.68	0.82
6	19	29	55	0.69	0.72
7	18	29	65	0.77	0.83
8	18	32	58	0.85	0.81
9	18	28	48	0.56	0.78
10	18	29	53	0.82	0.88
11	20	31	60	0.85	0.88
12	19	30	70	0.82	0.88
13	18	31	45	0.66	0.81
14	18	31	55	0.65	0.84
15	18	35	70	0.89	0.88
16	18	29	75	0.61	0.85
17	19	32	67	0.73	0.86
18	19	25	80	0.74	0.88
19	17	33	54	0.76	0.76
20	20	28	69	0.61	0.75
21	19	27	68	0.89	0.92
22	18	31	53	0.73	0.66
23	20	33	68	0.95	0.82
24	17	32	64	0.71	0.81
25	19	28	59	0.76	0.78
26	18	39	58	0.81	0.86
27	20	31	71	0.81	0.77
28	17	28	55	0.48	0.72
29	18	33	60	0.87	0.84
30	18	35	58	0.9	0.93

Table C.5.: Psychometric data for subjects in 'hard old'.

ID	Age in years	Vocabulary (Lehrl, 1977)	Digit-Symbol (Wechsler, 1964)	reading span	fig. memory
1	65	30	33	0.4	0.61
2	69	34	40	0.37	0.72
3	74	36	35	0.32	0.68
4	76	31	45	0.35	0.64
5	68	34	47	0.68	0.73
6	66	32	42	0.29	0.57
7	68	33	52	0.42	0.69
8	66	33	44	0.35	0.6
9	70	31	54	0.48	0.7
10	69	34	54	0.55	0.67
11	65	32	58	0.5	0.64
12	65	33	41	0.19	0.69
13	65	32	55	0.65	0.75
14	67	34	39	0.4	0.65
15	67	33	47	0.39	0.67
16	67	34	54	0.23	0.58
17	65	32	49	0.58	0.69
18	66	34	38	0.58	0.67
19	69	33	35	0.37	0.72
20	67	35	47	0.5	0.75
21	65	31	60	0.53	0.7
22	71	33	39	0.39	0.66
23	75	34	58	0.56	0.71

Table C.6.: Psychometric data for subjects in 'frequent'.

ID	Age in years	Vocabulary score (Lehrl, 1977)	Digit-Symbol-Test (Wechsler, 1964)
1	25	28	73
2	21	31	63
3	21	33	62
4	22	33	54
5	24	27	71
6	21	33	59
7	20	35	57
8	23	33	70
9	21	30	56
10	24	-	-
11	27	30	73
12	30	32	52
13	19	31	56
14	36	33	61
15	23	34	60
16	24	35	60
17	21	26	89
18	22	33	54
19	22	35	47
20	22	33	53
21	23	35	57
22	26	29	47
23	27	35	58
24	21	30	63
25	35	35	43
26	27	33	83
27	20	27	56
28	20	31	82
29	28	32	69
30	24	27	60

Table C.7.: Psychometric data for subjects in 'proofreading'.

ID	Age in years	Vocabulary score (Lehrl, 1977)	Digit-Symbol-Test (Wechsler, 1964)
1	21	32	62
2	21	29	66
3	21	31	62
4	20	29	54
5	20	33	59
6	22	33	82
7	23	31	48
8	29	35	58
9	27	31	69
10	21	30	60
11	24	31	55
12	23	32	63
13	21	31	75
14	20	30	63
15	27	32	68
16	21	32	63
17	23	31	59
18	22	32	65
19	27	30	60
20	21	-	-
21	21	33	72
22	20	30	79
23	24	31	68
24	26	35	57
25	22	30	49
26	21	32	62
27	32	31	73
28	22	33	67
29	22	35	53
30	20	30	63

D. Output of lmer-Analysis for Various Group Comparisons

Table D.1.: Final LMM fitting log gaze duration for the samples 'original young' and 'hard young'.

Linear mixed-effects model fit by maximum likelihood			
RANDOM EFFECTS:			
Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0142989	0.119578
sentence id	(Intercept)	0.0016781	0.040965
subject id	(Intercept)	0.0132475	0.115098
Residual		0.0734312	0.270982
number of obs: 3948, groups: wid, 504; sn, 144; id, 54			
FIXED EFFECTS:			
	Estimate	Std. Error	t-value
(Intercept)	5.9245995	0.0300218	197.34
trial	-0.0004057	0.0001088	-3.73
vocabulary	-0.0124083	0.0060463	-2.05
pc1	0.0329241	0.0088280	3.73
pc2	-0.0166906	0.0114274	-1.46
condition(cnd)	0.0232783	0.0348903	0.67
frequency(n)	-0.0637593	0.0098872	-6.45
frequency(n+1)	-0.0243115	0.0063319	-3.84
predictability(n)	-0.0283618	0.0086406	-3.28
predictability(n+1)	0.0103939	0.0066728	1.56
1/length(n)	-0.6277035	0.1438315	-4.36
1/length(n)*freq(n)	0.6720729	0.0835288	8.05
1/length(n)*freq(n+1)	0.2271170	0.0766927	2.96

Table D.2.: Final LMM fitting log single fixation duration for the samples 'original young' and 'hard young'.

Linear mixed-effects model fit by maximum likelihood			
RANDOM EFFECTS:			
Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0098712	0.099354
sentence id	(Intercept)	0.0018903	0.043477
subject id	(Intercept)	0.0124744	0.111689
Residual		0.0849521	0.291465
number of obs: 21738, groups: wid, 550; sn, 144; id, 54			
FIXED EFFECTS:			
	Estimate	Std. Error	t-value
(Intercept)	5.341e+00	2.827e-02	188.89
trial	-4.053e-04	6.971e-05	-5.81
vocabulary	-1.273e-02	5.893e-03	-2.16
condition (cnd)	1.712e-03	3.322e-02	0.05
pc1	1.737e-02	5.815e-03	2.99
pc2	-3.109e-02	1.238e-02	-2.51
poly(frequency(n))1	-5.016e+00	1.350e+00	-3.71
poly(frequency(n))2	-2.805e+00	1.136e+00	-2.47
poly(frequency(n))3	-5.317e+00	6.398e-01	-8.31
predictability(n)	-1.979e-02	3.566e-03	-5.55
1/length(n)	2.360e-02	8.704e-02	0.27
length 6,7 letters	-4.981e-02	1.154e-02	-4.32
frequency(n-1)	-3.664e-02	3.887e-03	-9.43
predictability (n-1)	-1.379e-02	3.794e-03	-3.63
1/ length(n -1)	3.296e-01	4.296e-02	7.67
frequency(n+1)	-2.135e-02	3.841e-03	-5.56
predictability(n+1)	1.168e-02	3.540e-03	3.30
1/length(n+1)	1.030e-01	4.143e-02	2.49
incoming sacc.ampl.	3.286e-02	1.379e-03	23.83
rel. fix. position	-1.097e-01	1.542e-02	-7.12
(rel. fix. position)^2	-3.204e-01	3.686e-02	-8.69
outgoing sacc.ampl.	1.318e-02	1.553e-03	8.49
length*freq.(n)	1.963e-01	6.383e-02	3.07
freq.(n-1)*freq.(n)	1.284e-02	2.048e-03	6.27
freq.(n)*freq.(n+1)	7.703e-03	2.153e-03	3.58
cnd*trial	2.946e-04	9.804e-05	3.00
cnd*poly(freq(n))1	-1.278e+00	7.796e-01	-1.64
cnd*poly(freq(n))2	2.860e+00	7.413e-01	3.86
cnd*poly(freq(n))3	-1.489e-01	6.102e-01	-0.24
cnd*1/length(n)	1.415e-01	6.398e-02	2.21
cnd*predictability(n+1)	8.428e-03	3.351e-03	2.52
cnd*incoming sacc.ampl.	-6.967e-03	1.747e-03	-3.99
cnd*rel.fix.position	7.418e-02	1.878e-02	3.95
cnd*outgoing sacc.ampl.	1.332e-02	2.184e-03	6.10
cnd*length(n)*freq(n)	-1.642e-01	4.677e-02	-3.51

Table D.3.: Final LMM fitting log single fixation duration for the samples 'original old' and 'hard old'.

Linear mixed-effects model fit by maximum likelihood

RANDOM EFFECTS:

Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0068952	0.083037
sentence id	(Intercept)	0.0019077	0.043678
subject id	(Intercept)	0.0185159	0.136073
Residual		0.0832194	0.288478

number of obs: 18551, groups: wid, 550; sn, 144; id, 55

FIXED EFFECTS:

	Estimate	Std. Error	t-value
(Intercept)	5.407e+00	3.439e-02	157.23
trial	-2.958e-04	6.496e-05	-4.55
vocabulary	-2.557e-02	1.460e-02	-1.75
condition (cnd)	2.269e-03	4.279e-02	0.05
pc1	1.157e-03	7.808e-03	0.15
pc2	-3.334e-02	1.301e-02	-2.56
poly(frequency(n))1	-5.055e+00	1.058e+00	-4.78
poly(frequency(n))2	1.469e+00	9.097e-01	1.61
poly(frequency(n))3	-1.861e+00	5.506e-01	-3.38
predictability(n)	-2.011e-02	3.631e-03	-5.54
1/length(n)	1.122e-01	7.092e-02	1.58
length 6, 7 letters	-4.318e-02	1.033e-02	-4.18
frequency(n-1)	-3.539e-02	4.154e-03	-8.52
predictability(n-1)	-7.337e-03	3.879e-03	-1.89
1/length(n-1)	2.466e-01	4.423e-02	5.57
frequency(n+1)	-1.690e-02	4.006e-03	-4.22
predictability(n+1)	6.548e-03	3.101e-03	2.11
1/length(n+1)	1.766e-01	4.178e-02	4.23
incoming sacc.ampl.	2.722e-02	9.937e-04	27.40
rel.fix position	-2.667e-02	1.186e-02	-2.25
$(\text{rel.fix.position})^2$	-1.679e-01	3.644e-02	-4.61
outgoing sacc.ampl.	8.937e-03	1.231e-03	7.26
freq(n-1)*freq(n)	9.341e-03	2.036e-03	4.59
1/length(n)*freq(n)	1.584e-01	5.393e-02	2.94
freq(n)*freq(n+1)	5.925e-03	2.129e-03	2.78
cnd*trial	3.935e-04	1.113e-04	3.54
cnd*poly((freq(n))1	2.626e+00	6.482e-01	4.05
cnd*poly((freq(n))2	5.285e-01	6.577e-01	0.80
cnd*poly((freq(n))3	-1.512e+00	6.485e-01	-2.33
cnd*freq(n-1)	8.438e-03	3.588e-03	2.35
cnd*freq(n+1)	-1.342e-02	3.494e-03	-3.84
cnd*incoming sacc. ampl.	-5.996e-03	1.589e-03	-3.77
cnd*outgoing sacc.ampl.	7.864e-03	2.017e-03	3.90

Table D.4.: Final LMM fitting log gaze duration for the samples 'original old' and 'hard old'.

Linear mixed-effects model fit by maximum likelihood			
RANDOM EFFECTS:			
Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.00868958	0.093218
sentence id	(Intercept)	0.00090735	0.030122
subject id	(Intercept)	0.02213126	0.148766
Residual		0.06906424	0.262801
number of obs: 3107, groups: wid, 497; sn, 144; id, 55			
FIXED EFFECTS:			
	Estimate	Std. Error	t-value
(Intercept)	5.886229	0.028136	209.20
condition(cnd)	0.116457	0.042298	2.75
1/length(n)	-0.138734	0.149075	-0.93
frequency(n)	-0.068106	0.008898	-7.65
frequency(n+1)	-0.019465	0.005510	-3.53
predictability(n)	-0.012633	0.008128	-1.55
predictability(n+1)	0.010815	0.006024	1.80
length 6,7 letters	-0.046697	0.016175	-2.89
1/length(n)*freq(n)	0.357388	0.078577	4.55
cnd*1/length(n)	-0.467159	0.146423	-3.19

Table D.5.: Final LMM fitting log single fixation duration for the samples 'original' and 'frequent'.

Linear mixed-effects model fit by maximum likelihood

RANDOM EFFECTS:

Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0077223	0.087876
sentence id	(Intercept)	0.0016059	0.040074
subject id	(Intercept)	0.0131439	0.114647
Residual		0.0834532	0.288883

number of obs: 29638, groups: wid, 550; sn, 144; id, 63

FIXED EFFECTS:

	Estimate	Std. Error	t-value
(Intercept)	5.271e+00	2.188e-02	240.88
trial	-3.342e-04	4.099e-05	-8.15
pc1	1.489e-02	5.256e-03	2.83
pc2	-4.283e-02	1.180e-02	-3.63
condition (cnd)	1.034e-01	3.011e-02	3.43
poly(frequency(n))1	-3.864e+00	1.319e+00	-2.93
poly(frequency(n))2	-1.544e+00	1.079e+00	-1.43
poly(frequency(n))3	-4.880e+00	5.467e-01	-8.93
predictability(n)	-2.113e-02	3.101e-03	-6.81
1/ length(n)	1.899e-01	7.006e-02	2.71
length 6,7 letters	-4.954e-02	1.004e-02	-4.93
frequency(n-1)	-3.092e-02	3.416e-03	-9.05
predictability(n-1)	-1.400e-02	3.311e-03	-4.23
1/ length(n-1)	2.012e-01	3.770e-02	5.34
frequency(n+1)	-1.908e-02	3.392e-03	-5.63
predictability(n+1)	9.921e-03	2.777e-03	3.57
1/ length(n+1)	1.497e-01	3.603e-02	4.16
incoming sacc.ampl.	2.461e-02	1.019e-03	24.14
rel. fix. position	-1.370e-01	1.319e-02	-10.38
(rel. fix. position)^2	-4.670e-01	4.164e-02	-11.22
outgoing sacc.ampl.	1.088e-02	1.162e-03	9.36
length*freq.(n)	9.536e-02	5.139e-02	1.86
freq.(n-1)*freq.(n)	1.649e-02	1.781e-03	9.26
freq.(n)*freq.(n+1)	6.186e-03	1.936e-03	3.20
length*pred.(n+1)	-6.387e-02	2.477e-02	-2.58
cnd*incoming sacc.ampl.	3.849e-03	1.404e-03	2.74
cnd*rel.fix.position	4.320e-02	1.884e-02	2.29
cnd*(rel.fix.position)^2	1.027e-01	5.837e-02	1.76
cnd*outgoing sacc.ampl.	7.568e-03	1.596e-03	4.74

Table D.6.: Final LMM fitting log gaze duration for the samples 'original' and 'frequent'.

Linear mixed-effects model fit by maximum likelihood			
RANDOM EFFECTS:			
Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0118180	0.108711
sentence id	(Intercept)	0.0014388	0.037932
subject id	(Intercept)	0.0134622	0.116027
Residual		0.0638231	0.252632
number of obs: 5007, groups: wid, 521; sn, 144; id, 63			
FIXED EFFECTS:			
	Estimate	Std. Error	t-value
(Intercept)	5.8535369	0.0241348	242.54
pc1	0.0263919	0.0080858	3.26
pc2	-0.0170981	0.0105520	-1.62
condition(cnd)	0.0580516	0.0339583	1.71
trial	-0.0004627	0.0001259	-3.68
frequency(n)	-0.0449364	0.0087095	-5.16
frequency(n-1)	-0.0101946	0.0046129	-2.21
frequency(n+1)	-0.0161607	0.0050443	-3.20
predictability(n)	-0.0207809	0.0071711	-2.90
1/length(n)	-0.5866613	0.1292970	-4.54
length 6,7 letters	-0.0498581	0.0156166	-3.19
1/length(n)*freq(n)	0.3996051	0.0731655	5.46
1/length(n)*freq(n+1)	0.1947536	0.0598686	3.25
cnd*trial	0.0003999	0.0001810	2.21

Table D.7.: Final LMM fitting log single fixation duration for the samples 'frequent' and 'proofreading'.

Linear mixed-effects model fit by maximum likelihood

RANDOM EFFECTS:

Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0116300	0.107843
sentence id	(Intercept)	0.0016585	0.040724
subject id	(Intercept)	0.0108883	0.104347
Residual		0.0928295	0.304679

number of obs: 27813, groups: wid, 550; sn, 144; id, 60

FIXED EFFECTS:

	Estimate	Std. Error	t-value
(Intercept)	5.389e+00	2.236e-02	240.96
trial	-2.560e-04	3.737e-05	-6.85
pc1	1.946e-02	5.413e-03	3.60
pc2	-1.721e-02	1.075e-02	-1.60
condition(cnd)	-2.913e-02	3.043e-02	-0.96
poly(frequency(n))1	-2.753e+00	1.554e+00	-1.77
poly(frequency(n))2	2.030e+00	1.294e+00	1.57
poly(frequency(n))3	-3.703e+00	7.129e-01	-5.19
predictability(n)	-1.877e-02	3.981e-03	-4.72
1/ length(n)	1.545e-01	8.888e-02	1.74
length 6,7 letters	-4.469e-02	1.206e-02	-3.71
frequency(n-1)	-3.917e-02	4.021e-03	-9.74
predictability(n-1)	-8.933e-03	3.563e-03	-2.51
1/ length(n-1)	2.211e-01	4.060e-02	5.45
frequency(n+1)	-1.591e-02	3.734e-03	-4.26
predictability(n+1)	6.773e-03	3.376e-03	2.01
1/ length(n+1)	1.782e-01	3.944e-02	4.52
incoming sacc.ampl.	2.752e-02	9.490e-04	29.00
rel. fix. position	-8.632e-02	1.451e-02	-5.95
(rel. fix. position)^2	-3.155e-01	4.595e-02	-6.87
outgoing sacc.ampl.	1.825e-02	1.461e-03	12.49
1/ length(n)*freq.(n)	2.224e-01	6.410e-02	3.47
freq.(n-1)*freq(n)	1.564e-02	1.977e-03	7.91
freq.(n)*freq.(n+1)	7.387e-03	2.110e-03	3.50
cnd*poly(freq(n))1	-4.834e+00	9.128e-01	-5.30
cnd*poly(freq(n))2	1.079e+00	8.173e-01	1.32
cnd*poly(freq(n))3	-2.434e+00	6.576e-01	-3.70
cnd*pred.(n)	-1.175e-02	4.155e-03	-2.83
cnd*1/ length(n)	-2.580e-01	6.011e-02	-4.29
cnd*freq.(n-1)	1.289e-02	2.942e-03	4.38
cnd*pred.(n+1)	-6.609e-03	3.170e-03	-2.09
cnd*rel.fix.position	1.244e-01	1.755e-02	7.09
cnd*(rel.fix.position)^2	1.507e-01	5.510e-02	2.73
cnd*outgoing sacc.ampl.	8.197e-03	2.076e-03	3.95
cnd*1/ length(n)*freq(n)	2.413e-01	4.386e-02	5.50

Table D.8.: Final LMM fitting log gaze duration for the samples 'frequent' and 'proof-reading'.

Linear mixed-effects model fit by maximum likelihood			
RANDOM EFFECTS:			
Groups	Name	Variance	Std.Dev.
word id	(Intercept)	0.0159653	0.126354
sentence id	(Intercept)	0.0009298	0.030493
subject id	(Intercept)	0.0157904	0.125660
Residual		0.0976938	0.312560
number of obs: 6654, groups: wid, 528; sn, 144; id, 60			
FIXED EFFECTS:			
	Estimate	Std. Error	t-value
(Intercept)	5.947e+00	2.904e-02	204.78
pc1	3.081e-02	8.991e-03	3.43
pc2	-2.397e-02	1.213e-02	-1.98
condition(cnd)	1.239e-01	4.035e-02	3.07
trial	-4.553e-05	1.594e-04	-0.29
frequency(n)	-6.085e-02	1.186e-02	-5.13
predictability(n)	-1.386e-02	1.047e-02	-1.32
1/length(n)	-5.512e-01	1.866e-01	-2.95
length 6,7 letters	-7.623e-02	1.687e-02	-4.52
frequency(n-1)	-2.436e-02	7.010e-03	-3.48
1/length(n-1)	1.584e-01	8.551e-02	1.85
frequency(n+1)	-2.121e-02	6.074e-03	-3.49
predictability(n+1)	1.354e-03	6.134e-03	0.22
1/length(n)*freq.(n)	3.253e-01	1.010e-01	3.22
1/length(n)*freq.(n+1)	2.993e-01	6.882e-02	4.35
cnd*trial	-5.128e-04	1.804e-04	-2.84
cnd*freq.(n)	-5.584e-02	1.084e-02	-5.15
cnd*pred.(n)	-2.157e-02	1.082e-02	-1.99
cnd*1/length(n)	-1.321e+00	1.883e-01	-7.02
cnd*1/length(n)*freq.(n)	7.315e-01	9.728e-02	7.52

Throughout its empirical research history eye movement research has always been aware of the differences in reading behavior induced by individual differences and task demands. This work introduces a novel comprehensive concept of reading strategy, comprising individual differences in reading style and reading skill as well as reader goals. In a series of sentence reading experiments recording eye movements, the influence of reading strategies on reader- and word-level effects assuming distributed processing has been investigated. Results provide evidence for strategic, top-down influences on eye movement control that extend our understanding of eye guidance in reading.

Christiane Wotschack studied Patholinguistics at the University of Potsdam. The present study was conducted at the Department of Cognitive Psychology at the University of Potsdam.

ISSN 1869-3822
ISBN 978-3-86956-021-2
9 783869 560212