Investigating online reading with eye tracking and EEG: The influence of text format, reading task and parafoveal stimuli on reading processes

Jaana Simola

Psychology, Institute of Behavioural Sciences
University of Helsinki, Finland

Academic dissertation to be publicly discussed by due permission of the Faculty of Behavioural Sciences at the University of Helsinki in Lecture Hall VI, Fabianinkatu 33, on the 7th of December, 2011, at 12 noon.

UNIVERSITY OF HELSINKI
Institute of Behavioural Sciences
Studies in Psychology 81: 2011

Supervisors: Associate Professor Kenneth Holmqvist
Cognitive Science
Lund University, Sweden

Docent Ilpo Kojo
Institute of Behavioural Sciences
University of Helsinki, Finland

Professor Jussi Saarinen
Institute of Behavioural Sciences
University of Helsinki, Finland

Reviewers: Professor Ralph Radach
General and Biological Psychology
University of Wuppertal, Germany

Professor Simon Liversedge
School of Psychology
University of Southampton, UK

Opponent: Professor Thierry Baccino
Laboratory “Human and Artificial Cognition”, CHArt/LUTIN
University of Paris VIII, France

ISSN 1798-842X
ISSN-L 1798-842X
ISBN 978-952-10-7357-1 (pbk.)
ISBN 978-952-10-7358-8 (PDF)
http://ethesis.helsinki.fi/
Unigrafia
Helsinki 2011

Contents

Abstract 5
Tiivistelmä 6
Acknowledgements 7
List of original publications 9
Abbreviations 10
1 Introduction 11
1.1 Reading process 11
1.2 Eye movements and attention during reading 12
1.3 Perceptual span 16
1.4 Text format 18
1.5 Models of eye movement control in reading and scene viewing 19
1.6 Attention allocation to task-irrelevant stimuli 24
2 Aims of the study 25
3 Methods 27
3.1 Participants 27
3.2 Eye tracking 27
3.2.1 Data acquisition 28
3.2.2 Data analysis 28
3.2.3 Modeling with Hidden Markov Model (HMM) 30
3.3 Electroencephalography (EEG) and eye fixation related potentials (EFRPs) 33
3.3.1 Data acquisition 35
3.3.2 Data analysis 35
3.4. Behavioral methods 37
4 Experiments and results 38
4.1 Study I: Visual field differences in parafoveal processing of words 38
4.1.1 Stimuli and procedure 39
4.1.2 Results 40
4.2 Study II: The effect of vertical arrangement of text on reading 44
4.2.1 Stimuli and procedure 44
4.2.2 Results 45

4.3 Study III: The effect of animated and abruptly appearing ads on online reading 47
4.3.1 Stimuli and procedure 47
4.3.2 Results 49
4.4 Study IV: Alternating processing states during information search tasks 55
4.4.1 Stimuli and procedure 56
4.4.2 Results 57
5 Discussion 63
5.1 Parafoveal processing of words across visual fields 63
5.2 Text format 65
5.3 Attention to parafoveal pictorial stimuli 67
5.4 Reading processes in information search tasks 69
5.5 Future directions 71
6 Conclusions 74
7 References 76

Abstract

Research on reading has been successful in revealing how attention guides eye movements when people read single sentences or text paragraphs in simplified and strictly controlled experimental conditions. However, less is known about reading processes in more naturalistic and applied settings, such as reading Web pages. This thesis investigates online reading processes by recording participants' eye movements.

The thesis consists of four experimental studies that examine how location of stimuli presented outside the currently fixated region (Study I and III), text format (Study II), animation and abrupt onset of online advertisements (Study III), and phase of an online information search task (Study IV) affect written language processing. Furthermore, the studies investigate how the goal of the reading task affects attention allocation during reading by comparing reading for comprehension with free browsing, and by varying the difficulty of an information search task.

The results show that text format affects the reading process, that is, vertical text (word/line) is read at a slower rate than a standard horizontal text, and the mean fixation durations are longer for vertical text than for horizontal text. Furthermore, animated online ads and abrupt ad onsets capture online readers' attention and direct their gaze toward the ads, and distract the reading process. Compared to a reading-for-comprehension task, online ads are attended to more in a free browsing task. Moreover, in both tasks abrupt ad onsets result in rather immediate fixations toward the ads. This effect is enhanced when the ad is presented in the proximity of the text being read. In addition, the reading processes vary when Web users proceed in online information search tasks, for example when they are searching for a specific keyword, looking for an answer to a question, or trying to find a subjectively most interesting topic. A *scanning* type of behavior is typical at the beginning of the tasks, after which participants tend to switch to a more careful *reading* state before finishing the tasks in the states referred to as *decision* states. Furthermore, the results also provided evidence that left-to-right readers extract more parafoveal information to the right of the fixated word than to the left, suggesting that learning biases attentional orienting towards the reading direction.

Tiivistelmä

Aikaisemmissa lukututkimuksissa on selvitetty lukijan tarkkaavaisuuden ja katseen ohjautumista, kun ärsykkeinä on käytetty yksittäisiä lauseita tai lyhyitä tekstejä. Sen sijaan soveltavissa ympäristöissä, kuten Internetissä, katseen ohjautumista ja lukuprosesseja on tutkittu vähemmän. Tässä väitöskirjatyössä tutkittiin katseen ohjautumista, kun koehenkilöt suorittivat Internet ympäristölle tyypillisiä lukutehtäviä. Lisäksi ensimmäisessä osatyössä tutkittiin, miten opittu lukusuunta vaikuttaa tarkannäön (fovean) ulkopuolella esitetyn teksti-informaation prosessointiin. Osatyöt koostuivat kokeista, joissa tarkasteltiin tekstin esitystavan sekä tekstin ympärillä esitettyjen kuvien ja sanojen vaikutusta lukuprosessiin. Lisäksi tutkittiin lukuprosessin vaihtelua ajallisesti tehtävän edetessä tai tehtäväohjeen muuttuessa.

Tulokset osoittivat, että kapeapalstaista (sana/rivi) tekstiä luettiin hitaammin kuin normaalia vaakasuorilla riveillä esitettyä tekstiä ja että katseen pysähdysten (fiksaatioiden) keskimääräiset kestot olivat pidempiä kapeapalstaista tekstiä luettaessa. Lisäksi havaittiin, että tekstin ympärillä esitetyt mainokset häiritsevät lukemista, sillä lukuprosessi keskeytyi lukijan katseen ohjautuessa mainokseen. Lisäksi mainosten ilmestymisaika ruudulle oli ajallisesti yhteydessä hetkeen, jolloin katse siirtyi kohti mainosta. Mainoksia katsottiin enemmän tilanteessa, jossa koehenkilöt saivat vapaasti tutustua sivustoon verrattuna vaativampaan lukutehtävään, jossa koehenkilöitä pyydettiin vastaamaan tekstin sisältöä koskeviin kysymyksiin. Tulokset osoittivat myös, että lukuprosessit vaihtelevat tiedonhakutehtävien eri vaiheissa, kun koehenkilöt etsivät lauselistojen joukosta joko tiettyä avainsanaa, vastausta kysymykseen tai henkilökohtaisesti kiinnostavaa aihetta. Tehtävän alussa koehenkilöiden silmänliikkeiden piirteet viittasivat *silmäilyyn*, jonka jälkeen koehenkilöt siirtyivät huolellisempaan *lukemiseen.* Huolellisen lukemisen jälkeen silmänliikkeet muuttuivat jälleen koehenkilöiden siirtyessä lukuprosessista *päätöksenteko*-prosessiin, johon tehtävät lopetettiin. Lisäksi tulokset osoittivat, että opitulla lukusuunnalla oli vaikutusta tarkkaavaisuuden ohjautumiseen, sillä vasemmalta oikealle lukevat koehenkilöt hahmottivat enemmän informaatiota keskellä esitetyn sanan oikealta puolelta vasempaan puoleen verrattuna.

Acknowledgements

This thesis has been carried out in different laboratories, and I am very thankful to have had the privilege of working with many people while completing the studies. First, I thank Associate Professor Kenneth Holmqvist, Docent Ilpo Kojo, and Professor Jussi Saarinen for their supervision. Without the inspiration of, and introduction to the field of eye movement research by, Ilpo Kojo, I probably would have never started this thesis. I extend great gratitude to Kenneth Holmqvist for his constructive criticism, questions and discussions, which have helped me to finish the thesis. Further, I thank Jussi Saarinen and Professor Kimmo Alho for smoothly and efficiently organizing the official matters at my home institute, the former Department of Psychology at the University of Helsinki.

A great thanks goes to all my co-authors. I thank Jari Laarni for patiently teaching me a great deal about research work. I am also very grateful to my colleagues Jarmo Kuisma and Anssi Öörni for productive collaboration, Professor Liisa Uusitalo for support and helpful discussions, and Professor Jukka Hyönä for his expertise and insightful comments on our research. I also thank Jarkko Salojärvi for teaching me concepts related to modeling and for his comments on this thesis. Further thanks go to Associate Professor Magnus Lindgren for his advice during the writing of the manuscript of Study I. It has been a great pleasure to work with all of you.

Study I was conducted during my research visit at the Humanities Laboratory, Lund University, Sweden. Studies II-IV were carried out in the Aalto University School of Economics (former Helsinki School of Economics) both at the Center for Knowledge and Innovation Research (CKIR) and at the Department of Marketing (Study III). I finalized the studies and wrote the thesis at Cognitive Science, Institute of Behavioural Sciences, University of Helsinki. I thank the directors of these laboratories, Seija Kulkki (CKIR), Professor Henrikki Tikkanen (Department of Marketing), Professor Sven Strömqvist (Humanities Laboratory) and Professor Christina Krause (Cognitive Science) for providing excellent research environments, which have made all the scientific work possible. During the course of my PhD-studies, I also had the opportunity to work at the Advanced Magnetic Imaging (AMI) Centre at the former Helsinki University of Technology. I thank all

employees of these laboratories. A special thanks goes to Docent Simo Vanni for sharing a great deal of intellect and teaching me important research skills.

The research described in this thesis was supported by grants from the Academy of Finland, Helsingin Sanomat Foundation and NordForsk. I am grateful for the opportunity to be a full-time doctoral student during most of my thesis work. I am also grateful for my research visit to Lund University, which was a great experience and which has proved to be productive also in terms of future collaboration. I thank the official reviewers of this thesis, Professor Simon Liversedge and Professor Ralph Radach for constructive and positive comments on the manuscript. I also thank Professor Thierry Baccino for agreeing to serve as my opponent.

A warm thanks goes to the members of the zoo club: Devin Terhune, Anne-Cécile Treese, and Gerd Waldhauser for inspiring discussions, friendship, and for making my time in Lund enjoyable. I especially miss our evenings at Ariman. I also thank my other colleagues in Lund: Jana Holsanova, Richard Andersson, Marcus Nyström, Richard Dewhurst, Petter Kallioinen, Paulina Lindstörm, Nils Holmberg, Roger Johansson, and Philip Diderichsen for company and for commenting my papers on the article seminars. I extend further thanks to my colleagues at Cognitive Science: Alina Leminen, Katja Irvankoski, Otto Lappi, Mona Moisala, Sanna Korpela, Jari Torniainen, and Teemu Peltonen. Working and sharing the hours in the lab has been instructive, interesting, and fun.

I offer deep thanks to my friends Vappu Rossi, Liina Hemminki, and Johanna Kolo for empathy and presence during the bright and dark moments in the course of my studies. Special thanks belong to too many people to be listed here. I thank all of you for the shared company, wine, food, ultimate frisbee and travels during these years.

This thesis is dedicated to my parents Soile and Hannu Simola, to my sisters Leena and Sanna Simola, and to my grandmothers Airi Valio and Aili Simola. Thank you for ever-present support, and I am sorry that my father never saw this happen. Finally, with overwhelming gratitude, I thank Heli for being on my side.

Helsinki, November, 2011 Jaana Simola

List of original publications

This thesis is based on the following publications, referred to by their roman numerals.

I Simola, J., Holmqvist, K., & Lindgren, M. (2009). Right visual field advantage in parafoveal processing: Evidence from eye-fixation related potentials. *Brain and Language*, 111, 101–113.

II Laarni, J., Simola, J., Kojo, I., & Näsänen, R. (2004). Reading vertical text from a computer screen. *Behaviour & Information Technology*, 23, 75–82.

III Simola, J., Kuisma, J., Öörni, A., Uusitalo, L., & Hyönä, J. (2011). The impact of salient advertisements on reading and attention on web pages. *Journal of Experimental Psychology: Applied*, 17, 174–190.

IV Simola, J., Salojärvi, J., & Kojo, I. (2008). Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cognitive Systems Research*, 9, 237–251.

The articles are reprinted with the kind permission of the copyright holders.

Studies I and IV: Elsevier (http://www.elsevier.com)
Study II: Taylor & Francis (http://www.tandfonline.com)
Studies III: American Psychological Association (http://www.apa.org)

Abbreviations

ANOVA	analysis of variance
dHMM	discriminative hidden Markov model
EEG	electroencephalography
EFRP	eye fixation related potential
ERP	event related potential
GEE	generalized estimated equations - model
HMM	hidden Markov model
LH	left hemisphere
LSD	least significant difference
LVF	left visual field
RH	right hemisphere
RVF	right visual field
SWIFT	Saccade generation With Inhibition by Foveal Targets
WPM	words per minute

1 Introduction

1.1 Reading process

Reading matters in our society; it determines an individual's ability to study and earn a living and act as a fully functioning member of society. Investigating the reading process is crucial to theories of reading, but it also helps in understanding how different environments affect the reading process and what constraints they set on our abilities to read. Prior eye movement research on reading has been successful in revealing how people recognize words and how attention guides eye movements in relatively simple settings, that is, when people read single sentences or text paragraphs presented on an otherwise blank computer screen (see Radach, Huestegge, & Reilly, 2008). Compared to a fairly good understanding of the reading process in such simple settings, relatively little is known about reading in more complex and applied settings such as on Web pages.

The Web environment is likely to set new challenges for the reader because the information presented around the text differs from the traditional reading conditions. That is, the reading tasks in Web pages are heavily influenced by the information presented outside the current eye fixation position (*parafoveal information*). The Web page layouts may force the text to appear in narrow columns that diverge from the traditional format of text presented in horizontal rows. Furthermore, the online texts are often surrounded by pictorials that may be animated, include sound, or appear abruptly. The purpose of these design cues is to capture readers' attention and to direct their focus away from the text. Other types of reading tasks than reading for comprehension are also common in online environments. For example, scanning the pages in order to find relevant information is a typical online task that incorporates both reading and visual search processes. Thus, the Web introduces a multimodal environment where traditional reading of the rows from top to bottom rarely occurs, but the fragmented texts may be entered and read in different order and with varying goals and strategies. Because online reading is an everyday task for many people, it is crucial to investigate reading in the Web environment.

The goal in this thesis was to present the participants with as naturalistic online reading tasks as possible, but to retain the control of an experimental design. Because online reading tasks are heavily influenced by parafoveal information, **Study I** introduces the fundamental concepts related to information processing in reading by examining parafoveal information extraction both to the left and to the right of fixation. The following studies examined how the format of the text being read (**Study II**), animated or abruptly appearing pictorial stimuli (**Study III**), and the phase and the goal of an information search task (**Study IV**) affect online reading. Eye movements were recorded in all studies. In addition, **Study I** combined eye movement and electroencephalography (EEG) recordings to measure eye fixation related potentials (EFRP). EFRP is a type of event related potential (ERP) measuring electrical activity in the brain as a response to eye fixations (e.g. Baccino & Manunta, 2005). The technique is relatively novel, and contrary to traditional ERPs, it allows the observers to move their eyes freely during the task.

The thesis begins with an introduction to what is known about eye movements and attentional allocation during reading of single sentences or text paragraphs. Concepts related to visual constraints on reading such as perceptual span are introduced in detail, followed by an introduction to how different text formats, reading processes and surrounding pictorial stimuli affect reading. In the following chapters, the aims of the study, experimental methods and results are presented and discussed. Finally, the results are set in a more general context of online reading.

1.2 Eye movements and attention during reading

Reading is a demanding activity, which depends on the dynamic integration between visual information processing, word recognition, attention and oculomotor control. Vision has an important role in reading, because the neurophysiological properties of the oculomotor system set constraints for reading. For example, as you read this, your eyes are moving by a sequence of rapid eye movements (*saccades*) and periods when the eyes are relatively stable (*fixations*), and your brain is converting visual images of letter strings into words, together with their meanings.

The fact that we make fixations and saccades to read indicates that the number of letters that can be acquired during a single fixation is limited. This results from the non-uniform distributions of cones and ganglion cells in the retina, which show higher density of cells in the center of the retina (*fovea*) compared to peripheral retina, and a higher cell density along the horizontal than along the vertical meridian (Curcio & Allen, 1990). The distribution of ganglion cells is crucial because it sets the upper limit on the proportion of information that is transmitted to higher cortical areas (Curcio & Allen, 1990).

The acquisition of visual information (e.g., word identification) occurs in fixations. Several processes need to be completed within a fixation (duration around 200–250 ms in reading) before the eyes can move to the next word (see Sereno & Rayner, 2003). During the first 50 ms, the visual information is transmitted from the retina to higher cortical areas where the lexical processing can begin, (i.e., *the eye-to-brain lag*, VanRullen & Thorpe, 2001). The lexical processing is supposed to be in progress within the first 100–200 ms to meet the time line for *programming the next saccade* (Sereno & Rayner, 1998). The terminal portion of a fixation is reserved for *saccade latency*, the time needed to encode the location of the next saccade and to initiate that saccade (Findlay & Walker, 1999; Rayner, 2009). It has been proposed that two separate pathways are concerned with the spatial (e.g., saccade length) and the temporal (e.g., fixation duration) parameters of eye movements (Findlay & Walker, 1999), allowing the processes within a fixation to occur at least partly in parallel. The decision of *where* to move the eyes is strongly influenced by low-level text properties, such as word length and spaces between the words. Whereas, the decision of *when* to move the eyes is driven by lexical properties of the fixated word, for example, word frequency and predictability from the context (reviewed in Rayner, 2009).

The vision is actively suppressed during saccades in order to obscure the motion of images as they sweep across the retina when the eye moves. This phenomenon is called *saccadic suppression*, during which the sensitivity for seeing stimuli declines around 25 ms before saccade onset and recovers to normal levels around 50 ms after saccade (Ishida & Ikeda, 1989; Morrone & Burr, 2009). Although new information is not acquired during saccades, processes devoted to word recognition and word identification continue. That is, the post-saccadic processing times are shorter after a long than a short saccade, suggesting

that participants are able to use the extra time during the long saccade to reduce the time needed to perform a lexical decision or word identification (Irwin, 1998). In addition, saccades are motor events that need to be planned and executed. The period of saccade latency, that is, the minimum amount of time needed to detect a target and to move the eyes to its location lasts around 175–200 ms (reviewd in Reichle, Pollatsek, & Rayner, 2006). The typical saccade durations are around 30–50 ms depending on their length (Rayner, 2009). The average saccade length for Finnish text is estimated to be around 10–11 characters (Hyönä & Niemi, 1990), which differs from the typical saccade length of approximately eight characters for English words. This difference is due to the greater length of Finnish words, but compared to English the saccade lengths are equal in terms of words (Hyönä & Niemi, 1990).

A typical sequence of eye movements, a *scanpath*, during reading consists of one fixation per word, and mostly the eyes move forward in the text. Occasionally, the readers also skip or refixate words. The probability of skipping or refixating a word depends strongly on word length: short words (e.g., function words) are skipped more often than long words (e.g., Hautala, Hyönä, & Aro, 2011; Rayner, Sereno, & Raney, 1996), while longer words are refixated more often than short words before the eyes leave the word (Vergilino-Perez, Collins, & Doré-Mazars, 2004). For example, 4% of six-letter monospaced Finnish words were skipped, while around 20% of four-letter monospaced words were skipped (Hautala, et al., 2011). It has been suggested that words are skipped because they have been sufficiently processed on the previous fixation (e.g., Rayner, Ashby, Pollatsek, & Reichle, 2004). On the other hand, words are refixated in order to compensate visual acuity limitations when not all the letters of a word fit in the fovea area, or in order to correct a mislocated initial saccade landing position (Vergilino-Perez, et al., 2004). Studies have also shown that refixations occur due to difficulties in cognitive processing during the first fixation (Rayner, et al., 1996). Moreover, as people read, they also make backward saccades (*regressions)* due to oculomotor errors, difficulties in comprehending the text, or difficulties with word identification (reviewed in Vitu & McConkie, 2000). Hyönä and Niemi (1990) reported that the regression frequencies varied between 12–15% when participants read the same Finnish texts three times.

Studies have shown that as the difficulty of the text increases, fixation durations, and frequency of fixations and regressions increase, while saccade lengths decrease (Hyönä & Niemi, 1990; Rayner, 1998, 2009). Previous research has also indicated that eye movements are sensitive to word characteristics. For example, when word length is controlled, longer eye fixations have been reported on low-frequency than on high-frequency words (Rayner, 1998, 2009). Further, eye fixation times are longer on words that are unpredictable from the text context compared to words that are predictable (e.g., Rayner, et al., 2004). Together, these findings suggest that processes related to ongoing word recognition, comprehension and cognitive load are the major determinants of eye movement behavior. These results have been taken as an evidence for the *eye-mind link* assumption, which proposes that where an observer is looking at a certain time reflects, at least partly, what is being processed in his/her mind at that time (Just & Carpenter, 1980). The spatial and temporal relationship between eye movements and cognitive processing cannot, however, be captured solely by such a simple principle (Radach & Kennedy, 2004). Numerous studies have shown that there is a substantial amount of preprocessing of next word, and that processing may *spill over* from one word to the next word (reviewed in Rayner, 1998, 2009).

Moreover, studies have shown that viewers can allocate visual attention without moving their eyes (Posner, Snyder, & Davidson, 1980). These *covert* attention shifts enhance processing of stimuli at the attended location compared to unattended locations (Brefczynski & DeYoe, 1999; Simola, Stenbacka, & Vanni, 2009). In complex tasks like reading, however, attention and eyes are usually directed to the same location, because saccadic eye movements are rapid and easy to produce in order to direct the high resolution foveal region to the location of interest. Empirical evidence also supports a close link between covert attention and eye movements. For instance, Deubel and Schneider (1996) demonstrated a coupling between saccade preparation and spatial attention which was indexed by enhanced discriminability at the saccade target location preceding the saccade execution. These results support a close relationship between attention and the oculomotor system. As a consequence, the eye movement data provide an indicator of the way in which *overt* visual attention (eye location) is distributed across the stimulus display.

1.3 Perceptual span

The area from which useful information for reading (e.g., information about word length) can be gathered during a fixation is called the *perceptual span* (for a summary see Rayner, 1998, 2009). The perceptual span is considered to be a consequence of visual and attentional constraints, and it is strongly asymmetrical towards the reading direction. That is, readers of left-to-right scripts (e.g., English and Finnish) acquire more information to the right of fixation than to the left. Experiments using a gaze-contingent moving-window paradigm have demonstrated that the perceptual span extends 3–4 letters to the left and up to 14–15 letters to the right of fixation (McConkie & Rayner, 1975; Rayner, 1998). This rightward bias is reversed for readers of right-to-left scripts (e.g., Hebrew) with an asymmetry towards the left of fixation (Pollatsek, Bolozky, Well, & Rayner, 1981). Häikiö, Bertram, Hyönä, and Niemi (2009) showed that the *letter identity span* (the number of letters readers can identify during a fixation) of Finnish readers was comparable to the span found for English readers (i.e., nine characters to the right of fixation for Finnish readers, versus 7–8 characters for English readers).

Characteristics of the writing system can also influence the overall size of the perceptual span. For example, for readers of Chinese, the perceptual span extends about one character to the left and 2–3 characters to the right of fixation (Rayner, 2009). Moreover, Osaka and Oda, (1991) showed that Japanese readers read vertical text as efficiently as horizontal text, and that their perceptual span in the vertical direction (roughly 5–6 characters) corresponds to their horizontal span size of around 5–7 characters (see Naoyuki Osaka, 1992). Furthermore, their results suggested an asymmetry in the vertical direction, with the perceptual span extending a few characters above and around 4–5 characters below the fixation. Yu, Park, Gerold, and Legge (2010) showed that also for English readers, the perceptual span was asymmetric in vertical direction, suggesting a lower visual field advantage for reading of vertical rotated text or text with downward cascade of letters (“marquee” style). A lower visual field advantage has been reported in many visual tasks, possibly indicating a greater attentional resolution in the lower visual field (He, Cavanagh, & Intriligator, 1996; Simola, et al., 2009).

Ojanpää, Näsänen and Kojo (2002) investigated the vertical *word identification span* (the area from which words can be identified during a single fixation) in a word search task by measuring the number of fixations needed to recognize a target word in a vertical word list. The longest list that could be processed (at a probability of 0.79) during a single fixation was approximately 4–5 words. When the vertical span was compared to the horizontal span, the results indicated that the two-dimensional word identification span was elongated to the horizontal direction (about 10 characters horizontally and about 4–5 characters vertically). This is in agreement with the anatomical results of the sensory limitations for visual perception (Curcio & Allen, 1990).

The visual field is typically divided into three regions: fovea, parafovea and periphery. The fovea covers the central 2° of vision, the parafovea extends out to 5° to either side of the center of fixation, and the remaining area is termed the periphery (e.g., Liversedge & Findlay, 2000). In terms of character spaces, when the font size and reading distance are normal, the fovea extends approximately 6–8 characters around the fixation, while the parafovea extends up to 15 characters to the right of fixation (when reading from left to right) (Häikiö, et al., 2009). No information relevant for reading is extracted from the periphery.

Previous research suggests that parafoveal information is acquired primarily from the location that is about to be fixated next (Henderson, Pollatsek, & Rayner, 1989). Thus, useful information for reading is not only confined to the foveated information, but usually a word has been viewed parafoveally on the previous fixation. The characteristics of the parafoveal word determine whether or not we need to make a saccade to identify that word. For example, short words or functional words are often processed in the parafovea and are therefore skipped (Rayner, 1998, 2009). The difficulty of the fixated word also affects the amount of attentional resources that can be allocated to the parafoveal words. For instance, when the fixated word is difficult, readers get little or no parafoveal information from the word to the right of fixation (Henderson & Ferreira, 1990). Previous research has shown that parafoveal information can influence reading in different ways. First, *parafoveal-on-foveal effects* describe the influence of parafoveal words on the fixation durations on the

fixated (foveal) word. Second, *parafoveal preview benefits* reflect shortening of the processing times on words that have been parafoveally visible prior to fixation.

1.4 Text format

In Western languages, text is typically presented in horizontal lines that are read from left to right and from top to bottom. In Semitic languages, such as Hebrew or Arabic, the text is arranged and read from right to left. Additionally, vertical presentation of text is typical for Asian languages. For example, Japanese text can be written either horizontally or vertically. Furthermore, the visual features of the text, such as the font difficulty, the spacing between letters, words and lines, word length, letter size and case affect reading rate and eye movements during reading (O'Regan, Lévy-Schoen, & Jacobs, 1983; Pelli et al., 2007; Yu, et al., 2010). The classic curve shows that reading rates rise fast by increasing text size, but stays constant after reaching the maximum rate at a critical print size (e.g., Pelli, et al., 2007). Increasing the line spacing in the vertical word search task, resulted in enhanced search times, number of fixations and saccade amplitudes (Ojanpää, et al., 2002), whereas increased character spacing were associated with shorter saccades during reading (O'Regan, et al., 1983).

Previous studies have shown that the marquee text was read slower than texts that were rotated 90° to the left or to the right, while the rotated texts were read at slower than horizontal text (Byrne, 2002; Yu, et al., 2010). It was suggested that disrupting the normal orthogonal relationship between word and letter orientation affects parallel processing of letters within a word, and results in slower reading of marquee text. Moreover, Koriat and Norman (1985) showed that lexical-decisions became slow and less accurate when words and non-words were presented in vertical rotation exceeding 60°, especially when the transformations disrupted the whole-word features. Although, reading vertically presented text is slower for normal readers, it has been suggested that people who have to use their peripheral vision to read (such as macular degeneration patients) might benefit from vertically presented text (Feng, Jiang, & He, 2007; Yu, et al., 2010).

Among the few studies that have examined the effects of vertical arrangement of words on reading, Coleman and Kim (1961) found that in a normal reading situation horizontal

text was read faster than vertical texts, but the comprehension scores did not differ between different formats. When single sentences were presented with a tachistoscope, a vertical text was reproduced more accurately than other text formats. However, Coleman and Hahn (1966) failed to find any advantage for the vertical format.

In the word search task, Ojanpää, et al. (2002) showed that search times did not differ between vertical and horizontal word lists, but there where fewer fixations and shorter saccades for vertical than for horizontal lists. These results suggested that although the word identification span is smaller in the vertical direction, more words were processed in vertical lists during a single fixation. Taken together, previous studies suggest that a vertical presentation of text might be an alternative way to present text, which might better utilize the vertical word identification span. Also, Western people have some practice in reading vertically presented text. For example, in small screen devices the text often appears in couple of words per line, and newspaper or Web page layout may force the text to appear in narrow vertical columns. However, previous reading studies mostly concern reading of horizontally presented text while other types of text formats have been less studied.

1.5 Models of eye movement control in reading and scene viewing

Models of eye movements in reading provide a way to computationally test the theories of how perceptual, cognitive and motor processes interact to determine when and where the eyes move during reading[1]. The *cognitive models* of reading assume that eye movements are driven by attention and language related processes. These models can be divided into two categories: *serial-attention* models and *attention-gradient* models.

The serial models suggest that attention is allocated serially to one word at a time. According to the most developed serial model, E-Z Reader (Reichle, et al., 2006; Reichle, Warren, & McConnel, 2009), completion of an early stage of lexical processing (‘familiarity check’) causes the oculomotor system to begin programming a saccade to the

[1] An overview of several of these models is presented in the 2006 special issue in *Cognitive Systems Research*.

next word. The completion of a second stage (‘lexical access’) disengages attention from the currently fixated word to the next word. Although, the word processing is strictly serial, saccade programming occurs in parallel with the comprehension process in reading, because these processes are supported by distinct systems (Findlay & Walker, 1999).

In contrast to the serial models, the parallel, attention-gradient models assume that attention is distributed as a gradient that covers more than one word at a time, supporting parallel lexical processing of words. That is, the parafoveal words may influence recognition of the fixated word. The most famous parallel model, SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) assumes a random timer that initiates saccade programming. Saccade target selection is object-based, with the target objects being words. The word with the highest activation has the highest probability of being selected as the next target. Difficulties with lexical processing can delay the saccades that otherwise move the eyes forward.

The serial and parallel models of reading disagree on whether parafoveal information can influence the processing of the currently fixated word (the parafoveal-on-foveal effects). The parallel models assume that processing is distributed across several words, that is, the processing rate is highest for the foveal word and it decreases to parafoveal words, but at least the word to the right of fixation is processed parafoveally (Engbert, et al., 2005). The parallel models also suggest that semantic information can be extracted from the parafoveal words. In contrast, the serial models assume a strict serial processing, indicating that one word is attended and processed at a time and that word meanings are not accessed parafoveally (Reichle, Liversedge, Pollatsek, & Rayner, 2009; Reichle, Warren, et al., 2009). However, it has been suggested that information about low-level features from words within the perceptual span can be processed concurrently with the fixated word, and that an early stage of orthographic processing can be completed in parallel on several words (Reichle, et al., 2006).

As an alternative to cognitive models described above, *primary oculomotor control* models assume that eye movement control is based on low-level visual information. These models also suggest that the driving force behind eye movements in reading is a ‘move forward’ strategy, which is independent of lexical processing (see Deubel, O'Regan, &

Radach, 2000). For example, Reilly and O'Regan (1998) showed that eye movements during reading were best explained by a model that targeted the longest word in the right parafovea within a window of 20 characters to the right from the current fixation. Further supporting the idea of the 'dumb' (move forward) eye movement strategy, Vitu, O'Regan, Inhoff and Topolski (1995) reported that saccade landing site distributions and skipping rates were very similar to normal reading in a task where participants scanned through lines of 'z'-letters spaced out like a normal text. However, Vitu et al. claimed that a 'dumb' strategy possibly underlies eye movement control in reading, but this strategy can be modulated by ongoing linguistic processing.

Other researchers have also proposed models that are based on interplay of oculomotor and cognitive factors. The assumption is that the oculomotor and cognitive approaches define the extremes of a continuum, and that a successful model can accommodate both cognitive and oculomotor aspects (see Radach & Kennedy, 2004). An example of such a model is Glenmore (Reilly & Radach, 2006), which suggests that saccades are programmed towards low spatial frequency word-objects on the basis of their length and eccentricity from the current fixation. During the fixation, however, processing is strongly modulated by ongoing cognitive processing (e.g., by word frequency). Glenmore also assumes that higher level processing on the sentence and discourse level can affect eye movements during reading.

Moreover, some models of eye movements in reading emphasize the anatomical and neurophysiological properties of the visual system (McDonald, Carpenter, & Shillcock, 2005; Whitney, 2001). These models are based on research indicating that information in the right visual field (RVF) projects directly to the left hemisphere (LH), while the left visual field (LVF) projects to the right hemisphere (RH). Because for most people the cortical structures underlying language processing are lateralized to LH, information from the LVF/RH needs to be transferred to the LH. An opposite pattern is required for RH language dominant readers. The standard neural model of reading suggests that by 180–200 ms post-stimulus, the transcallosal information from the RH reaches the visual word form area, a system located in the left inferior temporal region, devoted for processing of letter strings (Cohen et al., 2000; Cohen et al., 2002).

Computational models of reading often use tailored parameter values based on specific hypotheses on the reading process, and the models are evaluated by their capability to reproduce low-level variables, such as oculomotor (fixation landing-site distributions) and word difficulty (word frequency, length or predictability) effects. In addition, the computational models assume that reading is driven mainly by visual and lexical information (however see Reichle, Warren, et al., 2009). Thus, the models do not typically consider higher-level intra-individual factors such as reading intention, motivation or global reading strategy, which are widely assumed to affect reading (Radach & Kennedy, 2004). Due to these reasons relatively little is known how such top-down factors affect reading, and possibly therefore, it is often assumed that the same reading process (e.g., reading for comprehension) is maintained throughout the task.

Among the few studies that have investigated the effects of top-down factors on reading, Radach et al. (2008) showed that the task and text format affected eye movements during reading. Their results suggested that word viewing times were shorter when readers had to answer detailed comprehension questions as opposed to verifying which word among multiple choices was present in the text. However, during the more demanding comprehension task, an overall increase in refixation frequency was observed. Moreover, Radach et al. (2008) presented the texts either as a whole passages or as single sentences. The total viewing times for words was longer for passages as compared to sentences, but first-pass viewing times were shorter for passages than for sentences. These observations suggested that passages were read with quick first-pass reading followed by re-reading. Overall, the study by Radach et al. (2008) demonstrated that readers dynamically adapt to changes in reading conditions.

In addition, Carver (1990) has argued that readers can use different processes in order to accomplish their goals, and that the ongoing process is adjusted based on task instructions or the difficulty of the text (also within the task). Carver distinguished five reading processes based on variations in reading rates (words per minute, wpm). The *scanning* state is performed at 600 wpm and is used, for example, when the reader is looking for a particular word from the text. Another rapid process *skimming* (450 wpm) is typical in situations where the reader wants to get an overview of the text without reading the whole

text. The *'rauding'* (300 wpm) process corresponds to the normal reading with the aim of comprehending the text content. *Learning* (200 wpm) is a slow process used for acquisition of new knowledge, while *memorizing* (138 wpm) is the slowest process involving continuous checks of whether the ideas encountered are remembered later.

The division of eye movement trajectories into focused and overview behavior is a well-known distinction in eye movement research (Buswell, 1935). The differences in eye movement trajectories are assumed to reflect different processing states within the specific tasks. For example, supporting a two-level model of attention, studies have demonstrated systematic changes in fixation durations and saccade amplitudes during the time course of scene inspection (Unema, 2005), or as a response to critical events (Velichkovsky, Rothert, Kopf, Dornhöfer, & Joos, 2002). In the two-level model, a *preattentive* / *ambient* processing state is characterized by relatively long saccades and short fixations. This processing state is used for localizing targets from the visual periphery, beyond the foveal and parafoveal regions. Whereas, in the *attentive / focal* processing state, longer fixations are related to short saccades. The saccades are initiated mainly within the parafoveal region where objects can be focally attended and identified during the relatively long fixations.

Moreover, Liechty, Pieters and Wedel (2003) have distinguished different processing states from eye movement data collected in an advertisement viewing task. The existence and relative prevalence of separate attention states over the time course of stimulus exposure was tested using a hidden Markov model (HMM). The HMM segmented the eye movement data into two hidden states reflecting different covert attention states: a *global* and a *local* state. The global state was characterized by short fixation durations and saccades that directed the gaze to nonadjacent (global) locations in the grid that was overlaid on the stimulus. Whereas, long fixation durations along with saccades directed to adjacent (local) cells were common for the local attention state. Furthermore, their results suggested that instead of an orderly global-to-local attention sequence, the participants switched dynamically back and forth between the attention states. In summary, the results presented above, suggest that eye movements can reflect not only the spatial distribution but also the level (or depth) of attention.

1.6 Attention allocation to task-irrelevant stimuli

The online reading process may be disrupted when a salient stimulus captures the reader's overt attention. Theories of visual attention suggest that salient objects attract attention, and that directing attention away from the most salient location of a visual scene requires voluntary effort (Itti & Koch, 2000). Previous studies have demonstrated that changes in the visual field can capture attention involuntarily in a stimulus-driven, *bottom-up* manner. Theeuwes (1994) showed that in a singleton search task, participants could not override the stimulus driven activation although they were told to do so. Moreover, Theeuwes and Burger (1998) observed that, when the salient element was unpredictable and changed from trial to trial, it interfered with the visual search task.

Attention can also be allocated voluntarily based on individuals' current goals, referred to as *top-down* or goal-directed attention. For example, top-down attention allows readers to actively maintain attention on the text or to shift attention voluntarily from one display region to another. The ability to selectively ignore distractors depends on the presence of an attentional set for target and distractor properties (Theeuwes & Burger, 1998). Studies considering the Web environment have suggested that top-down control of attention is capable of overriding the bottom-up attentional capture by salient low-level visual features (e.g., color, orientation, luminance, or motion), and that Web users mainly rely on top-down strategies that help them ignore, for example, the online ads (Drèze & Hussherr, 2003; Stenfors, Morén, & Balkenius, 2003).

Previous research has also shown that overt fixations rarely occur on irrelevant stimuli, such as the online ads, suggesting that the ads affect the performance of the primary task covertly (Burke, Hornof, Nilsen, & Gorman, 2005; Day, Shyi, & Wang, 2006; Drèze & Hussherr, 2003; Hong, Thong, & Tam, 2004). For example, longer search times have been reported in the presence of the ads (Burke, et al., 2005; Hong, et al., 2004). In contrast, some studies showed shorter responses when ads were presented (Day, et al., 2006), possibly suggesting that peripheral ads may increase the participants' level of arousal, which in turn motivates them to increase processing resources. The studies presented above build up a background for the investigation of attentional control during online reading when the text is surrounded by animated or abruptly appearing ads.

2 Aims of the study

The aim of this thesis was to investigate reading in the Web context by using eye tracking and EEG recordings along with behavioral methods. More specifically, this thesis examines how the location of the stimuli presented outside the currently fixated region, text format, animation and abrupt onset of online advertisements, as well as the phase of an information search task affect written language processing in applied reading settings.

Study I examined parafoveal processing of words located either to the left or to the right of the currently fixated word. Previous research has mostly investigated how parafoveal information is processed at the reading direction. Study I introduces the fundamental concepts of information processing during reading. The online reading tasks presented in Studies II-IV differ from the traditional reading conditions mostly in terms of the parafoveal information acquisition. Thus, Study I forms the basis on which the applied work can be developed. The study was designed to test whether attention, long-term learning, or the lateralization of brain structures account for the possible visual field differences in parafoveal processing of words. Furthermore, the study tested the predictions based on parallel and serial models of reading by investigating what type of lexical information is acquired parafoveally. That is, whether the parafoveal effects occur at pre-lexical or at lexical level of word processing (Reilly & Radach, 2006). Therefore, three semantic conditions were included: the stimulus words were either semantically associated, non-associated or the target stimulus was a non-word. To find out whether parafoveal processing differs between the visual fields, the parafoveal words were presented either in the left visual field (LVF) or in the right visual field (RVF). Previous behavioral and brain imaging studies have consistently reported a processing advantage for words presented in the RVF (reviewed in Chiarello, Liu, & Shears, 2001; Ellis, 2004).

Study II investigated the effect of vertical arrangement of text on reading, since Web page layouts often make the text appear in vertically arranged narrow columns. The question was whether reading vertical text benefits from the omission of the horizontal eye movements, the relatively short saccades needed to fixate a single word per line, and the finding that more words can be processed from the vertical than from the horizontal format

during a single fixation (Ojanpää, et al., 2002). The study also included a condition where words longer than 10 characters were hyphenated, because previous studies indicate that long words, such as compound words, are often refixated (Hyönä, Bertram, & Pollatsek, 2004). Moreover, a centered vertical condition was presented, because previous research has shown that the *optimal viewing position*, from which the words are the most easily identified, is close to the center of words (O'Regan & Jacobs, 1992). Centering the words was expected to allow the readers to shift their gaze more easily to the optimal viewing position, resulting in faster reading of centered words.

Study III investigated whether parafoveally presented Web advertisements interfere with online reading, and whether the ads attract more attention during an easy browsing task, than during a more engaging reading task. The online advertisements were presented either above or to the right of a central text. The tested salient features of the ads were animation (Experiment 1) and abrupt onset of an ad (Experiment 2). Moreover, Experiment 3 examined the effect of task orientation on attention toward the ads. The first hypothesis was that if top-down control of attention fails and the ads capture bottom-up attention, the ads are overtly fixated during reading. The second hypothesis was that the ads are covertly attended while the eyes remain on the text with the interference showing up indirectly in the reading performance measures. The third hypothesis suggested that the ads are totally ignored while attention is focused on the primary reading task.

Study IV analyzed the whole sequence of eye movements with a discriminative hidden Markov model (dHMM) in order to gain an insight into how processing alternates within reading tasks. Instead of fixed model parameters, the eye movement parameters were learned from the data. To do so, the data was divided into two subsets: training and testing data. The best model was selected using the training data, and its capability to generalize was tested using the testing data. The inference of processing states was based on comparison of the estimated model parameters to previous eye movement studies.

3 Methods

3.1 Participants

Participants were volunteers with normal or corrected-to-normal vision, and they all provided written informed consents prior to recordings. Table 1 summarizes the participant information across the studies. Participants in **Study I** were all right-handed native speakers of Swedish, and participants in **Studies II-IV** were native speakers of Finnish.

Table 1. Participants in Studies I-IV

Study	**N**	**Females**	**Age (mean) in years**
I	30	17	20–41 (28)
II	8	6	22–31 (24)
III: Exp 1	28	14	19–49 (35)
III: Exp 2	30	15	20–58 (34)
III: Exp 3A III: Exp 3B	32 30	16 15	18–53 (37) 19–54 (29)
IV	10	6	23–29 (26)
Total	168	89	18–58 (30)

3.2 Eye tracking

An advantage of the eye tracking method is that eye movements can be measured as a normal part of reading, and tracking the eyes does not disrupt the ongoing process of reading. The eye movement data can be recorded at a very high temporal resolution, providing a good moment-to-moment indication of cognitive processing during reading (e.g., Liversedge & Findlay, 2000). Video-oculographic eye-trackers measure the center points of pupil and corneal reflex. The eye is illuminated by an infrared (IR) led and an eye camera collects samples of eye images at a high rate. Computationally, the pupil is taken to be the darkest area in the video, and the corneal reflex the bright reflex in the cornea, the second outmost layer in the eye. Pupil and corneal reflex systems use the fact that when the

eye moves, pupil moves faster than the corneal reflex, and thus the relative positions of the pupil center and corneal reflex change (e.g., Holmqvist et al., 2011). The eye-tracker reads the coordinates of both the pupil center and the corneal reflex, and calculates the gaze position on the basis of their relation. Combined use of pupil and corneal reflex provides tolerance to head versus camera movements because the corneal reflex offers a reference point within the eye image. The illumination of the eyes by infrared light allows measurement of the eye coordinates irrespective of varying lightning conditions. The relation between the gaze positions and the stimulus coordinates is established in the calibration procedure where a participant is asked to fixate a set of points for which the precise position is known.

3.2.1 Data acquisition

In **Study I**, participants' eye movements were recorded using a Hi-Speed system eye-tracker (SensoMotoric Instruments, SMI, Teltow / Berlin, Germany), which samples the eye positions at 500 Hz. This sampling frequency was selected in order to use the same sampling frequency for acquisition of both eye movement and EEG data sets. **Study II** used a head-mounted iView RED eye-tracking system by SMI. In **Studies III** and **IV**, eye movements were recorded with a 1750 remote eye-tracking system by Tobii (Danderyd, Sweden), in which the cameras are integrated in the monitor. The Tobii system allows a relatively large degree of head movement, and therefore provides a distraction-free test environment (i.e., no chin or forehead rests were necessary). Both SMI iView RED and Tobii 1750 systems sample the positions of the eyes at 50 Hz sampling frequency.

3.2.2 Data analysis

In order to calculate fixation and saccade metrics from the raw eye coordinate samples, eye movement events (fixations, saccades and blinks) need to be extracted from the raw data. In **Study I**, the event detection was done with a saccade velocity-based algorithm (Smeets & Hooge, 2003), which identifies a saccade when a given velocity threshold is exceeded. Events that are not identified as saccades are assumed to be fixations. Using a velocity-based algorithm requires a high sampling frequency. Because saccade events are short

(around 30–50 ms), and the sampling frequency was low (50 Hz) in **Studies II-IV**, the fixations in these studies were detected using a dispersion-based algorithm. The dispersion algorithm detects fixations by finding data samples that land within a given dispersion threshold for a specified minimum duration of time, and assumes that the rest of the data are saccades.

In **Study I**, a prime word was presented in the middle of the screen along with a parafoval target word, which was located either to the left or to the right of the prime word. Three eye movement measures were calculated on the central prime words: the mean of first fixation and gaze durations (the sum of fixation durations before the eyes left the prime word), and the total reading times (sum of all fixation durations). In **Study II**, the number of fixations per word and per line, as well as the mean fixation duration and the number of regressions were extracted from the eye movement data. In **Study III**, the eye movements were analyzed separately for the text and the ad regions. For the ad regions, the number of eye entries, the total number of fixations, the time of first entry and the total dwell time (summed duration of all fixations) was calculated. For the text region, in addition to the measures listed above, also the mean fixation duration and the number of regressions was analyzed.

In **Studies I** and **II**, the eye movement results were analyzed with a repeated-measures analysis of variance (ANOVA). If the sphericity assumption was violated in **Study I**, the reported *p* values were corrected according to the Greenhouse–Geisser procedure. Newman–Keuls test was used for post hoc testing in **Study II**. In both studies, follow-up ANOVAs and paired t-tests were conducted to test the differences between factor levels if the ANOVAs were significant. In **Study III**, the eye movements landing on the text region were analyzed using an ANOVA. For the eye movements detected on the ads, a Generalized Estimating Equations (GEE) model was used, due to the non-normal distributions of these variables. The frequency measures (i.e., the number of entries and fixations) were analyzed with a negative binomial distribution using a log link, and a gamma regression using an inverse link was used for the other eye movement measures. Least significant difference (LSD) test was used for the post hoc comparisons in **Study III**.

3.2.3 Modeling with Hidden Markov Model (HMM)

HMMs are probabilistic models that associate alternating statistical properties of a signal stream by switching of an unobserved hidden state (Rabiner, 1989). HMMs attempt to describe the full process of how data is being created, and HMMs are applied in a case where the statistical properties of the signal change over time. The model explains these changes by a switch of a hidden (unobservable, latent) state within the model. The underlying assumption is that the signal can be characterized as a parametric random process, and that the parameters of the processes can be estimated in a precise, well-defined manner.

A classical example of an HMM is a scenario where a person is performing a coin tossing experiment behind a curtain and does not tell what s/he is doing exactly; s/he will only tell you the result of each coin flip. Thus, a sequence of hidden coin tossing experiments is performed, and the observation sequence consists of a series of heads and tails (i.e., the number of distinct observation symbols, *M*, per state is two in this case[2]). Given the scenario, the problem of interest is how to build an HMM to explain the physical mechanism that accounts for the observed sequence of heads and tails. The first problem is to decide what the states in the model correspond to (e.g., a distinct biased coin), and then to decide *N*, the number of states (e.g., coins) in the model. Generally, the states in an HMM are interconnected in a way that any state can be reached from any other state, and a state transition probability distribution (*A*) describes the transitions between the states, that is, the probability of switching coins. A complete model specification requires two additional parameters: the observation symbol probability distribution in a state (*B*), that is, the bias of the coin, and the initial state distribution (π), that is, the probability of initiating the coin tossing with a given coin.

Before the model can be useful in real-world applications, we must solve, which model best matches the actual observations, and how do we find the most probable state sequence. A maximum likelihood estimate of the model parameters can be obtained with Baum-Welch algorithm. The observation sequence used to adjust the model parameters is called a

[2] HMMs can also use continuous observation vectors.

training sequence, and the best model is selected based on how well a model is able to predict an unknown (test) observation sequence. The most probable state sequence is found by the Viterbi algorithm (Forney, 1973), which is a special case of dynamic programming algorithms.

In order to study how processing states alternate within information search tasks, a discriminative HMM (dHMM) was learned from the eye movement data collected in **Study IV.** Participants performed three different types of information searches: finding a word among a list of titles (W), finding a sentence that answers a question (A), and choosing a subjectively most interesting topic (I). In addition to the eye movement data, the only information given to the dHMM was the task type of the learning data. A first order Markov property was assumed for the transitions between states, that is, the transition to the next state depended only on the current state. Pieters, Rosbergen and Wedel (1999) have shown that eye movements follow this property. Maximum likelihood parameter values were obtained with the Baum-Welch algorithm, and the most probable path through the model was obtained using the Viterbi algorithm.

A generative model can be converted to a discriminative model by optimizing the conditional likelihood of the model via Bayes formula. Compared to a fully discriminative model (e.g., logistic regression), the converted model still has benefits of a generative model, that is, easier interpretation of the model parameters (for description of the model differences, see Salojärvi, Puolamäki, & Kaski, 2005b). Discriminative training of the HMM is carried out by assigning a set of "correct" hidden states in the model to always correspond to a certain class c (task type in our case), and then maximizing the likelihood of the state sequences that go through the "correct" states for the training sequence. The parameters of a discriminative HMM (dHMM) are optimized with a discriminative Expectation-Maximization algorithm (for derivation of the algorithm, see Salojärvi, Puolamäki, & Kaski, 2005a).

Modeling with HMMs was carried out in a data-driven fashion, that is, the best model topology and parameter values were learned from the training data. The generalization capability of the model was then tested with the test set. Thus, the eye movement

trajectories of **Study IV** were randomly split into training (67 % of the data) and testing data (33% of the data).

The number of hidden states used for modeling the task types (W-A-I) was determined with a 6-fold cross-validation among different state configurations (2-2-2, 2-2-3, 2-3-3, 3-3-3, 3-3-4, 3-4-4, 4-4-4). The number of hidden states was decided by comparing the mean of perplexities of validation sets (see Robertson, Kirshner, & Smyth, 2004). In addition, the paired perplexity values for the 2-3-3, 3-3-3 and 3-3-4 hidden state configurations were compared with Wilcoxon signed rank test. The perplexity of the 3-3-3 model was significantly better than the perplexity of the 2-3-3 model, while the 3-3-3 and 3-3-4 models did not differ. Since the data did not support any preference of the 3-3-4 model, the less complex, 3-3-3 model was chosen. Thus, the dHMM that best fitted the data segmented each information search task into three hidden states (Figure 1).

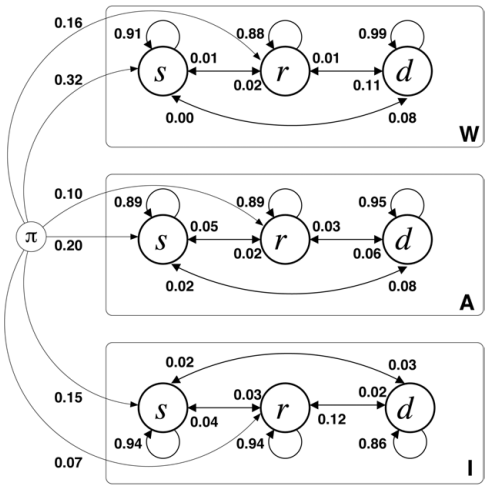


Figure 1. The topology and the transition probabilities of a discriminative hidden Markov model, dHMM. The circles denote the hidden states of the model, and the arrows show the transitions between states, along with their probabilities. The beginning of the sequence is denoted by π, and the capital letters denote the information search task type (W=word search, A=question-answer, I=subjective interests). Small letters within the circles denote the interpretations given for the hidden states (*s=scanning, r=reading, d=decision*).

For the time series model (dHMM), four features of each fixation were computed. These features are listed below with the corresponding modeling distribution:

1) Logarithm of fixation duration in *ms* (one-dimensional Gaussian)

2) Logarithm of outgoing saccade length in *pix* (one-dimensional Gaussian)

3) Outgoing saccade direction (four directions + a fifth state indicating a trial ending, Multinomial)

4) Indicator of whether the currently fixated word has been previously fixated (Binomial)

In the current implementation, we used an HMM that emitted the fixation durations by changing the time scale of the HMM into fixation counts. Thus, instead of having an HMM that is in a state for a certain time, we had an HMM that is in a state at a certain fixation, which has a certain duration.

3.3 Electroencephalography (EEG) and eye fixation related potentials (EFRPs)

EEG records the differences in electrical potentials between two scalp locations as a function of time (Rugg & Coles, 1995). The EEG recorded from the scalp reflects post-synaptic (dendritic) synchronous activity of pyramidal cells that have radial or tangential orientation, allowing the summation of the electrical fields of individual neurons to yield a dipolar field. As EEG records electrical activity directly related to neuronal activity, it can reflect the time course of neuronal activity at a millisecond scale. However, due to attenuation and distortions by the tissues between the electrodes and the source of the neuronal activity, the EEG signal does not carry accurate information of its spatial source location in the brain. Embedded in the EEG are the relatively small neural responses associated with specific sensory, cognitive, or motor events. These responses can be extracted from the background EEG by averaging together multiple responses belonging to the same stimulus condition. These averaged responses are called event-related potentials (ERP).

In order to prevent overlap between different cognitive processes, the stimuli in the ERP paradigms are presented in isolation with unnaturally long interstimulus intervals between them. However, presenting the stimuli (e.g., words) isolated by long intervals prevents the

study of processes that would occur under normal conditions, such as spillover effects during reading (Baccino, 2011). In the EFRP -technique, participants' EEG and eye movements are recorded simultaneously. EEG and eye tracking have comparable temporal resolutions, which is an advantage in the data analyses. That is, instead of analyzing responses related stimulus events, EEG analyses can be time-locked to eye fixations on specific stimulus regions (Baccino, 2011; Baccino & Manunta, 2005; Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, in press; Hutzler et al., 2007). Due to saccadic suppression, little or no useful visual information is acquired during a saccade (Morrone & Burr, 2009). Thus, the fixation onset provides a natural time-locking point to study information processing during unconstrained viewing situation (Dimigen, et al., in press). Averaged potentials aligned to fixation onsets are called eye-fixation-related potentials (EFRPs) (Baccino, 2011; Baccino & Manunta, 2005) or alternatively fixation-related potentials (FRPs)[3] (Dimigen, et al., in press; Hutzler, et al., 2007).

Saccades are typically treated as artifacts in the EEG analyses, because the movement of the eyelids and the rotation of the eyeball's corneoretinal dipole produces fluctuating electrical fields which propagate to the EEG electrodes and contaminate the recording of brain activity (Berg & Scherg, 1991; Rugg & Coles, 1995). However, these potentials attenuate with increasing distance to the eyes (Picton et al., 2000). To avoid eye movement related artifacts, EFRP analyses can be restricted to the fixation period before the saccade when the eye is relatively still (e.g., Baccino & Manunta, 2005). When the analysis is restricted to the fixation period, it is possible to analyze the early ERP components such as the N1 or P2 which have been shown to be sensitive to word characteristics (Baccino & Manunta, 2005; Sereno & Rayner, 2003).

When the time window of the EFRP exceeds the fixation duration, a careful correction of corneoretinal and myogenic eye movement artifacts (that does not eliminate the genuine brain activity) is a necessary precondition for the EFRP analyses (Baccino, 2011; Dimigen, et al., in press). For example, there is a discrepancy between the typical fixation durations

[3] Potentials aligned at saccade onsets are called saccade-related potentials (see Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (in press). Co-registration of eye movements and EEG in natural reading: Analyses & review. *Journal of Experimental Psychology: General*.)

during reading (200–250 ms) and the latency of the N400 component. N400 is a negative wave around 400 ms post-stimulus and a robust measure of semantic processing in psycholinguistic ERP research (Kutas & Hillyard, 1980). Thus, in normal reading situations, the eyes have already left the word when the N400 peaks (Sereno & Rayner, 2003). Other factors that need to be taken into account when analyzing EFRPs in natural viewing conditions (e.g., during reading) are the varying degree of temporal overlap between the potentials evoked by target fixations and the background EEG activity as well as the temporal overlap between the potentials elicited by successive fixations (Baccino, 2011; Dimigen, et al., in press).

3.3.1 Data acquisition

The fixation duration on a particular word does not tell when within a fixation the different stages of word recognition are accomplished (Sereno & Rayner, 2003). To elucidate the stages of processing that occur within a fixation, **Study I** included co-registration of eye movements and EEG. The EEG data were recorded with a 128-channel HydroCel Geodesic Sensor Net™ connected to an AC-coupled, 128-channel, high-input impedance amplifier (300 MΩ, Net Amps™, Electrical Geodesics Inc., EGI, Eugene, USA). Amplified analog voltages were high-pass filtered (0.1 Hz) and digitized at 500 Hz. Individual sensors were adjusted until impedances were less than 50 KΩ. Electro-oculography was monitored with sensors placed on the outer canthus and infraorbital ridge of each eye. Synchronization of the EEG and eye movement data was obtained by the stimulus presentation software (E-Prime, Psychology Software Tools Inc., Pittsburgh, USA), which was sending a synchronizing signal to both data sets as soon as a stimulus was presented on the screen.

3.3.2 Data analysis

The data were analyzed with Net Station (EGI) software. Amplified voltages were off-line referenced to linked mastoids and filtered with a 0.3–40 Hz band pass filter. Remaining artifacts were removed automatically with ±140 μV rejection level. ERPs time-locked to word pair presentation were epoched with a window of -100 to 300 ms and baseline corrected relative to a 100 ms pre-stimulus interval. ERP amplitudes and peak latencies

were analyzed over eight channel groups located near the standard electrode sites of the 10/20 system (Jasper, 1958) (Figure 2): frontal area (F3 and F4), central area (C3 and C4), parietal area (P3 and P4) and occipital area (O1 and O2). At frontal, central and parietal electrode sites, the mean amplitudes and peak latencies were calculated for the N1 (70–120 ms) and P2 (140–280 ms) components. At occipital sites, the following components were calculated: P1 (90–140 ms), N1 (140–200 ms) and P2 (200–280 ms).

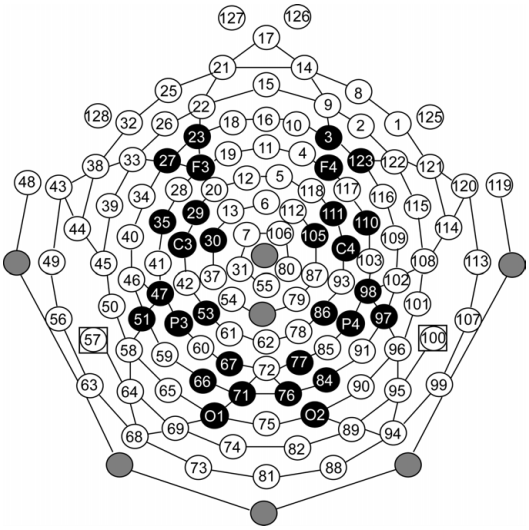


Figure 2. Geodesic sensor net layout depicting the numbered electrode sites. Black electrode clusters show the sites around the standard electrode positions of the 10/20 system that were included in the EFRP analyses.

Eye movement and EEG responses were combined offline to analyze the EFRP responses. Trials were excluded if the first fixation was not detected on the prime word or if the first fixation or gaze duration were shorter than the EEG epoch length (300 ms). Due to these criteria, refixations on the prime words were possible before the eyes moved to the target. A set of analyses confirmed that such refixations did not confound the results. That is, the proportion of refixations did not differ between word conditions. The within-word saccade amplitudes were around 1.5°, producing deflections around 24 μV at the frontal electrodes (Luck 2005). These voltages fade off as the distance between the eyes and the electrode site increase, and thus they were unlikely to distort EEG data at the occipital sites.

Moreover, the within-word saccade amplitudes, saccade directions and the refixation durations did not differ between conditions that were critical for the results.

The grand average EFRPs showed different components at occipital sites as compared to frontal, central and parietal sites, and therefore separate statistical testing was performed for the occipital site. All EFRPs were analyzed with repeated-measures ANOVA, and the Greenhous–Geisser correction was applied when appropriate. If the ANOVAs gave significant effects, then follow-up ANOVAs or paired samples t-tests were undertaken to test the differences between different levels of factors.

3.4. Behavioral methods

In **Study I,** the behavioral (semantic judgement) responses were collected to restrict the eye movement and EFRP analyses to trials where participants agreed with the pre-classification of the stimulus categories. Furthermore, the response time data were collected in **Study I**. The behavioral data were analyzed with repeated-measures ANOVAs in **Studies I** and **II**. In **Studies II** and **III**, reading rate was measured as the number of words covered within a certain time (words per minute, wpm), and comprehension accuracy was measured as the proportion of correct answers to text content questions. Furthermore, post-experimental questionnaires investigating participants' perception and attitudes towards the stimuli were collected. In **Study II**, participants gave preference scores for each text format. In **Study III**, after participants finished reading the texts, they were asked to rate whether they had paid attention to the online ads, and whether the ads distracted them during reading. In **Study III**, the effect of participants' self reported attention and distraction on eye movement measures was tested by adding the self-report measures to the ANOVA and GEE models. Further, least significant difference (LSD) test was used for the post hoc comparisons.

4 Experiments and results

4.1 Study I: Visual field differences in parafoveal processing of words

The study examined how the characteristic of the parafoveal words presented either to the left or to the right of the fixated word affected the processing of the currently fixated foveal word (i.e., the parafoveal-on-foveal effects). Three semantic conditions were included, that is, the stimulus words were either semantically associated, non-associated or the parafoveal stimulus was a non-word. Previous studies have demonstrated that words presented in the right visual field (RVF) are processed faster and more accurately than words presented in the left visual field (LVF) (Chiarello, et al., 2001; Ellis, 2004). It has been suggested that this asymmetry results either from the *perceptual learning* effects related to the reading direction, *attentional* asymmetry, or the *structural* (cerebral) asymmetry for language processing.

The *perceptual learning* account suggests that the visual training associated with the regularity of reading eye movements biases word recognition within a restricted horizontal region close to the fovea toward the side of the reading direction (Dehaene, Cohen, Sigman, & Vinckier, 2005). Studies have shown that perceptual learning occurs at early stages of visual processing providing a mean for rapid and efficient recognition of words (Gilbert, Sigman, & Crist, 2001; Sigman & Gilbert, 2000). According to this account, a RVF advantage is predicted because the Swedish-speaking participants are trained to read from left-to-right (Nazir, Ben-Boutayab, Decoppet, Deutsch, & Frost, 2004).

According to the *attention* account, an attentional asymmetry is produced by the pattern of constant left-to-right fixations during reading of alphabetical languages (Inhoff, Pollatsek, Posner, & Rayner, 1989). The attention account, however, is also able to predict similar responses for the RVF and the LFV presentations, because studies have shown that parafoveal information is extracted primarily from the location which is about to be fixated next (Henderson, et al., 1989; Inhoff, et al., 1989).

The *structural* account assumes that information to the right of fixation benefits from direct access to the left cerebral hemisphere, which for most individuals is dominant in language processing. The structural account proposes that information extracted from the RVF is directly transferred to the left hemisphere, and that around 200 ms post-stimulus, information from both visual fields converge in the visual word form area, in the left inferior temporal region (Cohen, et al., 2000). After that the responses have similar topographies independent of the originally stimulated visual field.

4.1.1 Stimuli and procedure

The stimuli comprised 288 nouns as central prime words. For the prime words, 96 semantically associated and 96 non-associated nouns, along with 96 non-words were selected as parafoveal targets (Figure 3). The semantic relatedness between the primes and targets was assessed by 80 participants, and further confirmed with latent semantic analysis (LSA) (Landauer & Dumais, 1997). The participants' ratings and LSA indexes showed that the target words were more related to the primes in the semantically associated condition than in the non-associated condition. The non-words consisted of random letter strings. Because random generation can produce pronounceable letter combinations, 80 participants assessed how easy it was to pronounce the non-words. These ratings suggested that the non-words were more difficult to pronounce than words. The prime and target words were presented as pairs that were balanced in word length allowing up to two letters difference. Furthermore, the conditions were balanced on word frequency and orthographic neighborhood density. In each condition, half of the stimuli were presented in the LVF and half in the RVF with the visual fields counterbalanced across participants, allowing each word to be seen equally often in both visual fields.

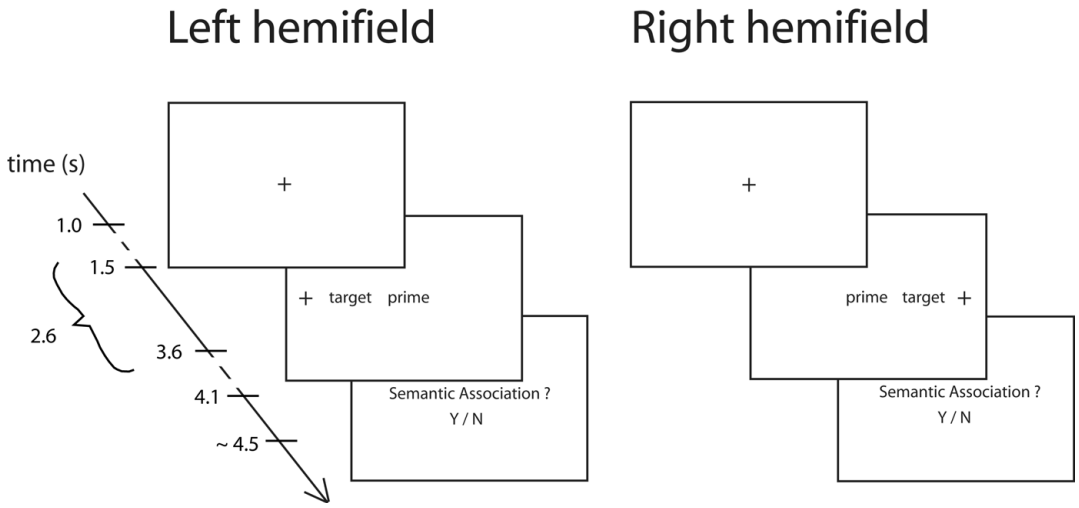


Figure 3. Schematic display of the stimuli used in Study I. After the central fixation, the word pair was visible for 2.6 s. Subsequently, participants were asked to make a semantic association judgment. The average trial length was 4.5 s, and the next trial began after the response was given.

Figure 3 illustrates the example trials. Followed by a central fixation, the central prime word and the parafoveal target word appeared on the screen. Participants were asked to read the prime, proceed to the target, and then to move their eyes to the cross on the side of the target. To investigate parafoveal-on-foveal effects, all analysis considered processing of the prime words. Because previous studies using eye tracking did not find differences between left-to-right and right-to-left reading (see Inhoff, et al., 1989), the EFRP technique was used to obtain a finer analysis of the visual field differences.

4.1.2 Results

The response data was evaluated against the pre-classification of words into semantic conditions, and marked correct when they were in accordance with the pre-classification[4]. The word conditions differed in percentage of errors, suggesting greater amount of errors in the non-associated than in the associated condition, and more errors in the associated than

[4] All eye movement and EFRP analyses were restricted to trials with correct responses.

in the non-word condition (Table 2). Further, the response times were longer for non-associated than for associated or non-word pairs.

The eye movement results indicated longer first fixation durations on the primes in the non-associated condition compared to the non-word condition when the target was presented in the RVF (Table 2). Furthermore, the total reading times of the primes were longer for semantically associated and non-associated conditions as compared to the non-word condition when the target was presented in the RVF, suggesting longer reading times when the targets were orthographically legal words (Table 2).

Table 2. Response and eye movement results (in ms) as a function of word conditions and the visual fields (LVF / RVF). The eye movement measures correspond to the processing of the prime words. Reading time is calculated as the sum of all fixation durations on the prime word.

	Errors (%)	***Response time***	***First fix dur***	***Gaze dur***	***Reading time***
Associated					
LVF	2.74	715.27	318.70	416.67	673.01
RVF	2.92	678.64	324.74	422.46	673.24
Non-associated					
LVF	5.97	799.31	331.98	415.73	688.43
RVF	6.63	771.78	338.35	436.63	652.56
Non-word					
LVF	.46	736.41	333.15	408.97	639.45
RVF	.80	699.11	318.90	408.55	602.00

The EFRPs measured at occipital sites showed greater P1 amplitudes for the LVF than for the RVF targets in the left hemisphere. Moreover, the occipital N1 responses were enhanced over the left hemisphere (Figure 4), possibly reflecting processing taking place at the visual word form area in the left inferior temporal region (Cohen, et al., 2000). The differences between word conditions in occipital P2 responses approached significance. The follow-up analyses indicated that the word conditions differed significantly when the targets were presented in the RVF. These results suggested greater P2 responses when the target was a semantically associated or non-associated word compared to the non-word

condition. Thus, in the RVF, the parafoveal processing of non-words differed from the processing of orthographically legal words at 200–280 ms post-stimulus, while the semantic association did not result in any parafoveal-on-foveal effects (Figure 4). No such differences were observed for the targets presented in the LVF. The topographic maps show the differences between the RVF and LVF presentations in scalp distributions of the word–non-word effects (Figure 5).

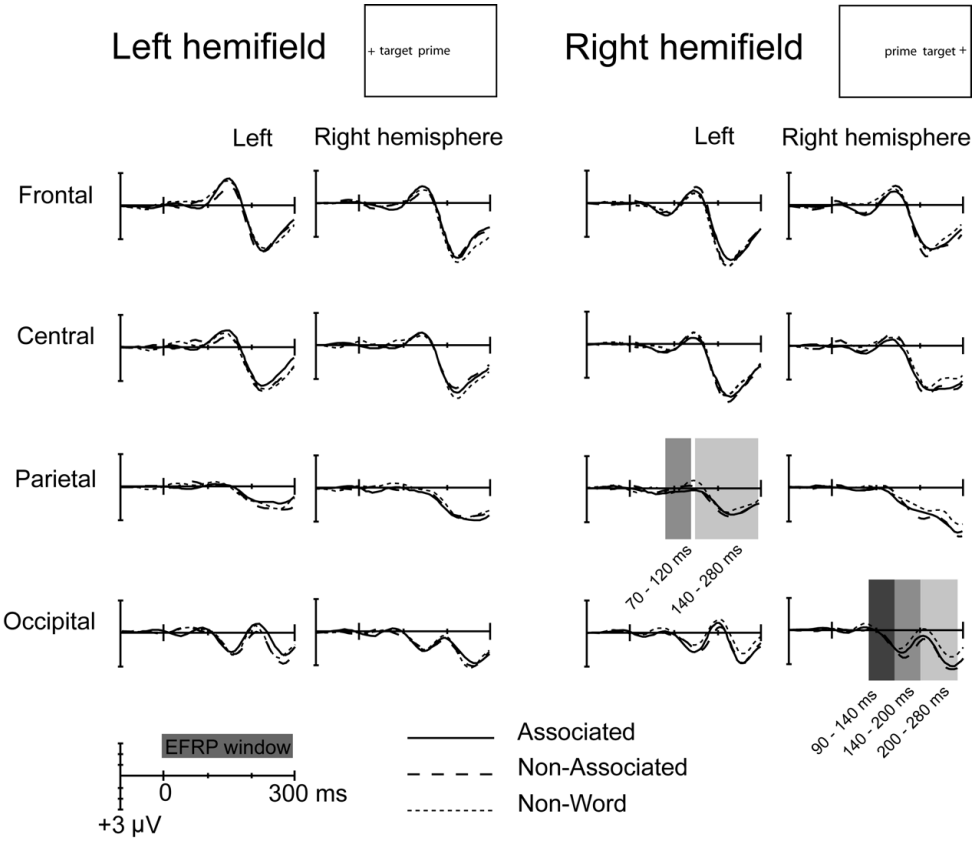


Figure 4. Grand average eye-fixation-related potentials (EFRPs) elicited by the Associated, Non-associated and Non-word conditions presented to the left and right visual fields at representative electrode sites over right and left hemispheres (see Figure 2). A 15 Hz filter was used for data plotting.

Fronto-parietally, the N1 responses at left frontal and left central sites were stronger for LVF targets. Furthermore, the RVF targets elicited stronger N1 responses over the frontal right site compared to the responses elicited by the LVF targets. These observations suggested ipsilateral effects, that is, greater N1 responses at the hemisphere of the

stimulated visual field. In contrast, the P2 responses at left frontal and left central sites were larger for RVF targets, and the P2 responses at right frontal site were enhanced for LVF targets, suggesting contralateral effects. Moreover, the P2 responses at left central site were larger for associated and non-associated targets presented in the RVF as compared to responses elicited by target words presented in the LVF. Also at parietal sites, the P2 responses were stronger for RVF than for LVF targets, and the amplitudes were emphasized at right hemisphere.

To summarize, the behavioral results in Study I did not differ between visual fields. However, the first fixation durations and the total reading times were longer when the RVF parafoveal targets were words as compared to when the targets were non-words. Moreover, measuring the EFRPs allowed a finer analysis of the visual field differences. These results demonstrated that a bilateral occipital P2 response at 200–280 ms post-stimulus differentiated processing of orthographically legal words from non-word processing when the target stimuli were presented in the RVF. No such effect was observed for targets presented in the LVF.

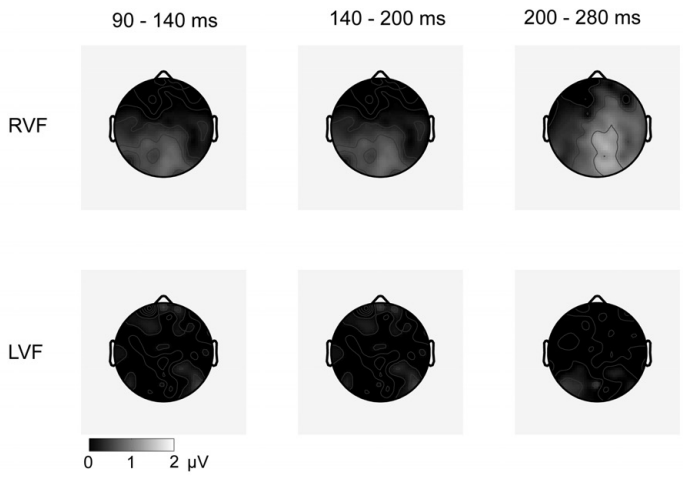


Figure 5. Topographic maps illustrating the scalp distributions of the relative amplitude differences between parafoveal word versus non-word processing in the three time windows used in the analyses for the occipital

effects. The top row shows the responses when the target word was presented in the right visual field (RVF), and the bottom row shows responses elicited by the left visual field (LVF) targets.

4.2 Study II: The effect of vertical arrangement of text on reading

The purpose of Study II was to compare eye movement patterns during vertical and horizontal reading. One prediction was that reading vertical text would be faster due to omission of horizontal saccades, and because previous research has demonstrated fewer fixations and shorter saccades in word search from vertical than from horizontal lists (Ojanpää, et al., 2002). Moreover, it was hypothesized that vertical reading would benefit from shorter saccades needed for word-by-word fixations and the decreased saccade planning times associated with shorter saccade lengths (Viviani & Swensson, 1982).

An alternative hypothesis was that because the Finnish participants have learned and practiced reading in horizontal direction, vertical reading involving planning and making upright eye movements would be slower than reading of horizontal lines. Furthermore, it is well know that useful information for reading is extracted from words to the right of the currently fixated word. In vertical reading, this useful parafoveal information is lost. However, the readers may obtain parafoveal information in the vertical direction (Ojanpää, et al., 2002), but since the visual acuity falls off more rapidly in the vertical than in the horizontal direction (Curcio & Allen, 1990), and because the readers have not learned to utilize parafoveal information in the vertical direction, it is possible that they cannot benefit from the vertical presentation of text. Slower reading of vertical text would create usability problems on Web pages where the text appears in narrow columns.

4.2.1 Stimuli and procedure

The vertical texts were either left aligned or centered and the words (longer than 10 characters) were either hyphenated or not (Figure 6). In addition, a fifth, standard horizontal reading condition was included. The material comprised online magazine articles of around 1100 words. In the horizontal condition, the average length of the texts was 7 pages. In the vertical one-word-per-line conditions, the average length was 34 pages, and in the

hyphenated formats the texts were on average 44 pages. All participants read two texts of each condition. The effect of practice was studied by comparing reading performance between the first and the second reading sessions. The order of presenting the conditions and the texts was counterbalanced across participants.

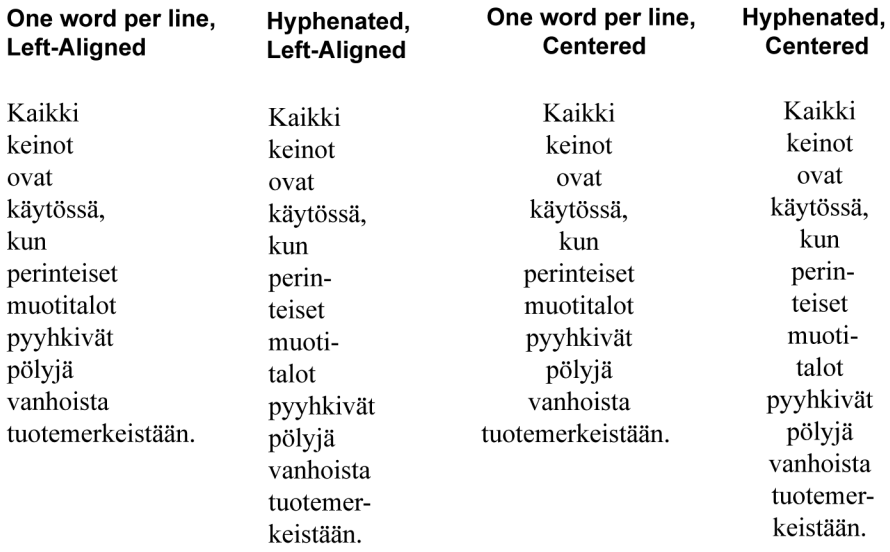


Figure 6. Examples of the four vertical text formats used in Study II.

4.2.2 Results

The results showed that vertical texts were read at slower rate measured as words per minute (wmp) than horizontal text, but there were no differences between the hyphenated and the one-word-per-line conditions (Table 3). Moreover, the first reading session was faster than the second one for each text format. Comprehension accuracy did not differ between the text formats, but the reading efficiency as Speed x Percentage Correct on the Comprehension Test (Jackson & McClelland, 1979) was better for the horizontal than for the one-word centered format. Because the number of pages for vertical text were 5–7 times larger, it is possible that reading time was longer for vertical text because participants spent more time in turning the pages.

The preference scores (Table 3) indicated that the horizontal and the one-word-per-line formats were equally preferred, whereas the horizontal format was preferred over the left aligned hyphenated format. Generally, the hyphenated formats were preferred less than other formats. In their post-experiment reports, participants told that they would prefer vertical formats for skimming and rapid reading. They thought that hyphenation was disturbing because it made the text appear in fragments. Preference for left aligned and centered formats varied between participants, some of them preferred the left aligned format while the others preferred the centered texts.

Table 3. Summary of the reading performance measures in Study II.

	Horizontal	***One word per line***		***Hyphenated***	
		Left-Aligned	Centered	Left-Aligned	Centered
Reading rate (wpm)	215	188	192	164	161
Comprehension	.90	.88	.88	.86	.93
Efficiency	193	165	169	140	160
Preference	5.1	4.6	4.9	2.9	3.2

Eye movement results showed that the mean fixation duration was longer for the vertical formats than for the standard horizontal text (Table 4). The text format did not affect the number of fixations per word, the number of fixations per line, or the number of regressions.

Table 4. Summary of the eye movement measures in Study II.

	Horizontal	***One word per line***		***Hyphenated***	
		Left-Aligned	Centered	Left-Aligned	Centered
No. of fix / word	1.2	1	1.1	1.3	1.2
No. of fix / line		1	1.1	.90	.90
Mean fix dur	209	279	274	300	290
No. of regressions	.09	.07	.08	.08	.08

4.3 Study III: The effect of animated and abruptly appearing ads on online reading

Study III investigated whether animation or onset of advertisements on Web pages capture attention and distract reading, because human visual system has been shown to allow priority to behaviorally urgent events, such as peripheral motion (Franconeri & Simons, 2003; Theeuwes, 1994). As a consequence, it was expected that the ads might influence online reading in three possible ways: i) ads are efficiently avoided by top-down attentional processes, ii) ads attract visual attention and are overtly fixated, causing a disruption to the ongoing reading process, or iii) covert attention to ads disrupts reading with the influence showing up indirectly in the reading performance while the eyes remain fixated on the text region. The salient feature of the ads in Experiment 1 was animation, Experiment 2 tested the effect of abrupt ad onset on reading, and Experiment 3 assessed the effect of task-orientation on attention toward online ads.

4.3.1 Stimuli and procedure

The stimuli comprised 32 authentic Web pages each containing a central text, an ad above the text and another ad to the right of the text (Figure 7). The texts were shortened versions (approximately 100 words) of the online magazine articles used in Study II. The advertisements were 64 full-color ads depicting 16 different topics. To control for the effects of ad content, four different versions of each topic were professionally designed. That is, the same topic was presented above and to the right of the text either in static or animated versions, but different versions of the same topic never appeared simultaneously on a stimulus Web page. The static versions were representative frames of the corresponding animated ads.



Figure 7. A black-and-white version of a stimulus Web page used in Study III. The actual stimulus pages were colored. The frames indicate regions of interest in the eye movement analyses: the text, the ad above the text, and the ad to the right of the text. No frames were presented on the actual stimuli.

In Experiment 1, both ads were simultaneously visible and presented under the following four conditions: both ads were static (S+S), the ad above was static and the ad to the right was animated (S+A), the ad above was animated and the ad to the right was static (A+S), and both ads were animated (A+A). The expectation was that a high degree of animation (the A+A condition) would attract the most attention. In Experiment 2, either the ad above or the ad to the right of the text appeared after a random delay of 0–12 s, or either of the ads was presented throughout the trial. Experiment 3 tested the effect of task orientation on attention toward the ads under four Web page conditions: no ads were presented (baseline), two ads were presented throughout the trial (the ad above the text was static and the ad to the right was animated, S+A), only an animated ad above the text was presented after a random delay of 0–12 s (A+blank), and only an animated ad to the right of the text was presented after a random delay of 0–12 s (blank+A).

In Experiments 1, 2 and 3A, the participants were instructed to read the texts for comprehension, and a four-choice question about the text content was presented after each text. Thus, attention was primarily directed to the text, while the ads were considered as secondary stimuli to the reading task. In Experiment 3B, the task was to browse the pages according to participants own interest, and no comprehension questions were presented. In

all experiments, the participants read the texts or browsed the pages at their own pace while their eye movements were recorded.

4.3.2 Results

In Experiment 1, the condition where the ad above the text was static and the ad to the right was animated (S+A) showed a clear pattern of increased attention toward the ads. This was indexed by increased dwell times (summed fixation durations) on both ads and by greater number of fixations toward the ad on the right in the S+A condition (Table 5). Contrary to the expectation that a high degree of animation would attract the most attention, the results indicated that the ads were attended the most when a moderate amount of animation was presented. Further, the results showed no effects of the ad conditions on eye movements during reading the texts (Table 5). Participants' self-report measures were consistent with their eye movement results, suggesting longer dwell time on the ad above and increased number of fixations on the ad to the right for participants who reported attention to ads.

Table 5. Reading performance and eye movement measures for the text, the ad above and the ad to the right of the text in Experiment 1 of Study III, as a function of ad conditions: *S+S* = both ads were static, *S+A* = the ad above was static and the ad to the right was animated, *A+S* = the ad above was animated and the ad to the right was static, *A+A* = both ads were animated. Time of entry is calculated as seconds from the trial onset.

Ad condition	***S+S***	***S+A***	***A+S***	***A+A***
Text				
Comprehension	.88	.89	.85	.87
Reading rate (wpm)	216.91	220.13	197.41	194.40
No. of entries	1.93	2.09	2.27	1.92
No. of fixations	93.81	98.17	100.60	95.56
No. of regressions	18.20	19.33	20.17	18.65
Time of entry (s)	.58	.62	.70	.61
Mean fixation duration (ms)	208.19	214.38	211.84	210.33
Ad above text				
No. of entries	.40	.28	.50	.23
No. of fixations	.86	.80	.96	.51
Time of entry (s)	15.35	16.98	15.72	13.10
Dwell time (s)	.92	1.56*	.88	.75
Ad to the right of text				
No. of entries	.48	.77	.64	.59
No. of fixations	.74	1.55*	1.02	1.03
Time of entry (s)	13.69	15.39	13.26	15.90
Dwell time (s)	.80	1.44*	.90	.87

* indicates a significant difference between the S+A –condition and the other conditions

Experiment 2 tested the effect of abrupt ad onsets on reading and attention toward the ads. The results indicated poorer comprehension when the ad was presented to the right compared to when the ad was presented above the text (Table 6). Further, the onset of the ad on the right impaired reading comprehension more than the onset of the ad above. However, the reading rate was slower when the ad above was presented as compared to

when the ad on the right was presented. Moreover, reading comprehension improved when the ad above appeared abruptly compared to when it was present throughout the trial.

Table 6. The effect of abrupt onset of ads on reading performance and eye movement measures for the text in Experiment 2 of Study III.

	Ad above the text		***Ad to the right of the text***	
	Onset	No-onset	Onset	No-onset
Comprehension	.93	.89	.85	.89
Reading rate (wpm)	178.04	182.86	200.25	211.30
No. of entries	2.16	1.97	2.22	2.40
No. of fixations	108.30	106.22	109.64	108.62
No. of regressions	23.25	22.70	24.20	23.36
Time of entry (s)	.44	.46	.38	.47
Mean fix dur (ms)	260.66	260.56	265.77	264.40

The eye movement results on the ads showed an association between the onset time of the ad on the right and the time when it was fixated for the first time (r = .482, p < .001). No such effect was observed for the ad above the text, but an abrupt onset increased the number of eye entries and the number of fixations toward the ad above (Table 7). Abrupt onset did not affect the eye movements toward the ad on the right. Participants' self-report results showed that reported attention was associated with increased number of entries to the text and with longer mean fixation durations during reading. Furthermore, participants who reported paying attention to ads made more entries and fixations to both ads compared to participants who did not report attention to ads. These results suggest that overt gaze behavior was consistent with participants' self-reports.

Table 7. The effect of abrupt onset of ads on eye movements to ads in Experiment 2 of Study III.

	Ad above the text		***Ad to the right of the text***	
	Onset	No-onset	Onset	No-onset
No. of entries	.36*	.21	.48	.53
No. of fixations	.76*	.54	1.27	1.30
Time of entry (s)	15.94	12.49	15.69	13.92
Dwell time (ms)	765.86	873.56	1125.72	1263.47

* indicates a significant difference between the onset and no-onset conditions

Experiment 3 investigated the effect of task-orientation on attention toward online ads by a comparison of two tasks. In Experiment 3A, the participants read the texts for comprehension. In Experiment 3B, they browsed the pages according to their own interest. The task had an effect on eye movements during reading, with faster reading rates and fewer fixations and regressions in the free browsing than in the reading task (Table 8). Furthermore, in the S+A condition, the text was entered later during free browsing than during reading. The results suggested more superficial reading strategy in the browsing than in the reading task. These results are also in line with earlier findings on the relationship between eye movements and text memory (Hyönä, Lorch, & Kaakinen, 2002; Hyönä & Nurminen, 2006).

Table 8. Eye movement measures for the text in Experiment 3 of Study III, as a function of task (Reading: Exp. 3A vs. Browsing: Exp. 3B) and the studied ad conditions: *Baseline* = no ads were presented, *S+A* = the ad above was static and the ad to the right was animated (same condition as in Experiment 1), *A+blank* = an animated ad above the text appeared after a random delay between 0–12 s, *blank+A* = an animated ad to the right appeared after a random delay of 0–12 s.

Ad condition	***Baseline***	***S+A***	***A+blank***	***blank+A***
Reading rate (wpm)				
Reading	173.73	169.16	169.74	163.60
Browsing	282.37*	291.39*	283.59*	287.09*
No. of fixations				
Reading	109.55	110.32	111.44	113.40
Browsing	82.45*	78.62*	81.92*	81.77*
No. of regressions				
Reading	23.33	23.00	23.14	24.12
Browsing	17.31*	16.49*	17.52*	17.52*
Time of entry (s)				
Reading	.34	.40	.37	.36
Browsing	.39	.92*	.52	.42
Mean fixation duration (ms)				
Reading	277.76	281.61	276.48	282.17
Browsing	258.74	261.73	262.21	262.01

* indicates a significant difference between the reading and free browsing tasks

Participants paid more attention to the ads in the free browsing than in the reading task. This was indicated by increased number of eye entries and fixations as well as longer dwell times on the ads in the browsing than in the reading task (Table 9). Moreover, participants' self-reported attention was associated with the number of entries and fixations as well as the dwell times on both ads. In the free browsing task, the onset time of the ad above (r = .256, p = .011) and the onset time of the ad to the right of the text (r = .520, p < .001) correlated with the time when the first fixations were directed to the ads. In the reading

task, the onset time of the ad above did not affect the time when the first fixations were detected on the ad, but the onset time of the ad to the right was associated with the time of first fixations to it (r = .246, p = .014). Figure 8 illustrates that the ads attracted more fixations during the first five fixations after an ad onset compared to later occurring fixations.

Table 9. Eye movement measures for the ads in Experiment 3 of Study III, as a function of task (Reading: Exp. 3A vs. Browsing: Exp. 3B) and the studied ad conditions: *A+blank* = only the ad above (animated) appeared after a random delay of 0–12 s, *S+A* = the ad above was static and the ad to the right was animated, *blank+A* = only the ad to the right (animated) appeared after a random delay of 0–12 s.

	Ad above the text		***Ad to the right of the text***	
	A+blank	S+A	blank+A	S+A
No. of entries				
Reading	.19	.13	.68	.47
Browsing	.91*	.84*	1.42*	1.37*
No. of fixations				
Reading	.40	.29	1.25	.95
Browsing	2.30*	2.15*	4.56*	4.21*
Time of entry (s)				
Reading	12.27	20.62	19.25	20.69
Browsing	17.30	11.56	14.11	8.90
Dwell time (ms)				
Reading	882.51	680.96	1015.09	882.69
Browsing	1698.12*	1184.82*	2309.41*	2057.30*

* indicates a significant difference between the reading and free browsing tasks

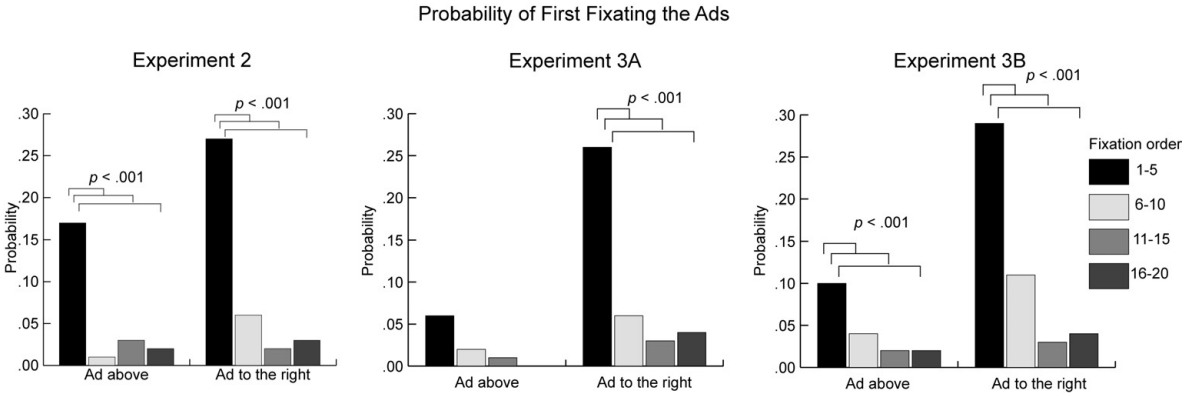


Figure 8. The probability of first fixating the ads at different ordinal fixation positions after abrupt onsets of the ads. In Experiment 2 and 3A the task was to read the texts for comprehension, and in Experiment 3B participants browsed the pages according to their own interest. Brackets indicate significant differences in Z-tests for two proportions.

4.4 Study IV: Alternating processing states during information search tasks

A typical task in the Web environment is to search for information on a specific topic. Reading processes in such search tasks most likely differ from the reading process when single sentences are being read, and which is the typical scenario described by the eye movement models of reading. Previous research investigating online information search behavior has shown that users spend most of the time fixating the first and the second search engine results before they make a selection, while users who selected the lower ranked documents had typically viewed more results overall (Granka, Joachims, & Gay, 2004). This observation suggests, that users scan the search engine listings from top to bottom. Moreover, Aula, Majaranta and Räihä (2005) demonstrated that scanning strategies varied according to users' experience with computers. More experienced users tended to fixate only a few relevant-looking results before they made a selection, whereas less-experience users scanned also the irrelevant results below the document that was selected. However, instead of reporting measures that are typical for reading research, the studies above analyzed only the total time spent fixating on each search engine result. Such

measures do not allow making inferences about the actual online reading behavior during information search.

Furthermore, eye movement studies on reading typically analyze how word features determine when and where the eyes move, and report summary statistics (e.g., mean fixation duration, saccade landing site). Therefore, such studies do not account for variations in the time course of the reading task. In Study IV, the whole sequence of eye movements was analyzed to gain an insight into how processing alternates within the information search tasks. A reverse inference approach was adopted to infer hidden cognitive states from the eye movement behavior (see for discussion Poldrack, 2006).

The relation between eye movements and cognitive states was modeled with a discriminative hidden Markov model (dHMM). The dHMM was used to map the changes in statistical patterns of eye movements to changes in the hidden states of the model. A hypothesis on the cognitive states corresponding to the hidden states was made then by comparing the model parameters to literature on eye movement research where the cognitive state is known. The best model topology, that is, the number of hidden states, was found by comparing several possible model topologies with cross-validation.

4.4.1 Stimuli and procedure

The stimuli were 500 online newspaper titles (revised to grammatical sentences) divided into 50 lists of 10 sentences. To control for previous topic knowledge, three general topics were selected: Finnish homeland news (20 lists), foreign news (20 lists) and business/finance news (10 lists). The tasks represented simple online information search tasks where the user is scanning listings returned by a search engine (e.g., Google) to find a topic of interest. The task types were selected to fit the possible practical implementation in a proactive information retrieval application.

The level of complexity of the searched topic varied. The target could be a word (W), an answer to the question presented before the list was displayed (A), or the most subjectively interesting title in the list (I). To minimize stimulus-driven factors on processing, the same stimuli were presented in all task types. Number of fixations, mean fixation duration and saccade lengths were calculated across repetitions, to control for the possible effects of

presenting the same lists three times during the experiment. These analyses revealed no effects of repetition.

The trial structure was identical across the tasks. First, an assignment was presented, followed by a list of sentences. Participants were instructed to view the sentences until they found a relevant line. Eye movements were recorded for this period. After finding the relevant line, the same sentences were presented with the line numbers, and participants typed in the number corresponding to the selected line. Each participant conducted a total of 150 assignments.

4.4.2 Results

The 9-state, 3-3-3, dHMM achieved the classification accuracy of 60.2% for the testing data, that is, 27 % above the pure change level (33.3%). The dHMM could predict the word search and subjective interest tasks (Table 10), but separating the question-answer task was more difficult. This was possibly because the difficulty of the question-answer tasks varied, that is, the search behavior in easy tasks may have been similar to the word search tasks, while the more difficult tasks may have resembled the behavior in the subjective interest tasks. Moreover, the perplexity for the dHMM was 2.32, which was significantly better than the perplexity for logistic regression, 2.42, suggesting that the time series of the eye movement data contained relevant information for determining the task type, which supports the results by Robertson et al. (2004).

Table 10. Confusion matrix showing the number of assignments classified by the dHMM into the three task types (columns) vs. their true task type (rows). The percentages denote column- and row-wise classification accuracies.

	Prediction		
	W (70.0%)	*A* (50.0%)	I (57.5%)
W (78.9%)	142	22	16
A (35.5%)	43	54	55
I (62.8%)	18	39	96

Discriminative training of the HMM models a subset of the data (here the task type) as well as possible. As a result, other variables of the data are modeled less accurately. A way to interpret the parameter values is to compare the conditional and maximum likelihood parameter values. If these values do not differ considerably from each other, the model is close to the 'true/correct' model. In Study IV, the parameters of the discriminative and joint density models were roughly the same, suggesting that the model uses the information about the task type fairly well. In the following, the parameter values of the 9-state dHMM were compared to eye movement literature on reading and other cognitive tasks (see Poldrack, 2006).

The parameters for the three hidden states reflected relatively similar behavior across the three task types (Table 11). With a combined probability of 67%, each three tasks began from a state termed as *scanning* (see also Figures 1 and 9). The parameters suggested rather long saccades to almost random direction, and relatively short fixation durations (about 135 ms). On average, the scanning state took 2.8 s from the beginning of the task (Table 12).

The second state was termed as *reading* because the parameters suggested frequent forward saccades with the average fixation duration of 200 ms, which is typical for reading (Table 11). The saccade lengths were approximately 10 characters, corresponding to the average word length in Study IV (9.9 characters). Moreover, the percentage of regressions was 12–15%, that is, the typical amount of regressions during reading of Finnish texts (Hyönä & Niemi, 1990). There was a tendency for being in the reading state before switching to the decision state (Figure 9).

The third and final states were characterized by frequent forward and backward saccades (Table 11). The percentage of regressions was 20–30 %, that is, almost twice the amount usually observed for reading. Saccade lengths were approximately 10.7 characters, corresponding to the average word length. In 78–86 % of the cases, fixations landed on previously fixated words, and the average fixation durations were 175 ms. In sum, the third states reflected re-reading of the previously seen lines, and the participants ended the assignments almost always while they were in the third states (Figure 9). Therefore, these states were termed as *decision* states.

Table 11. Discriminative HMM parameter values for the 9-state model.

		Scanning	***Reading***	***Decision***
Beginning of the task (%)	*W*	32	16	0
	A	20	10	0
	I	15	7	0
Ending the task (%)	*W*	1	0	18
	A	0	0	16
	I	0	0	7
Fixation duration (ms)	*W*	134	199	171
	A	134	205	177
	I	134	200	176
Saccade length (pix)[a]	*W*	166	132	132
	A	160	133	137
	I	160	128	133
Saccade direction (%)				
Forward	*W*	31	61	39
	A	37	63	33
	I	41	61	37
Upward	*W*	22	6	6
	A	21	5	14
	I	21	7	15
Backward	*W*	19	15	20
	A	16	12	27
	I	13	13	30
Downward	*W*	28	18	17
	A	27	20	10
	I	26	19	11
Previous fixations (%)	*W*	23	24	78
	A	28	26	86
	I	27	24	86

[a]160 pix approximates to 13 letters. Mean word length was 9.9 characters.

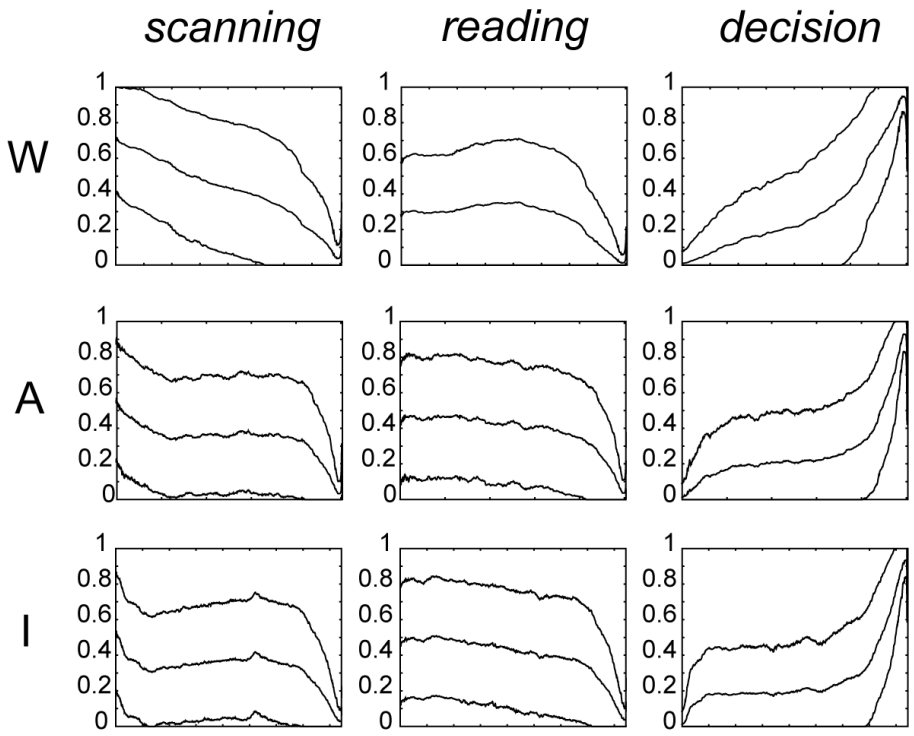


Figure 9. The probabilities (y-axis) of being in different HMM states (*scanning*, *reading*, or *decision*) as a function of time. W = word search, A = question – answer and I = subjective interest tasks. The plots show mean probabilities along with the 66 % confidence interval. For example, in the word search condition (W), the participants began the assignments form the *scanning* state with a probability of 70 %.

Participants spent more time in the scanning and reading than in the decision states (Table 12). However, the times spent in each of the states did not differ considerably across the task types, except the finding that the decision times were longer in the question-answer and subjective interest task than in the word search. In the word search task, the time in decision state corresponds to the duration of making the decision, because participants did not switch back to the scanning and the reading states, unlike in other tasks. The time to reach the decision state also increased with task complexity, suggesting longer times before reaching the decision state for the question-answer and the subjective interest tasks.

Table 12. Expected mean dwell times in scanning, reading and decision states across the task types (W=word search, A=question – answer, I=subjective interest), plus the total times and times before and after reaching the decision state and mean percentages of prevalence of the states. The values are computed from the observation trajectory which was segmented using the Viterbi algorithm on dHMM.

		Task type		
		W	***A***	***I***
Total time (s)		4.1	8.5	11.6
Time to decision (s)		3.4	6.1	8.0
Time after reaching decision (s)		0.8	2.5	3.6
Time in states (s)	*Scanning*	2.2	2.8	3.4
	Reading	4.3	6.1	6.2
	Decision	0.7	1.4	1.8
Percentage in states	*Scanning*	51	47	47
	Reading	33	38	38
	Decision	16	15	15

Comparisons of the associated transition probabilities (Figure 1) showed that participants continued in the same state for several fixations, indicating that processing states operated on time scale longer than one fixation. Transitions between states suggested that in the word search condition participants rarely switched back from the decision state, whereas these transitions occurred with 5% probability in the question-answer and with 14 % probability in the subjective interest condition (Figure 1). In the word search and question-answer tasks, participants switched from scanning to decision (80 % probability) more often than to reading (20 % probability). In the subjective interest task, transitions from decision to reading were common (86 % probability).

The probabilities for being in one state at a given time varied between one and zero (Figure 9), suggesting that the states were not mutually exclusive but rather reflect a mixture of ongoing processing which is optimal for the task (Yang & McConkie, 2005). This is in accordance with previous research suggesting that eye movements are generated

through multiple competing processes (Findlay & Walker, 1999). Eye movement trajectories showed that the decision state was adopted when the participants approached the relevant line (Figure 10). In the word search task, the trajectories indicated mostly scanning, whereas in the other tasks, the lines were read word by word, but the state of processing varied, possibly depending on whether the line was relevant for the task or not.

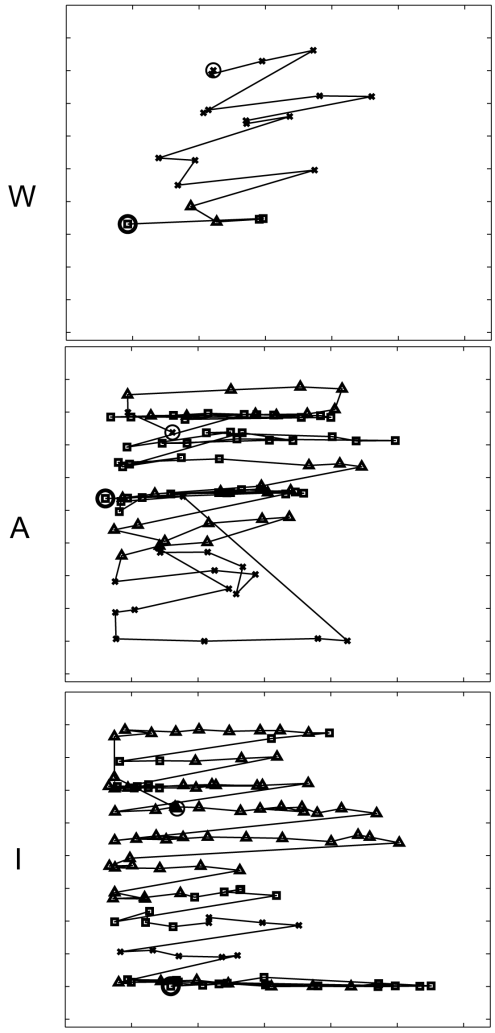


Figure 10. Eye movement trajectories in the word search (W), question-answer (A) and subjective interest (I) conditions. '×' denotes the scanning state, 'Δ' –reading state and '□' – decision state. The beginning of a trajectory is marked with a circle and ending with two concentric circles.

5 Discussion

5.1 Parafoveal processing of words across visual fields

Study I showed evidence for a RVF advantage in parafoveal processing, suggesting that more information is extracted parafoveally to the right of fixation than to the left. The P2 EFRP responses measured bilaterally at occipital sites 200–280 ms post-stimulus were stronger for orthographically legal words than for non-words when the parafoveal stimuli were presented in the RVF. No such effect was measured for the LVF stimuli. In addition, the P2 responses measured at left central site were stronger for RVF than LVF targets in the associated and non-associated word conditions, and the first fixation durations on the foveal words were longer when the parafoveal RVF stimulus was a non-associated word than a non-word.

Study I allowed investigations of the origins of the RVF advantage. The *structural* account assumes that the RVF advantage in word recognition occurs because the RVF projects directly to the left hemisphere, which is specialized for language processing for most individuals. The occipital N1 responses (at 140–200 ms post-stimulus) were enhanced over the left hemisphere, which most likely corresponds to the processing taking place at the visual word form area, a system located in the left inferior temporal region, which is specifically devoted to the processing of letter strings (Cohen, et al., 2000; Cohen, et al., 2002). The structural account suggests that information presented in LVF/RH must be transferred to the LH, and that all processing that follows the visual word form area, are identical irrespective of the originally stimulated visual field. Contrary to this, Study I indicated that non-word responses departed from responses to orthographically legal words at occipital sites between 200 to 280 ms post-stimulus. This exceeds the typical time course of word processing suggested by the structural account.

The behavioral and most eye movement results in Study I did not differ between visual fields, supporting the *attention account*, which assumes that information is extracted equally from both visual fields. The attention account was formulated on the basis of previous eye movement studies which have shown that parafoveal information is primarily

extracted from the location which is about to be fixated next (Henderson, et al., 1989; Inhoff, et al., 1989). The attention account can, however, explain the RVF advantage, because even though attention and targeting of saccades are tightly linked together (Deubel & Schneider, 1996), the shift of attention towards the saccade goal occurs only around 50 ms before saccade onset (Doré-Mazars, Pouget, & Beauvillain, 2004). Thus, attention may have been focused more toward the right during saccade planning, enhancing parafoveal processing of the RVF. Moreover, the words presented in the RVF are likely to benefit from the correct activation pattern, that is, the decreasing left-to-right activation pattern as result of visual acuity and attention decrements toward the visual periphery (Whitney, 2001).

The results of Study I are also compatible with the *perceptual learning* account, suggesting that the visual training associated with long-term reading at a particular direction can generate visual field differences (Nazir, et al., 2004). The results show that parafoveal information relevant for reading is primarily extracted from RVF, that is, from the normal reading direction of the participants. Thus, the results suggest that the RVF advantage in parafoveal processing can also be explained by an attentional mechanism that biases attention toward reading direction. A similar explanation has also been proposed before (Ducrot & Grainger, 2007; Eviatar, 1995).

Previous research suggests that the *structural*, *attentional* and *perceptual learning accounts* may not be mutually exclusive. For example, the pure effect of perceptual learning has been somewhat difficult to demonstrate, since readers of right-to-left scripts do not show a constant LVF advantage (Nazir, et al., 2004), as the perceptual learning account would predict. Also, studies comparing readers with typical LH language lateralization and atypical RH language lateralization provide further evidence for the structural account (Brysbaert, 1994; Hunter, Brysbaert, & Knecht, 2007; Nazir & Huckauf, 2008). These studies have demonstrated the RVF advantage for readers with LH language lateralization, while the readers with RH language lateralization show an opposite pattern.

The difference observed in parafoveal processing of orthographically legal and illegal words is consistent with reading studies showing parafoveal-on-foveal effects at the level of orthographic familiarity (Starr & Inhoff, 2004; White, 2008). The results are in line with

parallel reading models assuming that attention can be allocated to several words at a time (Engbert, et al., 2005). In contrast, models assuming serial attention shifts in reading (Reichle, et al., 2006; Reichle, Warren, et al., 2009) do not support parafoveal-on-foveal effects. However, serial attention models suggest that low-level information within the perceptual span can be processed in parallel with the foveal word, and that orthographic processing can be completed on several words in parallel. Such pre-lexical processing state can explain the differences in P2 responses between the words and non-words in the RVF.

Contrary to the results by Baccino and Manuta (2005), Study I did not support semantic effects. Using an otherwise similar paradigm except that the targets were presented only in the RVF, Baccino and Manuta showed that parafoveal word form information was processed within 119 ms from the stimulus onset. A subsequent P2 EFRP-component between 200–230 ms differentiated semantically associated and non-associated words from each other at electrode sites extending from frontal to occipital areas. The different effects are possibly due to differences in experimental designs. In Study I, the direction of the next saccade was randomly determined at the onset of each word pair stimulus. Previous studies have shown that precueing of saccade target position led to saccade latencies that were about 40 ms shorter compared to when the saccade target was unknown (reviewed in Findlay & Walker, 1999). Therefore, saccade programming most likely took longer in Study I compared to the study by Baccino and Manuta (2005). This delay may have prohibited parafoveal extraction of the semantic information. However, this is still a minor concern, because the main comparisons in Study I were between the word conditions within each visual field, and no latency differences were observed at occipital sites between visual fields or the word conditions.

5.2 Text format

In Study II, the vertical texts were read at slower rate and less efficiently than the standard horizontal text. There were no differences in the number of fixations and number of regressions between the horizontal and vertical texts, but the fixation durations were longer for the vertical formats than for the standard horizontal text format. The vertical texts comprised of greater number of pages than the horizontal text. The difference in the

number of pages may explain, at least, partly the differences in reading rates, because frequent page shifts take time and the readers need to move their eyes from the bottom of the page to the top of the next page.

The benefits of the vertical presentation may also have been obscured by the participants' experience in horizontal reading. Previous studies have shown that practice in reading increases the size of *the letter identity span*, which in turn correlates with increased reading speed (Häikiö, et al., 2009). Furthermore, the findings by Osaka and Oda (1991) showed an equal performance for Japanese readers in horizontal and vertical reading. These findings together suggest that practice in vertical reading might increase the reading speed and make reading in the vertical direction more efficient.

One possibility is that vertical reading was slow because the readers' *word identification span* (Ojanpää, et al., 2002) and visual acuity (Curcio & Allen, 1990) are reduced in the vertical direction. The visual-span hypothesis proposes a causal link between the size of the *visual span* (the area from which information can be acquired without any help of linguistic information) and the reading speed. This theory is described in the Mr. Chips model (Legge, Hooven, Klitz, Mansfield, & Tjan, 2002), suggesting that the mean saccade lengths decrease along with decreasing visual span size.

Contrary to the assumption that the visual span is limited by visual acuity (Curcio & Allen, 1990), Pelli et al. (2007) proposed that the limits of the visual span are determined by *crowding*. That is, the border between uncrowded (central) and crowded (peripheral) visual field is not fixed on the retina, but it depends on spacing between the objects being viewed and on their distance from fixation (eccentricity). According to Levi and Carney (2009) crowding is the main bottleneck that affects reading and object recognition in peripheral vision. Objects that can be easily identified in isolation seem jumbled when they appear in clutter due to inappropriate integration of features into an object in which they do not belong.

In addition, Feng, Jiang, and He (2007) found that crowding was significantly stronger for horizontally than for vertically arranged configurations, and confirmed that the asymmetry in crowding was not a property of low-level sensory processing. They suggested that horizontal reading might explain why the attentional mechanism more likely

integrates the horizontally arranged items into single units, but the effect is less likely to occur for vertical configurations. Thus, the potential value of vertical text may be related to the diminished crowding in vertical direction (Feng, et al., 2007; Yu, et al., 2010).

A possible explanation for the slow performance in vertical reading might be the readers' inability to extract useful parafoveal information in the vertical direction. Furthermore, since reading in the vertical direction is not as common as horizontal reading, the foveal processing load may have been greater during vertical than during horizontal reading. This was possibly reflected by the longer fixation durations observed for vertical reading. Greater foveal processing load may also be linked with diminished parafoveal processing, because previous studies have shown that increased foveal processing load is associated with decreased parafoveal information extraction (Henderson & Ferreira, 1990).

5.3 Attention to parafoveal pictorial stimuli

Study III investigated how online readers allocate attention toward peripheral advertisements when the combination of animated and static ads varied (Experiment 1), when the ads appeared abruptly on the screen (Experiment 2), and when the nature of the primary task was manipulated (Experiment 3). The results of all experiments showed that the effect of ads on reading was mostly accompanied by direct fixations to ads. This finding is in line with the view that attentional capture by ads is primarily related to mechanisms of overt attention, and runs counter to studies, suggesting that online ad processing occurs peripherally via the covert attention mechanism (Burke, et al., 2005; Day, et al., 2006; Drèze & Hussherr, 2003).

Moreover, the findings that the ad conditions had only a slight effect on reading eye movements and that participants' self-reports about experienced attention were consistent with the eye movement results further support the hypothesis that distraction by ads occurs through overt fixations toward the ads rather than as covert processing of ads. However, it is possible that some intricate events may be time locked to the appearance of the ads, and that such subtle events are buried in the variance of the global analyses reported in Study III. For example, the results of Experiment 2 suggested that reading comprehension decreased when the ad was presented to the right compared to when the ad was presented

above the text, and that the onset of the ad on the right impaired reading comprehension more than the onset of the ad above. These observations could also be explained by covert processing of the ad information to the right of the text, which did not elicit an overt eye movement.

Experiment 1 showed that a combination of one static and one animated ad increased attentional capture. This was an unexpected finding, since it was originally expected that two animated ads would produce the most attentional capture. Thus, it is likely that when two animated ads are presented simultaneously, they are equally salient (both contain motion) and compete equally for readers' attentional resources. This increased saliency is likely to be accompanied by increased attentional resources being invested in ignoring the task-irrelevant stimuli. As a consequence, two animated ads are ignored by top-down attention as effectively as two static ads, but when only one of the ads is animated, it is individually more salient and more likely captures attention.

The results of Experiment 2 showed an association between the ad onset time and the time when the ad to the right was fixated for the first time during a reading task. No such effect was observed for the ad above, but the abrupt onset increased the number of entries and fixations toward the ad above. Furthermore, both ads were fixated more often during the first five fixations after the ad onsets. These results suggest that abrupt onset captures attention immediately, especially when it occurs in the proximity of the text. Although, the abrupt onset in the visual periphery captured attention less immediately, attention was drawn to the ad above more often when it appeared abruptly. This was indicated by increased number of eye entries and fixations toward the ad above when it appeared abruptly.

The reading results of Experiment 2 further suggested that comprehension was better when the ad above appeared abruptly compared to the condition when the ad was present throughout the trial. The abrupt onset in the periphery possibly increased attention toward the text, resulting in improved reading comprehension. Prior studies have reported similar findings, suggesting that peripheral ads might increase the participants' level of arousal and result in increased processing resources invested in the primary task (Burke, et al., 2005; Day, et al., 2006).

In Experiment 3, participants paid more attention to ads when the task was to browse the pages according to their own interest compared to the reading for comprehension task. Moreover, the participants adopted a more superficial reading strategy in the free browsing than in the reading task. In the free browsing task, the abrupt onset time of both ads was associated with the time when the participants first fixated the ads. In the reading task (similarly to Experiment 2), such an effect was observed only for the ad to the right of the text. This result suggests that when participants are engaged in a reading task, they are more capable of overriding the attentional capture by peripheral abrupt onsets compared to when the abrupt onset occurs in the proximity on the main task area. Interestingly, the results of Experiment 3 indicate that the mere presence of the ads did not interfere with reading, because the baseline condition where no ads were presented was not beneficial in any of the reading measures. This finding further supports our conclusion that the interference by ads occurs primarily through overt fixations rather than through covert attention to ads.

The finding that the ad to the right of the text was attended more often than the ad above suggests that animation or abrupt onset attract more attention when presented in the proximity of the text compared to when the ad is presented in the periphery. Moreover, Experiment 2 showed that reading comprehension was impaired more when the ad was presented to the right than above the text. Most likely, the ad on the right was attended more than the ad above, because when reading the text from left to right, the readers approach the ad on the right each time they reach the end of the line. Thus, the ad on the right enters their perceptual span at least occasionally (Rayner, 1998). On the contrary, the further the readers advance in the text, the longer their gaze is from the ad above the text.

5.4 Reading processes in information search tasks

Study IV adopted a reverse inference approach (see Poldrack, 2006) by comparing the model parameters to eye movement behavior reported in existing studies where the cognitive processing state of the observer is supposedly known. This approach was used to make hypotheses about hidden cognitive states in everyday information search tasks, where participants were asked to perform a simple word search, find an answer to a question, or to

find a subjectively most interesting topic. The dHMM model suggested that participants shifted their reading processes reflected in their eye movement patterns as they proceeded in the tasks.

A *scanning* state was typical in the beginning of the tasks. The dHMM parameters for the scanning states indicated long saccades with no preference for direction, accompanied with short fixations. The second states were labeled as *reading* because of frequent forward saccades, with the distance corresponding to an average word length. Also, the mean fixation duration and number of regressions were in accordance with the previous studies of reading (Hyönä & Niemi, 1990). The third states were termed as *decision* because the parameters suggested rather careful processing of the sentences. Saccades landed almost always on previously fixated words, and the saccade lengths corresponded to an average word length. Furthermore, the amount of regressions was twice the amount usually observed for reading.

Previously, Liechty et al. (2003) have adopted a similar approach by modeling eye movement data in order to identify different processing states in an advertisement viewing task. The processing states discovered in Study IV shared several similarities with their findings. The scanning state and their global processing state were both characterized by long saccades and relatively short fixations. Short saccades and long fixations were typical for the attentive processing state in Liechty et al. (2003) and the reading and decision states in Study IV. However, Study IV suggests a finer structure by segmenting the attentive processing into two (i.e., reading and decision) states.

The HMMs are designed for reverse inference tasks, and thus they differ from the models of reading (e.g., Engbert, et al., 2005; Reichle, et al., 2006; Reilly & Radach, 2006) that are models of forward inference. That is, they describe how perceptual and cognitive processes drive eye movements, whereas our model tries to make inferences about cognitive processes given the eye movements. A potential concern regarding the comparisons of dHMM parameters to previous studies on reading (e.g., Rayner, 1998, 2009) is that participants may have altered their processing states also in the previous tasks. However, as pointed out by Hyönä et al. (2002), many reading studies have treated differences in reading strategies as a nuisance, and attempted to investigate reading under

simplified conditions, for example, by presenting single sentences. Therefore, it is likely that previous studies reflect rather 'pure' types of reading processes.

Both oculomotor and cognitive models can explain the results of Study IV. A strategy-tactics model (O'Regan & Jacobs, 1992) suggests that readers can adopt either careful or risky global strategies that influence their eye movements. The assignment presented before the word lists may have biased participants to use a certain eye movement strategy on the forthcoming task. Thus, it is possible that a global eye movement strategy was adopted for each task beforehand. On the other hand, the task types differed in transition sequences between the processing states, possibly suggesting that the processing state was also adjusted based on the current tasks (Radach et al., 2008; Carver, 1990).

Despite the controversial theoretical views, the results of Study IV have relevance in practical applications. The finding that the eye movement patterns differ within a task may be applicable in a proactive information retrieval application. Such an application can search for more documents on a specific topic after detecting eye movement behavior indicating observers' interest on that topic (see Puolamäki, Salojärvi, Savia, Simola, & Kaski, 2005).

5.5 Future directions

This thesis opens several ideas for future research. For example, in Study I the possible confound resulting from the experimental design in which the direction of the next saccade was randomly determined at the onset of each word pair stimulus could be tested by using a blocked design (or by cuing the next saccade direction). Such a design may help to compare the results with Baccino and Manuta (2005) who showed that parafoveal word form information was processed much earlier (at 119 ms post-stimulus) than in the present study (around 200–280 ms post-stimulus). An improved design may also help to test the predictions based on reading models (e.g., Engbert, et al., 2005; Reichle, et al., 2006).

The EFRP time window of 300 ms used in Study I may have been too short to reveal the semantic parafoveal-on-foveal effects, because the saccade programming most likely took more time than in the study by Baccino and Manuta. They showed that a P2 EFRP-component between 200–230 ms differentiated semantically associated and non-associated

words from each other. Improved artifact correction methods (Baccino, 2011; Dimigen, et al., in press) would allow investigating the EFRPs in longer time windows than 300 ms. Moreover, the visual field differences in parafoveal information extraction should be tested under more naturalistic reading conditions, for example, when participants are reading whole sentences or text chapters. As pointed out by Rayner et al. (1996) reading is a complex task, and attempting to generalize from eye movement behavior when single words are read to eye movement control in reading may be somewhat hazardous.

Study II suggested that vertical reading could be improved by practice. Thus, it would be interesting to train the participants in vertical reading, since previous research (e.g., Study I) suggests that practice improves information processing at reading direction. Moreover, besides the word search study by Ojanpää et al. (2002), there is little research on parafoveal processing of words presented in the vertical direction. Thus, investigating vertical parafoveal preview and parafoveal-on-foveal effects would be of interest in future studies. In such studies, the spacing between the lines is clearly a relevant factor. If the line spacing is small, it is possible that readers can extract more information from the next lines, but on the other hand, the effect of vertical crowding may increase. However, if the lines are too far away from each other, it becomes difficult for the readers to extract information from the next lines.

The results of Study III suggested that when two animated ads were presented simultaneously, they competed equally for the readers' attentional resources and were ignored by top-down attention as effectively as two static ads. In future studies, it would be of interest to test how increasing the amount of peripheral ads affects attention allocation across them, and whether the effectiveness of ads is cancelled out when many ads compete for the viewers' attentional resources simultaneously. In addition to the two ad formats (i.e., vertical and horizontal banners) investigated in the present study, the Internet Advertising Bureau (IAB) has standardized around 16 different online ad formats that vary in size. It would be of interest to know how users process these ad formats. Moreover, new advertising technologies have been developed. For instance, formats referred to as "over-the-page" units have features such as peeling back, floating over the page, and expanding or appearing between pages. Eye tracking could be used to examine how viewers process such

ad formats that cover the main task area on the screen. In comparison to Study III, which investigated the influence of peripheral ads, it would be of interest to study, for example, how fast readers are able to return to the topic after an appearance of such a 'focal' distractor.

Study IV showed that a hidden Markov model is able to reveal intra-individual and with-in task variations in information search tasks. This approach could be applied more widely to study top-down strategic modulations on reading. For example, applying the model to the reading tasks presented in this thesis, that is, to study vertical reading (Study II) and reading of texts that are surrounded by pictorial distractors (Study III) may reveal interesting features about the ongoing processes. It would be of interest to know whether the with-in task modulation is greater in the presence of peripheral ads, which may give indications of the possible covert processing of the ads. Moreover, it would be interesting to compare with-in task variations between the reading for comprehension and free browsing tasks. Further, it would be interesting to investigate whether readers adopt different reading strategies when reading vertical texts as compared to the normal horizontal reading. Also, the current HMM implementation did not take word features into account, but instead modeled global features (e.g., saccade length, and mean fixation duration) of the eye movement trajectory. Possibly, the model performance could have been enhanced if the word features were taken into account.

Moreover, as pointed out by Baccino (2011) one of the major difficulty in interpreting the eye movements is to determine whether a fixation represents deep (e.g., semantic) or more superficial processing. Thus, other indices of the processing may help to make more fine-grained interpretations of the underlying cognitive processing. EFRPs can provide a complementary measure to capture the cognitive processes under natural (ecologically valid) viewing tasks by showing when and in which order different cognitive processes occur (e.g., Dimigen et al., in press). Thus, the reading tasks presented in Studies II-IV may also benefit of co-registration of eye movements and EEG. For example, EFRP analyses might help to detect whether some attentional / semantic component can be associated to a fixation (see Baccino, 2011). Such information possibly helps in making inferences on the strategic, top-down, modulation of reading processes.

Besides the relative lack of previous experimental studies on online reading, there is an ongoing public debate on the possible long-term effects of online reading on cognition. This debate claims that applied environments give rise to new types of more superficial reading processes, and that the Internet might have detrimental effects on cognition that diminish the capacity for concentration and contemplation (Carr, 2010). These claims have not been experimentally tested, suggesting that online reading is not well understood currently.

6 Conclusions

The Internet has become a ubiquitous information search and communication channel. Due to an extensive user population and because Web pages are complex visual stimuli that incorporate a combination of textual, pictorial and multimedia content, it is important to understand how the Web users process online texts. Compared to a fairly good understanding of eye movement guidance in reading when people read simple sentences or text paragraphs, relatively little is known about eye movements in online reading situations. This thesis comprises of eye tracking studies investigating information processing during various reading tasks. **Study I** investigates how parafoveal information is extracted either to the left or to the right of the current fixation. **Studies II-IV** examine information processing in various online reading tasks in which the extraction of parafoveal information plays a crucial role.

This thesis demonstrated that text format affects reading performance and eye movements. The results showed that long-term learning of left-to-right reading affects the way in which parafoveal information is extracted from words presented either to the left or to the right of the foveal word. The results indicated that more parafoveal information was extracted from the normal reading direction of the participants, that is, to the right of the fixated word. The results also suggest that attention was directed more to the right than to the left, suggesting that the observed right visual field advantage was explained by a mechanism that biases attentional orienting a function of reading direction. Moreover, vertical text was read at slower rate than a standard horizontal text, and the mean fixation duration was longer for vertical formats than for horizontal text. It was assumed that these

differences result from the readers experience in horizontal reading, and that the differences may diminish with training in vertical reading.

Further, the results showed that animated online advertisements and abrupt ad onsets capture online readers' overt attention, and distract the reading process. Attentional capture by abrupt ad onsets was observed especially when the ads were presented to the right of the text being read. Taken together, the results suggest that when parafoveal stimuli (either words or pictorials) are presented to the right of the fixated region, they are attended to more than stimuli presented elsewhere on the screen. The most likely explanation for this finding is that at least occasionally the stimuli on the right enter the observers' region of effective vision, that is, the perceptual span.

Moreover, it was shown that the primary task of the observers affects the way in which attention is allocated across the screen. When the observers were asked to browse the Web pages according to their own interest, the ads were looked at more frequently and for longer periods compared to when they were reading the online texts for comprehension. The results further suggest that the processing states vary within the tasks when Web users are performing online information search tasks. For example, when they are searching for a specific keyword, looking for an answer to a question, or trying to find a subjectively most interesting topic. A *scanning* type of behavior was observed in the beginning of the tasks, after which participants tended to switch to a more careful *reading* state before finishing the tasks in the states termed as *decision* states.

7 References

Aula, A., & Räihä, P. M. a. K.-J. (2005). Eye-Tracking Reveals the Personal Styles for Search Result Evaluation. In M. F. C. a. F. Paternò (Ed.), *Proceedings of Human-Computer Interaction - INTERACT 2005* (pp. 1058 –1061).

Baccino, T. (2011). Eye movements and concurrent event-related potentials: Eye fixation-related potential investigations in reading. *The Oxford Handbook of Eye Movement.* S. Liversedge, I. D. Gilchrist and S. Everling. Oxford, Oxford University Press**:** 857–870.

Baccino, T., & Manunta, Y. (2005). Eye-fixation-related potentials: Insight into parafoveal processing. *Journal of Psychophysiology, 19*, 204–215.

Berg, P., & Scherg, M. (1991). Dipole modelling of eye activity and its application to the removal of eye artefacts from the EEG and MEG. *Clinical Physics and Physiological Measurement, 12*, 49–54.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience, 2*, 370–374.

Brysbaert, M. (1994). Interhemispheric transfer and the processing of foveally presented stimuli. *Behavioural Brain Research, 64*, 151–161.

Burke, M., Hornof, A., Nilsen, E., & Gorman, N. (2005). High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *Transactions on Computer-Human Interaction, 12*, 423–445.

Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Byrne, M. D. (2002). *Reading vertical text: Rotated vs. marquee* Paper presented at the Human Factors and Ergonomics Society, Santa Monica, CA.

Carr, N. (2010). *The Shallows: What the Internet Is Doing to Our Brains*. New York: W. W. Norton & Company.

Carver, R. P. (1990). *Reading Rate: A Review of Research and Theory*. San Diego, California: Academic Press, Inc.

Chiarello, C., Liu, S., & Shears, C. (2001). Does global context modulate cerebral asymmetries? A review and new evidence on word imageability effects. *Brain and Language, 79*, 360–378.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*, 291–307.

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain, 125*, 1054–1069.

Coleman, E. B., & Hahn, S. C. (1966). Failure to improve readability with a vertical typography. *Journal of Applied Psychology, 50*, 434–436.

Coleman, E. B., & Kim, I. (1961). Comparison of several styles of typography in English. *Journal of Applied Psychology, 45*, 262–267.

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *The Journal of Comparative Neurology, 300*, 5–25.

Day, R.-F., Shyi, G. C.-W., & Wang, J.-C. (2006). The effect of flash banners on multiattribute decision making: Distractor or source of arousal? *Psychology & Marketing, 23*, 369–382.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences, 9*, 335–341.

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

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*, 1827–1837.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (in press). Co-registration of eye movements and EEG in natural reading: Analyses & review. *Journal of Experimental Psychology: General*.

Doré-Mazars, K., Pouget, P., & Beauvillain, C. (2004). Attentional selection during preparation of eye movements. *Psychological Researchs, 69*, 67–76.

Drèze, X., & Hussherr, F.-X. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing, 17*, 8–23.

Ducrot, S., & Grainger, J. (2007). Deployment of spatial attention to words in central and peripheral vision. *Perception & Psychophysics, 69*, 578–590.

Ellis, A. W. (2004). Length, formats, neighbours, hemispheres, and the processing of words presented laterally or at fixation. *Brain and Language, 88*, 355–366.

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

Eviatar, Z. (1995). Reading direction and attention: Effects on lateralized ignoring. *Brain and Cognition, 29*, 137–150.

Feng, C., Jiang, Y., & He, S. (2007). Horizontal and vertical asymmetry in visual spatial crowding effects. *Journal of Vision, 7*, 1–10.

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

Forney, G. D. J. (1973). The viterbi algorithm. *Proceedings of the IEEE, 61*, 268–278.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics, 65*, 999–1010.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron, 31*, 681–697.

Granka, L., Joachims, T., & Gay, G. (2004). *Eye-tracking Analysis of User Behavior in WWW Search.* Paper presented at the Conference on Research and Development in Information Retrieval (SIGIR), Sheffield, UK.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102*, 167–181.

Hautala, J., Hyönä, J., & Aro, M. (2011). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research, 51*, 1719–1727.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*, 334–337.

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

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics, 45*, 196–208.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: a comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Hong, W., Thong, J. Y. L., & Tam, K. Y. (2004). Does animation attract online users' attention? The effects of flash on information search performance and perceptions. *Information Systems Research, 15*, 60–86.

Hunter, Z. R., Brysbaert, M., & Knecht, S. (2007). Foveal word reading requires interhemispheric communication. *Journal of Cognitive Neuroscience, 19*, 1373–1387.

Hutzler, F., Braun, M., Võ, M. L., Engl, V., Hofmann, M., Dambacher, M., et al. (2007). Welcome to the real world: Validating fixation-related brain potentials for ecologically valid settings *Brain Research, 1172*, 124–129.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition, 32*, 523–532.

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

Hyönä, J., & Niemi, P. (1990). Eye movements during repeated reading of a text. *Acta Psychologica, 73*, 259–280.

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*, 31–50.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology, 41 A*, 63–89.
Irwin, D. E. (1998). Lexical Processing during Saccadic Eye Movements. *Cognitive Psychology, 36*, 1–27.
Ishida, T., & Ikeda, M. (1989). Temporal properties of information extraction in reading studied by a text-mask replacement technique. *Journal of the Optical Society of America, 6*, 1624–1632.
Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.
Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General, 108*, 151–181.
Jasper, H. (1958). The 10-20 electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371–375.
Just, M. A., & Carpenter, P. A. (1980). A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review, 87*, 329–354.
Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 490–508.
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.
Legge, G. E., Hooven, T. A., Klitz, T. S., Mansfield, S. J., & Tjan, B. S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Research, 42*, 2219–2234.
Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology, 19*, 1988–1993.
Liechty, J., Pieters, R., & Wedel, M. (2003). Global and local covert visual attention: Evidence form a bayesian hidden markov model. *Psychometrika, 68*, 519–541.
Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*, 6–14.
Luck , S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578–586.
McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. S. (2005). An anatomically constrained, stochastic model of eye movement control in reading. *Psychological Review, 112*, 814–840.
Morrone, C., & Burr, D. (2009). Visual stability during saccadic eye movements. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences, 4th edition* (pp. 511–524). Massachusetts, USA: MIT Press.
Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition or printed words. *Brain and Language, 88*, 294–311.
Nazir, T. A., & Huckauf, A. (2008). The visual skill reading. In Grigorenko & Naples (Eds.), *Single word reading : Cognitive, Behavioral and Biological Perspectives* (pp. 25–42): Lawrence Erlbaum.
O'Regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 185–197.
O'Regan, J. K., Lévy-Schoen, A., & Jacobs, A. M. (1983). The effect of visibility on eye-movement parameters in reading. *Perception & Psychophysics, 34*, 457–464.
Ojanpää, H., Näsänen, R., & Kojo, I. (2002). Eye movements in the visual search of word lists. *Vision Research, 42*, 1499–1512.
Osaka, N. (1992). Size of saccade and fixation duration of eye movements during reading: Psychophysics of Japanese text processing. *Journal of the Optical Society of America, 9*, 5–13.
Osaka, N., & Oda, K. (1991). Effective visual field size necessary for vertical reading during Japanese text processing. *Bulletin of the Psychonomic Society, 29*, 345–347.
Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision, 7*, 1–36.
Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson JR, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria *Psychophysiology, 37*, 127–152.

Pieters, R., Rosbergen, E., & Wedel, M. (1999). Visual attention to repeated print advertising: A test of scanpath theory. *Journal of Marketing Research, 36*, 424–438.
Poldrack, R. A. (2006). Can cognitive processes be inferred form neuroimaging data? *Trends in Cognitive Sciences, 10*, 59–63.
Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual space for Hebrew readers. *Brain and Language, 14*, 174–180.
Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology: General, 109*, 160–174.
Puolamäki, K., Salojärvi, J., Savia, E., Simola, J., & Kaski, S. (2005). *Combining eye movements and collaborative filtering for proactive information retrieval.* Paper presented at the SIGIR '05: the 28th annual international ACM SIGIR conference on Research and development in information retrieval.
Rabiner, L. R. (1989). *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition.* Paper presented at the IEEE.
Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement contrtol 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. *European Journal of Cognitive Psychology, 16*, 3–26.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506.
Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 720–732.
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, 5*, 1188–1200.
Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences, 13*, 115–119.
Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research 7*, 4–22.
Reichle, E. D., Warren, T., & McConnel, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*, 1–21.
Reilly, R. G., & O'Regan, K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research, 38*, 303–317.
Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7*, 34–55.
Robertson, A., Kirshner, S., & Smyth, P. (2004). Downscaling of daily rainfall occurence over Northeast Brazil using a hidden Markov model. *Journal of Climate, 17*, 4407–4424.
Rugg, M. D., & Coles, M. G. (1995). *Electrophysiology of mind*. London: Oxford University Press.
Salojärvi, J., Puolamäki, K., & Kaski, S. (2005a). *Expectation maximization algorithms for conditional likelihoods.* Paper presented at the 22nd international conference on machine learning (ICML-2005), New York, USA.
Salojärvi, J., Puolamäki, K., & Kaski, S. (2005b). On discriminative joint density modeling. In J. Gama, R. Camacho, P. Brazdil, A. Jorge & L. Torgo (Eds.), *Machine Learning: ECML 2005, Lecture Notes in Artificial Intelligence 3720* (pp. 341–352). Berlin, Germany: Springer-Verlag.
Sereno, S. C., & Rayner, K. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. *NeuroReport, 9*, 2195–2200.
Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*, 489–493.
Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature neuroscience, 3*, 264–269.
Simola, J., Stenbacka, L., & Vanni, S. (2009). Topography of attention in the primary visual cortex. *European Journal of Neuroscience, 29*, 188–196.

Smeets, J. B. J., & Hooge, I. T. C. (2003). Nature of variability in saccades. *Journal of Neurophysiology, 90*, 12–20.

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

Stenfors, I., Morén, J., & Balkenius, C. (2003). Behavioural strategies in Web interaction: A view from eye-movement research. In R. R. J. Hyönä, H. Deubel (Ed.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 633–644). Amsterdam, The Netherlands: Elsevier.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 799–806.

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1342–1353.

Unema, P., Pannasch, S., Joos, M. Velichkovsky, B.M. (2005). Time-course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*, 473–494.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience, 13*, 454–461.

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception *Transportation Research Part F: Traffic Psychology and Behaviour, 5*, 145–156.

Vergilino-Perez, D., Collins, T., & Doré-Mazars, K. (2004). Decision and metrics of refixations in reading isolated words. *Vision Research, 44*, 2009–2017.

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

Vitu, F., 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 & Psychophysics, 57*, 352–364.

Viviani, P., & Swensson, R. G. (1982). Saccadic eye movements to peripherally discriminated visual targets. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 113–126.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception & Performance, 34*, 205–223.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8*, 221–243.

Yang, S.-N., & McConkie, G. W. (2005). New directions in theories of eyemovement control during reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 105–130). Great Britain:: Oxford University Press.

Yu, D., Park, H., Gerold, D., & Legge, G. E. (2010). Comparing reading speed for horizontal and vertical English text *Journal of Vision, 10*.