EYE MOVEMENT CONTROL IN TURKISH READING:
A CORPUS-ANALYTIC APPROACH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

FİGEN BEKEN FİKRİ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COGNITIVE SCIENCE

SEPTEMBER 2015

EYE MOVEMENT CONTROL IN TURKISH READING:
A CORPUS-ANALYTIC APPROACH

Submitted by **Figen BEKEN FİKRİ** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Science Department**, **Middle East Technical University by**,

Prof. Dr. Nazife Baykal
Director, Informatics Institute ________________

Prof. Dr. Cem Bozşahin
Head of Department, Cognitive Science ________________

Assist. Prof. Dr. Cengiz Acartürk
Supervisor, Cognitive Science, METU ________________

Examining Committee Members

Prof. Dr. Cem Bozşahin
Cognitive Science, METU ________________

Assist. Prof. Dr. Cengiz Acartürk
Cognitive Science, METU ________________

Prof. Dr. Deniz Zeyrek Bozşahin
Cognitive Science, METU ________________

Assoc. Prof. Dr. Bilal Kırkıcı
Foreign Language Education, METU ________________

Assoc. Prof. Dr. Selda Özdemir
Special Education, Gazi University ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Figen BEKEN FİKRİ

Signature:

ABSTRACT

EYE MOVEMENT CONTROL IN TURKISH:
A CORPUS-ANALYTIC APPROACH

Beken Fikri, Figen
M.S. Cognitive Science
Supervisor: Assist. Prof. Dr. Cengiz Acartürk

September 2015, 71 pages

In this thesis, the eye movement parameters during Turkish reading were examined in terms of the characteristics of the words. These eye movement parameters were first fixation landing position, first fixation duration, gaze duration, refixation probability and skipping probability. The text characteristics were specified by word length, launch site, word frequency, and word predictability. A corpus-analytic approach was employed in order to achieve a general picture of Turkish reading patterns. The analyses were conducted by using linear mixed model so that random effect variances - subject, sentence and word type - were controlled. The results of first fixation landing position showed that first fixations land around word centers except for sufficiently long words. Mean initial landing position values with respect to word length showed patterns consistent with those reported in previous, cross-linguistic studies. Furthermore, the preferred viewing location (PVL) effect was observed. The findings for first fixation duration were in good agreement with the inverted optimal viewing position (IOVP) effect. Finally, the general results of gaze duration, refixation probability and skipping probability were consistent with earlier findings.

Key Words: Turkish Reading, Eye Movements, Linear Mixed Model, Preferred Viewing Location, Inverted Optimal Viewing Position

ÖZ

TÜRKÇE OKUMADA GÖZ HAREKETLERİ KONTROLÜ:
DERLEM-ÇÖZÜMLEMELİ BİR YAKLAŞIM

Beken Fikri, Figen
Yüksek Lisans, Bilişsel Bilimler
Tez Yöneticisi: Yrd. Doç. Dr. Cengiz Acartürk

Eylül 2015, 71 sayfa

Bu tezde, Türkçe okumada göz hareketleri parametreleri sözcük özellikleri açısından incelenmiştir. Bu parametreler, gözün sözcük üzerindeki ilk sabitleme lokasyonu, ilk sabitleme süresi, toplam sabitleme süresi, tekrar sabitleme olasılığı ve sözcüğün atlanması olasılığıdır. Metin özellikleri ise sözcük uzunluğu, hedef sözcüğü önceleyen giriş mesafesi (İng. launch site), sözcüğün derlem sıklık değeri ve sözcüğün tahmin edilebilirliğidir. Türkçe okuma örüntülerinin genel bir resmini elde etmek için derlem-çözümlemeli bir yaklaşım uygulanmıştır. Analizler lineer karma model kullanılarak yapılmıştır. Böylece rastgele etki varyansları - katılımcı, tümce ve sözcük türü - denetim altına alınmıştır. İlk sabitleme lokasyonu analizi, uzun olanlar dışındaki sözcüklerde, ilk sabitlemenin sözcüğün merkezi etrafında olduğunu göstermiştir. Ortalama ilk sabitleme lokasyonu değerleri, önceki çalışmalarla uyumlu bulunmuştur. Ayrıca, tercih edilen görü lokasyonu (İng. preferred viewing location) etkisi gözlemlenmiştir. İlk sabitleme süresi ile ilgili bulgular, ters optimal görü lokasyonu (İng. inverted optimal viewing position) etkisi ile tutarlıdır. Son olarak, sözcük üzerindeki toplam sabitleme süresi, tekrar sabitleme olasılığı ve sözcüğün atlanması olasılığının genel sonuçları alanyazındaki çalışmalarla uyumlu bulunmuştur.

Anahtar Kelimeler: Türkçe Okuma, Göz Hareketleri, Lineer Karma Model, Tercih Edilen Görü Lokasyonu, Ters Optimal Görü Lokasyonu

To "Meine Herzallerliebste",
Kağan Tonyukuk FİKRİ

ACKNOWLEDGEMENTS

I would like to declare that this thesis has been conducted at METU Informatics Institute, Cognitive Science Department within the research project "The Investigation of Cognitive Processes in Reading: Development of a Corpus of Turkish Reading Patterns for Eye Movement Control Modeling" (No: 113K723) funded by The Scientific and Technological Research Council of Turkey (TÜBİTAK).

I would like to express my appreciation to my supervisor Assist. Prof. Dr. Cengiz Acartürk for his guidance, encouragement and support throughout the course of this work. I am also thankful to my committee members, Prof. Dr. Cem Bozşahin, Prof. Dr. Deniz Zeyrek Bozşahin, Assoc. Prof. Dr. Bilal Kırkıcı, and Assoc. Prof. Dr. Selda Özdemir for their insightful comments and suggestions.

I would like to mention that I am grateful to my parents Aslıhan and Fikret Beken, who always believed in me and always there for me whenever I needed them; their love is precious to me. I am also thankful to Oya Fikri, whose support was invaluable to me. I would like to thank Fulya, Esra, and Orhan for their emotional support; I am too lucky to have them.

I would like to thank all my friends Özge, Mehmetcan, Natalia, Burcu, Melda, Ece, Tunç, Gökhan, Serkan, and Tuğçe for their kindness. It was worth a lot to see their smiling faces everyday. I was lucky to have such friends. I also would like to thank Eda Arbaş and Hakan Güler for their sincere assistance during the thesis submission process.

Finally, I would like to express my gratitude to Kağan for his endless trust and love. He was the one who encouraged me to apply to Cognitive Science department three years ago. I am thankful to him for his emotional support during these years. Most importantly, I am thankful to him for being with me in this journey, as my best friend and my love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Meaning
FLP	First Fixation Landing Position
FFD	First Fixation Duration
GD	Gaze Duration
RP	Refixation Probability
SP	Skipping Probability
WL	Word Length
LS	Launch Site
WF	Word Frequency
WP	Word Predictability
RL	Right to Left
LR	Left to Right
LMM	Linear Mixed Model
GLMM	Generalized Linear Mixed Model
OVP	Optimal Viewing Position
PVL	Preferred Viewing Location
IOVP	Inverted Optimal Viewing Position

CHAPTER 1

1. INTRODUCTION

1.1. Reading Process

Reading is one of the sophisticated cognitive processes that has been studied within the fields of cognitive psychology and cognitive science as an information processing setting in which the reader extracts information from the written text by decoding the symbols and generates the meaning. In this context, reading can be defined as the coordination of word recognition processes as well as sentence-level syntactic and semantic processing.

Reading process can be investigated through different ways depending on the aspect that is desired to be studied; for instance, if the aim is to study reading comprehension, then it can be feasible to examine readers' performances on answering questions about the content of a text. Likewise, if the goal is to study the role of inner speech, different methods may be applied to determine its role. However, if the goal is to explore the cognitive processes occurring during normal silent reading, the component process of the study could be distorted by different techniques that readers do. Thus, choosing any task will not be plausible if the working process during silent reading of text is not known (Rayner, Pollatsek, Ashby, & Clifton Jr., 2012).

The favored method for investigating cognitive processes during silent reading is eye movement recording (Rayner et al., 2012), which has been widely used in the past thirty years to interpret cognitive processes during reading (Rayner, 1978a, 1998, 2009; as cited in Rayner et al., 2012). Eye movements in reading are at the center of research because they provide regular eye movement patterns (cf. fixations) where there is a link between these patterns and comprehension (Rayner, 1998; as cited in Hayhoe, 2004).

1.2. Purpose of the Study

The aim of this study is to investigate the eye movement patterns during normal silent reading of Turkish sentences by a set of eye movement measures. These measures are given below with their corresponding meanings.

- First Fixation Landing Position (FLP): The position of the first fixation on the word in terms of pixel distance or character distances.

- First Fixation Duration (FFD): The duration of first fixation on a word; expressed in ms.

- Gaze Duration (GD): The sum of all fixation durations on a word during first-pass reading; expressed in ms.

- Refixation Probability (RP): The probability for a word being fixated more than once during first-pass reading. This measure is usually expressed as a percentage of the refixated words across total number of words.

- Skipping Probability (SP): The probability for a word being skipped, usually expressed as a percentage of the skipped words across total number of words.

The analyses were conducted by considering the effects of word properties:

- Word Length: The number of characters in a word.

- Launch Site: The number of characters including the empty space character from the last fixation landing position to the beginning of the currently fixated word.

- Word Frequency: The count of the whole-word within BOUN Corpus (Sak, Güngör, & Saraçlar, 2008).

- Word Predictability: Predictability of a word within a sentence by seeing the previous part of the sentence until the word in question. Predictability was estimated by a cloze task (Taylor, 1953), and the score was the proportion of subjects who correctly guessed the word from preceding context (Nilsson, 2012).

In this study, a corpus-analytic approach was employed (Yan, Zhou, Shu, Yusupu, Miao, Krügel, & Kliegl, 2014).

In this section the goal of the study was explained. Also, the eye movement measures and the word properties were briefly introduced. These measures and properties will be discussed in a more detailed way in the upcoming sections (see Chapter 2, Section 2.2). In the following section, the hypotheses of the study will be presented.

1.3. Hypotheses & Motivation

In the context of reading research, eye movements during reading have been examined in various languages like English, German, and Finnish etc. In these studies, many classical findings were observed. The primary expectation is to observe these findings in Turkish, too. The hypotheses tested in this study can be summarized as follows:

- **i.** The first hypothesis is that word length has an effect on the first fixation landing positions, first fixation durations and gaze durations. It is expected that first fixations will tend to land around word centers except for long words. Gaze duration is expected to increase as the word length increases since the number of fixations will tend to increase.

- **ii.** The second hypothesis is that launch site will have an effect on the first fixation landing positions. As the launch site increases, first fixations will tend to land close to the word beginnings.

- **iii.** The third hypothesis is the effect of word frequency on first fixation landing position, first fixation duration and gaze duration. Fixation durations will increase on low frequency words comparing to high frequency words. It is also expected that low frequency words will tend to have more refixations.

- **iv.** The final hypothesis is the effect of word predictability. It is expected to observe the effect of word predictability on first fixation durations as well as skipping probabilities. More predictable words will tend to be skipped more.

The motivation of this study is that there is a lack of previous work on eye movement patterns during reading in Turkish. A corpus analytic approach has the potential to provide the general characteristics of Turkish reading patterns.

In the following chapter, eye movements in reading will be presented in detail under the sections of the relevant literature, basic characteristics, and the relation to cognitive processes.

CHAPTER 2

2. BACKGROUND

2.1. Eye Movements in Reading

2.1.1. Relevant Literature

Reading process is one of the research topics that have been investigated by analyzing eye movement patterns. In this section, a brief literature will be presented about the development of eye movement studies through the advancements of eye tracking technologies.

Early Studies of Eye Movements during Reading

According to Rayner (1998), *the first era* of eye movement research was between 1879 and 1920. He stated that many basic facts and issues about eye movements like "saccadic suppression (the fact that we do not perceive information during an eye movement), saccade latency (the time that it takes to initiate an eye movement), and the size of the perceptual span (the region of effective vision)" were considered in this era (Rayner, 1998, p. 372).

As Wade (2010) stated, in the book *The Psychology and Pedagogy of Reading* written by Edmund Burke Huey, eye movement behavior was described with the following words: "Eyes do not move continuously from left to right along the line, but proceed by a succession of quick, short movements to the end, then return in one quick, usually unbroken movement to the left" (Huey, 1908, p. 15; as cited in Wade, 2010). According to Huey (1908), these eye movements during reading were first described by the French ophthalmologist Louis Émile Javal in 1879 (Huey, 1908; as cited in Wade, 2010).

In 1879, Hering published his study about the spatial sense and eye movements which then leaded him to work on eye movements during reading which is published in the same year. His study was described as follows:

> *He attached a rubber tube to a cigar holder and listened to the sounds produced when it was placed on the eyelids: he "heard a surprisingly strong and whirring roar" (page 137). When he placed the device on the eye of curarised dogs or rabbits, the sounds ceased even though the blood flow continued. Having determined that the sounds reflected muscular contractions, Hering applied the technique to the eye lid of an open eye. Even when he tried to keep his eyes still, he heard the sounds: "Throughout one's observations, one hears quite short, dull clapping sounds, which follow each other at irregular intervals" (page 145). Hering was able to use his experience with afterimages to establish that the sounds were correlated with eye movements by comparing them: "every clapping sound corresponds to a displacement of the afterimage" (page 145). This was confirmed by movements of floaters, which occurred with the clapping sounds (Hering, 1879b; as cited in Wade, 2010).*

The relevance of Hering's study to reading is that he also applied his technique to reading during which he observed the clapping sounds very clearly and stated that the clapping sounds revealed the jerky movement of the eyeball (Hering, 1879b; as cited in Wade, 2010). Hence, the discontinuity of eye movements was described by Hering whose report was published in the same year as Javal described "saccades" during reading (Wade, 2010).

In sum, before the observation of the saccades, it was assumed that eyes sweep smoothly across the text. The observation of saccades was a milestone in the reading research area. This finding leaded to some questions about reading process like "Where does the eye stop?", "For how long?", "Why does it stop there?", and "Why does it regress at times?" (Paulson & Goodman, 2009).

Further Studies in Reading by Development of Eye Tracking Technologies

In the 20th century, currently available eye tracking technologies started to appear in the eye movement research field. An early eye tracker was built by Edmund Huey. He used a kind of contact lens associated with an aluminum pointer that was moving when the eyes moved. He studied regressions and saccades (Huey, 1908).

In the following years, which can be called *the second era*, reading research continued with the focus of behavioral approaches rather than eye movement studies which were rather few than those of behavioral studies. In 1958, Tinker made a review with a pessimistic note declaring that everything that could be found out about reading from the eye movements had been found (Tinker, 1958; as cited in Rayner, 1998). This opinion might be influential in those days so that between 1950s and mid-1970s there were not much research about eye movements (Rayner, 1998).

In the third era, researches about eye movements began in mid-1970s and eye movement data analysis methods were concerned by various studies (see Kliegl & Olson, 1981; Pillalamarri, Barnette, Birkmire, & Karsh, 1993; Scinto & Barnette, 1986; as cited in Rayner, 1998), and much has been found out about the essentials of eye tracking systems (see Deubel & Bridgeman, 1995a, 1995b; Muller, Cavegn, d'Ydewalle, & Groner, 1993; as cited in Rayner, 1998). According to Rayner, enormous technological proceedings have been achieved in this era. Hence, progressive methods like gaze contingent display change paradigm have been developed (McConkie, 1997; McConkie & Rayner, 1975; Rayner, 1975b; Reder, 1973; as cited in Rayner, 1998). Finally, Rayner (1998) stated that by means of the advancements of general theories of language processing, it became conceivable to utilize eye movement records for the investigation of underlying cognitive mechanisms of reading (Rayner, 1998).

To sum up, reading studies have continued since the 19^{th} century. The studies were accelerated by the advancements of currently available eye tracking technologies in the 20^{th} century. This allowed reading researchers to test and observe many other hypotheses beyond basic characteristics of eye movements during reading. In the following section, these basic characteristics will be explained in more detail.

2.1.2. Basic Characteristics

In this section, the basic characteristics of eye movements during reading will be presented. In the previous section, it was mentioned that eyes do not make continuous movements during reading; rather, they make quick short movements (Huey, 1908). In order to better understand this behavior, it is worth to provide brief information about the general basic types of eye movements.

The Types of Eye Movements and Their Functions

In this section, measurements of eye movements during reading will be presented. Before moving on with this topic, it would be better to describe the eye movement types and their functions briefly.

The four basic eye movements are defined in the book of *Neuroscience* (Purves, Augustine, Fitzpatrick, Hall, LaMantia, McNamara, & Williams, 2004) as follows:

- *“Saccades* are rapid, ballistic movements of the eyes that abruptly change the point of fixation.” (p. 457)

- *“Smooth pursuit movements* are much slower tracking movements of the eyes designed to keep a moving stimulus on the fovea.” (p. 457)

- *“Vergence movements* align the fovea of each eye with targets located at different distances from the observer.” (p. 458)

- *“Vestibulo-ocular movements* stabilize the eyes relative to the external world, thus compensating for head movements.” (p. 458)

Here, the basic types of eye movements were explained briefly. Before moving on with the characteristics of eye movements during reading, it could be better to gain insight about the anatomy of retina and visual acuity.

The Visual Field and Acuity

Eyes make frequent saccades because of visual field acuity limitations (Rayner, 1998). The visual field is divided into three regions: foveal, parafoveal, and peripheral. Visual field is determined by the anatomical structure of the retina which has two different cell types namely cones and rods. Cone cells provide visual acuity, while the rod cells have functions like providing night vision. Cone cells are the majority in the fovea; whereas, peripheral region is predominated by rod cells. In the parafovea, both the cone and rod cells exist. In Figure 1, the distribution of the cones and rods across the visual field are displayed (Rayner et al., 2012).

The foveal region has the highest acuity, which extends out to 2 degrees of visual angle and corresponds to 6 to 8 characters depending on the font size. The parafoveal region extends out to 5 degrees of visual angle, meaning 15 to 20 characters in reading (Wotschack, 2009). The region beyond parafovea is considered as peripheral region.

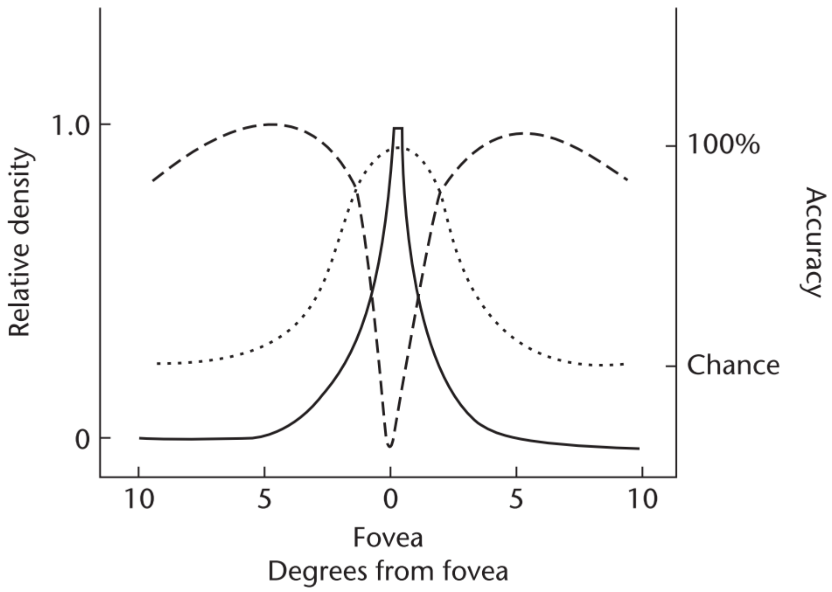

Figure 2.1. Relative density of cones (solid line) and rods (dashed line) across the visual field. Dotted line shows the accuracy of identifying a target word exposed briefly to the right or left of fixations (Rayner et al., 2012).

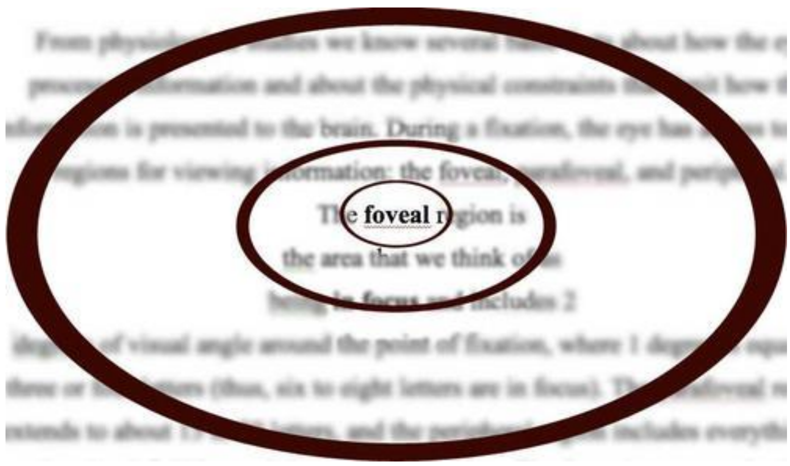

Figure 2.2. The foveal, parafoveal, and peripheral region from the inside out (Encyclopedia of Educational Technology. Eye Movement. http://eet.wikispaces.com/Eye+movement).

The anatomy of the retina and the visual field information indicates that readers' eyes move over the text during reading in order to bring the text into the region of central vision (Wotschack, 2009).

All in all, saccadic eye movements are resulted from physiological properties of the eyes. During each fixation, the acquired number of letters is limited due to visual field acuity. Since fovea has the highest visual acuity, saccades are made during reading to place fovea over the text.

Fixations and Saccades

There are two basic components of eye movements during reading: fixations and saccades (Rayner, 2009; Reichle, Rayner, & Pollatsek, 2003). Saccades are rapid continuous eye movements. Saccades take about 30 ms., whereas, during fixations, eyes remain relatively still for about 200-300 ms (Rayner, 1998). The fixation range changes from 100 ms to 500 ms (Sereno & Rayner, 2003). A usual saccade length is between 1 and 20 characters with the average of 7 to 9 characters. Fixations and saccades are conducted during reading because of the limitation on the number of letters that can be acquired during a single fixation (Simola, 2011).

The visual input during reading is taken only during the course of fixations, and the reader is literally blind during saccades. However, saccades carry the eyes from one word to the next word, shift the fixation of currently fixated word, skip the next word, and move back to an early word (Kliegl & Engbert, 2013). This blindness during saccades is referred as *saccadic suppression* phenomenon, which was described by Bridgeman et al. (Bridgeman, G., Hendry, D., & Stark, L., 1975).

During a fixation, some processes are completed before moving on to the next word (Sereno & Rayner, 2003). Visual processing from the retina to brain takes 50 ms (*eye-to-brain lag*; VanRullen & Thorpe, 2001). Then, the lexical processing is expected in the first 100-200 ms for programming the next saccade (Sereno & Rayner, 1998). After that, the fixation terminates with the *saccade latency* during which the location of the next saccade is encoded (Findlay & Walker, 1999; Rayner, 2009).

To sum up, the two basic types of eye movements during reading are fixations and saccades. Due to saccadic suppression, information is gained only during fixations (Rayner, 2009). Next, the questions of where and when to move the eyes will be discussed in turn in the light of the findings discovered until today.

2.1.3. Where to Move the Eyes

The *optimal viewing position (OVP)* in single word recognition has been reported as at the center of the word (O'Regan & Lévy-Schoen, 1987; as cited in Wotschack, 2009). There are two reasons to call this position as optimal: The first reason is that word recognition is the fastest at this location; and second one is that the more the first fixation landing position deviates from the optimal viewing position, the more likely that refixations happen (Wotschack, 2009).

In contrast to the optimal viewing position, *preferred viewing location (PVL)* indicates that eyes tend to land slightly to the left of the word centers during continuous reading (Rayner, 1979; as cited in Yan et al., 2014). The previous studies showed that preferred viewing location curves form a Gaussian distribution (cf. Yan et al., 2014). This effect has been replicated in many other studies, which showed a consistency with the original finding (cf. McConkie, Kerr, Reddix, & Zola, 1988; McConkie, Kerr, Reddix, Zola, & Jacobs, 1989). McConkie et al. (1988) described it as "simply the maximum point in a distribution of all fixations on the word, which is referred to as the composite distribution" (McConkie et al., 1988, p. 6). The variance between preferred viewing locations and optimal viewing positions can be attributed to saccadic range error and oculomotor variability (McConkie et al., 1988). Moreover, refixation probability decreases at the preferred viewing locations (Wotschack, 2009).

In the previous studies, it was reported that low level, orthographic text properties like word length and spaces between the words are strongly responsible for the spatial movements of the eyes (reviewed in Rayner, 2009). It was shown that when spaces between the words are removed from English texts, the preferred viewing location distribution no more has a Gaussian shape, it rather decreases and distinctly and linearly from the beginning to the end of a word (Rayner, Fischer, & Pollatsek, 1998; Rayner & Pollatsek, 1996; as cited in Yan et al., 2014). Also, similar results were found in Chinese studies (Li, Liu, & Rayner, 2011; Yan, Kliegl, Richter, Nuthmann, & Shu, 2010; as cited in Yan et al., 2014).

2.1.4. When to Move the Eyes

The question of when to move the eyes is decided by the lexical properties of the fixated word, such as word frequency and predictability of a word from the sentential context (reviewed in Rayner, 2009). Previous studies have shown that fixation duration is affected by the difficulty of word (Liversedge & Findlay, 2000; Rayner, 1998; Yang, 2012).

Vitu and his colleagues observed that fixation durations during continuous reading vary with the fixation landing positions in such a way that fixation durations have the highest value around the word center and decrease at the ends of the word, which is called *inverted optimal viewing poisiton (IOVP)* effect (Vitu, McConkie, Kerr, & O'Regan, 2001). Nuthmann and colleagues has argued this effect in terms of saccadic error which is stated as causing more mislocated fixations at the end of the words (Nuthmann, Engbert, & Kliegl, 2005, 2007; as cited in Wotschack, 2009).

In this study, where to move the eyes and when to move the eyes have been considered by analyzing the effects of word length, launch site, and word frequency as well as word predictability. For the spatial analysis, only first fixation landing position has been included in the analysis. Second fixation durations or regressions were considered as beyond the scope of this analysis. For the temporal analysis, first fixation duration and gaze duration during first-pass reading have been observed. Similarly, fixation durations during regressions were not included in the analysis. These eye movement measures will be presented in the next section.

2.2. Eye Movement Measures in Reading Research

In reading studies, eye movements are measured with respect to certain characteristics of saccadic movements. In Figure 3, these measures are visualized. As it is seen in the figure, a word is called *refixated* when it gets more than one fixation. *Regression* happens when the eyes jump back to the previously passed words. Finally, a word is called *skipped* if it receives no fixation during reading (Wotschack, 2009).

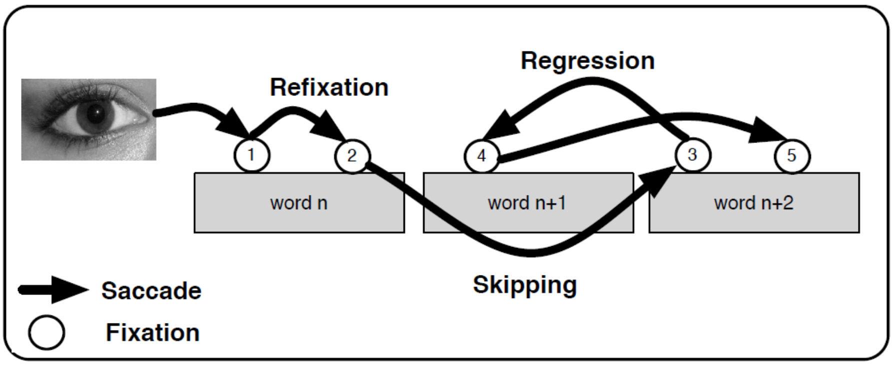

Figure 2.3. Types of saccadic eye movements: The circles 1 - 5 represent fixations in a sequential order; fixations 1, 2, and 3 are first-pass fixations, fixations 4 and 5 are second-pass fixations; the sum of fixations 1 and 2 is the first-pass gaze duration on

word n (as well as its total reading time); the sum of fixations 3 and 5 is the total reading time on word n+2; fixation 3 is a regression origin fixation, fixation 4 is a single fixation regression goal; word n+1 has been skipped in first-pass reading (Wotschack, 2009).

There are also different types of processing time measures, which are considered through first pass and second pass reading depending on the analysis. The first pass reading includes the initial forward fixations within a sentence; whereas, second pass reading covers all the fixations during first pass reading as well as the fixations after a regressive eye movement (Wotschack, 2009).

In this study, first fixation landing position (FLP), first fixation duration (FFD), and gaze duration (GD) have been examined by the effects of word length, launch site, word frequency, and word predictability.

2.3. Eye Movements and Cognitive Processes during Reading

What can eye movements tell us about cognitive processes? Before considering this question in the scope of reading studies, it could be better to consider it in general. Going back to the earlier studies than reading, we come across with Yarbus' (1967) study. He investigated the relationship between eye movements and cognitive processes by means of natural viewing and scene perception. The conclusion was that eye movements are driven by their relevance to the task rather than the saliency of the objects (Yarbus, 1967; as cited in Wotschack, 2009). Relying on this result, it can be expected that eye movement patterns on identical reading material may vary depending on the reading intention (Wotschack, 2009).

The relationship between eye movements and attention is the primary reason behind the importance of eye movements for information processing researches. It has been reported that visuospatial attention affects eye movements, which are not random (Hoffman, 1998). Accordingly, analysis of the eye movements during reading is particularly important because it allows measurement of cognitive processes related to attention. The methods used in reading research from past to present has shifted from behavioral experiments to modeling based on eye movements.

In addition to behavioral experiments and brain imaging methods, computational models for eye movements have become a focus of interest among researchers who investigate cognitive processes. In these models, the regularities in eye movement patterns during reading have been evaluated by considering the characteristics of the words.

In summary, the most significant benchmarks that make the reading process observable, can be revealed by examining the eye movements that occur during reading. Eye movement recordings are also important in terms of providing objective data compared to alternative behavioral research methods. Therefore, modeling of the cognitive processes within the framework of information processing theories is carried out by the modeling of the eye movements during reading. These models are tested and developed by the analysis of data obtained by the experiments. In the following section, eye movement modeling in reading will be discussed.

2.4. Models of Eye Movement Control in Reading

The development of a model of eye movement control during reading is one of the goals of eye movement research in reading. According to Reichle (2006), one critical distinction about models concerns the extent to which on-going cognitive processes like word identification influence the moment-to-moment decisions about when to move the eyes (Reichle, 2006). As stated by Reichle, the early models of eye movement control could be separated into two types: *cognitive-control models*, in which it is noted that eyes move owing to the completion of some cognitive event which is in general word identification (Just & Carpenter, 1980, 1987; Morrison, 1984; Rayner & Pollatsek, 1989; Thibadeau, Just, & Carpenter, 1982; as cited in Reichle, 2006); and *oculomotor-control models,* in which the oculomotor system is responsible for the eye movements providing the persistency of the progress through the text (O'Regan, 1990, 1992; Suppes, 1990, 1994; as cited in Reichle, 2006).

In the context of reading research, current computational models of eye movement control considers the issue by combining both lower-level oculomotor and higher-level cognitive viewpoints of processing which expectedly directs eye movements in reading (e.g., SWIFT, Engbert, Nuthmann, Richter, & Kliegl, 2005; E-Z Reader, Reichle, Pollatsek, Fisher, & Rayner, 1998) (Wotschack, 2009). It is further stated that the models so far have considered the issue at the word-level and have not gone beyond like syntactic parsing or reading intentions or reading strategies (see Reichle, Warren, & McConnell, 2009, for a first attempt) (Wotschack, 2009).

In the scope of the current computational models, it is worth noting that these models are based on a set of assumptions. As a matter of fact, there is a variation among these models in terms of the assumptions they make which causes new and interesting predictions about the eye movements of readers in particular circumstances (Reichle, 2006). For instance, the E-Z Reader model assumes that the oculomotor system's saccade programming is caused by an early stage of lexical processing, whereas SWIFT hypothesized that complication with lexical processing can cause an inhibition of saccades that are otherwise started at random intervals (reviewed in Reichle, 2006). The assumptions of the models are not only based on the distinction of the lower-level oculomotor or higher-level cognitive processes, but also Reichle (2006) suggested that the role of attention in these models lead to a range of assumptions (Reichle, 2006). For example, according to the Competition/Interaction model and SHARE, attention has a minor role on the guidance of eye movements of the readers. On the contrary, in E-Z Reader model, Glenmore, and SWIFT, the assumptions include the effect of attention (reviewed in Reichle, 2006).

In the last decade, the debate among the eye movement control models, which can predict the eye movements during reading, has been focused on whether the relation between eye movements and cognitive processes is straightforward or not (Henderson, 2012). In this context, the E-Z Reader model (Reichle, Rayner, & Pollatsek, 2003) and the SWIFT model (Henderson, 2012; Richter, Engbert, Kliegl, 2006; Schad & Engbert, 2012) differ from each other in terms of the assumptions about cognitive processes they made. In particular, the SWIFT model assumes that multiple words are processed at the same time. On the other hand, the E-Z Reader model predicted that multiple words can be processed in one eye fixation; however, it also predicted that this assumption did not require the words to be processed simultaneously. Therefore, in the E-Z Reader model, it is assumed that one word is processed at each time, but at each fixation more than one words are processed.

In this section, the eye movement models were presented briefly. In the next section, modeling with linear mixed model will be introduced.

2.5. Modeling with Linear Mixed Model (LMM)

Linear mixed model is a statistical model which has the power of dealing with inter-dependencies by controlling random effects. This allows assuming a different baseline response value for each factor; i.e., the individual differences with respect to each factor can be modeled by assuming different random intercepts for each response. In fact, the reason why linear mixed model called “mixed” is because both fixed and random effects are controlled in the same model (cf. Bates, 2007).

In this study, linear mixed model analysis was conducted. The fixed and random effects were as follows:

- Fixed effects
    - Word length
    - Launch site

    - Word frequency
    - Word predictability

- Random effects
    - Subjects
    - Sentences
    - Word types

In this thesis *lme4* and *LMERConvenienceFunctions* packages were used to fit models. The details can be found in Chapter 3, Section 3.7.

In the following section, the corpus-analytic approach and its properties will be introduced.

2.6. The Corpus-Analytic Approach

The corpus-analytic approach has been widely used in the field of linguistics (Römer, 2006). A corpus based analysis has the following properties (Biber, Douglas, Conrad, Susan, Reppen, & Randi, 1998; Conrad, 1990; as cited in Wang, 2005, p. 506):

- "It is empirical, analyzing the actual patterns of use in natural texts."

- "It utilizes a large and principled collection of natural texts as the basis for analysis."

- "It makes extensive use of computers for analysis, using both automatic and interactive techniques."

- "It depends on both quantitative and qualitative analytical techniques, especially functional interpretations of language use."

In this study, a corpus analysis was conducted. The sentences were chosen from METU Turkish Corpus (Say et al., 2002) on a random basis except some eliminations (see Chapter 3, Section 3.2). These eliminations were due to technical limitations (e.g. limited number of characters in single-line sentences) and some assumptions (e.g. independency of the sentences from the former ones). The corpus-analytic approach allowed to observe the eye movement variables in a natural environment.

CHAPTER 3

3. METHODOLOGY

In this chapter, the information about participants, materials, and apparatus will be presented. Moreover, the procedure, eye movement data analysis, and data selection will be explained in detail. Finally, linear mixed modeling used in this study will be described in the last section of this chapter.

3.1. Participants

Thirty-seven students from Middle East Technical University, who were native Turkish speakers, participated in the experiment. Seventeen of these participants were female and twenty of them were male, whose ages were between 18 and 27 (M=21.6, SD=2.36). Before the experiment started, the participants read and signed an informed consent form where they were acknowledged about the study, and they agreed to participate in the experiment knowing that they could leave the experiment any time they liked. The participants were also given a demographic data form where they were asked about their personal information and language use. According to the answers, none of the participants were diagnosed with language disorders like dyslexia, stuttering, etc. Only one participant noted that he had just little speech impairment in the past[1]. None of the participants were wearing contact lenses during the experiment. Two participants reported that they had astigmatism and myopia correspondingly. Although, they did not wear glasses during the experiment, their data were not considered as reliable, so the data from those participants were removed from further analyses. One participant used glasses during the experiment;

[1] A post-check revealed no difference in average results of this data.

those data were also not included in the analysis. All these removed data were from male participants; hence, there were seventeen male participants left, and overall thirty-four data took part in the analysis. After the removal of the subjects' data, new age scale ranged from 18 and 27 (M=21.47, SD=2.22). The experiment sessions lasted for about 30 minutes and each participant received 20 TRY for participation.

Another independent group of seventy native Turkish students from Middle East Technical University participated in predictability test study to determine the predictability of words within the sentences. Thirty-three of these participants were female and thirty-seven of them were male, whose ages were between 19 and 36 (M=23.83, SD=3.53). This experiment lasted for about 20 minutes and each participant received 20 TRY for participation in this study.

3.2. Materials

A total of 120 sentences were chosen from METU Turkish Corpus, which is a collection of 2 million words of post-1990 written Turkish samples (Say, Zeyrek, Oflazer, & Özge, 2002). The selection of the sentences was done based on some criteria as presented below:

- The sentences consisted of nine words (M=9.0, SD=0.0). The sentence character sizes were not greater than 73 in order to fit into a line on the screen. Courier New font type and 18 font size were used. Thus, each character corresponded to 0.46 degrees of visual angle.

- There were no hyphenated words or numbers as well as punctuation marks within the sentences except the full stops at the end.

- There were not abbreviations like "ODTÜ" in the sentences.

- Question and interjective sentences were avoided.

- After applying the above criteria, there were around a thousand sentences left. These sentences were checked manually (i.e. read one-by-one) and the ones which were obviously related to the former ones were avoided.

The chosen sentences were comprised of 1080 tokens of 775 words (types). Word length varied from 1 to 16 (M=5.6, SD=2.5). The percentages of words with the length of less than 5, between 5 and 10, and above 10 letters were 36%, 60%, and 4% respectively. Suffix number varied from 0 to 5. The frequencies of the complete word forms were taken from BOUN Web Corpus (Sak et al., 2008). Additionally, 3 paragraphs were chosen from Barış Bıçakçı's "Şehir Rehberi" which is composed of short passages of stories.

3.3. Apparatus

Eye movements of the subjects were recorded monocularly with an EyeLink 1000 system (1000 Hz) where tower mount was used. Using the tower mount reduces the head movements substantially by the help of its forehead rest and chin rest parts. Sentences possessed just one single line on the screen, left-aligned, and were displayed each one in turn at the 1/3 vertical position from the top of the screen of a 17 inch VGA Monitor (1024 x 768 resolution) controlled by a computer running at 3.0 GHz under a Windows XP environment. Subjects were seated 62 cm in front of the screen with the head positioned on a forehead rest and chin rest. Texts were displayed with 18 pt fonts where each character corresponds to 0.46 degrees of visual angle and approximately 14.03 pixels.

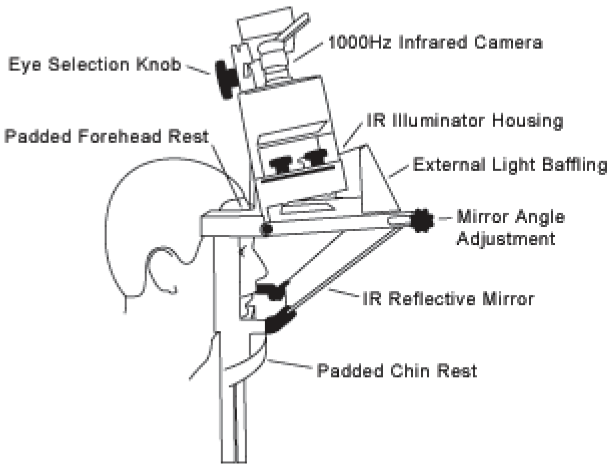

Figure 3.1. EyeLink tower mount (Source: http://www.sr-research.com/mount_tower.html)

3.4. Procedure

All participants were properly informed about the study and they participated in the experiments with the acknowledgement of the use of their responses in academic publications with the assurance of the confidentiality of their personal identities. The study was approved by the Middle East Technical University ethics committee.

Eye Tracking Experiment

The participants who participated in eye tracking experiments were instructed to read the sentences silently for comprehension with their normal reading speeds. They were informed that after some trials there would be yes-or-no questions related to the sentences. They were enabled to reply via EyeLink Button Box's left and right buttons. The left button was associated with the answer “Yes” which was displayed on the left side of the screen; and, the right button was associated with the answer “No” which was displayed on the right side of the screen. The participants accurately answered 96% (SD=5%) of the questions.[2] Also, they were informed that there would be breaks during the experiment and they could have a rest as long as they wanted and they could continue when they were ready. In the experiment, in addition to the 120 sentences, 7 practice trials, 3 paragraphs, and 18 questions were used; hence, the experiment was comprised of overall 148 trials. The paragraphs were displayed in the beginning, in the middle and at the end of the experiment.

The experiment started with general instructions and then the subjects were calibrated with a standard nine-point grid and the validation of calibration accuracy was checked for each subject. Then, a bull’s eye was displayed on the left side of the screen, and when the participant fixate on this mark for 500 ms, the sentence was displayed such that the first letter of the sentence was at the same position with the bull’s eye. Another bull’s eye was used on the lower right side of the sentence screen so that the participants looked at this mark after they read the sentences and again after 500 ms, the blank screen with the bull’s eye displayed on the left side followed.

The experiment started with the practice sentences; each participant read seven practice trials where five of them were sentences and two of them were questions; then, one paragraph was displayed and after the participant read the paragraph, the experiment continued with 46 trials comprised of sentences and random questions; then, the first break session came. After that, the 23 trials were displayed which is followed by the second paragraph; and then, again 23 trials were displayed and the second break session followed. After the break, the experiment continued with 46 trials and then the last paragraph was displayed and experiment was terminated.

[2] Five of the participants’ answers were not recorded due to technical issues. So, this estimation is based on 29 participants.

After each break session, calibration and validation of calibration accuracy were carried out before proceeding with the experiment. The sentences and the paragraphs were randomized in themselves for each participant; so, the sentences and the paragraphs were counterbalanced across subjects.

Predictability Experiments

In order to collect predictability data, a cloze task was applied. each sentence was divided into seven, because the first and the last words were excluded from the analysis; hence, predictability scores for these words were not taken into account which leaded to seven words left in each sentence. Consider the sentence "Yıllarca yaşamış bir geminin mutlaka anlatacak bir şeyleri vardır" meaning "*A ship that has existed for years has definitely something to tell*". The sentence was divided as follows:

1. Yıllarca …
   For years ...
2. Yıllarca yaşamış …
   Existing for years ...
3. Yıllarca yaşamış bir …
   A(n) ... existing for years
4. Yıllarca yaşamış bir geminin …
   Of a ship that has existed for years ...
5. Yıllarca yaşamış bir geminin mutlaka …
   A ship that has existed for years definitely ...
6. Yıllarca yaşamış bir geminin mutlaka anlatacak …
   A ship that has existed for years definitely ... to tell
7. Yıllarca yaşamış bir geminin mutlaka anlatacak bir …
   A ship that has existed for years definitely a(n) ... to tell

All the sentences were divided as such, and every sentence in each condition put together; hence, constituted an experiment session. So, there were overall seven separate experiment sessions – one for each condition –, which had 120 trials. Each participant was presented a different set of 120 trials.

The participants were instructed to fill the blanks with a single word with respect to the sentential context depending on their intuitions and not thinking about the word for a long time. Each condition was applied to a different participant. So, overall 70 leaded to each word being tested by 10 participants. For each participant, correctly predicted words were scored by 1, and others were scored as 0. The scoring was applied by comparing the whole-word results of participants with the actual words.

This assumption was subject to future work. For each word, the predictability score was the proportion of the subjects who had predicted the word in question correctly.

So far, the eye tracking experiments and predictability experiments were described. Besides, predictability score assumption and calculation were explained briefly. The analysis of eye tracking experiment will be illustrated in the following section.

3.5. Eye Movement Data Analysis

Data analysis was carried out via the EyeLink Data Viewer which is a software tool that is used for viewing, filtering, and processing EyeLink gaze data. It provides auto segmented interest areas[3] (see Figure 3.2) for each word and different kinds of reports for various variables.

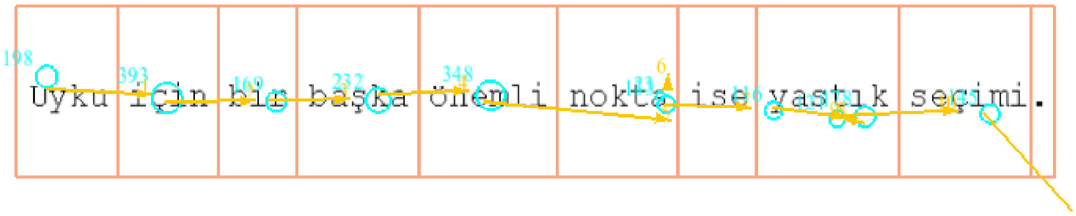

Figure 3.2. A sample eye movement data in DataViewer. The circles indicates *fixations* where the numbers are *fixation durations*; the lines are *saccades*; and the rectangles are *interest areas*.

In this study, *Fixation Report* and *Interest Area Report* outputs from DataViewer were used. From the Interest Area Report, first fixation durations (FFDs) and gaze durations during first pass (GDs) for each interest area were obtained directly. They are named in the Data Viewer as IA_FIRST_FIXATION_DURATION and IA_FIRST_RUN_DWELL_TIME respectively. The report also provided IA_FIRST_FIXATION_X which gives the pixel position of the first fixation along the x-axis within the current interest area. The following calculations were carried out on this value to get the first fixation landing positions (FLPs) in terms of characters: First of all, the empty characters on the left side of the sentences along x-axis, which was 42 pixels, was subtracted from the fixation position value. Then, this new value was divided by 14.03 which corresponds to one character in pixel. Thus, the pixel value had been converted into character unit. After that, for each word, the

[3] Data Viewer allows to create interest areas by segmenting the sentence with respect to words, so that each word falls into one interest area. Thus, the fixation properties within the interest areas can be automatically obtained by the *Interest Area Report*.

sum of word lengths and the number of spaces before the current word were subtracted. The final value became a fractional number, and it does not make sense to talk about fractional character indices; so, it was carefully checked and observed that rounding up these values gave the landed exact character position; hence, they were rounded up in order to get the FLP value.

Similarly, launch site (LS) value is calculated through the Interest Area Report variable IA_FIRST_SACCADE_AMPLITUDE. It is defined in the Data Viewer as the amplitude of the first saccade entering into the current interest area in degrees of visual angle. First, this value was converted into pixel unit by using the screen width resolution, screen width dimension and the subject's eye position distance to the monitor. After that, this pixel value is divided by 14.03 in order to become in character unit. Then, the FLP value of the current word was subtracted, and the result was rounded up which became the launch site in character units.

Interest Area Report also provided whether the current word had been skipped or not during first pass reading. The words which have blinks or not during first pass reading have been obtained via Fixation Report.

In the analysis, log10 transformation was applied to word frequency and log2 transformation was applied to word length values, by following the practice in the literature (Kliegl, Nuthmann, & Engbert, 2006; as cited in Yan et al., 2014). Also, natural logarithm transform was applied to first fixation duration and gaze duration. This transformation removed the skewness in the distributions.

3.6. Data Selection

Some exclusion criteria were applied on the data before proceeding with the analyses. First of all, the first words and the last words were removed from the analysis. This is a common practice in reading research (Yan et. al., 2014). Secondly, all the data were checked manually for any data loss or offset issues; and, the trials with data loss or offset were also removed. Also, first and last fixated words during first pass reading were detected manually and not included in the analysis. Furthermore, the words containing blinks, the words with first fixation durations (FFDs) shorter than 60 ms or longer than 600 ms or gaze durations (GDs) longer than 800 ms were also removed. Fixations on the space before the words were not included in the analysis. The frequencies of the words which were not found in the BOUN Web Corpus (Sak et al., 2008) were assigned as 0. They were removed from the analysis since logarithm of 0 is not defined (recall that log10 transformation was employed to word frequencies; cf. Section 3.5). Finally words with launch sites less than 0, which means regression, were not included in the analysis.

The data selection criteria were presented here briefly. Next section covers how linear mixed model approach was applied to this data.

3.7. Linear Mixed Model (LMM) Analysis

In the preceding chapter, linear mixed models were introduced briefly (see Chapter 1, Section 1.5). In this section, this modeling method will be considered as part of the analysis.

In this study, linear mixed model (LMM) and generalized linear mixed model (GLMM) approaches were used in order to control the random effects structure while defining the fixed effects. The analyses were carried out to observe the effects of the following variables:

- Word length
- Launch site
- Word frequency
- Word predictability

These effects were observed on the eye following eye movement parameters based on linear mixed models (LMMs):

- First fixation landing position (FLP)
- First fixation duration (FFD)
- Gaze duration (GD)

Note that observed FLP results were added to the FFD and GD models since FFD follows FLP (Yan et al., 2014).

The following parameters were analyzed by generalized linear mixed models (GLMMs):

- Refixation probability (RP)
- Skipping probability (SP)

Note that, FLP is added as an additional covariate to RP model. Launch site was not included in the SP model, because skipped words do not have launch sites.
During the analyses, variance components for subjects, sentences, and word types were taken into account. Also the variance components and correlation parameters

for subject-related experimental main effects were included in the models whenever they were significant.

The random effects were specified as follows:

- Random intercepts were included for subjects, sentences, and word identities.
- Random slope factors were only tested by subject variation.
    - Word length and word frequency features were all stable within the sentences and words; hence, it would not make sense to put these variables into the model by sentence and by word identity variation.
    - It is also unnecessary to put the launch site variable as a random slope factor into sentence variation, since launch site is calculated through the words, not by sentences.

The analyses were carried out with the *lmer* and *glmer* functions of the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015), and the *fitLMER.fnc* function of the *LMERConvenienceFunctions* package (Tremblay & Ransijn, 2015).

In this study, *fitLMER.fnc* function was applied to the linear mixed model and generalized linear mixed model analyses.

To apply *fitLMER.fnc* function, an initial model was fitted by the *lmer*/*glmer* function by including all the fixed effects as interaction terms and one random effect, namely the intercept for subjects. Then, this model was tested with the all possible modest random effects by the *fitLMER.fnc* function which returns the optimal model with the necessary fixed and random effects.

As stated by Tremblay and Ransijn (2015), the *fitLMER.fnc* function followed three steps as follows:

i. First, one of the *bfFixefLMER_F.fnc* or the *bfFixefLMER_t.fnc* was called depending on whether the analysis required modeling with linear mixed model or generalized linear mixed model. The analysis of first fixation landing position (FLP), first fixation duration (FFD) and gaze duration (GD) were based on linear mixed models; thus, for model fitting process, *bfFixefLMER_F.fnc* was chosen and the fixed effects were determined by the F value and back-fitted properly. For refixation probability (RP) and skipping probability (SP) analysis, generalized linear mixed model was used since these variables had binary values; thus, *bfFixefLMER_t.fnc* function was chosen to fit the fixed effects in the models. Fitting started with considering highest order interaction model terms. The model term with the highest ANOVA p-value or with the lowest t/z-value was identified based on the backfitting method of F or t respectively. If this p-value was higher; or t/z-value is lower than the threshold, the model term was removed and a new model was fitted. Then, these two models were compared by log-likelihood ratio test. If the comparison did not yield a significant result, i.e., if keeping the term did not increase the model fit, then it was removed from the model; otherwise, it was kept. This was repeated for each model term. After evaluating all the highest-order interaction terms, the process moved on with the second highest order interaction, and applied the same steps that had been applied to the highest-order interaction terms. This procedure continued until all the main effects had been evaluated. Note that if a term was not significant itself within the model, but it was part of a higher order interaction, it was kept in the model.

ii. In the second step, the *ffRanefLMER.fnc* was called to forward-fit the random effects. This function checked each random effect one by one removing from the model and comparing the model to that of the full model by log-likelihood ratio test. If the comparison became significant, then the random effect was kept in the model.

iii. In the third and last step, the fixed effects were checked and back-fitted again because of the possibility that the inclusion of certain random effects might affect the significance of certain fixed effects.

After all the steps applied, *fitLMER.fnc* function returned the optimal fitted model.

In this chapter, the methodological issues were provided. The information about the participants, materials and apparatus have been provided. Also, the experimental procedures, data analysis and data selection have been explained in detail. Finally, information about linear mixed models and how they were employed in this study was described. Next section covers the results of the LMM and GLMM analyses.

CHAPTER 4

4. RESULTS

In this chapter, the results of the models for first fixation landing position (FLP), first fixation duration (FFD), and gaze duration (GD) as well as refixation probability (RP) and skipping probability (SP) were reported individually. Before moving on to the results, it would be nice to recall the meanings of these variables as follows.

- First Fixation Landing Position (FLP): The position of the first fixation on the word in terms of pixel distance or character distances.
- First Fixation Duration (FFD): The duration of first fixation on a word; expressed in ms.
- Gaze Duration (GD): The sum of all fixation durations on a word during first-pass reading; expressed in ms.
- Refixation Probability (RP): The probability for a word being fixated more than once during first-pass reading. This measure is usually expressed as a percentage of the refixated words across total number of words.
- Skipping Probability (SP): The probability for a word being skipped, usually expressed as a percentage of the skipped words across total number of words.

4.1. Findings for First Fixation Landing Position

This section starts with the outcomes and the comparison of the mean first fixation landing position values found in this study with the results found in previous studies. Then, the significance of word length, launch site, and word frequency effects will be discussed. Then, the results will be discussed for the model which has the word predictability as an additional covariate besides all fixed effects including refixation rate.

4.1.1. Comparison with the previous research

The results of first fixation landing position values in this study have been compared to the previous research. Mean FLP values by word length in the previous studies are shown in Table 4.1 in comparison with the present study. It is seen that the results of the present study are in a good agreement with the previous studies (Yan et al., 2014; Nuthmann & Kliegl, 2009; Deutsch & Rayner, 1999; Rayner et al., 1996; Rayner, 1979; Fig. 4.1).

Table 4.1. Mean first fixation landing position (FLP) values by word length in the present and previous studies (LR: Left-to-Right; RL: Right-to-Left)

		Word Length						
		3	4	5	6	7	8	9
LR	Present Study	2.1	2.7	3.2	3.5	3.6	3.9	4.1
RL	Yan et al. (2014)	2.0	2.4	2.7	3.2	3.3	3.8	4.1
LR	Nuthmann and Kliegl (2009)[4]	1.7	2.1	2.5	2.9	3.0	3.2	3.6
RL	Deutsch and Rayner (1999)	2.3	2.7	2.8	3.2	3.5	3.8	-
LR	Rayner et al. (1996)	-	-	2.8	3.1	3.6	4.1	-
LR	Rayner (1979)	2.1	2.4	2.8	3.2	3.5	3.8	-

The comparison showed that the outcomes found in the present study are consistent with the previous studies, which suggest that:

[4] In this study, median landing positions of single fixations were reported.

- The first fixations landed around word center.
- Mean landing position was slightly to the right of the word center for the words with length of 3 to 7.
- These landing positions started to move towards the beginning of words for the words with length of 8 and greater.
- An increase in fixation landing positions were observed with the incremented word lengths as the previous studies suggested.

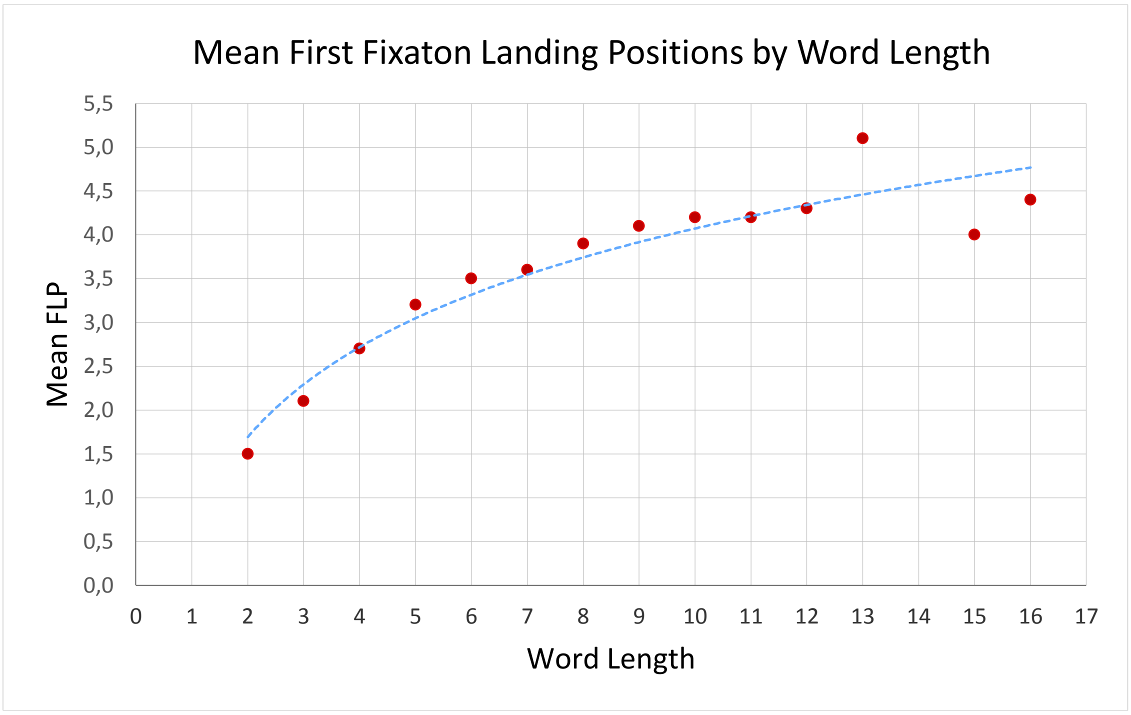

Figure 4.1. Mean first fixation landing position (FLP) values by word length

In Figure 4.1, mean FLP values are displayed by word length. In general, the points constitute a logarithmic increase until word length of 12; however, after 12, the pattern fluctuates. It should be noted that the number of words with length greater than 12 were relatively few (see Appendix A; Table A.2). This could be the reason for the pattern of very long words.

Figure 4.2 displays the preferred viewing locations with respect to word length. It is seen that the most frequently fixated location in 8- to 12-letter words was to the left of the word centers. For 2- to 7-letter words, preferred viewing location was to the right of the word centers. Since the number of 13- to 16-letter words were limited in the data set, the PVL curves for these words is hard to be interpreted here.

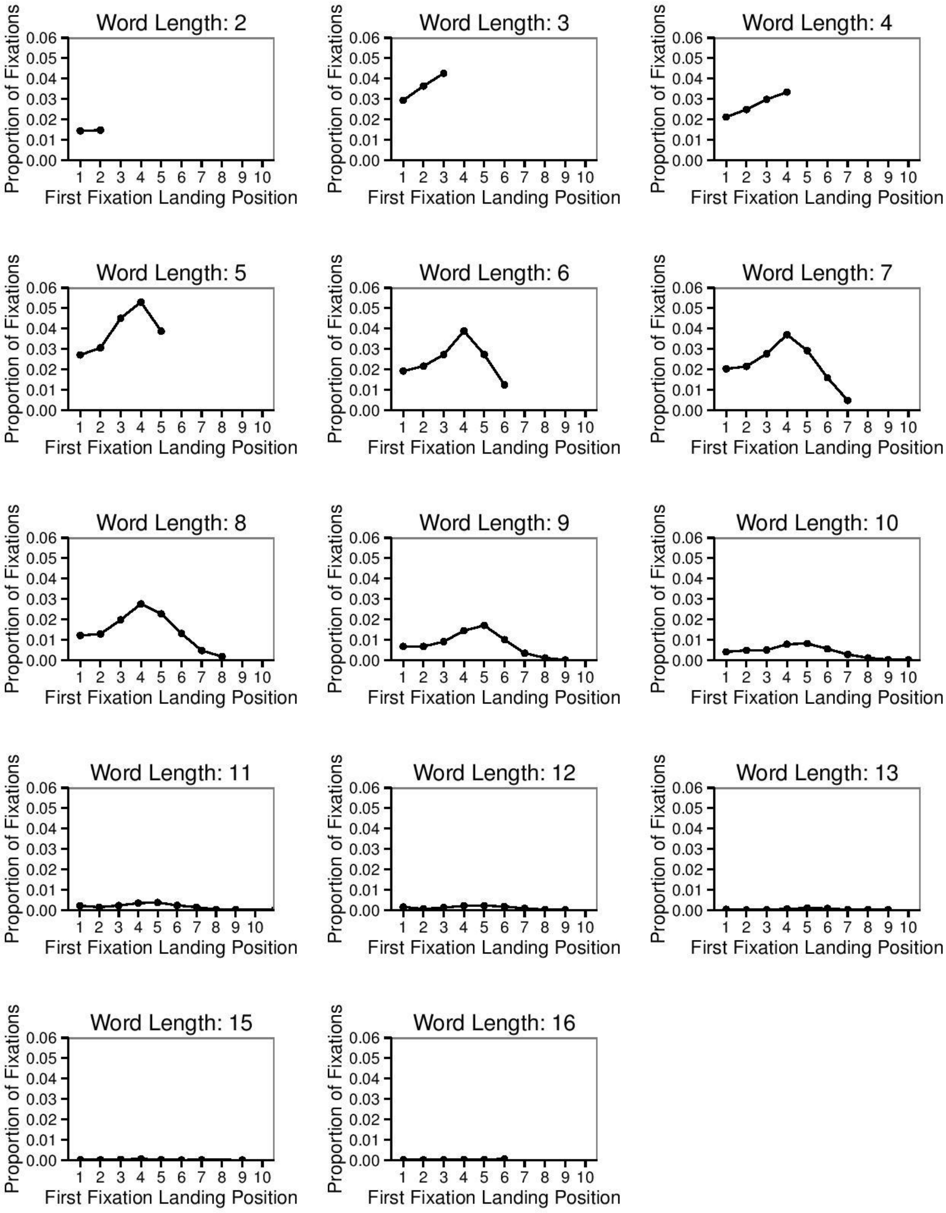

Figure 4.2. Preferred viewing location (PVL) curves for words with respect to the word length

4.1.2. Effects of word length, launch site, and word frequency on first fixation landing positions

In the previous section, first fixation landing position values were considered and compared to the previous studies. In this section, the effects of word length, launch site, and word frequency will be given.

The linear mixed model (LMM) analysis results for first fixation landing position (FLP) can be summarized as follows:

- Word length (b=1.126, SE= .067, t=16.73) and launch site (b=-.222, SE= .004, t=-50.86) has the largest significant effects on FLPs. Landing positions increased with increasing word length, but decreased with increasing launch sites.

- The interaction between these two variables has also been found significant (b=-.068, SE= .009, t=-7.21), which is illustrated in Figure 4.3. It can be observed that while word length was increasing, FLP value also increased; whereas, while launch site was increasing; FLP value decreased substantially.

- The effect of word length faded away while launch site was increasing considerably. The estimated and observed means shows a different pattern for very large launch sites, especially for the launch sites after ten, which may suggest saccade targeting may require the eyes to be within around certain number of letters like ten (Yan et al., 2014; see Fig. 4.3).

- Finally, the significant effect of word frequency on FLP showed that FLP values decreased with decreasing word frequency (b=.051, SE= .023, t=2.16). However, the frequency values were surface (i.e. whole-word) frequencies. When the words get longer, the frequency values drop. So, this effect may also be attributed to the word length effect (see Fig. 4.4).

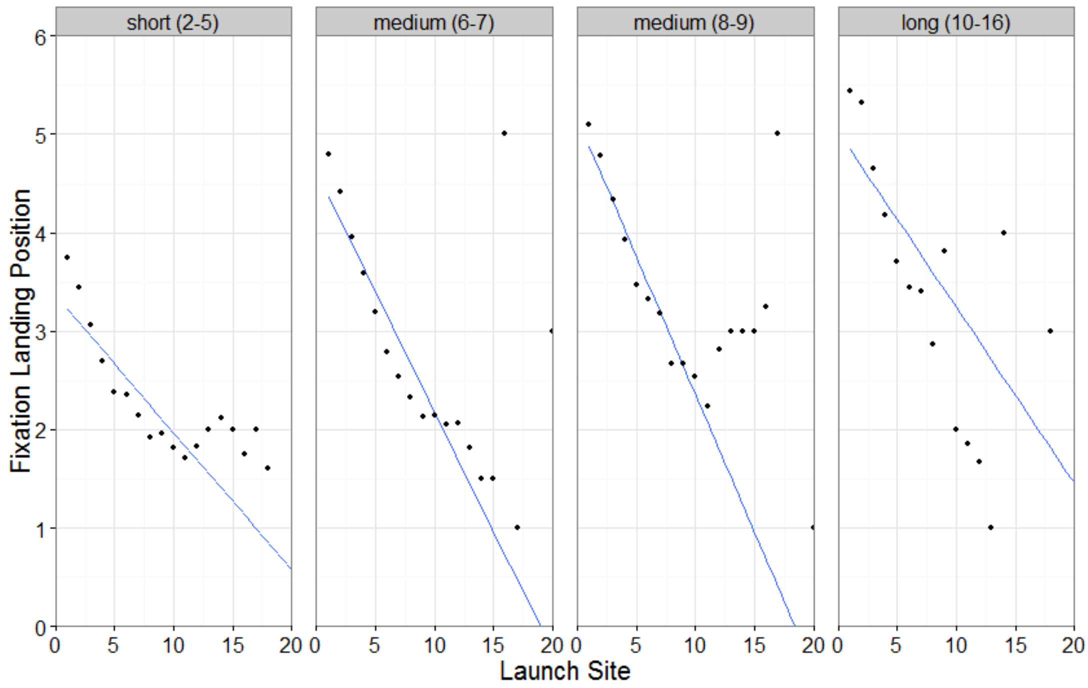

Figure 4.3. Effects of word length and launch site on FLP. Word lengths are grouped into four categories (short words: 2-5 letters; medium words: 6-7 letters; medium words: 6-7 letters; long words: 10-16 letters). All variables were specified as continuous in LMM estimates. Points are observed means.

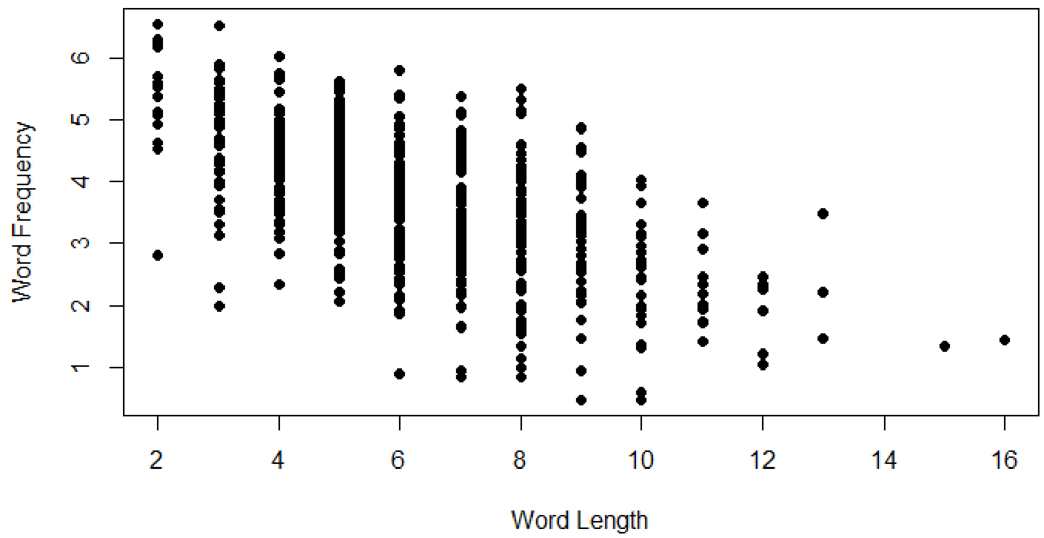

Figure 4.4. The relationship between word frequency and word length

LMM estimates of random effects showed that there is variability in FLP due to word types, sentences and subjects, as well as subject related experimental main effects of word length and word frequency (see Appendix B, Table B.1).

To sum up, word length, launch site and the interaction of these two effects were observed significantly; initial landing positions increased with increasing word length and decreasing launch sites. Word frequency positively affected FLP; however, this observation could be the result of relation between word length and word frequency (see Fig. 4.4).

4.1.3. The effect of word predictability on first fixation landing position

Predictability score was also added to the model as an additional covariate and checked whether it has a significant effect on first fixation landing position (FLP).

- It was found that there was a significant main effect of predictability on FLP (b=.235, SE=.117, t=2.00).

- The effects of word length (b=1.126, SE=.067, t=16.68), launch site (b=-.224, SE=.004, t=-51.27).

- Word frequency (b=.043, SE=.024, t=1.79) became insignificant.

The results showed that word predictability had an effect on initial landing positions. The other main effects, except word frequency, were conserved although predictability values were added to the model.

4.2. Findings for First Fixation Duration

In the previous section, the results of the models for first fixation landing position were presented. In this section, the results of the effects of word length, launch site, and word frequency on first fixation duration will be presented. Then, the effects of refixation rate and predictability scores on first fixation duration will be discussed in turn.

4.2.1. Effects of word length, launch site, word frequency, and first fixation landing position on first fixation duration

In the first fixation duration model, first fixation landing position was used as an additional covariate besides word length, launch site, and word frequency, since FLP is followed by FFD (Yan et al., 2014).

The findings are as follows:

- First fixations were longer when they landed close to the word center (b=.03, SE=.002, t=16.05; Fig. 4.5), which is consistent with the inverted optimal viewing position (IOVP) effect which has been reported in many other reading studies (e.g. Hyöna & Bertram, 2011; Nuthmann et al., 2005; Vitu, McConkie, Kerr, & O'Regan, 2001; as cited in Yan et al.).

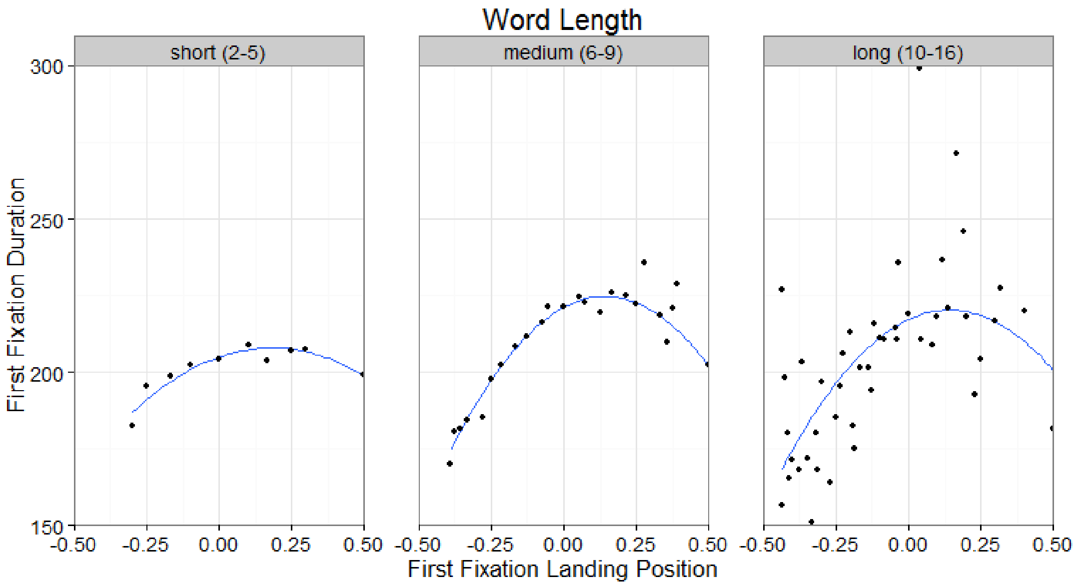

Figure 4.5. Effect of FLP on FFD with respect to word length. Points are observed means. Note that 0.0 represents word center on the FLP axis.

- First fixation durations were longer on less frequent words (b=-.026, SE=.004, t=-5.82).

- A positive significant relation was found between first fixation durations and launch sites (b=.009, SE=.001, t=10.29). This means when launch site increased, first fixation durations also increased.

- The effect of word length was found significant where first fixation duration was higher for short words; and lower for long words (b=-0.019, SE=0.009, t=-2.02) which may be explained by high refixation rates for long words (see Fig. 4.6).

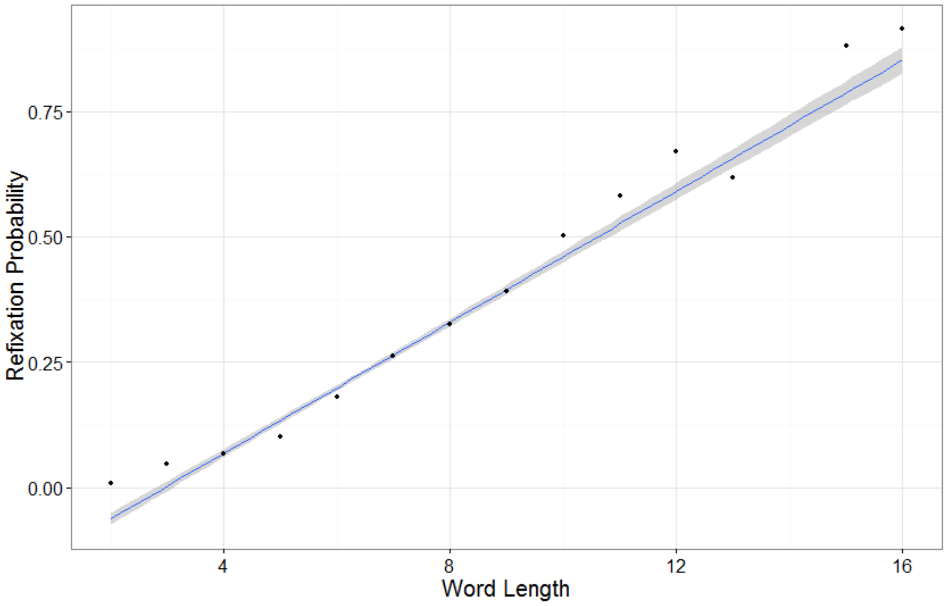

Figure 4.6. The relation between word length and refixation probability

In summary, the inverted optimal viewing position effect was observed clearly; fixation durations were higher around the word centers than the word beginnings or ends. Also, significant effects of word length, launch site and word frequency were found; durations increased with increasing word length and launch site, but decreased with word frequency. In the previous section, the relation between word length and word frequency was mentioned (see Fig. 4.4); while word length was increasing, word frequency decreased. It is seen than FFD results are consistent with this relation.

4.2.2. The effect of word predictability on first fixation duration

Predictability score was also added to the model as an additional covariate. Comparing to the model results found in section 4.2.1, the current results are given below:

- It was found that word predictability has a significant effect on FFD (b=-.084, SE=.022, t=-3.81) where FFD decreased with increasing predictability scores which is a logical outcome that shows readers eyes spend less time with fixating on more predictable words.

- The effect of first fixation landing position (b=.03, SE=.002, t=16.37), word length (b=-.018, SE=.009, t=-1.92), launch site (b=.009, SE=.001, t=10.64) and word frequency (b=-.022, SE=.004, t=-5.00) remained significant.

In summary, it was observed that predictability affected significantly first fixation durations. Addition of predictability scores to the model did not change the effects of first fixation landing position, word length, launch site and word frequency.

4.3. Findings for Gaze Duration

In this section, similar to the previous section, the effects of word length, launch site, and word frequency on gaze duration will be discussed first. Recall that gaze duration is the sum of all fixation durations on a word during first-pass reading. After that, the model with predictability scores in turn, will be discussed.

4.3.1. Effects of word length, launch site, and word frequency on gaze duration

In the model fit for gaze duration, similar to first fixation duration model, first fixation landing position is used as an additional covariate besides word length, launch site, and word frequency (Yan et al., 2014).

- It was found that gaze durations were longer when first fixations landed close to the word beginning (b=-.025, SE=.002, t=-11.62; see Fig. 4.7).

- Gaze durations were longer on less frequent words (b=-.041, SE=.006, t=-7.26).

- Gaze durations increased with increasing launch sites (b=.011, SE=.001, t=9.83) and word length (b=.211, SE=.018, t=11.48).

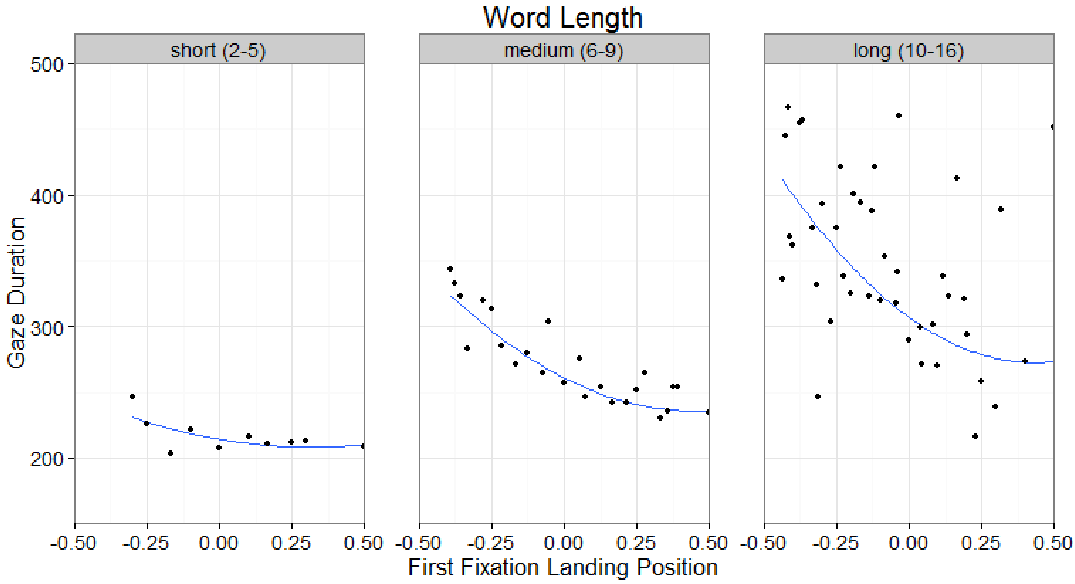

Figure 4.7. Effect of FLP on GD with respect to word length. Note that 0.0 represents word center on the FLP axis.

In Figure 8, it is seen that gaze duration was longer for medium and long words. In general, gaze duration increased when first fixation landing position was around the beginning of the word, and decreased around word centers and ends.

4.3.2. The effect of word predictability on gaze duration

The effect of word predictability on gaze durations was also checked. Here is the results which are given in relation to the findings discussed in section 4.3.1:

- The effect of the predictability of the words on gaze duration was found significant (b=-.123, SE=.028, t=-4.43); gaze durations decreased with increasing word predictability.

- The effect of word frequency was again found significant (b=-.036, SE=.006, t=-5.74).

- The effects of launch site (b=.011, SE=.001, t=10.24) and first fixation landing position (b=-.024, SE=.002, t=-11.23) remained significant.

- Finally, in agreement with the previous model results (cf. section 4.3.1), word length effect remained significant (b=0.213, SE=0.017, t=12.80).

To sum up, it was observed that predictability affected gaze durations significantly; the more predictable the word, the less time eyes spent on it. All the main effects used in the previous model remained significant.

4.4. Findings for Refixation Rate

The effects of word length, launch site, and word frequency as well as predictability scores on refixation probability were analyzed with generalized linear mixed models (GLMMs).

The GLMM results showed that:

- The effect of word frequency was significant where refixation rate increased with decreasing word frequency (b=-.152, SE=.028, t=-5.37). In Figure 4.8, data distribution across word frequency and refixation probability has been visualized.

- Refixation rate increased with increasing word length (b=1.929, SE=.067, t=29.00) as expected and mentioned above (see Fig. 4.6).

- Refixation rate increased with increasing launch sites (b=.172, SE=.011, t=15.38).

- Finally, it was found that refixation rate is higher for the words with low predictability scores (b=-.49, SE=.149, t=-3.29).

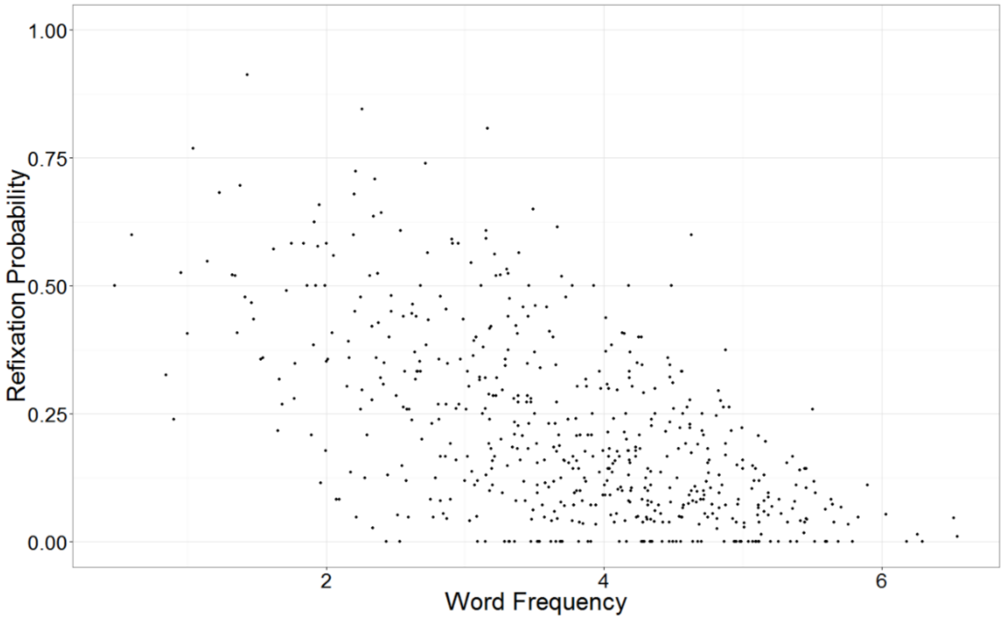

Figure 4.8. The relation between word frequency and refixation probability

4.5. Findings for Skipping Rate

The effects of word length, word frequency and word predictability were considered. Note that launch site values are not accessible for skipped words, so launch site was not added to the model. Similar to refixation rate analysis, a generalized linear mixed model (GLMM) analysis was performed. The results can be summarized as follows:

- Word length influenced skipping significantly (b=-2.94, SE=.10, t=-28.19); skipping rate increased with the decrease of word length.
- Word frequency effect was found significant (b=.33, SE=.06, t=6.03); frequent words tend to be skipped more.
- Finally, it was found that skipping rates increased with increasing predictability scores (b=.49, SE=.23, t=2.11).

4.6. Summary of the Findings

In this chapter, the findings about the eye movement parameters have been presented in detail. These eye movement parameters were first fixation landing position, first fixation duration, gaze duration, refixation probability, and skipping probability. These parameters were analyzed by the effects of word length, launch site, word frequency, and word predictability. All these findings have been summarized in Table 4.3.

Table 4.2. The summary of the findings presented with respect to the positivity or negativity of the effects.

	Eye Movement Parameters				
Fixed Effects	FLP	FFD	GD	RP	SP
Word Length	Positive	Negative	Positive	Positive	Negative
Launch Site	Negative	Positive	Positive	Positive	-
Word Frequency	No effect[5]	Negative	Negative	Negative	Positive
First Fixation Landing Position	-	Positive	Negative	-	-
Word Predictability	Positive	Negative	Negative	Negative	Positive

[5] This effect was significant and positively affected FLP without the effect of predictability.

CHAPTER 5

5. DISCUSSION & CONCLUSION

In this thesis, eye movement parameters during Turkish reading have been analyzed in terms of the characteristics of the text (in particular words). These eye movement parameters were first fixation landing position, first fixation duration, gaze duration, refixation probability and skipping probability. The text characteristics were specified by the word length, launch site, word frequency and word predictability. These parameters were determined based on the previous research and the findings in the literature. In order to achieve a general picture of Turkish reading patterns, a corpus-analytic approach was employed. Hence, the readers' behaviors were observed in a natural environment since the sentences were selected on a random basis and except for some critical eliminations, not a single change have been employed to the sentences. Furthermore, linear mixed model analyses were conducted. These analyses helped to control the subject, sentence and word type related variances for the eye movement parameters.

To begin with, the results for the first fixation landing position were in a good agreement with the previous studies (Yan et al., 2014; Nuthmann & Kliegl, 2009; Deutsch & Rayner, 1999; Rayner et al., 1996; Rayner, 1979; see Table 1). Mean landing positions reflected a logarithmic increase with respect to word length (see Fig. 4.1). It was observed that fixation landing position was around word center. It increased with word length. However, considering the relative position of the first fixation with respect to the word center, it was observed that the fixations tend to move towards the word beginning. The reason for this pattern could be the parafoveal preview effects: The longer the word, the more refixations are likely to be conducted. As a result, readers' eyes may tend to land towards the word beginning when word length increases.

In addition to mean landing positions, preferred viewing locations were considered. Furthermore, the fixations were close to the word beginning when launch site increased. This was an expected result which was one of the hypotheses of this study. Word frequency effect was also found significant; first fixation landed towards word beginning for low frequency words. The reason for this pattern could be that surface (i.e. whole-word) frequency values were used in this study. This means frequency values decreased when the word length increased (see Fig. 4.4). Hence, this pattern could be simply the result of word length effect rather than word frequency.

The last effect considered in this study was the predictability effect. This was added as an additional covariate. The results showed that predictability has an effect on the first fixation landing positions; fixations landed further into the words with increasing predictability scores. Furthermore, in each of the FLP models, the largest effect was found for word length. This result was in line with the generally accepted assumption that where to move the eyes are mainly based on word length (reviewed in Rayner, 2009; as cited by Yan et al., 2014).

Findings for first fixation duration yielded that first fixations were longer when they landed close to the word center, which is the so-called inverted optimal viewing position effect that has been reported in many other reading studies (e.g. Hyöna & Bertram, 2011; Nuthmann et al., 2005; Vitu, McConkie, Kerr, & O'Regan, 2001; as cited in Yan et al.; see Fig. 4.5).

As hypothesized in the beginning of this study, first fixations remained longer on less frequent words. Also, first fixation durations were longer with increasing launch sites. The duration was longer on short words. The reason for this could be the high refixation rates for long words (see Fig. 4.6). Lastly, it was observed that first fixation durations decreased with increasing word predictability. This means readers spent less time on more predictable words. In the latter model, all other effects remained the same.

The other eye movement parameter considered in this study was gaze durations. Recall that gaze durations were the sum of all fixation durations during first-pass reading. The results showed that gaze durations were longer when the reader’s eyes landed close to the word beginning. Durations were also longer on long words. These results could be attributed to the increasing refixation rates for long words (see Fig. 4.6). Similar to the first fixation durations, gaze durations were also found longer on less frequent words. Finally, word predictability was added as an additional covariate. It was found that gaze durations decreased when the predictability of the word increased. All the other effects including word length conserved their effects.

Finally, refixation probability and skipping probability analyses yielded consistent and expected results.

To sum up, this thesis showed that eye movement parameters are affected by certain properties of words during reading in Turkish. In general, the results were in a good agreement with the previous findings in the literature. Both preferred viewing locations and inverted optimal viewing positions effects have been observed in this study. Refixation and skipping rate analyses yielded expected results. The effects of word length, launch site, word frequency and word predictability have been observed on the eye movement parameters.

There were also certain limitations of the study that need to be considered. The first limitation relates to the number of participants. Although it is not known how the increased number of participants will affect the eye movement analyses, this limitation should be addressed in the future studies. For the predictability experiment, it could be better to have more participants. Recall that overall seventy students participated in the predictability experiment. However, the sentences had to be broken into seven parts and each of them was applied to one person. This ended up with ten predictability tests for each word. Since this study was conducted in a restricted time interval and participant number decreased in time, it was decided to stop at the number of seventy participants. Predictability scores could have been more reliable if there were more participants for each condition.

Another limitation could be due to the corpus-analytic approach. This approach was suitable for the goal of the current study. It allowed to get a general picture of Turkish reading patterns. Although it is very useful in many aspects, especially providing a natural environment, it is limited with the words it contains and word properties were not manipulated in this study.

Besides these limitations, it should be noted that the fitted mixed models were decided by *fitLMER.fnc* function (cf. *LMERConvenienceFunctions*; Tremblay & Ransijn, 2015). This function automatically checked all the fixed and random effects by applying back-fitting and forward-fitting algorithms respectively. In this study, this was helpful since there were too many parameters and fixed and random effects. However, it could be better to check them manually in the light of the expectations and data distribution.

In conclusion, this study provided a general picture of eye movement patterns during Turkish reading. Additional analyses could be carried out by further experimental designs. It should be noted that the effect of morphological complexity (i.e. number of suffixes) was beyond the course of this study. In order to observe the effect of morphological complexity, a target-word approach should be considered (e.g. applying orthographic manipulations to the target word by varying the number of suffixes). The effect of suffixes should be considered with respect to their types (e.g. inflectional or derivational, etc.). In this study, there were various suffix types (e.g.

inflectional, derivational, plural, tense, etc.). However, there were not enough number of suffixed words for each type of suffixes; so that classifying the suffixed words by their suffix types would not yield reliable results. In order to observe the suffix effect, they should be controlled in the experiment. Moreover, bigram frequencies of the morphemes could be considered. In addition to these approaches, the predictability test (i.e. the cloze task) and word frequencies should be considered at the suffix level in future studies.

REFERENCES

Bates, D. (2014). Computational methods for mixed models.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Rcpp, L. (2015). Package ‘lme4’.

Bıçakçı, B. (2014). *Baharda yine geliriz* (8. Baskı. ed.). İstanbul: İletişim.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719-722.

Deubel, H., & Bridgeman, B. (1995). Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research, 35*(4), 529-538.

Deubel, H., & Bridgeman, B. (1995). Perceptual consequences of ocular lens overshoot during saccadic eye movements. *Vision Research, 35*(20), 2897-2902.

Deutsch, A., & Rayner, K. (1999). Initial Fixation Location Effects in Reading Hebrew Words. *Language and Cognitive Processes*, *14*(4), 393-421.

Engbert, R., & Krugel, A. (2010). Readers Use Bayesian Estimation for Eye Movement Control. *Psychological Science*, *21*(3), 366-371.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review, 112*(4), 777-813.

Eye Movements and Sensory Motor Integration. (2004). In D. Purves, G. Augustine, D. Fitzpatrick, W. Hall, A. LaMantia, J. McNamara, et al. (Eds.), *Neuroscience* (3rd ed., pp. 457-458). Sunderland, Masachusets: Sinauer Associates.

Findlay, J., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences Behav. Brain Sci., 22*(04), 661-674.

Hayhoe, M. (2004). Advances in Relating Eye Movements and Cognition. *Infancy, 6*(2), 267-274.

Henderson, J. (2012). Introduction to “Computational Approaches to Reading and Scene Perception”. *Visual Cognition, 20*(4-5), 357-359.

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

Hyönä, J., & Bertram, R. (2011). Optimal viewing position effects in reading Finnish. *Vision Research, 51*(11), 1279-1287.

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

Just, M., & Carpenter, P. (1987). *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.

Kliegl, R., & Olson, R. (1981). Reduction and calibration of eye monitor data. *Behavior Research Methods & Instrumentation, 13*, 107-111.

Kliegl, R., & Engbert, R. (2013). Evaluating a Computational Model of Eye-Movement Control in Reading. *Models, Simulations, and the Reduction of Complexity*.

Li, X., Liu, P., & Rayner, K. (2011). Eye movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research, 51*(10), 1146-1156.

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

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

Mcconkie, G., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*(6), 578-586.

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

Mcconkie, G. (1997). Eye Movement Contingent Display Control: Personal Reflections and Comments. *Scientific Studies of Reading, 1*(4), 303-316.

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

Nilsson, M. (2012). Computational Models of Eye Movements in Reading: A Data-Driven Approach to the Eye-Mind Link. Retrieved August 15, 2015, from http://uu.diva-portal.org/smash/get/diva2:484385/FULLTEXT01.pdf

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45*(17), 2201-2217.

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

Nuthmann, A., & Kliegl, R. (2009). Preferred viewing locations: A validation and an extension. *Perception, 38*(6), 901-902.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading.

Paulson, E., & Goodman, K. (1999). Influential Studies in Eye-Movement Research. Retrieved August 12, 2015, from http://readingonline.org/research/eyemove.html

Pillalamarri, R., Barnette, B., Birkmire, D., & Karsh, R. (1993). Cluster: A program for the identification of eye-fixation-cluster characteristics. *Behavior Research Methods, Instruments, & Computers, 25*(1), 9-15.

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

Rayner, K. (1978). Eye movement latencies for parafoveally presented words. *Bulletin of the Psychonomic Society Bull. Psychon. Soc., 11*(1), 13-16.

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

Rayner, K., & Pollatsek, A. (1996). Reading unspaced text is not easy: Comments on the implications of Epelboim et al.'s (1994) study for models of eye movement control in reading. *Vision Research, 36*(3), 461-465.

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

Rayner, K., Fischer, M., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*(8), 1129-1144.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506.

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

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

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr., C. (2012). Introduction and Preliminary Information. In *Psychology of reading* (2nd ed., p. 20). New York, N.Y.: Psychology Press.

Reder, S. (1973). On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behav. Res. Meth. Instru. Behavior Research Methods & Instrumentation, 5*(2), 218-228.

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

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

Reichle, E., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences Behav. Brain Sci., 26*(04), 445-476. http://dx.doi.org/10.1017/S0140525X03000104

Reichle, E., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences Behav. Brain Sci., 26*(04), 445-476.

Reichle, E. (2006). Computational models of eye-movement control during reading: Theories of the “eye–mind” link. *Cognitive Systems Research, 7*(1), 2-3.

Richter, E., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7*(1), 23-33.

Sak, H., Güngör, T., & Saraçlar, M. (2008). Turkish language resources: Morphological parser, morphological disambiguator and web corpus. In*Advances in natural language processing* (pp. 417-427). Springer Berlin Heidelberg.

Say, B., Zeyrek, D., Oflazer, K., & Özge, U. (2002, August). Development of a corpus and a treebank for present-day written Turkish. In *Proceedings of the eleventh international conference of Turkish linguistics* (pp. 183-192).

Schad, D., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*(4-5), 391-421.

Scinto, L., & Barnette, B. (1986). An algorithm for determining clusters, pairs or singletons in eye-movement scan-path records. *Behavior Research Methods, Instruments, & Computers, 18*(1), 41-44.

Sereno, S., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*(11), 489-493.

Taylor, W. L. (1953). "Cloze procedure": a new tool for measuring readability. *Journalism quarterly*, 30, 415-433.

Team, R. C. (2015). R: A language and environment for statistical computing. Vienna, Austria; 2014. *URL http://www. R-project. org*.

Thibadeau, R., Just, M., & Carpenter, P. (1982). A model of the time course and content of reading. *Cognitive Science, 6*(2), 157-203.

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

Tremblay, A., & Ransijn, J. (2015). Package ‘LMERConvenienceFunctions’.

Vanrullen, R., & Thorpe, S. (2001). The Time Course of Visual Processing: From Early Perception to Decision-Making. *Journal of Cognitive Neuroscience, 13*(4), 454-461.

Vitu, F., Mcconkie, G., Kerr, P., & O'regan, J. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41*(25), 3513-3533.

Wade, N. (2010). Pioneers of eye movement research. *I-Perception, 1*(2), 33-68. doi:10.1068/i0389

Wang, S. (2005). Corpus-based approaches and discourse analysis in relation to reduplication and repetition. *Journal of Pragmatics, 37*(4), 505-540.

Wotschack, C. (2009). *Eye movements in reading strategies how reading strategies modulate effects of distributed processing and oculomotor control*. Potsdam: Universitätsverlag.

Yan, M., Kliegl, R., Richter, E., Nuthmann, A., & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *The Quarterly Journal of Experimental Psychology, 63*(4), 705-725.

Yan, M., Zhou, W., Shu, H., Yusupu, R., Miao, D., Krügel, A., & Kliegl, R. (2014). Eye movements guided by morphological structure: Evidence from the Uighur language. *Cognition, 132*(2), 181-215.

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

APPENDICES

Appendix A

Table A.1. Number of fixations broken down by word length and morphological structure

	Number of Suffixes					
Word Length	0	1	2	3	4	5
2	454	0	0	0	0	0
3	1632	61	0	0	0	0
4	1187	520	0	0	0	0
5	2048	974	19	0	0	0
6	595	1173	522	0	0	0
7	290	1438	716	0	0	0
8	283	544	854	107	0	0
9	19	150	745	157	0	0
10	23	49	359	184	0	0
11	0	26	146	98	0	0
12	0	0	48	118	0	0
13	0	0	0	40	0	20
15	0	0	0	0	25	0
16	0	0	0	0	23	0

Table A.2. Means (standard deviations) for eye movement measures by word length and morphological structure

WL	N	FLP	FFD	GD	RP	SP
2	15	1.5 (0.5)	195 (71)	196 (73)	1 (9)	77 (42)
3	46	2.1 (0.8)	200 (66)	207 (75)	5 (21)	54 (50)
4	66	2.7 (1.1)	203 (65)	216 (83)	7 (25)	30 (46)
5	118	3.2 (1.3)	202 (65)	219 (85)	10 (30)	12 (33)
6	90	3.5 (1.5)	214 (69)	246 (99)	18 (38)	5 (21)
7	94	3.6 (1.6)	212 (69)	262 (114)	26 (44)	2 (13)
8	68	3.9 (1.7)	212 (66)	276 (120)	33 (47)	1 (7)
9	43	4.1 (1.7)	213 (68)	293 (131)	39 (49)	1 (9)
10	25	4.2 (1.9)	209 (65)	314 (142)	50 (50)	0 (4)
11	12	4.2 (2.0)	201 (62)	322 (139)	58 (49)	0 (0)
12	7	4.3 (2.0)	201 (58)	368 (163)	67 (47)	1 (11)
13	3	5.1 (2.1)	213 (69)	378 (154)	62 (49)	0 (0)
15	1	4.0 (1.8)	211 (44)	478 (169)	88 (33)	0 (0)
16	1	4.4 (1.7)	182 (43)	403 (146)	91 (29)	0 (0)

Note. WL: word length; N: number of word types; FLP: first fixation landing position; FFD: first fixation duration; GD: gaze duration; RP: refixation probability; SP: skipping probability. The values for FLP, FFD, GD and RP were the means except skipped words.

Appendix B

Linear mixed model estimates for first fixation landing position, first fixation duration, gaze duration, refixation probability and skipping probability

Table B.1. Linear mixed model estimates of first fixation landing position (FLP)

	Estimate	SE	t-Value
Fixed Effects			
Grand Mean (GM)	3.317	0.063	53.07
Word Length (WL)	1.126	0.067	16.73
Launch Site (LS)	-0.222	0.004	-50.86
Word Frequency (WF)	0.051	0.023	2.16
WL x LS	-0.068	0.009	-7.21
WL x WF	0.070	0.031	2.27
LS x WF	0.016	0.004	4.03
WL x LS x WF	-0.033	0.004	-8.26

	Variance	SD	Correlation Parameters		
Variance Components					
Word - GM	0.171	0.41			
Sentences - GM	0.007	0.08			
Subjects - GM	0.113	0.34			
Subjects - WL	0.078	0.28	0.89		
Subjects - WF	0.001	0.03	-0.78	-0.98	
Subjects - WLxWF	0.005	0.07	0.26	-0.21	0.40
Residual	1.480	1.22			

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency. Log Likelihood: -25752.9; REML Deviance: 51505.8; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34.

Table B.2. Linear mixed model estimates of first fixation landing position (FLP) with predictability added as an additional covariate

	Estimate	SE	t-Value		
Fixed Effects					
Grand Mean (GM)	3.316	0.062	53.07		
Word Length (WL)	1.126	0.067	16.68		
Launch Site (LS)	-0.224	0.004	-51.27		
Word Frequency (WF)	0.043	0.024	1.79		
Word Predictability (WP)	0.235	0.117	2.00		
WL x LS	-0.068	0.010	-6.99		
WL x WF	0.064	0.032	2.03		
LS x WF	0.023	0.004	5.61		
WL x WP	0.114	0.250	0.46		
LS x WP	-0.154	0.023	-6.80		
WL x LS x WF	-0.029	0.004	-7.14		
WL x LS x WP	-0.167	0.045	-3.72		
	Variance	**SD**	**Correlation Parameters**		
Variance Components					
Word - GM	0.171	0.41			
Sentences - GM	0.007	0.08			
Subjects - GM	0.113	0.34			
Subjects - WL	0.078	0.28	0.88		
Subjects - WF	0.001	0.03	-0.71	-0.87	
Subjects - WLxWF	0.005	0.07	0.27	-0.21	0.38
Residual	1.474	1.21			

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency; WP: word predictability. Log Likelihood: -25723.6; REML Deviance: 51447.3; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34.

Table B.3. Linear mixed model estimates of first fixation duration (FFD)

	Estimate	SE	t-Value
Fixed Effects			
Grand Mean (GM)	5.295	0.020	259.67
Word Length (WL)	-0.019	0.009	-2.02
Launch Site (LS)	0.009	0.001	10.29
Word Frequency (WF)	-0.026	0.004	-5.82
First Fixation Landing Position (FLP)	0.030	0.002	16.05
WL x WF	0.025	0.006	4.52
WL x FLP	0.013	0.004	3.28
LS x FLP	0.002	0.001	4.27
WF x FLP	-0.006	0.002	-3.90

	Variance	SD
Variance Components		
Word - GM	0.005	0.07
Sentences - GM	0.001	0.03
Subjects - GM	0.013	0.11
Residual	0.069	0.26

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency. Log Likelihood: -1750.4; REML Deviance: 3500.9; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34.

Table B.4. Linear mixed model estimates of first fixation duration (FFD) with predictability added as an additional covariate

	Estimate	SE	t-Value
Fixed Effects			
Grand Mean (GM)	5.296	0.020	259.82
Word Length (WL)	-0.018	0.009	-1.92
Launch Site (LS)	0.009	0.001	10.64
Word Frequency (WF)	-0.022	0.004	-5.00
First Fixation Landing Position (FLP)	0.030	0.002	16.37
Word Predictability (WP)	-0.084	0.022	-3.81
WL x WF	0.025	0.005	4.58
WL x FLP	0.013	0.004	3.26
LS x FLP	0.002	0.001	4.09
WF x FLP	-0.006	0.002	-3.51
LS x WP	0.017	0.005	3.70

	Variance	SD
Variance Components		
Word - GM	0.005	0.07
Sentences - GM	0.001	0.03
Subjects - GM	0.013	0.11
Residual	0.069	0.26

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency; WP: word predictability. Log Likelihood: -1736.2; REML Deviance: 3472.3; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34.

Table B.5. Linear mixed model estimates of gaze duration (GD)

	Estimate	SE	t-Value
Fixed Effects			
Grand Mean (GM)	5.413	0.026	209.38
Word Length (WL)	0.211	0.018	11.48
Launch Site (LS)	0.011	0.001	9.83
Word Frequency (WF)	-0.041	0.006	-7.26
First Fixation Landing Position (FLP)	-0.025	0.002	-11.62
WL x WF	-0.055	0.007	-7.96
LS x FLP	0.003	0.001	5.23
WF x FLP	0.010	0.002	6.16

	Variance	SD	Correlation Parameters
Variance Components			
Word - GM	0.009	0.09	
Sentences - GM	0.001	0.04	
Subjects - GM	0.021	0.15	
Subjects - WL	0.007	0.08	0.51
Residual	0.104	0.32	

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency. Log Likelihood: -4972.6; REML Deviance: 9945.2; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34.

Table B.6. Linear mixed model estimates of gaze duration (GD) with predictability added as an additional covariate

	Estimate	SE	t-Value	
Fixed Effects				
Grand Mean (GM)	5.413	0.026	209.63	
Word Length (WL)	0.213	0.017	12.80	
Launch Site (LS)	0.011	0.001	10.24	
Word Frequency (WF)	-0.036	0.006	-5.74	
First Fixation Landing Position (FLP)	-0.024	0.002	-11.23	
Word Predictability (WP)	-0.123	0.028	-4.43	
WL x WF	-0.056	0.007	-8.15	
LS x FLP	0.003	0.001	4.98	
WF x FLP	0.011	0.002	6.69	
LS x WP	0.024	0.006	4.12	
	Variance	SD	Correlation Parameters	
Variance Components				
Word - GM	0.008	0.09		
Sentences - GM	0.001	0.04		
Subjects - GM	0.021	0.15		
Subjects - WL	0.005	0.07	0.56	
Subjects - WF	0.000	0.01	-0.15	-0.43
Residual	0.104	0.32		

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency; WP: word predictability. Log Likelihood: -4945.5; REML Deviance: 9891.0; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34.

Table B.7. Linear mixed model estimates of refixation probability (RP)

	Estimate	SE	z-Value	Pr(>|z|)	
Fixed Effects					
Grand Mean (GM)	-1.937	0.154	-12.60	< 0.000	***
Word Length (WL)	1.929	0.067	29.00	< 0.000	***
Launch Site (LS)	0.172	0.011	15.38	< 0.000	***
Word Frequency (WF)	-0.152	0.028	-5.37	0.000	***
Word Predictability (WP)	-0.490	0.149	-3.29	0.001	**
WL x LS	0.045	0.025	1.77	0.077	.
WL x WF	-0.145	0.037	-3.88	0.000	***
LS x WF	-0.043	0.010	-4.17	0.000	***
LS x WP	0.129	0.055	2.33	0.020	*
WL x LS x WF	0.048	0.015	3.24	0.001	**

	Variance	SD
Variance Components		
Sentences - GM	0.045	0.21
Subjects - GM	0.754	0.87

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency; WP: word predictability. Log Likelihood: -6235.5; REML Deviance: 12471.0; N of observations: 15647, N of word type: 581; N of sentences: 120; N of subjects: 34. (Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)

Table B.8. Linear mixed model estimates of skipping probability (SP)

	Estimate	SE	z-Value	Pr(>|z|)	
Fixed Effects					
Grand Mean (GM)	-2.242	0.122	-18.37	< 0.000	***
Word Length (WL)	-2.940	0.104	-28.19	< 0.000	***
Word Frequency (WF)	0.329	0.055	6.03	0.000	***
Word Predictability (WP)	0.485	0.230	2.11	0.035	*
WL x WF	0.273	0.079	3.43	0.001	***

	Variance	SD
Variance Components		
Word - GM	0.3602	0.6002
Subjects - GM	0.4126	0.6423

Note. WL: log2 word length; LS: launch site; WF: log10 word frequency; WP: word predictability. Log Likelihood: -6646.2; REML Deviance: 13292.4; N of observations: 20526, N of word type: 581; N of subjects: 34. (Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)

Appendix C

Table C.1. Stimuli sentences and the questions with the related answers

Sentence ID	Sentence	Answers
1	Acının yoğurduğu ve olgunlaştırdığı bir hayat ve şiirdir onunki.	-
2	Adanın öbür köylerinden kasabaya gelip iş yeri açanlar oldu.	-
3	Ağabeyim de ben de aşağı yukarı o yaşta evlendik.	-
4	Ağaç gibi dikilip duruşum sonunda aklını başına getirdi anlaşılan.	-
5	Aileyi bir araya getiren en önemli günler dinî bayramlardı. Question: Aileyi bir araya getiren en önemli günler millî bayramlar mıdır?	No
6	Alt ön dişlerimden sağlam sandığım ikisi de sallanıp durmakta.	-
7	Altı yaşıma girdiğimde hızlı bir şekilde okuyup yazmaya başladım.	-
8	Ana cadde ve sokaklar araç ve yaya trafiğine kapatıldı. Question: Ana cadde ve sokaklar mı trafiğe kapatıldı?	Yes
9	Annemi ikna etmek daha uzun sürerse ne yapardım bilemiyorum.	-
10	Arabanın durmasıyla tipik bir tamirci çırağı arabanın önüne koştu.	-
11	Arabanın içi iyice havalansın diye park yerinde biraz dolaştı.	-
12	Arabayla yarım saat süren yolu eşek sırtında gitmişler saatlerce.	-
13	Arka koltuğa oturur oturmaz başını geriye yaslayıp gözlerini yumdu.	-
14	Arkadan gelen fillerin durmaya hiç mi hiç niyetleri yoktu. Question: Arkadan gelen hayvanlar fil midir?	Yes
15	Arkamı döner dönmez hayatımda gördüğüm en güzel gözleri gördüm.	-
16	Aşağıya indiğimizde ortalık her zamanki gibi soğuk ve sessizdi.	-
17	Babamın bu konuda benden yana olacağını adım gibi biliyordum.	-
18	Bağımsız bir uğraşısı olmadığı için amcam onu yanına almıştı.	-
19	Bahar kokusuyla birlikte kil ve çürümüşlük kokusu genzime doldu.	-

20	Bankanın kasa bölümüne girmek için iki kat aşağı indik. Question: Bankanın kasa bölümü iki kat yukarıda mıdır?	No
21	Belki de ilk kez bir eylemin örgütlenmesinden haberdar oluyordu.	-
22	Belli bir nezaket içinde bu mesajların topluma yansıması gerekir.	-
23	Ben her zaman senin saf bir çocuk olduğunu düşünmüşümdür.	-
24	Beş yıl sonra karısını ve iki çocuğunu yanına aldırmış. Question: Üç yıl sonra mı karısını ve iki çocuğunu yanına aldırmıştır?	No
25	Beyin ölümü teşekkül ettiği için yaşaması söz konusu değildir.	-
26	Bir dakika içinde biten uygulama için cerrahi işlem gerekmiyor.	-
27	Bir insan çalışır ve çalıştığının karşılığı neyse onu alır.	-
28	Biraz yürüyüp iki katlı eski bir evin önünde duruyoruz. Question: Önünde durduğumuz ev iki katlı mıdır?	Yes
29	Birbirimizden nasıl olmuş da kopmuş ve iki yabancıya dönüşmüştük.	-
30	Birden aklıma o zamana kadar düşünmediğim bir şey gelmişti.	-
31	Bizim şu anda bin düşünüp bir adım atmamız lazım.	-
32	Bu hafta borsada yatay ve aşağı bir piyasa bekliyorum.	-
33	Bu karışık dünyada senin yaşadığını bilmek beni mutlu ediyor.	-
34	Buğuya karışmış bin bir çeşit ot kokusu genzimi yakıyor.	-
35	Bundan sonra ne istediğimi iki kere düşünmeye karar verdim.	-
36	Burada evler sık sık iklim değişmesinden dolayı tamir görmüş.	-
37	Butikte işler ilerleyince bana da maaş vermeyi teklif etti.	-
38	Bütçesinin üçte birini eğitime ayıran yoksul ülkeler de var. Question: Bütçesinin beşte birini mi eğitime ayıran ülkeler var?	No
39	Büyüklerin bizi sinemadan ayırması bize atılan en büyük dayaktı.	-
40	Cep telefonu kapalı ya da kapsama alanı dışında olabilir.	-
41	Ceza almasa bile ortalığı birbirine katıp karıştırmaya devam etti.	-
42	Daha iki metre inmiştim ki tepemden sular boşalmaya başladı.	-

43	Dalgaların arasında suya girebilen bir karınca sürüsü vardı sanki.	-
44	Derinde duvar ya da başka türde bir sınır görünmüyordu.	-
45	Dev uçağın kaptan pilotunun kadın olması da dikkat çekti.	-
46	Durumu protesto eden bazı gazeteciler ise salonu terk etti.	-
47	Dürüst bir insanken otuz yaşından sonra yalan söylemeye başladım. Question: Dürüst bir insanken otuz yaşından sonra mı yalan söylemeye başladım?	Yes
48	Eğitim yardımı almak isteyen öğrenci sayısında da patlama oldu.	-
49	Ev işlerinin asıl etkisi depresif bir ruh hali yaratmasıydı.	-
50	Gazete dışında her türlü dergi ve kitap da satarlardı. Question: Gazete dışında dergi de satarlar mıydı?	Yes
51	Gazeteyi şöyle bir karıştırdı ve ölüm ilanları sayfasını açtı. Question: Gazeteyi karıştırdıktan sonra iş ilanları sayfasını mı açtı?	No
52	Grubun herhangi bir parti veya kuruluşla organik ilişkisi yoktu.	-
53	Hapishanede geçirdiğim bu üç yılın sonunda bir umut belirdi.	-
54	Hemen hemen her sokağın ortasında bir kanal yer alıyor.	-
55	Henüz iki saattir yoldaydık ve ben altı kez düşmüştüm. Question: İki kez mi düşmüştüm?	No
56	Her akşam yemekten sonra tüm okulu okuma salonuna toplardı.	-
57	Her bağımlının aynı zamanda bir satıcı olduğu gerçeğini bilin.	-
58	Her derin uykudan uyanışımda bir sonraki uykuyu merak ediyorum.	-
59	Her geçen gün daha çok insanlaştık biz bu okulda.	-
60	Her hafta perşembe günleri genç bir kadın ziyaretime gelir. Question: Her hafta perşembe günleri mi genç bir kadın ziyaretime gelir?	Yes
61	Her yaşam gibi yanılsama üzerine kurulmuş bir yaşamdı bu.	-
62	Her yıl binlerce öğrenci öğretmen olma sevdasıyla mezun oluyor.	-
63	İnce belli iki cam bardak masanın üstünde karşılıklı duruyordu. Question: Masada ince belli iki cam bardak mı duruyordu?	Yes
64	İşimi yaptıktan sonra paramı alacak ve her şeyi unutacaktım.	-
65	İşverenler elbette ki bu felakete karşı birtakım önlemler aldı.	-
66	Kabul etmek gerekir ki günümüz toplumu bağımlı bir	-

	toplum.	
67	Kadının titrek sesi beyaz bir çığlık gibi uzayıp gidiyor.	-
68	Kâğıt mendil kutusu da her zaman koltuğun yanında dururdu. Question: Kâğıt mendil kutusu her zaman masada mı dururdu?	No
69	Kahveleri içtikten sonra babam yorgun olduğunu söyledi ve gitti.	-
70	Kamu kuruluşlarına ait çok sayıda dernek de faaliyet gösteriyor.	-
71	Kazanın üzerinden nerede ise bir yılı aşkın zaman geçmişti.	-
72	Kederi arttıkça arttı ve akıttığı gözyaşları dillere destan oldu.	-
73	Kendisini rahatsız eden her şey bir anda önemini yitirmiştir.	-
74	Kente yabancı olanlar bizden daha iyi biliyorlar böyle yerleri.	-
75	Kısa bir şaşkınlığın ardından toparlanıp içeriye buyur ettik onu.	-
76	Kiminin orduda depocu ya da subay bir tanıdığı vardır.	-
77	Kimse bana bu kötü büyüyü bozacak sihirli sözcüğü fısıldayamadı.	-
78	Kişilik belli bir yaşa dek gelişmemişse artık hiç gelişmez.	-
79	Klasik müzik bizim kültürümüzü çok yansıtan bir şey değil.	-
80	Konuşmalardan anladım ki bugün prova her zamankinden uzun sürmüş.	-
81	Koşunca her şey değişirdi ve toprak renkten renge girerdi.	-
82	Kulakları o denli büyük değil ama gözleri iri iriydi.	-
83	Lâkin şimdi hiç kimsenin bilmediği bir sırdan bahsediyordum ben.	-
84	Mali piyasalar açısından bu iki gelişme son derece kritik.	-
85	Medyaya kapalı olan davetin konukları arasında ben de vardım.	-
86	Mesleğini bildiğine ve onu iyi bir şekilde uyguladığına inanırdı.	-
87	Meydanda kim var kim yoksa çil yavrusu gibi dağıldı.	-
88	Muhteşem bir dünya kültür mirasını gezdiğimiz için çok mutluyuz.	-
89	Müthiş soğuk bir gündü ve tipi şeklinde kar yağıyordu.	-
90	Ne zaman ne yapacağı belli olmayan bir kadınla birlikteyim.	-
91	Nice yoksulluğa hep bu aşk yüzünden katlandığı günlerini aradı.	-
92	Okul tatil döneminde olduğu için biraz daha beklemem gerekiyordu.	-

93	Onlara bir kol saati ve bir oyuncak hediye ettik. Question: Onlara kolye mi hediye ettik?	No
94	Onu ikna etmek için şartları çok iyi tutmak gerekiyordu.	-
95	Onun benim için belirgin ama tanımsız bir kokusu vardı.	-
96	Onunla birlikte ifade veren genç bir kız daha vardı. Question: Onunla birlikte ifade veren yaşlı bir adam mı vardı?	No
97	Otobüsün sağ arka lastikleri ile kaporta kısmı hasar gördü. Question: Otobüsün kaporta kısmı hasar gördü mü?	Yes
98	Özene bezene kurulan her eve eninde sonunda ölüm girer.	-
99	Park edecek uygun yer bulana dek arabayı geriye aldı.	-
100	Perdeyi ne zaman aralayıp baksa bomboş bir sokak görüyor.	-
101	Renklere ve şekerlere bulanmış tatlı bir düş değildi bu.	-
102	Saat sekize geliyor ve biz daha hediyemizi bile alamadık.	-
103	Sabah çok erken yola çıkacağımız için geceyi kısa kestik.	-
104	Sabahları fırından taze ve çıtır simit alıp çaya yetişirdim.	-
105	Saçları seninkinden çok daha koyu bir çamurlu su rengindeydi. Question: Saçları çamurlu su renginde miydi?	Yes
106	Sayısı oldukça fazla olan kadın müşterilerle bu kız ilgileniyordu.	-
107	Seçim takvimine göre yapılacak her işlemin bir sırası vardır.	-
108	Sokaklar sırf hava almak için bile çıkılacak gibi değil.	-
109	Sonuçta annesi galip gelmiş ve evlerine hiç gündelikçi girmemişti.	-
110	Suyun ve sileceklerin gücü onları camdan tümüyle kazımaya yetmiyor.	-
111	Temiz deniz kıyılarına sahip olmak ne güzel diye düşündüm.	-
112	Tıp tarihi açısından bu son derece önemli bir olaydır.	-
113	Toplumsal yaşam için çok önemli bir duyguyu bütünüyle yitirdik.	-
114	Tüm süreci bir tek kişinin tamamlaması nesnelliği yok eder.	-
115	Yabancı olan biri için gübre kokusu hemen yüze çarpar.	-
116	Yağmur yağınca bu incecik yol bir denizin dibine döner.	-
117	Yalnızca mektup yazıp okumak için okuma yazma öğrenmeye değer.	-
118	Yarının yeni bir gün olduğundan emin olarak dudaklarını kıvırıyor.	-
119	Yaşama sevinci belki de en çok bu mevsimin adıydı.	-
120	Yıllarca yaşamış bir geminin mutlaka anlatacak bir şeyleri vardır.	-

Table C.2. Paragraphs used in the experiments

1	Bu berbat şehirde görüp görebileceğiniz en güzel şeyin terk edilmiş bir fabrikanın kara yıkıntısı olması saçma ya da gülünç mü? Değil! İnsana özgü bir yavaşlığı, sakarlığı hatırlatan tek şey bu yıkıntı çünkü. Şehirde otomobiller, yollar ve binalar, sonunda bütün sıcaklıkların evrenin ölgün sıcaklığıyla aynı olacağı bir geleceğe doğru son hızla gidiyor, uzanıyor, yükseliyor. Ama aralarında banka memuru sevgili dostum Tuğrul'un da bulunduğu sağlığına dikkat etmeyen, fazlasıyla hayalperest bazı insanlar var ki, onlar gece kurdukları saatin sabah çalışmamasını veya en iyisi geriye gitmesini gönülden dileyerek tatlı tatlı esniyorlar.
2	Günümüzden yaklaşık bin altı yüz yıl önce, bir Roma imparatorunun şehrimizi ziyaret etmesi vesilesiyle dikilen sütunun üzerinde bugün bir leylek yuvası var. O sütunu görünce insan ister istemez bazı yapılara bin yıl sonra üzerine leyleklerin yuva yapacağı beton yığınları gözüyle bakıyor. Hangisine yuva yapacak acaba leylekler? Baştan aşağı camla kaplı cephesinde gökyüzünü ve güneşi soğuk dikdörtgenler halinde yansıtan şu gökdelene mi? Yoksa, dev bir fabrikayı andıran şu alışveriş merkezine mi? Böyle düşüncelere de fazla kapılmaya gelmez! En iyisi ucuz marketlerin birinden alışveriş yapmak ve kendini ikramiye veya yakacak yardımı verilen ayların düzgün salınımına bırakmak.
3	Şehrin yüksek binalarından birine çıkıp aşağıya bakıyorum, her şehirde rastlanabilecek bir manzarayla karşılaşıyorum: Yüzlerce insan, bazen birbirlerinin yolunu keserek oradan oraya gidip geliyor… Ölümsüz gibi görünüyorlar. "Nedir bu?" diye soruyorum kendi kendime, anlamlandırmak gerekiyor, "Kâbus mu, şenlik mi?" Arka arkaya bir sürü karşıt anlamlı sözcük geçiyor aklımdan. Eksilerle artıların birbirini götürmesi gibi kalabalığın da bir matematiği var. Sıradanlık bu olmalı: Bütün karşıtlar birbirini götürüyor. Başka ne söyleyebilirim ki size?

Appendix D

Bilgilendirilmiş Onay Katılım Formu
(Informed Consent Form)

Bu çalışma, ODTÜ Enformatik Enstitüsü Bilişsel Bilimler Programı bünyesinde yürütülmekte olan 113K723 No.'lu TÜBİTAK projesi kapsamında düzenlenmektedir. Çalışma yürütücüsü Yrd. Doç. Dr. Cengiz Acartürk, araştırmacı Doç. Dr. Bilal Kırkıcı, danışmanlar Prof. Dr. Deniz Zeyrek ve Prof. Dr. Cem Bozşahin, bursiyerler çalışmanın gerçekleştirildiği dönem itibariyle yüksek lisans öğrencileri Figen Beken, Mehmetcan Fal ve doktora öğrencisi Özge Nilay Yalçın'dır.

Çalışma, Türkçe okuma alışkanlıklarının incelenmesi amacı ile gerçekleştirilmektedir. Çalışma boyunca gösterilecek materyal genel olarak kişisel rahatsızlık verecek içeriğe sahip değildir. Sizden beklenen, cihaz ekranında gösterilen yönergeleri takip etmektir. Ancak, katılım sırasında gösterilen materyalden ya da herhangi başka bir nedenden ötürü kendinizi rahatsız hissederseniz çalışmayı yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda çalışmayı yürüten kişiye, çalışmayı tamamlamadığınızı söylemek yeterli olacaktır.

Çalışmaya katılım bilgilendirilmiş onay (informed consent) esasına dayanmaktadır. Çalışma boyunca, bu form dışında sizden kimlik belirleyici hiçbir bilgi istenmemektedir. Bu formdaki kimlik bilgileri verilerle eşleştirilmemektedir. Cevaplarınız tamamıyla gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir; elde edilecek bilgiler bilimsel yayınlarda kullanılacaktır. Çalışma sonunda, varsa çalışmayla ilgili sorularınız cevaplanacaktır. Çalışma tamamlandığında katılım ücretiniz çalışma yöneticisi tarafından ödenecektir.
Katılımınız için şimdiden teşekkür ederiz.

(Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

Bu çalışmaya bilgilendirilmiş olarak katılıyorum ve istediğim zaman yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayınlarda kullanılmasını kabul ediyorum.

İsim Soyad	Tarih	İmza
	----/----/-----	

Appendix E

Dilbilgisel Artalan Anketi

<table>
<tr><td colspan="2">Kişisel Bilgiler</td><td colspan="3">Kod:</td></tr>
<tr><td>Soyadı</td><td colspan="2">Adı</td><td colspan="2">Bugünün Tarihi</td></tr>
<tr><td>Doğum Tarihi</td><td colspan="2">Kadın ( )</td><td colspan="2">Erkek ( )</td></tr>
<tr><td colspan="2">Telefon Numarası</td><td colspan="3">E-posta Adresi</td></tr>
<tr><td colspan="5">Şu anki mesleğiniz?</td></tr>
<tr><td>En yüksek tahsiliniz (veya muadili)
(lütfen yuvarlağa alınız)</td><td>Ortaokul</td><td colspan="2">Lise</td><td>Üniversite Derecesi</td></tr>
<tr><td></td><td>Mesleki Eğitim</td><td colspan="3">Diğer?</td></tr>
<tr><td colspan="5">Fakülteniz</td></tr>
<tr><td colspan="5">Bölümünüz</td></tr>
<tr><td>Sınıfınız</td><td colspan="4">Hazırlık ( ) 1. Sınıf ( ) 2. Sınıf ( )
3. Sınıf ( ) 4. Sınıf ( )</td></tr>
<tr><td>Lisede hazırlık okunuz mu?</td><td colspan="4">Evet ( ) Hayır ( )</td></tr>
<tr><td>Üniversitede hazırlık okudunuz mu?</td><td colspan="4">Evet ( ) Hayır ( )</td></tr>
</table>

<table>
<tr><td>Genel Sağlık Durumunuz</td><td></td><td></td><td></td></tr>
<tr><td>Yazarken hangi elinizi kullanıyorsunuz?</td><td>Sağ ( )</td><td colspan="2">Sol ( )</td></tr>
<tr><td>Tanısı konmuş herhangi bir dil</td><td>Hayır( )</td><td>Evet ( )</td><td>Varsa, lütfen ayrıntılandırınız.</td></tr>
</table>

bozukluğunuz var mı (disleksi, kekemelik gibi)?			
Çalışma sırasında gözlük kullandınız mı?	Hayır( )	Evet ( )	
Çalışma sırasında lens kullandınız mı?	Hayır( )	Evet ( )	

Hangi dil(ler)i, hangi sırayla öğrendiniz? (anadiliniz dahil)			
Dil	Hangi yaştan itibaren?	Ne kadar süreyle?	Öğrendiğiniz yer? (evde, okulda, başka) Lütfen belirtiniz.
1.			
2.			
3.			

Türkiye dışında başka ülkelerde yaşadınız mı?	Ne kadar süreyle?	Hangi sebeple? (okul, eğitim, vs.)
1.		
2.		
3.		

Dil kullanımınız
(Haftalık yüzde olarak)
Birinci sıraya lütfen günlük hayatınızda kullandığınız dilleri yazınız. Lütfen aşağıdaki tabloda yazılı olan kişilerle veya faaliyetler sırasında konuştuğunuz dillerin yaklaşık kullanım yüzdesini belirtiniz. Her sıradaki kullanım yüzdesinin toplamı %100 olmalıdır.

Aşağıda yazılı olan kişilerle hangi dilde	Dil ………	Dil …………	Dil …………	Dil …………	Dil …………

iletişim kurarsınız?	…..	…	…..	…..	…..
Eşinizle/partnerin izle					
Çocuklarınızla?					
Anne/babanızla?					
Akrabalarla?					
Arkadaşlarla?					
İşte/okulda?					
Hangi dilde TV izlersiniz?					
Hangi dilde müzik/radyo dinlersiniz?					
Hangi dilde gazete, kitap vs. okursunuz?					