

ANKARA YILDIRIM BEYAZIT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES



**DYSLEXIC READABILITY ANALYSIS OF TEXTS AND DIAGNOSIS
OF DYSLEXIA USING EYE MOVEMENTS**

M.Sc. Thesis by

Akın GÜNDÜZ

Department of Computer Engineering

May, 2019

ANKARA

DYSLEXIC READABILITY ANALYSIS OF TEXTS AND DIAGNOSIS OF DYSLEXIA USING EYE MOVEMENTS

A Thesis Submitted to

The Graduate School of Natural and Applied Sciences of

Ankara Yıldırım Beyazıt University

**In Partial Fulfillment of the Requirements for the Degree of Master of Science
in Computer Engineering, Department of Computer Engineering**

by

Akın GÜNDÜZ

May, 2019

ANKARA

M.Sc. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled “**DYSLEXIC READABILITY ANALYSIS OF TEXTS AND DIAGNOSIS OF DYSLEXIA USING EYE MOVEMENTS**” completed by **AKIN GÜNDÜZ** under the supervision of **ASSIST. PROF. DR. ÖZKAN KILIÇ** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Özkan KILIÇ

Supervisor

Assist. Prof. Dr. Gül TOKDEMİR

Jury Member

Assist. Prof. Dr. Hilal KAYA

Jury Member

Prof. Dr. Ergün ERASLAN

Director

Graduate School of Natural and Applied Sciences

ETHICAL DECLARATION

I hereby declare that, in this thesis which has been prepared in accordance with the Thesis Writing Manual of Graduate School of Natural and Applied Sciences,

- All data, information and documents are obtained in the framework of academic and ethical rules,
- All information, documents and assessments are presented in accordance with scientific ethics and morals,
- All the materials that have been utilized are fully cited and referenced,
- No change has been made on the utilized materials,
- All the works presented are original,

and in any contrary case of above statements, I accept to renounce all my legal rights.

Date: 29.05.2019

Signature:

Name & Surname: Akın GÜNDÜZ

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my supervisor, Assist. Prof. Dr. Özkan KILIÇ for his tremendous support and motivation during my study. His immense knowledge and precious recommendations constituted the milestones of this study. His guidance assisted me all the time of my research and while writing this thesis.

I must express my profound appreciations to my family for providing me emotional support, loving care and continuous encouragement throughout my years of study and through the period of writing this thesis. This accomplishment would not have been possible without them.

Finally, I would like to thank Mustafa Kemal Atatürk for proving that the cornerstone of education is really an easy system of reading and writing by making one of the most far-reaching language reforms in history.

2019, 29 May

Akın GÜNDÜZ

DYSLEXIC READABILITY ANALYSIS OF TEXTS AND DIAGNOSIS OF DYSLEXIA USING EYE MOVEMENTS

ABSTRACT

Although people with literacy training can easily and spontaneously read a piece of text, reading is a very complex activity. While letters and words are perceived by the visual system in the brain, cognitive faculties such as memory, language, reasoning, and such are also involved in reading activity. Literacy in the modern world is of utmost importance for managing one's daily life and contributing to the society as an educated person, as well as for individual and professional development. However, dyslexia, which is thought to affect around 10-15% of society, causes reading deficit for some, despite their normal intelligence. Albert Einstein, Picasso, Bill Gates, and Steven Spielberg are known to have suffered from dyslexia.

This thesis has made the readability of its texts measurable for a dyslexic reader with improved a novel readability index which named Gunduz Index. At the same time, the success of this index was measured as a result of the reading experiments.

By analysing the eye movements recorded during the reading experiments, it was estimated that the reader was a dyslexic or non-dyslexic reader with a success rate of up to 90%.

Another important point of the thesis is that this successful diagnosis on primary school students can be done before school. For this purpose, a game that testes the ability to establish the connection between geometric shapes and sounds has been created and the results are analysed. As a result, with the success rate of 70%, the person who played the game can be predicted to be dyslexic.

Keywords: Machine Learning, Natural Language Processing, Human Computer Interaction, Dyslexia, Readability, Eye Tracking

METİNLERİN DİSLEKTİK OKUNABİLİRLİK ANALİZİ VE GÖZ HAREKETLERİNİ KULLANARAK DİSLEKSİ TEŞHİSİ

ÖZ

Okuma, pek çok eğitimli insan tarafından hızlıca gerçekleştirilen bir etkinlik olmasının yanında oldukça karmaşıktır. Harfler ve kelimeler beynimizin görsel sistemi tarafından algılanırken hafıza, dil ve düşünme gibi diğer karmaşık bilişsel fakülteler de okuma etkinliğine katılır. Modern dünyada okuryazarlık, bireylerin yaşantılarını idare ettirmeleri, bulundukları topluma eğitimli bir birey olarak katkıda bulunmaları, kişisel ve mesleki gelişimleri için vazgeçilmezdir. Ancak, toplumun yaklaşık %10-%15'ini etkilediği düşünülen disleksi öğrenme bozukluğu, zekâ düzeyi normal olmasına rağmen bireylerin okuma güçlüğü çekmelerine neden olmaktadır. Albert Einstein, Picasso, Bill Gates, Steven Spielberg gibi tanınmış kişilerin bu probleminden şikâyetçi olduğu bilinmektedir.

Bu tezde, geliştirilen Gunduz İndeksi (dislektik okunabilirlik indeksi) sayesinde okuma metnlerinin dislektik bir okuyucu için uygunluğu ölçülebilir hale getirilmiştir. Aynı zamanda yapılan okuma deneyleriyle sonuçlar doğrulanmıştır.

Okuma deneyleri esnasında kaydedilen göz hareketleri analiz edilerek, okuyucunun dislektik bir okuyucu olup olmadığı 90% lara varan bir başarı oranı ile tahmin edilebilmiştir.

Tezin ulaştığı bir diğer önemli sonuç ise okuma deneyleri ile yapılan bu başarılı teşhisin, okul öncesinde de yapılabileceği bilgisidir. Bu amaçla geometrik şekiller ve sesler arasındaki bağlantıyı kurabilme becerisini test eden bir oyun oluşturulup alınan sonuçlar analiz edilmiştir. Sonucunda 72% başarı oranı ile oyunu oynayan kişinin dislektik bir okuyucu olup olmadığı tahmin edilebilmiştir.

Anahtar Kelimeler: Makine Öğrenmesi, Doğal Dil İşleme, İnsan Bilgisayar Etkileşimi, Disleksi, Okunabilirlik, Göz Takibi

CONTENTS

M.Sc. THESIS EXAMINATION RESULT FORM	ii
ETHICAL DECLARATION	iii
ACKNOWLEDGMENTS	iv
ABSTRACT	v
ÖZ	vi
CONTENTS	vii
NOMENCLATURE.....	ix
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER 1	1
INTRODUCTION.....	1
1.1 Aim of the Study.....	2
1.2 Review of Related Works	2
CHAPTER 2	15
BACKGROUND.....	15
2.1 Dyslexia.....	15
2.2 Reading Models.....	21
2.3 Readability.....	24
2.3.1 General Readability Rules	25
2.3.2 Language and Readability	26
2.3.3 Dyslexia and Readability	27
CHAPTER 3	29
METHODOLOGY	29
3.1 Coding Platforms.....	29
3.2 Method.....	29
3.2.1 Participants	29
3.2.1 Apparatus	29
3.2.3 Materials and Procedure	31
3.4.4 Experiments	43
CHAPTER 4	44

ANALYSE.....	44
4.1 Diagnosis of Dyslexia	44
4.1.1 Data.....	44
4.1.2 Preprocessing.....	45
4.1.3 Analysis	48
4.2 Dyslexic Readability Analysis.....	48
4.2.1 Algorithm.....	48
4.2.1 Data.....	53
4.2.1 Analysis	56
4.3 Shape Sound Connection.....	57
4.3.1 Data.....	57
4.3.2 Analysis	58
CHAPTER 5	60
RESULTS	60
5.1 Diagnosis Dyslexia	60
5.2 Dyslexic Readability Analysis	61
5.3 Shape Sound Connection.....	61
CHAPTER 6	64
CONCLUSION AND DISCUSSION	64
REFERENCES	68
APPENDICES	76
Appendix A - Sentences of the different difficulty levels	76
CURRICULUM VITAE	77

NOMENCLATURE

Acronyms

SVM	Support Vector Machine
TDK	Turkish Language Association
OS	Readability Score



LIST OF TABLES

Table 1.1 Overview of the studies	13
Table 3.1 Turkish alphabet according to the geometric shapes they contain	36
Table 3.2 Geometric shapes according to their origins	38
Table 3.3 Sentences of the different difficulty levels	42
Table 4.1 Dataset as a result of experiment.....	44
Table 4.2 The other dataset as a result of reading experiment	45
Table 4.3 Preprocessed eye tracking dataset	47
Table 4.4 Consonants in Turkish.....	50
Table 4.5 Output of the text scoring algorithm	53
Table 4.6 Average eye tracking data per sentence belong to non-dyslexic readers..	54
Table 4.7 Average eye tracking data per sentence belong to dyslexic readers	55
Table 4.8 Classic readability scores of sentences.....	56
Table 4.9 Result of first part of experiment	58
Table 5.1 Accuracy rates different sentences for dyslexia detection from SVM.....	60
Table 5.2 Correlation between scores and experiment results	61
Table 5.3 Percentage of time students spend on which stage.....	62
Table 5.4 Group Statistics	62
Table 5.5 Independent Samples Test.....	63
Table A.1 Sentences of the different difficulty levels (English)	76

LIST OF FIGURES

Figure 1.1 Steps of thesis	1
Figure 1.2 Accuracies when tested with different window sizes [2].	4
Figure 1.3 Classification results according to type of text [4].	6
Figure 1.4 Accuracies when tested with different window sizes [5].	8
Figure 2.1 E-Z Reader Model [61].....	21
Figure 2.2 The dual-route cascaded (DRC) [60].....	22
Figure 2.3 Parallel Distributed Processing Reading Model [62]	23
Figure 3.1 Tobii 4c Eye Tracker	30
Figure 3.2 Desktop mount with a chin and forehead rest.....	30
Figure 3.3 Total experiment system.....	30
Figure 3.4 First experiment screens	31
Figure 3.5 Second experiment screens.....	32
Figure 3.6 Third experiment screens.....	33
Figure 3.7 Fourth experiment screens.....	34
Figure 3.8 End of the first part of the experiment.....	35
Figure 3.9 First test screen example.....	39
Figure 3.10 Second test screen example	39
Figure 3.11 Third test screen example	40
Figure 3.12 Start second parts of experiment.....	41
Figure 4.1 Experiment preprocessed data	46
Figure 4.2 Dyslexic readers hate list	49
Figure 4.3 Text scoring algorithm functions.....	52
Figure 4.4 Shape and sound connection flow chart	59

CHAPTER 1

INTRODUCTION

Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction [1].

There is a hypothesis that dyslexic readers have difficulty in establishing the connection between letters and their corresponding sounds. This thesis is based on this hypothesis to further investigate if it is possible to detect dyslexia. Also, this thesis proposes new dyslexic readability index called Gunduz Index. The steps of the study are briefly as in Figure 1.1.

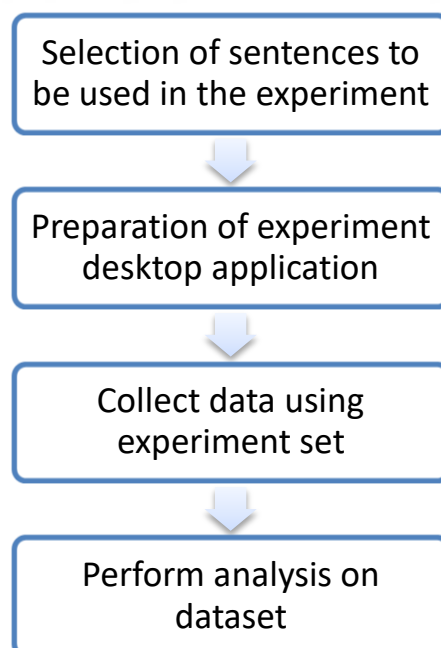


Figure 1.1 Steps of thesis

1.1 Aim of the Study

This thesis has three main aims: The first is to if it is possible to diagnose dyslexia in preschool periods by testing the connection between geometric shapes and sounds. The second aim is to diagnose dyslexia based on eye movements recorded during the reading. The third and last aim is to develop an algorithm/formula that will score texts based on its difficulty of reading for dyslexic children.

For these aims, it is hypothesized that,

- **H1:** It is possible to detect dyslexia in children by testing the time spent on the learnability of the connection between geometric shapes and corresponding phonetics with more than 70% accuracy.
- **H2:** It is possible to detect dyslexia in children from their eye movements with the highest 90% accuracy.
- **H3:** It is possible to propose new dyslexic readability index (Gunduz Index) for Turkish more successful than the existing readability indices.

1.2 Review of Related Works

There is a consensus in the literature that while reading, dyslexic people exhibit different eye movements compared to normal people. Dyslexic people display longer fixation durations and shorter saccades. Depending on the type of dyslexia, fixation order is disorganized with respect to normal readers due to the defects in their oculomotor planning. However, it will not be sufficient to examine eye movements only to detect dyslexia. There are two reasons for this. The first is that there is not enough literature because the scope of this research is too small. The second reason is that other studies using different methods on the same data will affect our approach to the problem positively. Therefore, the first common point in the literature survey was the use of eye movements recorded during reading. The second common point is that the methods used in the analysis are machine learning or deep learning methods.

Fixations and saccades can give different hints about the studies with eye movement on reading. For example, Ralf Biedert and Georg Buscher in their work “A Robust Real Time Reading-Skimming Classifier”, using the eye data recording while reading texts, try to determine whether some participants really read or not [2]. The researchers gathered eye movement data with Tobii 1750 eye tracking device from 12 university students whose age average was 24.6 and whose mother tongue was known as German. The participants were asked to read 8 different documents. A total of $12 * 8 = 96$ readings were obtained from the experiment. The results were first manually categorized by the two experts who had previously worked successfully in this area.

The classification is based on saccade distances as follows: If the line-base saccade distances below approximately 10 characters, then it belongs to a reading class. If the line base saccade distances are higher than 10 characters or the transitions are vertical instead of horizontal, they belong to a skimming class. With this method, experts have tagged the collected eye data, creating training and test sets manually. The training set includes 56% skimming labels, while the test set contains 51% skimming labels. Before future vectors were extracted from the training set, pre-processing operations were applied to achieve a more suitable representation. Besides these operations, there was also a denoising step with the median filter and converting the data into the document coordinate system. After the training set was obtained, future vectors were extracted with three different approaches. In the first approach, Character-Based Reading Detection, a traditional reading detection method, was performed and the words with the fixation points were magnetically mapped. The matching was done with line basis rather than word basis because there was too many vertical saccades in this data (baseline1). In the second approach, Normalized Saccade Vectors, in which each saccade represented in polar coordinates, were utilized. Feature vectors produced by this method consisted of the individual angles (θ) and normalized lengths (r) (baseline 2). The final approach was Contextual Saccade Shape, which had two main attributes to consider while drawing the future vector. Average angularity (h) and average forward speed (p). With this approach, the future vectors can be expressed as V_{css} ($V_{css} = h, p$). High

horizontality is observed as h approaches zero. For p , we can say that it represents the average length of all right-wingers.

Figure 1.2 drawn from the training set provides us with some good insights about the study. The best values for RBF-SVM were searched by testing different window sizes and different parameters as shown in Figure 1.2. Then we can see that the first results' feature vectors obtained by baseline 1 and baseline 2 approaches do not make big differences on total accuracy. As a second result, we can see that after setting the initial window size to 1, the accuracy first increases up to 3, and decreases afterwards. Moreover, the contextual saccade shape approach has the highest accuracy. The values used in the Linear Classification to be applied in the last phase of the work were determined by SVM.

Class selection was first formulated as, $\text{Class} = -2.97 + 5.36h + 0.17p$, then abbreviated as, $pl = p$. If the class value was less than 0, the reading belonged to the skimming class. Here, can be adjusted to improve the accuracy. The result of this linear model had 86% accuracy. This study has important findings in terms of reading habits and difficulties [2].

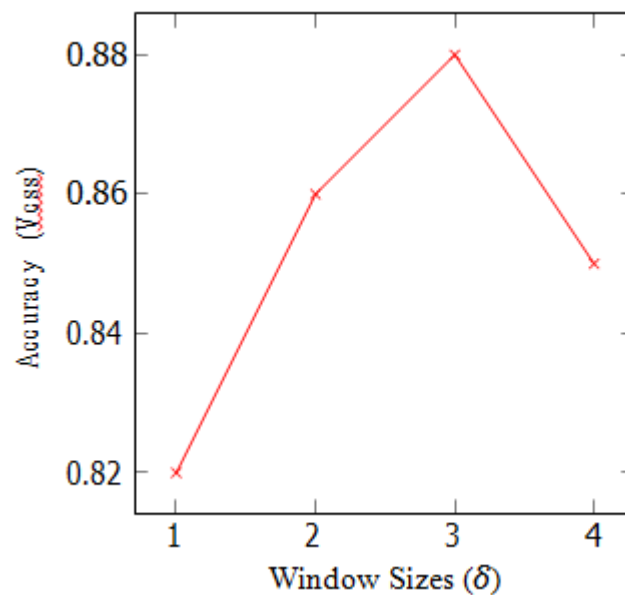


Figure 1.2 Accuracies when tested with different window sizes [2].

A similar article was published by Sepp Kollmorgen and Kenneth Holmqvist, “Automatically Detecting Reading in Eye Tracking Data” [3]. According to this article, if reading detection is performed only by observing eye movements, correct reading detection cannot be accomplished; therefore, only eye tracking is utilized. They describe the method in a way that the correct reading data can be obtained. The dataset used in this study was created with a random selection of 20 out of 96 participants who took part in a 3-layer manual test to ensure that they really read or did not. In this work, an instant eye tracking was done and a gaze point list was produced. The areas of interest were identified unlike the first paper. Therefore, for every 20 milliseconds of recording time, a gaze point was recorded. In addition to these gaze points, for each fixation point, $F_i = F(x_i, y_i, d_i, t_{0i})$ also there was a recording for the corresponding mean coordinates (x_i, y_i) , duration (d_i) and the start point (t_{0i}) . In the pre-processing, the start points were ignored and the values of each of the saccades were added. Four different learning models were constructed on this data set, namely, State HMM (Acc= 0.93), State HMM Simulated on line condition (Acc= 0.88), State HMM Unsupervised training (Acc= 0.87) and Neural Network (Acc= 0.88). This study together with HMM and Neural Network approaches allow for a stronger classification of work utilizing gaze data.

Studies involving eye movements are not only related to the detection of the words that are read or not read, but also about documents. One of these different studies is the article titled “I know what you are reading Recognition of Document Types Using Mobile Eye Tracking” by Kai Kunze, Yuzuko Utsumi, Yuki Shiga, Koichi Kise and Andreas Builing [4].

In this study, the type of a text was determined by using the eye movements taken from the participants in real-life settings, on the bus or in a coffee shop instead of reading a text on the computer screen. According to the authors, fixation patterns were very decisive features, and there were certain factors that affected these features. Some of these were the document type, font size, font type, reading task or goal, and the number and duration of distractions.

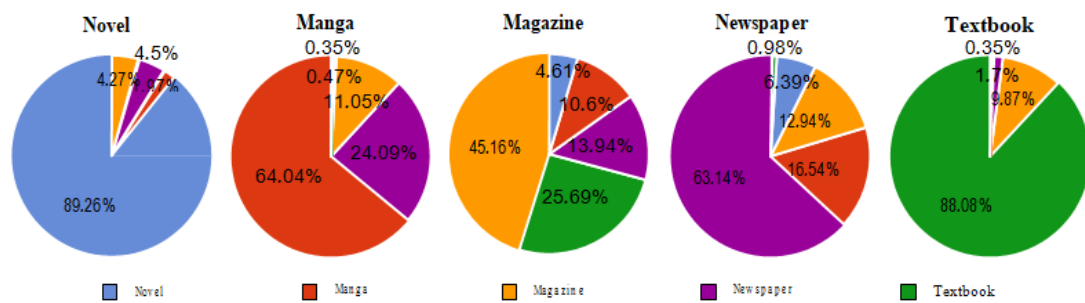


Figure 1.3 Classification results according to type of text [4].

The documents of this experiment were Japanese with 2 different writing styles: Yokogaki and Tategaki. The first one is written from right to left like English, while the other one is written in a top-to-bottom manner. The documents used in the experiment included both styles. A total of 8 Japanese participants wore a mobile head-mounted eye tracker SMI Eye Tracking Glasses. The participants were told that they could read 5 different kinds of Japanese texts starting from what they wanted. All participants engaged in reading for duration of time between 10 and 15 minutes.

While performing the document classification, the raw gaze data was first split according to the fixations and saccades (this gaze data contained 1-minute reading data). 4 different saccade directions were counted and Euclidean distances were computed. After that, saccade directions and finally slope over the fixations were calculated. These features were used as inputs to the J48 decision tree classifier.

Ultimately, more successful results were obtained in the text types with more regular fonts such as novels and textbooks. In magazines and manga categories, the classification was more difficult as in Figure 1.3. User dependent evaluations were observed as 99% accuracy, while 74% were observed as user-independent results [4].

For the studies mentioned above, it is very important to balance the age, gender, and reading abilities of the participants. Otherwise, certain patterns might deteriorate the classification task. For example, imagine that there are 15 dyslexic readers in an experiment with 100 participants. It will be a challenge to analysis the results properly because eye movement's patterns of dyslexic readers are very different from

the regular ones. The children learning reading or dyslexic readers shows longer fixation durations, with higher number of fixations, and shorter saccades and increased number of regressions than regular readers. Luz Rello and Miguel Ballesteros, for example, analysis this difference and tried to detect dyslexia from eye movement patterns in a research paper, “Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures” [5].

In this study [5], 92 Spanish participants read 12 different texts and their eye movements were recorded with a Tobii 1750 eye tracking device. 48 of the participants were known to have dyslexia. The texts used in the experiment had the same number of words, word length, font and size. They were composed of words that were usually not hard for dyslexic people to read. As a result, in order to understand which ones of the 1135 readings belong to the dyslexic participants, the researchers created 2 files for training and testing. Then, they ran the SVM binary classifier over the training set. This classifier was based on the following features: reading time, mean of fixations, the age of participants, number of visits, and number of fixations. Since the dataset was not very large, 10-fold cross-validation was used. This study achieved 80.18% accuracy as shown in Figure 1.4 in classification of dyslexic readers. When the age of the participants was neglected, the success rate was reported to be 76.39 %. This study is an important first step in the identification of special conditions such as dyslexia, where early diagnosis is very important. Supporting the work with the machine learning algorithm shows that dyslexic readers display some specific eye movement's patterns that are recognizable.

It is also of great importance to determine the appropriate reading material for such studies. In other words, the readability of the texts should be investigated and balanced across the participants to reduce possible confounding factors. In this area, an article entitled “Using gaze to predict text readability” by Ana V. Gonzalez-Garduno and Anders Sgaard focused on the contribution of readability models to the analysis of eye movement data [6].

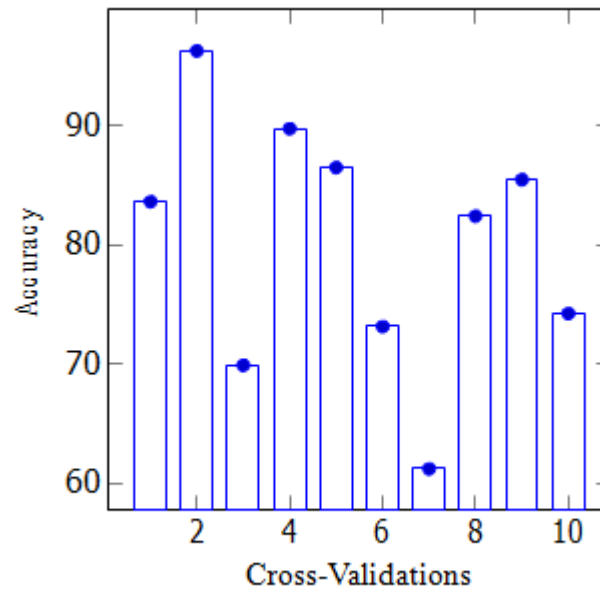


Figure 1.4 Accuracies when tested with different window sizes [5].

In this study, eye movements were taken from “The Dundee Treebank”, a work previously done by Maria Barrett, Eljko Agic and Anders Sgaard [7]. Every word in the Dundee Corpus is accompanied by per-word fixation durations. In addition to this, SimpleType and Standard Wikipedia were used for syntactic-level readability prediction, and OneStopEnglish corpus was used for evaluating. The features for prediction, such as syntactic and lexical properties were extracted from the dataset. For each sentence, a 33-element feature vector was extracted and used in order to predict an average first pass duration, regression path duration and total fixation duration.

Logistic/linear regression and Multilayer Perceptron (MLP) models were tested with extracted feature vectors. For 3-layer MLP architecture, the authors used the sigmoid activation function at the input and output layers while ReLu was utilized at the hidden layer [7]. For the Multi-task MLP model, the success rate average was 86% while the Multi-task LR model achieved only 78%. Taking these results into consideration, it can be concluded that eye movement data increased the single task

results. At the same time, the use of multi-task learning and deep learning algorithms was also one of the factors that increased the success rate.

Could we make a prediction about eye movements, even if we do not have a dataset with ready-to-read eye data like Dundee Corpus or an eye tracker to create? This question was tackled by Tanik Saikh, Srinivas Bangalore, Michael Carl and Sivaji Bandyopadhyay in their article “Predicting Source Gaze Fixation duration: a Machine Learning Approach” [8]. This study employed the Translation Process Research Database (TPR-DB). The dataset contained 3 articles translated from English to Danish by 24 translators. While these texts are being translated, eye movements of translators are recorded [9].

In the training task, the authors tried to extract feature vectors using these 3 news texts. Features used in this phase are the Parts of Speech (POS) tags, Unigram and bigram frequencies, First fixation duration, Target text number, Lexical Entropy, Perplexity, Syntactic entropy, Supertag, Translator Identity, and Degree of Polysemy.

The authors performed 3 different experiments on the feature vectors: Gaze per character level, gaze per word level and drafting gaze. In all experiments, gaze fixation duration was a dependent variable while other features were kept as independent variables and separated into classes using Support Vector Machine (SVM) with Weka tool. These classes were called G1, G2, G3, G4 and referred to the fixation durations that started from the lowest and gradually increased. As a result of the classification experiments, the total accuracy numbers were achieved as 42.6% for Gaze per character, 48.7% for Gaze per word level and 49.1% for Drafting Gaze. This work is a promising advancement in the development of intelligent automatic translation. Knowing how much time a user spends on a word can be useful for intelligent automatic translation algorithms.

We have been experiencing radical changes in the education field as well as in every field as technology develops. There are many courses available on the Internet. Can eye tracking data be used to create a more successful learning environment? Amir Shareghi Najari, Antonija Mitrovic and Kourosh Neshatian investigated this question

by comparing the behaviors of novices and advanced students through their eye tracking data in their work, “Utilizing Eye Tracking to Improve Learning from Examples.” [10]. In this work, they used the “Book Database” from the SQL Server Tutor. This database contains 6 different examples and each sample consists of text, the solution and an explanation.

22 undergraduate students who had used the same system in their schools participated in the study. They were presented with some examples that they had never seen before on a system with which the participants were familiar. During the experiments, students’ eye movements were recorded with a Tobii TX300 eye tracker. It was known that 12 of these students were novices and 10 of them were advanced. Their eye movements were examined in 4 different patterns: Pure reading, mixed reading, transferring, and scanning. Pure reading pattern means that the learner focuses on a particular area and reads it. Mixed reading means that if the student is reading a particular area, quickly look at another area and then return to the area he reads again. If he does not return to where he was, it is called transferring. If he does not focus on any area, it is called scanning. All eye movement data were separated into classes by a coding scheme called EGPA (Eye Gaze Pattern Analysis). According to the results of this classification, it was concluded that advanced students spent more time in the explanation section, while novice student’s eye movements were more complex and more pattern dependent.

A classification was further conducted using the patterns and behaviors as features. The authors tried to guess if a student was novice or advanced with machine learning algorithms such as W-J48 (Acc= 61.1%), W-Ladtree (Acc= 94.4%), W-BFTree (Acc= 72.2%), Rule Induction (Acc= 66.6%), WJRip (Acc= 77.8%), and Naive Bayes (Acc= 77.8%) employing 26 features (16 patterns and 10 behaviors). Note that “W” indicates that it was implemented with Weka tool. As it can be seen from the results, the best classification was achieved by the W- Ladtree. This study has important implications for developing educational models. Many online training platforms to enhance productivity can be developed, such as a training module that uses different methods according to a student’s learning profile.

A similar study was conducted by Shahram Eivazi and Roman Bednarik, “Predicting Problem-Solving Behavior and Performance Levels from Visual Attention Data” [11]. In this study, user behaviors in problem-solving tasks were analysed. For this purpose, eye movements and voices were recorded while 40 participants played an 8-tiles puzzle game. The first point was to classify the cognition states of participants. The classification categories were described as Cognitions, Evaluations, Plans and planning, Intentions, Concurrent move and Intention. Each feature vector was calculated from the eye movement data for the duration of the corresponding expression. For example, during a three- second coded state such as “evaluation”, all eye movement properties were calculated from this range.

The first prediction model over feature vectors was implemented using SVM. This model was based on learning the patterns of human cognition (five states) and eye movement features. These states were cognition, evaluation, planning, intention and concurrent moves. The model achieved a 75.84% accuracy rate. In addition to cognitive states prediction, all participants were divided into 3 groups as high-performance, medium-performance, and low performance. This distinction was made according to the duration of the game. Then, a classification was done with SVM over these 3 classes again. Ultimately, estimations were made at 87% accuracy using SVM classification together with eye movements. This is an indication that the participants’ eye movement data carry important information about their competence.

In addition to this, eye movements also contain important information about readers reading skills. With this in mind, Ya Lou, Yanping Liu, Johanna K. Kaakinen and Xingshan Li did a work, “Using support vector machines to identify literacy skills: Evidence from eye movements” in order to measure the literacy skills of the readers through the eye movements [12].

For this purpose, 67 undergraduate Chinese students were asked to study Chinese texts. These texts were related to dinosaurs and had 2600 words in total. After the texts were read, short questions were asked to ensure understanding. The participants’ eye movements were recorded during this experiment. Then, SVM

classification was performed on the data. The 24 features were selected from the previous studies. For example, the mean fixation duration, the number of forward fixations, or the number of regressive fixations, or reading rate were employed in the classification task.

With the use of different feature combinations, the lowest and highest accuracy were recorded as 54.1% and 80.3%, respectively. This study shows that eye movements contain important information about reading skills, and the SVM algorithm can classify eye movements according to the reader's expertise level.

Finally, let us look at a study that made predictions about the cognitive states of the brain by using eye movements again. This work was published as "Predicting Cognitive State from Eye Movements" by John M. Henderson, Svetlana V. Shinkareva, Jing Wang, Steven G. Luke and Jenn Olejarczyk [13]. In this study, there were 12 participants whose mother tongue was English. Participants performed 4 different tasks while looking at computer screens displaying 196 scene photographs and 140 texts. The tasks were scene search, scene memorization, reading, and pseudo reading.

Participants in scene memorization looked for the targets placed in the scenes from the real world. During the memorization of the scene, the participants tried to memorize what they saw. In reading, the participants read the paragraphs describing the latest news events. In pseudo-reading, participants were given fake texts made up from geometric shapes instead of letters. All participants performed all tasks consecutively in 3 days, while eye movements were recorded with the SR Research Eyelink 1000 eye tracker.

The authors chose some features to classify eye movements. These were the mean and standard deviation of fixation duration, the mean and standard deviation of saccade amplitude, the number of fixations per trial, and the three parameters, μ , σ and γ quantifying the shape of the fixation duration distribution with an ex-Gaussian distribution. These features were combined with naive Bayes classifiers trained using data for each participant in order to identify whether the participant performed reading, pseudo-reading, searching scenes, scenes or memorizing.

A number of different classifiers were tested on this data, either by changing feature vectors combinations or by changing the day of the training data retrieval. If it was trained and tested on samples taken on the same day, an accuracy of 80% was achieved. However, if the testing and the training data were taken on different days, the accuracy was reduced to 68%. This could also be the result of eye movements containing important information about the cognitive states and the cognitive states of the participants being influenced by external factors.

A summary of these studies can be found in Table 1.1.

Table 1.1 Overview of the studies

Study	Task	Tool/Dataset	Algorithm	Accuracy
“A Robust Real-time Reading-Skimming Classifier”	To determine whether the text is really read or not	Tobii 1750 Eye Tracker	SVM/Logistic Regression	86%
“Automatically Detecting Reading in Eye Tracking Data”	To determine whether the text is really read or not	SMI iView X eye tracking system with Polhemus head tracking	HMM/Neural Network	86%
“I know what you are reading Recognition of Document Types Using Mobile Eye Tracking”	To identify the type of text being read	SMI Eye Tracking Glasses	J48 decision tree	74% / 99%
“Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures”	To determine dyslexia from reading gaze data	Tobii 1750 Eye Tracker	SVM	80.18%
“Using gaze to predict text readability”	To determine the contribution of eye movements to readability models	The Dundee Tree-bank Dataset	Multi-Layer Perceptron/Logistic Regression	86%
“Predicting Source Gaze Fixation duration: a Machine Learning Approach”	To estimate gaze fixation durations	Translation Process Research Database (TPR-DB)	SVM	49.1%

Table 1.1 (Continued) Overview of the studies

Study	Task	Tool/Dataset	Algorithm	Accuracy
“Utilizing Eye Tracking to Improve Learning from Examples.”	To compare behaviors of novices and advanced students while studying examples.	Tobii TX300 Eye Tracker	Naive Bayes/W-Ladtree	77.8% 94.4%
“Predicting Problem-Solving Behavior and Performance Levels from Visual Attention Data”	To analyse user behavior during a problem- solving task	Tobii ET 1750 eye tracker	SVM	* 87%
“Using support vector machines to identify literacy skills: Evidence from eye movements”	To measure the literacy skills of the readers through the eye movements	No information about eye tracker brand	SVM	80.3%
“Predicting Cognitive State from Eye Movements”	To predicting cognitive state from eye movements	SR Research Eyelink 1000 eye tracker	Naive Bayes	80%

CHAPTER 2

BACKGROUND

2.1 Dyslexia

Although people with literacy training can easily and spontaneously read a piece of text, reading is a very complex activity. While letters and words are perceived by the visual system in the brain, cognitive faculties such as memory, language, reasoning, and such are also involved in reading activity. Literacy in the modern world is of utmost importance for managing one's daily life and contributing to the society as an educated person, as well as for individual and professional development. However, dyslexia, which is thought to affect around 10-15% of society, causes reading deficit for some, despite their normal intelligence. Albert Einstein, Picasso, Bill Gates, and Steven Spielberg are known to have suffered from dyslexia. To detect the condition in primary or elementary school children, initially a digit-span test and WISC-R (Wechsler Intelligence Scale for Children) intelligence test are applied for normal memory and intelligence. Visual tests should also be applied to check for vision impairment. The children who score normal on these tests, but poorly on the Weschler's reading tests (version 3 and 4), can be categorized as dyslexic. The children diagnosed with dyslexia can show progress with the help of special education techniques. The main problems with current methods of diagnosis are that they are slow, composed of multiple stages, and, in case of an error, cannot be repeated in an 8-month window. Furthermore, these techniques depend on the properties of English language. Since there is no eye-tracking corpus of Turkish dyslexic readers, it is unknown whether dyslexia have Turkish-specific properties in terms of eye-tracking data. Children who received poor education or suffer from other health problems are accidentally misclassified as dyslexic.

Dyslexia, (as a result of dementia or damage to the left lobe or visual cortex) may develop later (acquired dyslexia) and may also be congenital (developmental dyslexia). In the 19th century, dyslexia was thought to be a visual disorder (word

blindness) by Morgen, Hinshelwood and Orton. However, nowadays, it is known that it is caused by dysfunctions in the reading system in the human brain and it is a genetic disorder.

There are various types of dyslexia: mixing of letters or words that stand side by side (attention dyslexia), skipping letters or words to the left (ignored dyslexia), abnormalities of eye movements when reading to the right (hemianopic alexia), the problem of not being able to combine letters even though the letters are perceived (pure alexia). It is thought that dyslexia types are caused by simultaneous problems in visual and linguistic systems.

On the other hand, there are some types of phonological or semantic damages such as the inability to read the words that do not comply with the general rules of a language (superficial dyslexia), the inability to read the low frequency words (phonological dyslexia), and the misunderstanding of familiar words as very different words in the meaning of words but have similar typing (deep dyslexia).

Children with poor speech or writing skills tend to be academically unsuccessful and leave school [14,15,16]. Dyslexic children have problems not only learning their own language but also learning other languages [17]. For example, in Turkey the most widely taught foreign language is English. Due to deep opaque orthography, English is very difficult to dyslexic Turkish children, because it complicates the deep opaque orthography text-phoneme conversion [18].

If the problem of reading is experienced in children who do not have a problem of intelligence, who have received sufficient education, from similar socio-economic families and who have no problems in their senses, the most probable cause is dyslexia. This problem requires special training methods. Since they are not caused by intelligence, they can be successful and contribute to the society.

Studying visual-attention and oculomotor abilities that play a decisive role in reading development in preschool and primary school students, and comparison with normal readers will be useful for early diagnosis of children with a risk of dyslexia. On the other hand, oculomotor defects can be examined with other developmental disorders

(autism, distractibility, hyperactivity, schizophrenia, obsessive and compulsive disorder, etc.). The effects of dyslexia on visual-attention and oculomotor defects can be investigated. Eye tracking data in dyslexic individuals can contribute to the development of new methods by early educators as well as early diagnosis systems.

Dyslexia is an important problem not only for children but also for their families. Yıldız et al. (2012) conducted a study on the problems of families with dyslexic children [19]. In this study, families stated that teachers developed negative behaviors towards themselves, and that their children were subjected to insult, exclusion, psychological pressure and even violence by others. Unfortunately, family members may occasionally show these behaviors against dyslexic children. Another finding is that children are very difficult to diagnose because of lack of knowledge about dyslexia.

Primary education in Turkey is modified using a voice-based method in a reform carried out in 2005. According to this method, some letter-sound relationships and short forms that are meaningful are taught. Then, syllables, words and sentence structures are taught [20]. However, children with reading problems can be identified when they fail in second grade. There are some studies about the use of various methods by educators to diagnose dyslexia in children [21,22,23]. However, these methods are not always implemented [24,23,25].

In dyslexia studies using psychometric approach, the sentences are generally read to the participants and they are tested with some questions and shapes that they understand. In single word approaches, the effect of variables such as neighborhood, frequency and length on reading is examined. In such studies, eye tracking data is very valuable because it provides clues about instant cognitive processes [26]. Eye-integrity paradigm is used in experiments in which the image on the computer screen is changed according to the eye data of the participants. With the help of these studies, when the visual perception areas of the English readers were examined, it was determined that generally 14-15 characters left of the focus point were detected with 3-4 characters left [27,28]. It is thought that readers are collecting information from the words in the area (parafovea) adjacent to the fovea. Saccades determine the

word to be looked after. Focusing time is effective in determining when the next eye movement will be. While caution is on the figures and texts falling in the fovea area, it is thought that the oculomotor planning that determines the next saccade movement is influenced by the information in the fovea and parafovea forming the visual perception [29].

Children learning to read must develop linguistic skills such as phonological awareness, phonological analysis and writing. Dyslexia is thought to originate from a phonological deficit or visuo-attention deficit, or both [30,31]. Although there is no consensus on the cause of dyslexia in the literature, it has been reported that dyslexic individuals show different eye movements in linguistic and non-linguistic studies [32,33,34]. Prado, Dubois, and Valdois (2007) reported that visual-attention deficits were effective in the reading of dyslexic individuals, and they did not show a significant difference in visual processing compared to normal individuals [31].

This may be due to the intense use of visual-attention in reading. Reading comprehension is a basic skill and influences school education and development of individuals in future years [22,35]. There is a consensus in the literature that dyslexic eye screening data is unusual [36].

Many different eye movements are mentioned in the literature such as fixation, saccade, vertical saccade and regression [37]. In literature, it was seen that the most effective variables for the saccades and fixations were frequency and length of the word. On the other hand, the status of synonymity, whether it comes from the same sound root, whether it is a real word, the fact that it fits the spelling rules or is an exception, affects reading [38]. Regardless of the difficulty of writing, when dyslexic and normal readers' eye tracking data are examined, it is known that dyslexic children show more saccade and long-term fixation during reading [39].

The only study in Turkish, Bayram, Camnalbur and Esgin (2012) collected 15 follow-up data of dyslexia and 15 normal readers [40]. The eye tracking data of the readers who read real writings and false words were recorded. In Turkish, it is reported that dyslexic people have more and longer fixation and less regression. The word length for dyslexic children was found to be more correlated with the number

of fixations. At this point, studies with false words are important because for every child who learns to read, a word he does not know and for the first time will be perceived as a false word. For reading experiments, eye movement tracking is a very useful evaluation method because reading process provides objective and quantitative data on visual and attention capabilities [41,34]. Eye movements tell the researchers where, how many times and how long they are looked.

Differences in dyslexic eye movements are not the cause of dyslexia but the consequences of reading difficulties [42,26]. Differences in dyslexic eye tracking data can be observed when reading a full sentence [43], a single word [44] or a false word [45], as well as differences in non-reading activities [46,31]. Lobier, Zoubrinetzky and Valdois (2012) examined the non-linguistic eye tracking data of dyslexic people and tried to understand whether the source of the defect was the visual attention or the writing-sound transformation problem [47]. As a result of their experiments with non-language visual stimuli, they demonstrated the cause of dyslexia as defects in the visual attention interval. Although there is no consensus in the literature about whether the disease is due to tongue-dependent or oculomotor system, significant eye movement's differences have been recorded between the normal and dyslexic individuals [31,46]. Even Rello and Ballosteros (2015) reported that 80.18% of random readers successfully grouped as dyslexic or normal readers using eye tracking data and machine learning algorithms. This study underlines that eye tracking data is one of the most important indicators of dyslexia [5].

Jones et al. (2008) compared the healthy and dyslexic individuals in the eye tracking study using the phonological and visual information with the fast-automatic naming method [49]. The visual or phonological nature of the experimental content influenced eye movements. However, this effect was greater for the dyslexic group. They also determined that dyslexics were slower in nomenclature. De Luca et al. (1999) observed eye movements of dyslexia in non-language activities. They claimed that patients with superficial dyslexia were not very different. On the other hand, Trauzettel-Klosinski et al. (2002) noted that dyslexics in the 11-15 age group exhibited retrospective saccade movements more frequently than the control group [50]. Dyslexic eye tracking studies in the literature are generally performed in Indo-

European languages. Turkish is from the Ural-Altan language family and written in Latin letters. Dyslexic eye follow-up studies in the literature are generally performed in Indo-European languages. Turkish is from the Ural-Altan language family and written in Latin letters. There are letters ç, ğ, ı, ş, ü, and ğ in the alphabet instead of q, x and w. Due to the closeness between letters and sounds, it is one of the shallow writing languages. Therefore, for people who have dyslexia problem, Turkish has different differences compared to languages such as English and Swedish which have similar studies in this area [22].

Despite the recognition of dyslexia by Turkey Ministry of Education since 1997 it was only mentioned in the individual education plan. Teachers have generally learned about this by rehabilitation centers or school counselors. Teachers have been given a more in-depth education since 2009, but they are not sufficient. Diagnosis of dyslexia is done in state and university hospitals and this process cannot be performed in other hospitals. If a class teacher thinks that the student has dyslexia problem, the school will apply to one of the rehabilitation centers operated by the government on behalf of the child. These centers use the WISC-R or Stanford-Binet tests adapted for Turkish [51]. A special test is done for Turkish vocabulary reading, phonology and speed test [52]. According to the test results, the child can be admitted to the child psychiatry department of the university or state hospitals. The child psychiatrist makes the final diagnosis of the child. The comprehensive Disability Act (no. 5378, 2005), published in 2005, guarantees the rights of persons with disabilities to employment, education, daily life and transportation. However, dyslexia is not mentioned in this law. This includes dyslexia through additional decrees. These decrees advocate for children's right to education and give the rights to the children with dyslexia such as giving additional time and writing guidelines in examinations or giving oral right instead of written. They can use computers for written assignments or be exempted from reading aloud in the classroom.

2.2 Reading Models

Several models have been defined for how words are defined while reading. In some of these models eye movements are discussed. Some of these are Interactive-Activation [53], Activation-Validation [54], Multi-Layer [55], Multiple External Reading [56], Trace Memory [57], Coupler Double Process [58], Bayesian Reader [59] models. Apart from these models, the two most common reading models are the dual-route cascaded (DRC) [60] and the EZ Reader Model [61].

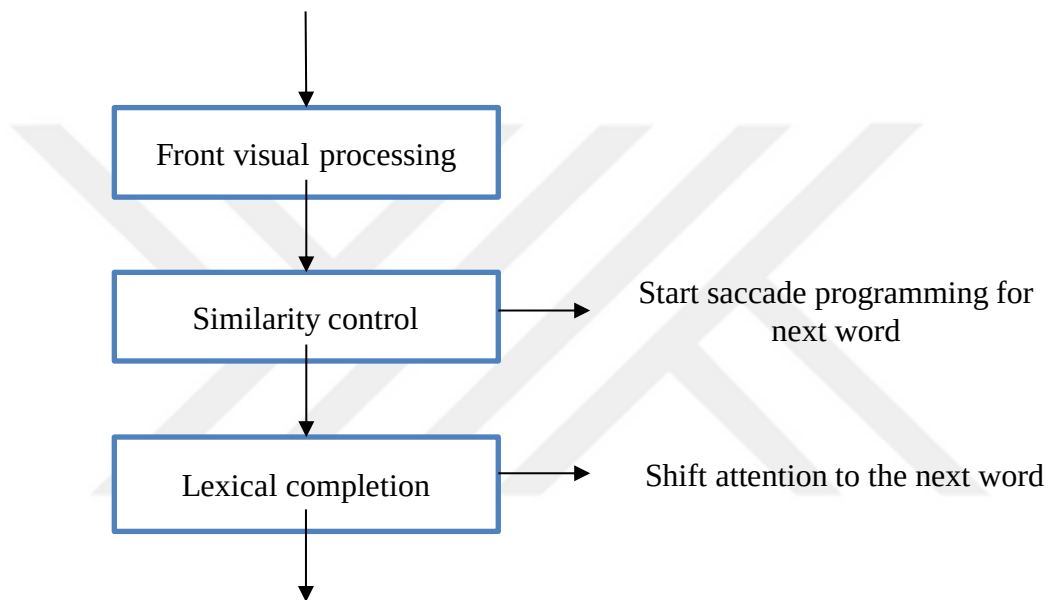


Figure 2.1 E-Z Reader Model [61]

The E-Z Reader Model takes reading as a series process and does not include high level linguistic and scientific operations. It is a model that ignores in-word regressions. It accepts the word perception stage as a signal to act. The second acceptance in the model is that the attention shifts from a word to word. According to this model shown in Figure 2.1, the reading begins with the visual characteristics of the writing fall on the retina. Meanwhile attention, it is not the word. The next stage is the word identification stage and the attention start with that word being over. Through visual information, all words that are similar to the word read in the literary words are activated. There is only visual similarity here, because this stage does not have access to the mental dictionary and the phonetic forms [61]. At this stage,

oculomotor planning is performed. With word boundaries, font sizes and information collected from the parafoveal area, the jump planning is done for the next saccade. In the final stage, a verbal identification is made. At this stage there is access to the mental dictionary and semantic knowledge is also used effectively. This is the stage where attention is most effective. When this process is completed, the attention will shift to the next word.

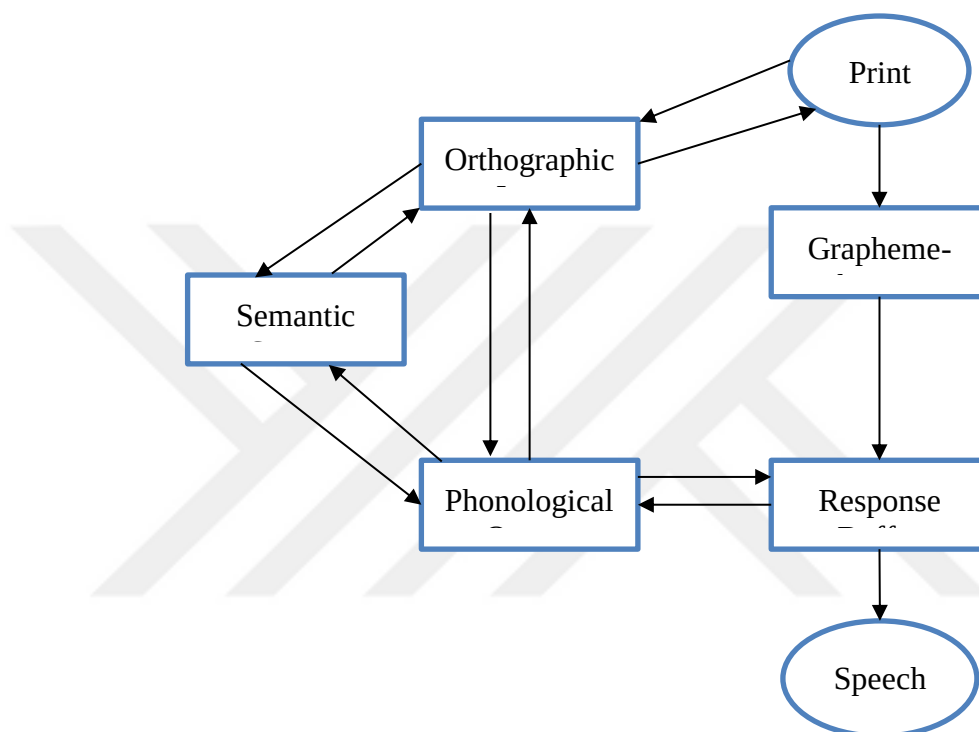


Figure 2.2 The dual-route cascaded (DRC) [60]

The dual-route cascaded (DRC) is shown in Figure 2.2 In this model, you can follow two different paths after reading an article. The first is the way directly to the mind dictionary and in this way the word visuals are matched with their meanings in the mental dictionary. The other way is the Text-to-Audio Unit Rule System that matches the text with sound. These rules are based entirely on statistical basis. The relationship between frequency and continuity determines how to read a text. For this reason, regular words are read faster than exceptions. The biggest advantage of this model is that it can explain the types of acquired dyslexia [60]. It was determined that patients with superficial dyslexia had defects in their access to the mental

dictionary, because the false words were observed to read more comfortable. On the other hand, it was observed that people with phonetic dyslexia deficits read regular and irregular words easily but could not read false words. This situation showed that there is a defect in accessing the second way, the Writing-Voice Unit Rule System.

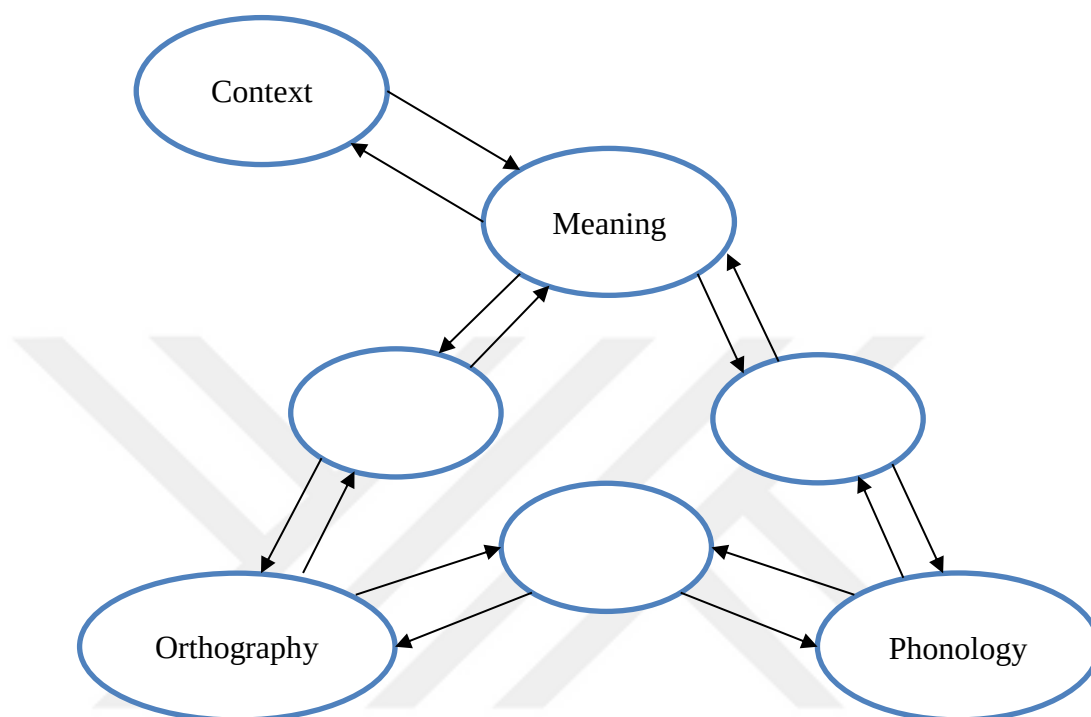


Figure 2.3 Parallel Distributed Processing Reading Model [62]

In the dual-route reading model, there are direct or indirect communication paths between the units. There is a competition between the data on both roads. Parallel Distributed Processing Reading Model is influenced by neural networks as shown in Figure 2.3 and assumes that there are similar connections between the units [62]. The Parallel Distributed Processing model recommends a single, connected model for reading both real and false words. There are no separate routes in this model. This model attempts to explain learning and wrong readings by assuming that there are activation weights between links. The basic problem of this model is that children ignore the need for prior knowledge and skills for reading [14]. It assumes that phonological data is recorded in a scattered and irregular way. In fact, children have already acquired a large phonological archive before learning to read in preschool. This model was developed based on data obtained from healthy and talented

individuals. However, the DRC model has been developed according to the data of individuals with reading problems [63]. Rastle and Coltheart (2006) compared the effects of reading studies with a comprehensive literature review. According to this, some of the loud reading activities necessitated serial communication and parallel distributed models could not explain this situation.

In this model, for example, false words can be read through the link between writing and phonology units, while false word readings cannot be made on the meaning. One of two ways will work. But in the DRC model they say false words or words that the reader didn't see before, activating both ways. This model is partially applied through artificial neural networks. For example, it has hidden neural connections between writing and phonology, and these links are invisible to other units even if they have their own weights.

This study will be based on the DRC model for dyslexic readers because this model is more descriptive than the E-Z Reader Model and Parallel Distributed Processing Model for different types of dyslexia.

2.3 Readability

The arguments used in communication from the past to the present played an important role in the quality of communication. When we look at the historical development of written communication, we can observe that it changes its shape according to the time, environment and purpose. As long as these arguments are chosen according to the environment, time and purpose of use, a quality communication can be mentioned.

If quality written communication is mentioned, the concept of readability has always been an important problem. In fact, legibility and intelligibility can be considered as two sub concepts while readability is considered as a top concept [66].

Environmental factors (quality of print, font size, distance between words and letters, colors etc.) are effective in measuring readability concepts, but structural factors are also effective. These structural factors can be sentence and word length, number of

syllables or word frequencies [67]. In order to formulate these measurements, it would be more logical to use contextual factors. As a matter of fact, the formulation studies carried out so far have been in this direction.

2.3.1 General Readability Rules

Many studies have been conducted so far to measure readability. This is a very large majority at the end of the study prepared for the basic formula was obtained in the English text. These formulas are the formulas that still preserve its topicality and still used by many text editors. Some; Flesch Reading Ease [68], Flesch Kincaid Grade Level, Coleman Liau Index [69], Gunning Fog Index [70], SMOG Index [71], ARI Index, RIX Index and LIX Index.

These formulas contain different parameters but work in similar logic. Among these, we can take a close look at the Flesch Reading Ease formula used by Microsoft in terms of ease of adaptation for Turkish.

The parameters used in the Flesch Reading Ease formula are the number of sentences, the total actual number and the total number of syllables.

$$FRE = 206.835 - (1.015 \times ASL) - (84.6 \times ASW) \quad (2.1)$$

FRE, *ASL* and *ASW* are denoted Flesch Reading Ease score, average syllable length and average sentence length respectively, in equation 2.1. As the score from this formula decreases, it is understood that the readability of the text becomes more difficult. In addition to the specified parameters, the constants in the formula are in fact the result of preliminary studies for English.

According to Flesch, the average sentence length for English texts is 14-15 words and the average word length is 1.4 syllables. Atesman concluded that the average word length in Turkish is 2.6 syllables and the length of the sentence is 9-10 words [72]. With this information, the formula was revised to obtain a structural readability formula for Turkish. *OS*, *a* and *b* are denoted readability score, average word length as a syllable and sentence length as a word respectively, in equation 2.2.

$$OS = 198.825 - (40,175 \times a) - (2.610 \times b) \quad (2.2)$$

As in the Flesch Reading Ease formula, it is understood that the level of readability decreases as the score decreases in the formula of Atesman.

2.3.2 Language and Readability

There is an important link between language and readability. The histories, grammar, structural and phonetic properties of language have been influential on the general readability rules. For example, Stanislas Dehaene explained how this linguistic difference affects readability by comparing English and Italian. In Italian, each letter corresponds to a phoneme, while in English the situation is more complex. Languages such as English, which do not have letters or letter groups that may correspond to each speech voice, have greater difficulties in legibility. In his book "Reading in the Brain", he summarizes the subject with the following sentences: "Even a simple letter X is unnecessary as it replaces the "KS" phonemes that already have their own writing. People in Turkey "taksi" they call. In a one-year period (1928-29), this country, which accepted the Latin alphabet, simplified its writing and taught how to read to three million people, is a good example of the applicability of the spelling reform." [73].

Another situation that has gained more importance in today's technology age is the dynamic of language. The language is used and changed in different ways with the changing living conditions of the generations using language. The words that are changed, combined or derived according to the subject of the agenda are negatively or positively influential on the readability. The main reason why this situation is more effective today is the widespread use of social media platforms.

As a result, cultures and geographies in which languages have been fed up to date have played a major role in the traditionalization of languages and the readability of languages in these traditions. Language reform can facilitate the lives of many people in many languages. Minimizing the age of reading and the duration of learning will make great contributions in all areas.

2.3.3 Dyslexia and Readability

Readability studies usually cover the majority. That is, texts are evaluated on the basis of the average reading capabilities of the general reader profile. This method is not unreasonable, of course. However, if we are talking about measuring readability, it is not possible to ignore the most disadvantageous profile, i.e., the dyslexic readers.

Reading books created for students who are trying to learn to read new are formed according to the results of general readability rules. If these rules were enough in a special situation like dyslexia, today dyslexic students would not have so much difficulty already in school life. For this reason, general readability formulas need to be revised again for the dyslexic students.

The multiplicity of long words is still a factor that affects dyslexic readability, but it is a bigger problem that the word begins with the same syllable followed by the word. For example, the successive use of the words “sari” and “sarma” is an important factor in the measurement of dyslexic readability.

There is a similar situation in short words. Existing readability rules consider short words readable. This assumption is partially correct. But, again, it had to be revised for dyslexic readability. For example, the word “buda” with the word “sada” have the same number of syllables but must have a different reading difficulty score, because the presence of letters in the word such as “b” and “d” in the form of mirror symmetry must be taken into account in the measurement of the readability.

Factors affecting readability are not simply formal factors. Many factors, such as font type, font size and inter-word spacing affect readability. A nice study on this subject is that Luz Rello is looking for the ideal writing parameters for dyslexic students [48].

As a result, we need to revise the available readability formulas for groups with different needs from the general audience, such as dyslexia. This revision should be adjusted according to the needs of the working group. In this way, it is inevitable that

a readability sequence will be created between the texts and a significant increase in the success rate will be achieved when going from an easy to difficult order.



CHAPTER 3

METHODOLOGY

3.1 Coding Platforms

There are 2 different technical parts of this study. The first one is the desktop application designed to collect the data set. I used the Microsoft Visual Studio for this section. Microsoft Visual Studio is an integrated development environment developed by Microsoft (IDE). I used C# for designing the experiment platform. I preferred it because it had a convenient interface to easily design experiment screens. In addition, the eye tracker I use for research purposes has a C# library to access gaze data, which is “Tobii Research SDK” and all documentation is available on the brand's own website [65].

For data analysis, Python was employed with the data collected in the first platform. The reason for using Python is the ease of access to the machine learning and natural language processing libraries.

3.2 Method

3.2.1 Participants

A total of 50 students participated as part of the study. 23 of them were dyslexic, while the rest were non-dyslexic. The ages of the students varied between 8 and 9 years and all of them were in the 2nd or 3rd grade. The required ethics committee and Ministry of National Education (MEB) permissions were also obtained.

3.2.1 Apparatus

Experiments were performed using the Tobii 4c Eye Tracker Figure 3.1. In fact, this eye tracker is built to be used in the gaming market. However, there is a license key to use it for academic purposes. You can only access raw eye movement data in this way. It is 17 x 15 x 335 mm and works at a frequency of 90 Hz.



Figure 3.1 Tobii 4c Eye Tracker

Participants used a desktop mount with a chin and forehead rest to reduce head movements during experiments Figure 3.2. Because head movements may reduce data quality, this device is used in almost all eye tracking applications. Especially as in this study, if your participants are children, this is much more important.



Figure 3.2 Desktop mount with a chin and forehead rest

Computer used with this experiment has a 2.80 GHz processor and display resolution of 1920x1080. Total experiment system is like Figure 3.3. The distance between the student and the computer screen is set to approximately 50 cm. By using this setting, two experiments were run.

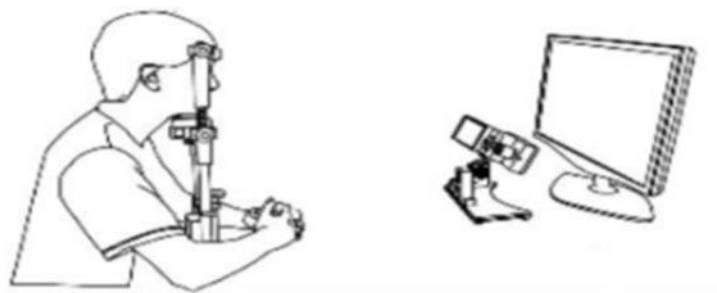


Figure 3.3 Total experiment system

3.2.3 Materials and Procedure

A desktop application was designed for use in experiments. This application consists of two main parts. Step by step all the stages are explained below along with screenshots and reasons.

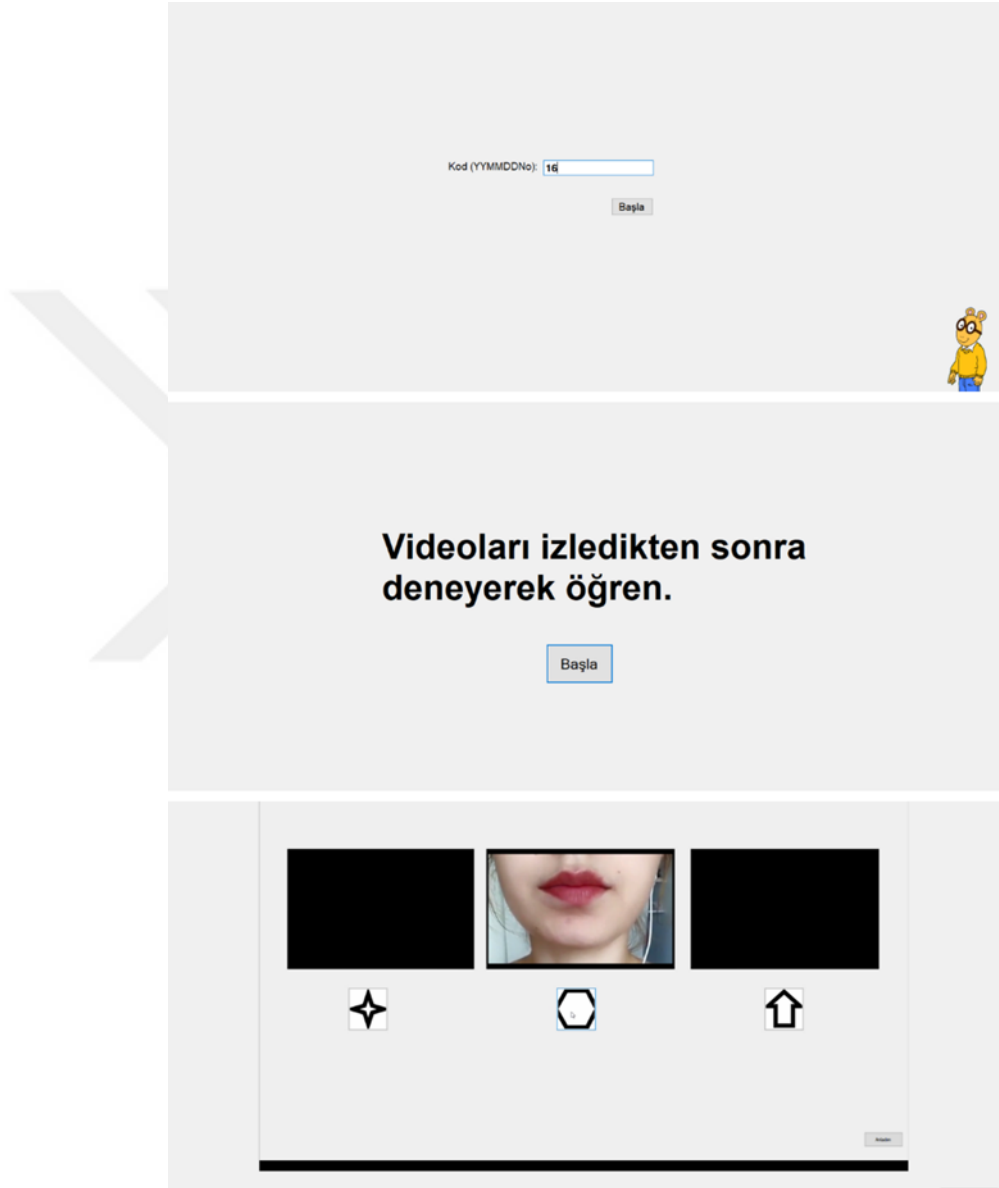


Figure 3.4 First experiment screens

Because we wanted the identity of the subjects to be confidential, the experiment starts on the login screen, which requires only the student ID. After login, there are

some instructions on how the student should proceed in the experiment. As a result of the experiments we designed while designing the system, we understood that written instructions would not be enough. Therefore, we have prepared video instructions Figure 3.4.

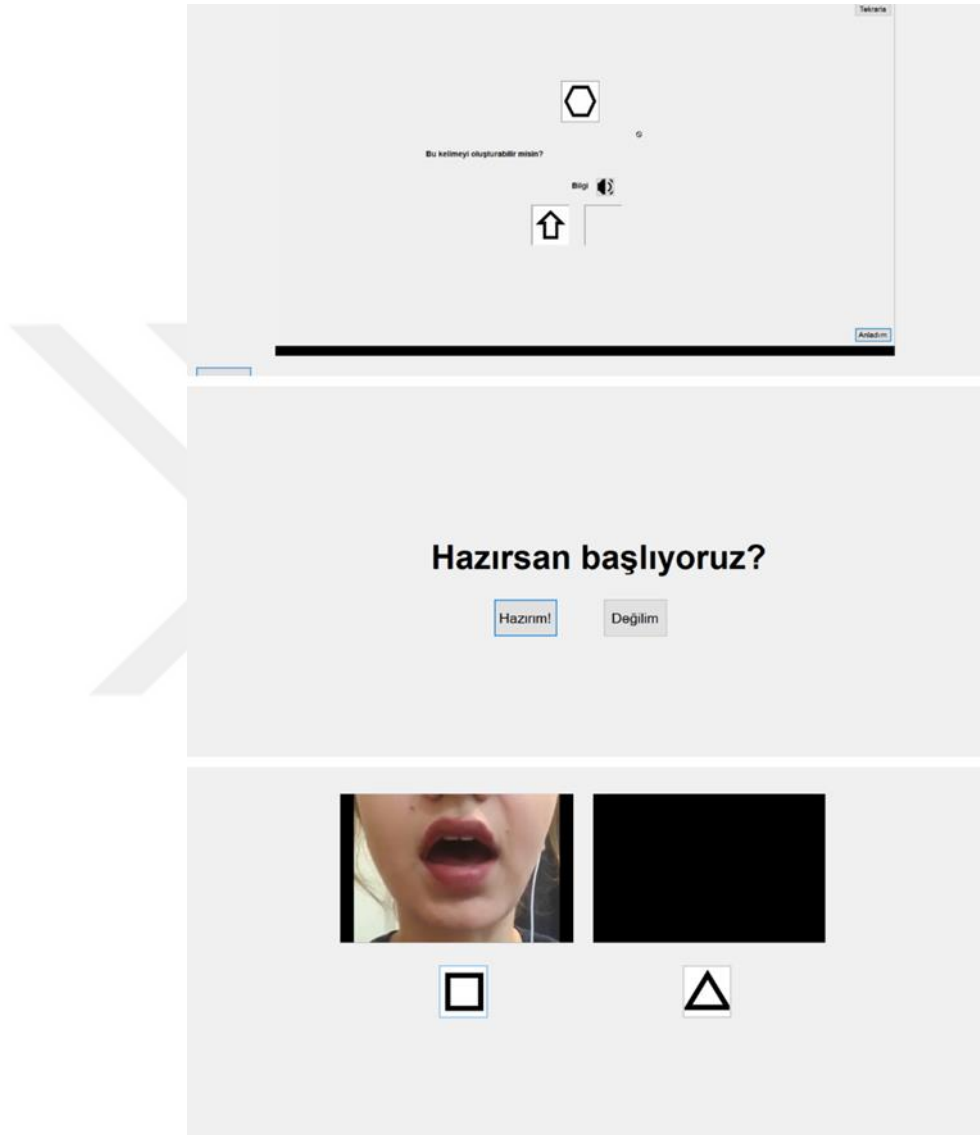


Figure 3.5 Second experiment screens

After the video instructions of both the learning and the test screen have passed, the first part of the experiment begins. Before starting all the sections, there is a question screen to see if the student is ready to begin. If student ready to start, experiment starts with a learning screen which contains 2 shapes and 2 videos. In this section, the

participant presses one of the buttons then watches and listen the corresponding syllable from video Figure 3.5.

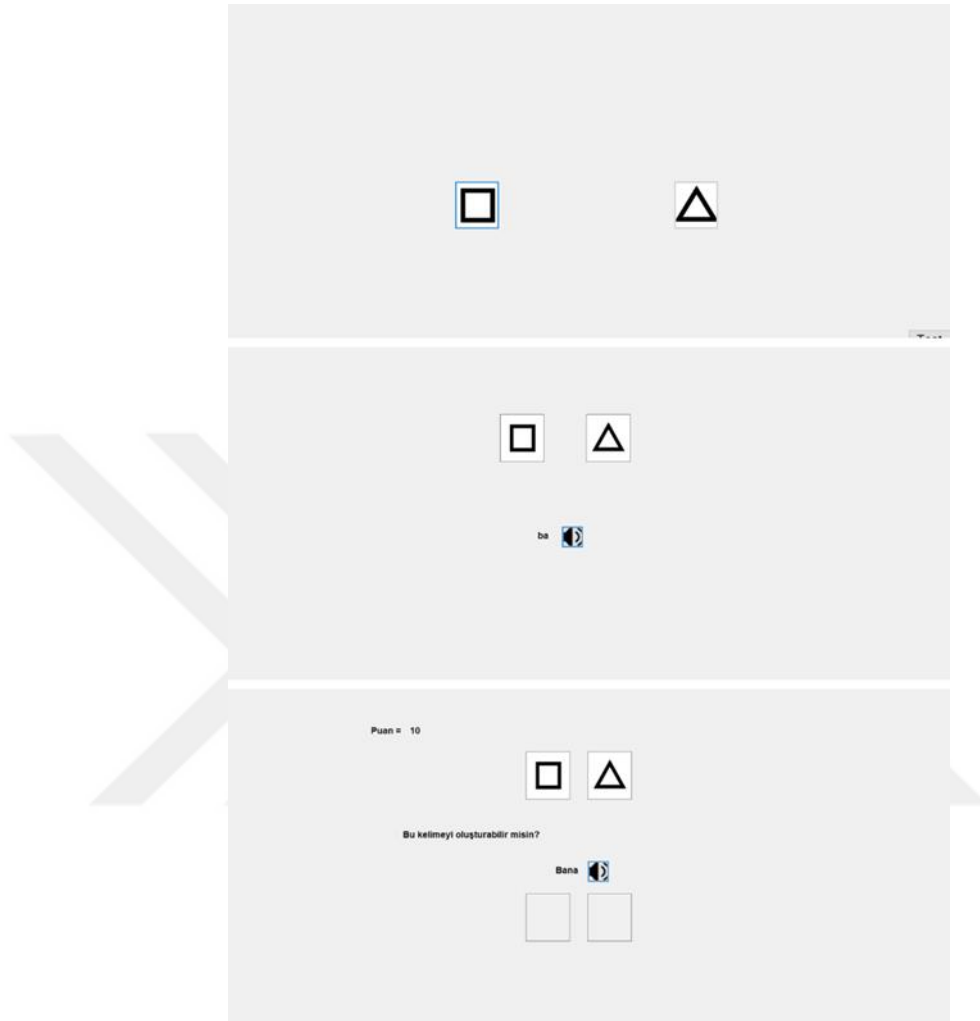


Figure 3.6 Third experiment screens

After you learn which geometric shape is related to which syllables from videos, the second screen comes with the same syllables without videos that is only sound and shapes. When the learning part is over, the test part starts.

On the first test screen, the participant is expected to choose which shape represents the syllable. If the participant wishes, he or she can hear the voice of this syllable by pressing the button next to the desired syllable. The second test screen is expected to construct a word from the two syllables learned. This can be done by dragging the

figures above to the boxes below. In all test stages, a score is kept in line with the correct answers given by the participants. This score can be seen in the upper left corner of the screen Figure 3.6.

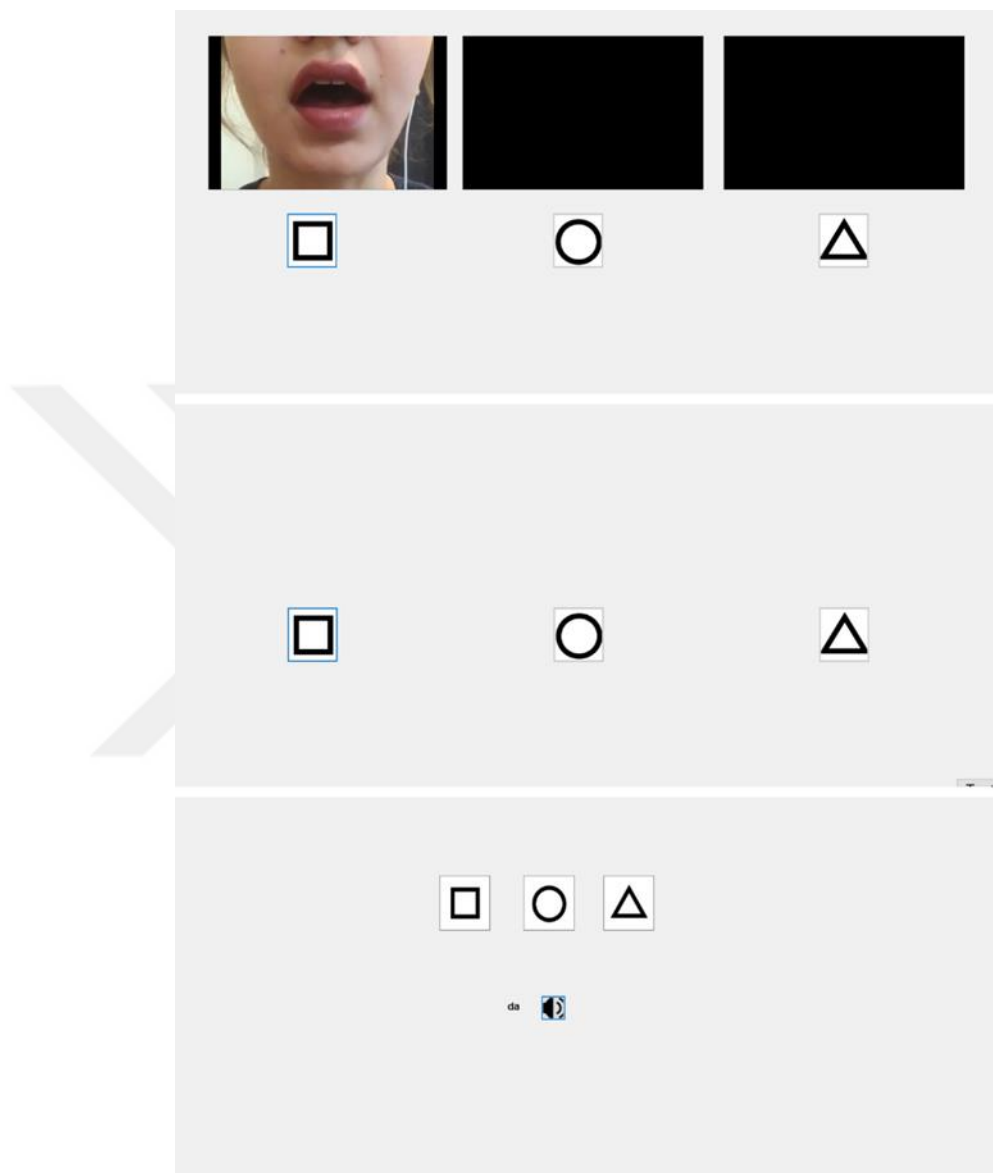


Figure 3.7 Fourth experiment screens

After the two-syllable learning and test parts are finished, the three-syllable section starts. This section proceeds with the same procedure except that there is an extra syllable. Two of these 3 syllables are previously learned syllables. A test system

from easy to difficult has been created in Figure 3.7. The final stage of the first part of the experiment is as shown in Figure 3.8.

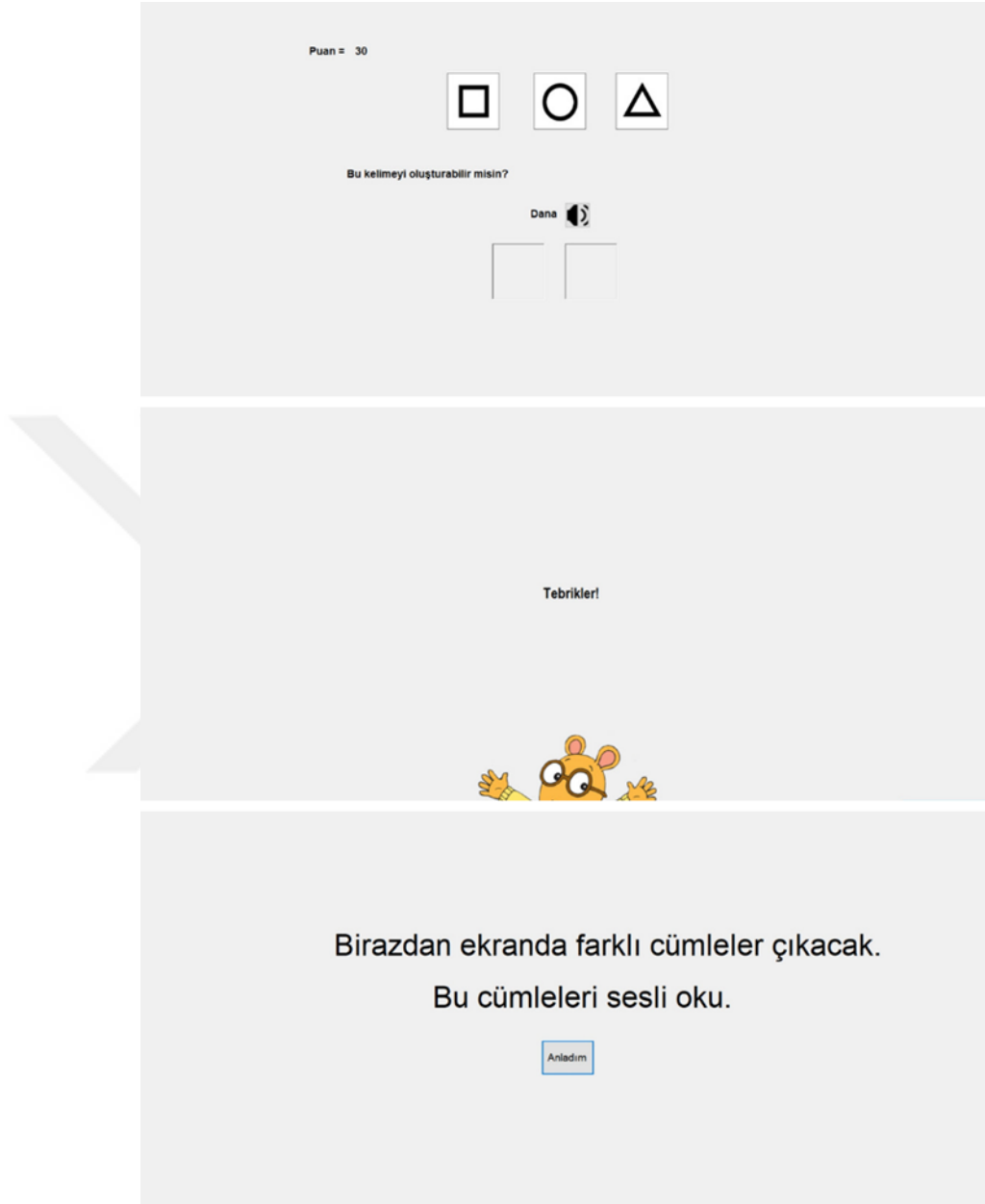


Figure 3.8 End of the first part of the experiment

In the first part of the experiment, there are 6 questions in which the children are displayed some symbols and expected to learn corresponding sounds associated with them. At this point, it is wise to mention why this part of the experiment is designed in this way.

One of the biggest problems of dyslexic students, as mentioned in the Background section, is the difficulty of establishing the connection between the letters and the sounds they represent. For this reason, the main purpose of this section is to analyse the relationship between letters and voices through syllables and geometric shapes hoping that this approach could be used with preschoolers in future. If the connection between these two relationships is similar as suggested, the test results of the two different experimental groups will be different. In other words, we proposed that dyslexic students will have more difficulty in establishing the connection between geometric shapes and syllables than non-dyslexic students. In order to understand this, we planned to use the correct answer numbers and test completion times in the learning and test sections.

The geometric shapes and syllables used in this section were not randomly selected. There are some hypotheses that are based on this choice. The visual representation of the letters in the Latin alphabet is entirely related to geometry. In other words, the shapes of the letters are formed by the combination of some geometric shapes.

Difficulty in reading letters with similar shapes is one of the most common problems of dyslexia. Therefore, we can say that the similarity of geometric shapes is the origin of the problem. Based on this logic, we can analyse the letters in the Turkish alphabet according to the geometric shapes they contain Table 3.1.

Table 3.1 Turkish alphabet according to the geometric shapes they contain

			
a	1	1	0
b	1	1	0
c	1	0	0
d	1	0	0
e	1	1	0
f	1	1	0
g	1	1	0

Table 3.1 (Continued) Turkish alphabet according to the geometric shapes they contain

			
ğ	1	1	0
h	1	1	0
ı	0	1	0
i	0	1	0
j	1	1	0
k	0	1	1
l	0	1	0
m	1	1	0
n	1	1	0
o	1	0	0
ö	1	0	0
p	1	1	0
r	1	1	0
s	1	0	0
ş	1	0	0
t	0	1	0
u	1	1	0
ü	1	1	0
v	0	0	1
y	0	0	1
z	0	1	1

It is possible to analyse geometric shapes with a similar way. Understanding their origins gives us information about how to use them. Only in this way hypothesis that we are defending will be understood Table 3.2.

Table 3.2 Geometric shapes according to their origins

			
	1	0	0
	0	1	0
	0	1	1
	1	1	0
	1	1	0
	0	1	1
	0	1	1
	0	1	1
	0	1	1
	0	1	1
	0	1	1
	1	1	0
	0	0	1
	0	0	1
	1	0	1

In order to prove the hypothesis that we defend in this part of the experiment, we put these geometric shapes into the experiment from simple to complex. This hierarchy was created according to the combination of geometric shapes. For example, the following 3 sample test screens can be examined.

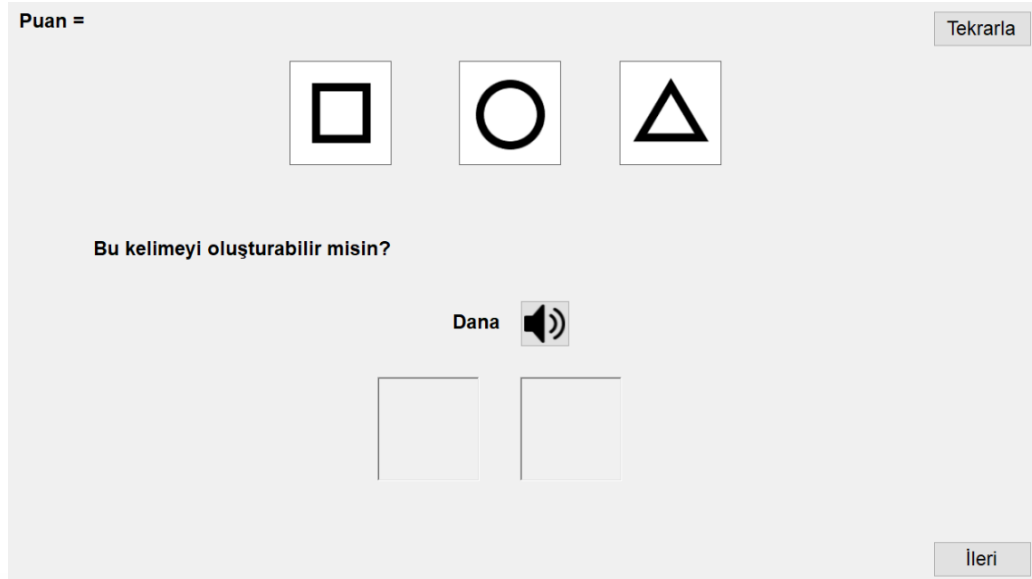


Figure 3.9 First test screen example

The simplest combination is between the three, because the formation codes are very different Figure 3.9.

$$\begin{aligned} \bigcirc &= 1\ 0\ 0 \\ \square &= 0\ 1\ 0 \\ \triangle &= 0\ 1\ 1 \end{aligned}$$

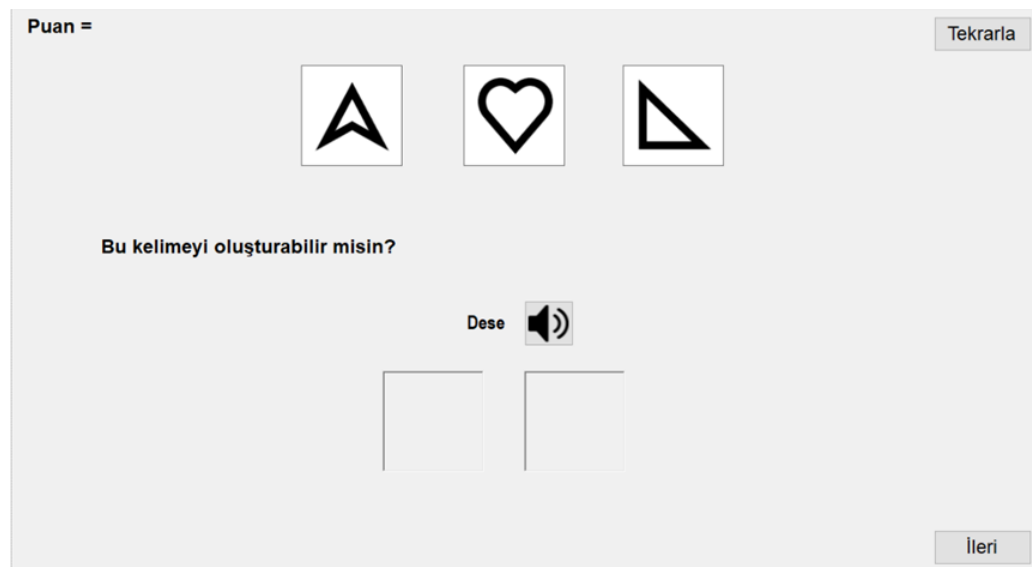


Figure 3.10 Second test screen example

The medium-level combination consists of shapes with more common features Figure 3.10.

$$\begin{aligned} \nearrow &= 0 \ 0 \ 1 \\ \heartsuit &= 1 \ 0 \ 1 \\ \blacktriangle &= 0 \ 1 \ 1 \end{aligned}$$

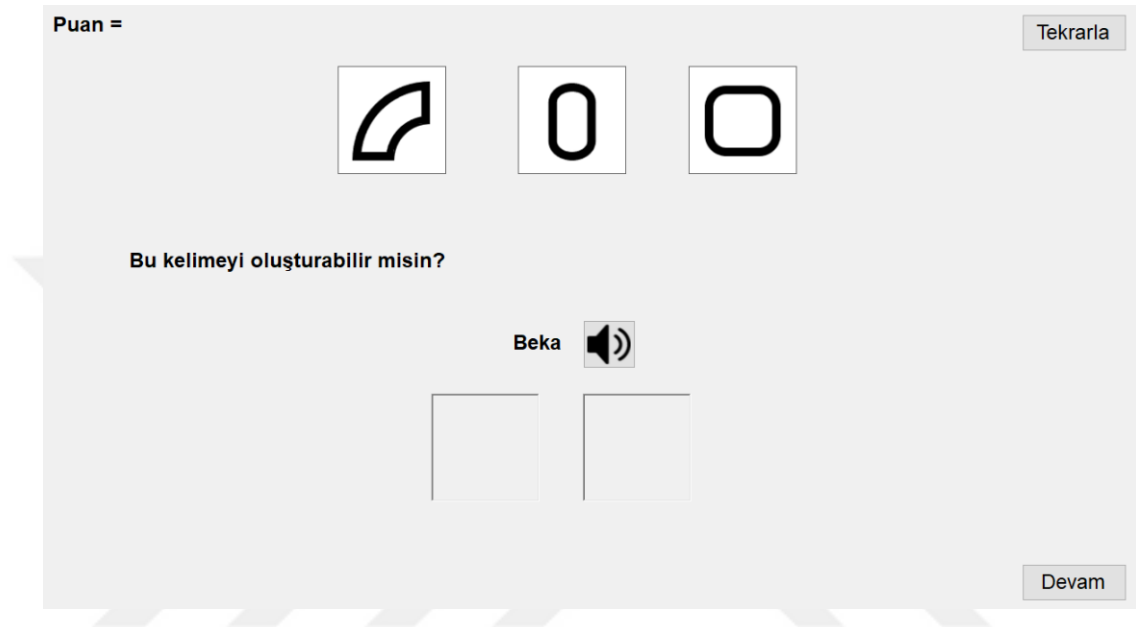


Figure 3.11 Third test screen example

The most complex combination is the closest geometric shapes Figure 3.11.

$$\begin{aligned} \curvearrowright &= 1 \ 1 \ 0 \\ \text{OO} &= 1 \ 1 \ 0 \\ \square &= 1 \ 1 \ 0 \end{aligned}$$



Figure 3.12 Start second parts of experiment

After the instructions on how to proceed the second part of the experiment is shown, the experiment starts with the screen where it is learned whether the participant is ready or not as illustrated in Figure 3.12. After that, participants are asked to read the sentence they saw on the screen. After reading this sentence, a simple question is asked about the content of the sentence. However, after answering this question as yes or no, the participant can switch to another sentence. In this way, participants read 15 sentences and answer 15 questions.

The design of this stage of the experiment is also based on some hypotheses. The first one is that dyslexia can be determined by eye movements. As mentioned in the Background section, there is a link between eye movements and dyslexia.

If a reader is having trouble reading, one can understand this by looking at eye movements. Based on this relationship we have prepared these 15 sentences in which we can analyse the difference between two groups of students. To understand this, we recorded the eye movements of the participants with the help of eye tracker while reading the sentences. The choice of the 15 sentences used in this part of the experiment was not random. These sentences are divided into 3 groups with different levels of difficulty. In order to do this, I have written a code that scales texts according to their dyslexic readability. You can read this scoring algorithm in more detail in the Text Scoring Algorithm section. The sentences of the different difficulty levels selected by this algorithm are as in Table 3.3.

Table 3.3 Sentences of the different difficulty levels

Difficulty Groups for Gunduz Index	Sentences
1- Easy	özge sarı topla oynadı
2- Easy	kemal sakince ikna oldu
3- Easy	gelen mektup kısa cevaplı
4- Easy	önceki koşul baki kaldı
5- Easy	köpek kafası başka yönde
6- Medium	demir derede olta atar
7- Medium	bade banada pudra sürdü
8- Medium	edip depodan buğday getir
9- Medium	aşçı mera menüsü açmış
10- Medium	atıl altı alet atıldı
11- Hard	sonunda asetonu asitle temizledim
12- Hard	destansı donatı dostlukla bitti
13- Hard	doğuş aşağıya doğru yuvarlandı
14- Hard	eminönünü gezenler memnun ayrıldı
15- Hard	bombeli bambulardan pompa yaptık

3.4.4 Experiments

Experiments with students with dyslexia were performed in "Ankara Dyslexia Association". The students who participated in the experiment were assured that they had no problems that could cause reading difficulties other than dyslexia.

Non-dyslexic student experiments were done at "Ankara Mutlu Primary School". The families of the students filled out a form before the experiments. This form measures whether students have difficulty reading. Data of students suspected of reading difficulties were excluded.



CHAPTER 4

ANALYSE

4.1 Diagnosis of Dyslexia

4.1.1 Data

We have seen how this data is obtained in the method section. The data set collected from the experiments is summarized in the Table 4.1.

Table 4.1 Dataset as a result of experiment

Trial ID	Sentence	TimeStamp	Eye	X	Y	Box
73	5. Cümle.	11:05:39.2817	Right	0,4213	0,6520	box2
74	5. Cümle.	11:05:39.2817	Left	0,4507	0,5917	box2
75	5. Cümle.	11:05:39.3507	Right	0,4099	0,6454	box2
76	5. Cümle.	11:05:39.3507	Left	0,4322	0,5844	box2
77	5. Cümle.	11:05:39.4232	Right	0,3643	0,6376	box2
78	5. Cümle.	11:05:39.4232	Left	0,3853	0,5649	box2
79	5. Cümle.	11:05:39.4948	Right	0,3014	0,6160	box2
80	5. Cümle.	11:05:39.4948	Left	0,3351	0,5513	box2
81	5. Cümle.	11:05:39.5682	Right	0,2680	0,6222	gap1
82	5. Cümle.	11:05:39.5682	Left	0,2990	0,5499	gap1
83	5. Cümle.	11:05:39.6382	Right	0,2466	0,6242	box1
84	5. Cümle.	11:05:39.6382	Left	0,2790	0,5500	box1
85	5. Cümle.	11:05:39.7125	Right	0,2526	0,6286	gap1

Table 4.1 (Continued) Dataset as a result of experiment

Trial ID	Sentence	TimeStamp	Eye	X	Y	Box
86	5. Cümle.	11:05:39.7125	Left	0,2788	0,5500	gap1
87	5. Cümle.	11:05:39.7889	Right	0,2526	0,6254	gap1
88	5. Cümle.	11:05:39.7889	Left	0,2802	0,5484	gap1
89	5. Cümle.	11:05:39.8589	Right	0,2515	0,6257	box1
90	5. Cümle.	11:05:39.8589	Left	0,2796	0,5484	box1
91	5. Cümle.	11:05:39.9343	Right	0,2463	0,6224	box1

As seen in the Table 5, the participant's eye movement data and some other important information were used as dataset. To understand eye movement data, the computer screen should be considered as a 2-D coordinate plane. Box is the word that the participant is looking at the screen. For example, box1 represents the first word and gap2 means the second space.

These data were collected in 15 different sentences for each participant. The data of each sentence is kept in different worksheets. As the worksheet 16, the finishing times and the correct numbers of the experiment are kept Table 4.2.

Table 4.2 The other dataset as a result of reading experiment

Student ID	Start Time	Finish Time	True Count	False Count
51	11:04:59.7001	11:07:41.5087	13	2

4.1.2 Preprocessing

The data obtained from the experiment contain raw eye movements. In other words, numerical data indicating which coordinate the eye is on the computer screen measured 90 times a second. It is not logical to use this numerical data without

converting it into eye fixations and gaze durations. This raw data must be pre-processed to create a meaningful data set.

The parameters that are important in the analysis of eye movement data were mentioned in the Background section. When dyslexic and normal readers' eye tracking data are examined, it is known that dyslexic children show more saccade and long-term fixation during reading [39]. Saccades and Fixations two important information to be used in the analysis.

A code was written using a python to extract this information from the raw data. For example, Euclidean distance was calculated between two consecutive coordinates to determine the fixation points. If these two points are too close to each other, it is called fixation point. This function also returns the fixation duration information Figure 4.1.

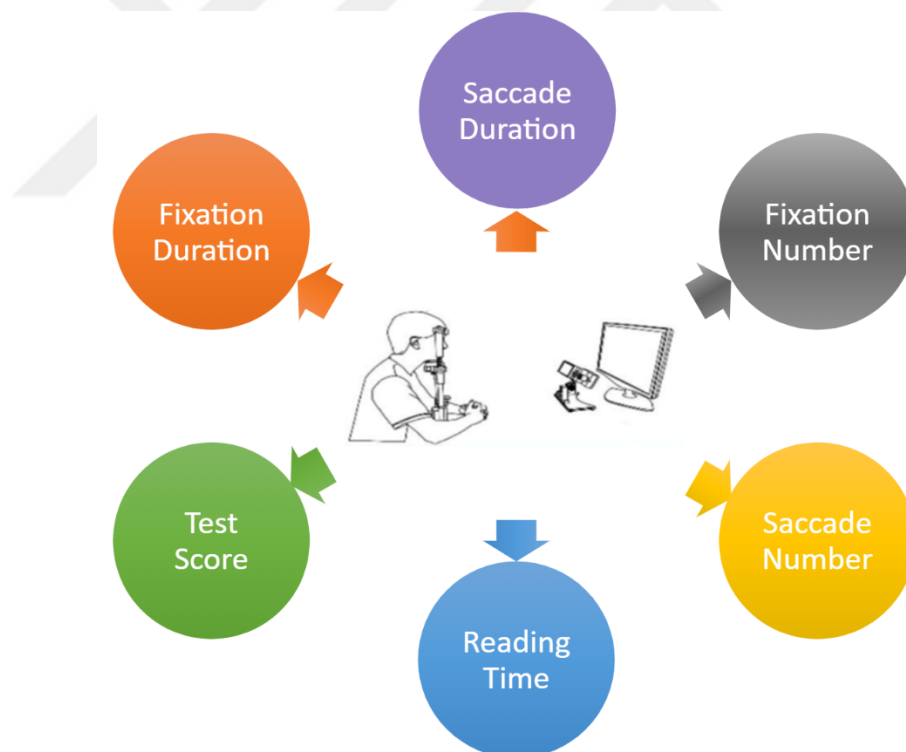


Figure 4.1 Experiment preprocessed data

Another function is detecting saccade points and saccade durations. To do this, it calculates velocity and acceleration in addition to Euclidean distance. It compares

these results with a threshold value and determines whether there is a saccade in the reading. For use in analysis, there is more information that needs to be added to the data set. These are shown in the Table 4.3 as well as reading, sentence number, student number and label. Label indicates which group the student is included in. If 1 is dyslexic and 0 is non-dyslexic group.

Table 4.3 Preprocessed eye tracking dataset

Fixation Number	First Fixation Duration	Fixation Duration	Saccade Number	Saccade Duration	Reading Time	Sentence Number	Student Number	Label
5	144.539. 990.234.	29.252.9 98.046.8	10	3.434.400. 634.765.62	75	2	1	0
5	77.139.9 90.234.3	2.744.09 9.853.51	7	308.380.02 9.296.875	78	2	10	0
6	900.599. 853.515.	4.895.09 9.609.37	12	549.719.92 1.875	107	2	11	0
8	684.100. 341.796.	2.178.30 0.537.10	12	5.108.500. 732.421.87	119	2	12	0
7	846.199. 951.171.	389.769. 970.703.	9	408.030.02 9.296.875	108	2	13	0
6	949.400. 390.625	399.200. 048.828.	11	4.728.800. 048.828.12	114	2	15	0
7	64.430.0 29.296.8	3.575.40 0.390.62	7	433.630.02 9.296.875	112	2	16	0
4	1.031.19 9.951.17	3.282.19 9.951.17	6	326.429.93 1.640.625	88	2	17	0
13	122.0	3.073.80 0.048.82	18	4.953.699. 951.171.87	146	2	40	1
4	523.500. 244.140.	189.889. 990.234.	14	1.921.200. 439.453.12	62	2	41	1
10	387.0	526.819. 970.703.	9	6.907.800. 048.828.12	132	2	42	1
11	1.097.30 0.048.82	13.257.5 99.609.3	28	15.437.798 .828.125	268	2	43	1
14	303.800. 048.828.	2.428.19 9.951.17	24	4.620.799. 560.546.87	121	2	44	1
5	362.0	10.547.0 01.953.1	5	3.347.599. 853.515.62	93	2	49	1
4	8.192.00 1.953.12	4.399.00 0.244.14	12	3.980.600. 341.796.87	111	2	50	1
16	5.482.99 8.046.87	398.839. 990.234.	22	419.780.02 9.296.875	115	2	51	1

4.1.3 Analysis

Diagnosing dyslexia from the preprocessed data set has now become a binary classification problem. Based on the test results, it should be estimated whether the student is dyslexic or not. It would be wise to do this with machine learning algorithms.

As seen in the studies in the literature, the most successful results on eye tracking data is Support-vector machine algorithm (SVM) Table 1.1. SVM is a supervised machine learning algorithm which can be used for classification or regression problems. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs. Simply put, it does some extremely complex data transformations, then figures out how to separate your data based on the labels or outputs you've defined.

Since students read 15 different sentences, a separate learning model should be created for each sentence. So, we have trained the individual models for all sentences and have received success rates separately.

4.2 Dyslexic Readability Analysis

4.2.1 Algorithm

The aim of this algorithm is to assign statistical scores to the sentences to be used in the experiment. The reason for this we want to see the difference between the two reading groups at maximum level. Therefore, it is wise to use analysed sentences rather than randomly selected sentences.

In Turkey, published books by the ministry of education for first and second grade students using a data set consisting of 100 sentences was created. From this data set, 15 sentences were selected which could make the difference between the two experimental groups the most obvious. The parameters of this selection, as shown in Figure 4.2, are things that can cause dyslexic readers to have difficulty reading.



Figure 4.2 Dyslexic readers hate list

The code written for these purposes consists of 6 basic functions. The first is a function that scores sentences with known readability formulas. As described in the Background section, these formulas generally use parameters such as word length or sentence length.

Another function examines the similarity of the first syllables. For example, if the words like "sadık" and "sabah" are the same as their first syllables, the function gives a high score to the sentence. If the words "sadık" and "sabah" are consecutive, this function gives a higher score to this sentence.

With a similar logic, there is another function for all word similarities in the sentence. This function calculates Levenshtein Distance (Edit Distance) for all words in the sentence. If the words that have high similarities are consecutive, this function gives a higher score to this sentence.

We need to measure for the remaining 3 functions is the co-existence of the letters. There is a Soundex Algorithm that can be easily calculated for English. Briefly, it is a phonetic algorithm that is used to index names by voice, as it is said in English. The general steps of the algorithm are as follows:

- Preserve the first letter of the name and drop all the “a, e, i, o, u, y, h, w” letters.
- b, f, p, v $\rightarrow$ 1
- c, g, j, k, q, s, x, z $\rightarrow$ 2
- d, t $\rightarrow$ 3
- l $\rightarrow$ 4
- m, n $\rightarrow$ 5
- r $\rightarrow$ 6

For example, using this algorithm we can index "Robert" and "Rubert" as "R163" and "R163". This result shows us that these two words are very similar to each other as phonetic. For this reason, it would be logical to edit this algorithm for Turkish.

Table 4.4 Consonants in Turkish

	STRONG		SOFT	
	Continuous	Discontinuous	Continuous	Discontinuous
Labial	f	p	m,v	b
Teeth	s, ş	ç,t	j,l,n,r,z	c,d
Palate		k	ğ,y	g
Throat	h	—	—	—

The fourth function measures the cacophony. Cacophony points to a situation in which there is a mixture of harsh and inharmonious sounds. In Turkish, consonant letters are divided into two according to the degree of tongue-tooth-mouth contact Table 4.4. These are continuous and discontinuous letters. For example, if the consonants in a word are mostly from the same group, it is difficult to pronounce the word. Examples of cacophony in Turkish can be given the words "nesterenleşmiş, koşullaştırılmışlık, olasılıklı". The rules of the algorithm that determine this difficulty are:

- f, ğ, h, j, l, m, n, r, s, ş, v, y, z $\rightarrow$ 1
- b, c, ç, d, g, k, p, t $\rightarrow$ 2
- a, e, ı, i, o, ö, u, ü $\rightarrow$ deleted

After indexing the letters in this way, it counts the same numbers that the word contains. If a word contains more than the same numbers, this function gives a higher score to this word.

The fifth function applies the same logic to the nodes of the letters. These word groups are labial, teeth, plate and throat as seen in the Table 4.4. The rules of the algorithm that determine this difficulty are:

- b, f, m, p, v $\rightarrow$ 1
- c, ç, d, j, n, s, ş, t, z $\rightarrow$ 2
- g, ğ, k, l, r, y $\rightarrow$ 3
- h $\rightarrow$ 4
- a, e, ı, i, o, ö, u, ü $\rightarrow$ deleted

Similar to the fourth function, after indexing the letters in this way, it counts the same numbers that the word contains. If a word contains more than the same numbers, this function gives a higher score to this word.

The last function finds the formal similarities in the word. As mentioned in the Background and Methodology sections, the use of letters with similar shapes makes

reading difficult for dyslexic readers. Therefore, it indexes the letters with similar shapes with the same number.

- b, p, d, h, g, ğ → 1
- m, n, a, e, r, u, ü → 2
- v, k, z, y → 3
- s, ş, ö, o, c, ç → 4
- ı, i, l, f, t, j → 5

Similar to the fourth and fifth functions, after indexing the letters in this way, it counts the same numbers that the word contains. If a word contains more than the same numbers, this function gives a higher score to this word. A brief visualization of the basic functions of this algorithm can be seen in Figure 4.3.

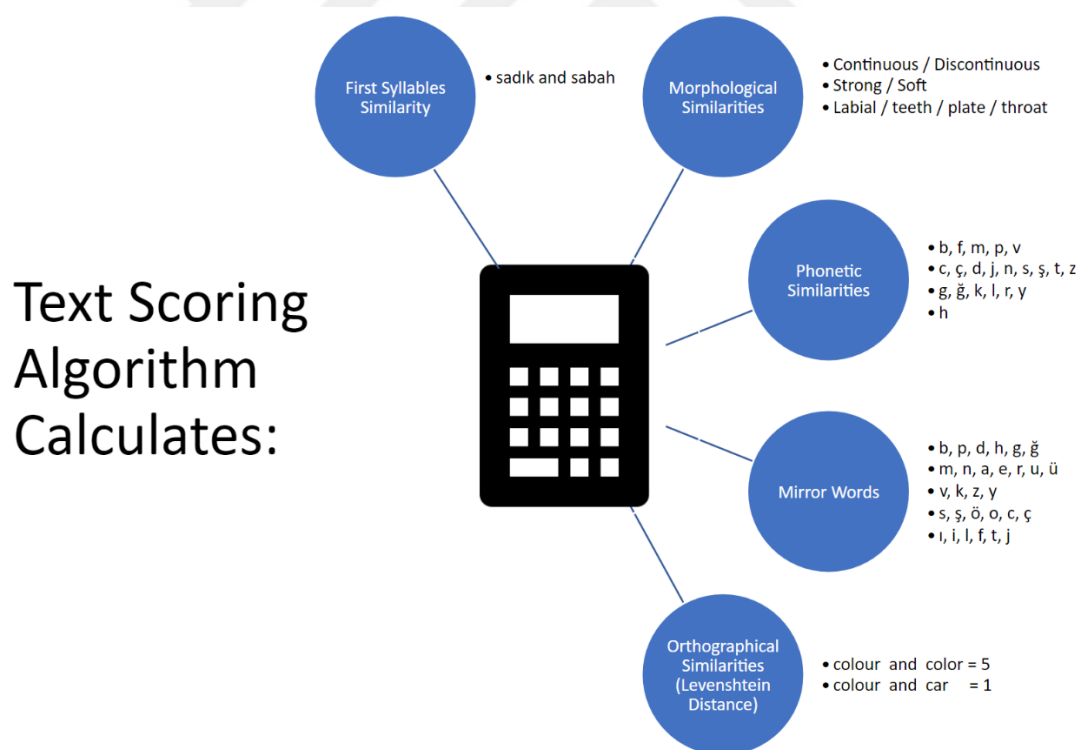


Figure 4.3 Text scoring algorithm functions

4.2.1 Data

As described in the Algorithm section, 15 sentences to be used in the reading experiment were scored. This dataset is as shown in the Table 4.5. Then, eye movement data will be used to verify the results from this algorithm.

Table 4.5 Output of the text scoring algorithm

Sentences	Mirrors	Levenshtein Distance	First syllable	Cacophony	Nodes	Gunduz Index
özge sarı topla oynadı.	0.6	0.48	0.0	0.6	0.2	1.4808
kemal sakince ikna oldu.	0.6	0.48	0.0	0.4	0.4	1.5068
gelen mektup kısa cevaplı.	0.6	0.48	0.0	0.6	0.6	1.5408
önceki koşul baki kaldı.	0.6	0.48	0.0	0.8	0.6	1.5448
köpek kafası başka yönde.	0.8	0.48	0.0	0.8	0.4	1.8328
demir derede olta atar.	0.8	0.48	0.2	0.4	0.2	1.9608
bade banada pudra sürdü.	0.8	0.42	0.2	0.8	0.4	1.9382
edip depodan buğday getir.	0.8	0.48	0.0	0.8	0.6	1.8628
aşçı mera menüsü açmış.	0.8	0.48	0.2	0.6	0.6	2.0248
atıl altı alet atıldı.	0.8	0.28	0.2	0.2	0.2	1.7548
sonunda asetonu asitle temizledim.	0.8	0.48	0.2	0.8	0.8	2.0588
destansı donatı dostlukla bitti.	0.8	0.48	0.2	0.8	0.8	2.0588
doğuş aşağıya doğru yuvarlandı.	0.8	0.44	0.0	0.8	0.8	1.8524
eminönünü gezenler memnun ayrıldı.	0.8	0.48	0.0	0.8	0.8	1.8928
bombeli bambulardan pompa yaptık.	0.8	0.48	0.0	0.8	0.8	1.8928

This data set was based on the hypotheses we defended. The data set should transform to measure the accuracy of these hypotheses. Therefore, we can add eye movement data from the experiment to the data set. Eye movements will be an independent label for the text scoring algorithm. In this way, we can measure how successful the algorithm is.

For each group of students, the average of the experimental results was taken for 15 different sentences. The non-dyslexic ones are shown in Table 4.6 and the dyslexic ones are shown in Table 4.7.

Table 4.6 Average eye tracking data per sentence belong to non-dyslexic readers

Sentences	Fixation Number	First Fixation Duration	Fixation Duration	Saccade Number	Saccade Duration	Reading Time
özge sarı topla oynadı.	0.0	0.0	0.0	0.0	0.0	128.25
kemal sakince ikna oldu.	5.468	730.996	3660.812	11.031	4496.628	109.75
gelen mektup kısa cevaplı.	4.093	877.449	2904.003	8.468	3693.434	76.687
önceki koşul baki kaldı.	4.125	992.937	3135.699	8.281	3817.809	73.218
köpek kafası başka yönde.	3.125	1550.465	3122.234	7.562	3688.815	61.437
demir derede olta atar.	2.968	1663.418	3267.121	7.937	3810.774	62.656
bade banada pudra sürdü.	4.687	1599.931	4789.790	9.875	5178.065	77.937
edip depodan buğday getir.	2.843	2102.734	4007.418	7.875	4081.540	57.656
aşçı mera menüsü açmış.	2.656	1894.131	3613.684	6.593	3561.528	50.468
atıl altı alet atıldı.	3.125	2139.868	4015.721	7.75	4202.696	54.968
sonunda asetonu asitle temizledim.	3.187	2355.040	4523.709	7.031	4137.821	55.281
destansı donatı dostlukla bitti.	3.0	2302.571	4380.684	6.687	4161.831	54.25
doğuş aşağıya doğru yuvarlandı.	2.156	2072.724	3155.428	4.406	2930.687	37.343
eminönünü gezenler memnun ayrıldı.	3.781	2655.828	5755.484	7.468	4708.203	61.75
bombeli bambulardan pompa yaptık.	3.562	2050.084	5553.768	6.75	4449.399	57.281

Table 4.7 Average eye tracking data per sentence belong to dyslexic readers

Sentences	Fixation Number	First Fixation Duration	Fixation Duration	Saccade Number	Saccade Duration	Reading Time
özge sarı toppla oynadı.	0	0	0	0	0	176.363
kemal sakince ikna oldu.	8.727	822.2136	5834.331	14.181	7321.295	150.5
gelen mektup kısa cevaplı.	7.136	1019.540	5102.763	11.909	6710.645	105.545
önceki koşul baki kaldı.	6.590	1303.722	5914.190	11.363	6232.813	100.636
köpek kafası başka yönde.	3.545	1703.668	4029.522	7.5	3854.345	64.772
demir derede olta atar.	4.363	1702.418	5057.795	8.227	4996.104	74
bade banada pudra sürdü.	5.045	2056.668	6567.463	10.136	6080.963	92.545
edip depodan buğday getir.	4.863	1428.640	5075.499	8.909	5516.699	70.363
aşçı mera menüsü açmış.	3.863	2510.913	4875.122	7.136	4213.186	65.090
atıl altı alet atıldı.	3.318	1486.118	3491.140	6.136	3883.954	45.545
sonunda asetonu asitle temizledim.	4.045	2086.968	5110.545	7.181	4849.159	59
destansı donatı dostlukla bitti.	4.409	1374.881	5253.231	6.727	4622.095	55.863
doğuş aşağıya doğru yuvarlandı.	3.318	1946.736	4368.690	5.909	4406.145	47.363
eminönünü gezenler memnun ayrıldı.	4.090	2165.699	5831.954	6.454	4870.863	58.272
bombeli bambulardan pompa yaptık.	3.727	2694.495	5951.181	6.363	4654.890	55.045

In addition, scores from other readability formulas mentioned in the Background section were added to the data set Table 4.8.

Table 4.8 Classic readability scores of sentences

Sentences	ARI	Flesch Reading Ease	Flesch Kincaid Grade Level	Gunning Fog Index	SMOG Index	Coleman Liau Index	LI X	RI X	Atesman
özge sarı topla oynadı.	2.942	12.425	12.52	11.6	8.477	4.6775	4.0	0.0	97.9912
kemal sakince ikna oldu.	4.120	12.425	12.52	11.6	8.477	6.15	29.0	1.0	97.9912
gelen mektup kısa cevaplı.	6.475	12.425	12.52	11.6	8.477	9.095	29.0	1.0	97.9912
önceki koşul baki kaldı.	4.120	12.425	12.52	11.6	8.477	6.15	4.0	0.0	97.9912
köpek kafası başka yönde.	52.974	12.425	12.52	11.6	8.477	7.6225	4.0	0.0	97.9912
demir derede olta atar.	2.942	12.425	12.52	11.6	8.477	4.6775	4.0	0.0	97.9912
bade banada pudra sürdü.	4.120	12.425	12.52	11.6	8.477	6.15	4.0	0.0	97.9912
edip depodan buğday getir.	6.475	12.425	12.52	11.6	8.477	9.095	29.0	1.0	97.9912
aşçı mera menüsü açmış.	2.942	12.425	12.52	11.6	8.477	4.6775	4.0	0.0	97.9912
atıl altı alet atıldı.	17.650	12.425	12.52	11.6	8.477	3.205	4.0	0.0	97.9912
sonunda asetonu asitle temizledim.	15.895	-93.325	27.27	41.6	1.395	20.875	79.0	3.0	47.7725
destansı donatı dostlukla bitti.	13.54	-29.875	18.42	31.6	12.486	17.93	54.0	2.0	77.9038
doğuş aşağıya doğru yuvarlandı.	12.362	-51.025	21.37	21.6	10.745	164.575	54.0	2.0	67.86
eminönünü gezenler memnun ayrıldı.	15.895	-72.175	24.32	31.6	12.486	20.875	79.0	3.0	57.8162
bombeli bambulardan pompa yaptık.	147.175	-29.875	18.42	21.6	10.745	194.025	54.0	2.0	77.9038

4.2.1 Analysis

A new dyslexic readability index (Gunduz Index) was generated from the scores obtained from the functions mentioned in the algorithm section. We used the test results shown in Table 4.6 to see how accurate this index is. For this purpose, we performed correlation analysis and used the Pearson approach.

It is a statistical method used to determine whether there is a linear relationship between two numerical measurements and, if any, the direction and intensity of this relationship. In the case of normal distribution of data, Pearson correlation coefficient and Spearman Rank correlation coefficient are preferred if the data are not distributed normally. In order to interpret a correlation coefficient, p value should be less than 0.05.

If the correlation coefficient is negative, there is an inverse relationship between the two variables. It means one of the variables increases and the other decreases. If the correlation coefficient is positive, one of the variables increases and the other increases.

We did the Pearson Correlation analysis using the Pandas library in Python. We have done this with Gunduz index, which is a self-extracting index, and Atesman, the most popular readability formula for Turkish.

4.3 Shape Sound Connection

4.3.1 Data

The data set used in this section was obtained from the first part of the experiment described in methodology. According to the hypothesis we have defend, 3 different difficulty levels have been solved by students. Student's answers to the questions were recorded during these sections. At the same time, how much time they spent on which part was recorded. You can see a part of this data set created in Table 4.9.

Table 4.9 Result of first part of experiment

Student Number	Label	Stage 1	Stage 2	Stage 3	Total Time	Total
1	0	105359	83682	50835	239876,00	0
2	0	112558	103479	106069	322106,00	2
3	0	59430	57032	46778	253240,00	12
6	0	17993	55378	245440	318811,00	4
8	0	16154	44995	219342	280491,00	5
10	0	8718	81143	143527	217801,00	5
11	0	9864	95680	112257	301228,00	6
12	0	15848	130153	155227	234161,00	4
58	1	60971	74934	76062	211967,00	2
59	1	13211	84918	121135	219264,00	3
60	1	42155	60957	100670	203782,00	7
61	1	2345	995	221075	224415,00	8
62	1	4442	74887	92068	171397,00	5
63	1	4284	16729	193893	214906,00	9
64	1	19354	27329	101307	147990,00	5
65	1	36862	54912	176987	268761,00	6

4.3.2 Analysis

It is known that children with reading difficulties have difficulty in establishing the connection between words and their voices. With this connection, it was one of the hypotheses that the connection between geometrical shapes and sounds was similar. In other words, the hypothesis we are defending would be difficult to connect

between geometric shapes and sounds if a student is having difficulty in linking letters to the sounds of letters.

Therefore, the first thing we tried on the data set was a classification algorithm. We used the test completion times and the number of correct answers as features. We have trained this data with the Support Vector Machine Algorithm Figure 4.4.

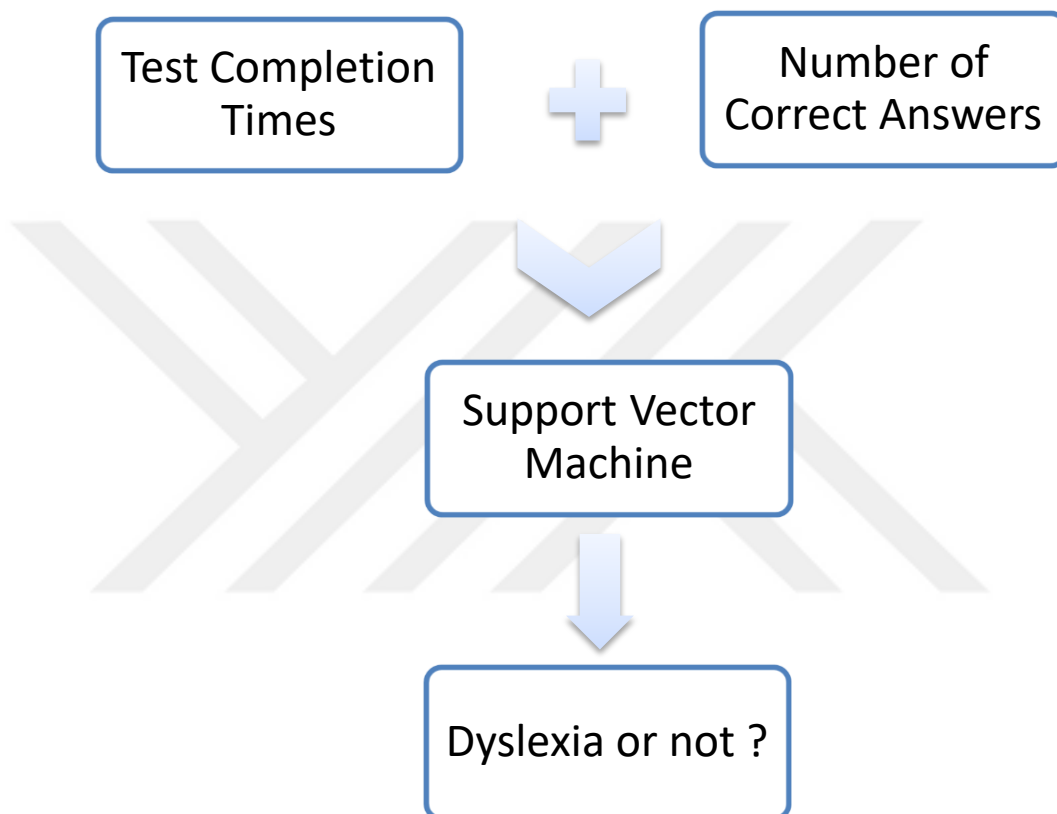


Figure 4.4 Shape and sound connection flow chart

Apart from that, another important information that can be extracted is the percentage of time students spend on which stage. For this purpose, we have normalized 3 different stage times by dividing the total time spent by students. T-test was applied to examine the difficulty hierarchy of geometric shapes defended in the hypothesis.

CHAPTER 5

RESULTS

5.1 Diagnosis Dyslexia

The success rates obtained from models trained on different sentences and different gaze characteristics are given in Table 5.1. Each success rate is taken from the SVM algorithm, which is based on the left-side sentence and above-eye characteristic.

Table 5.1 Accuracy rates different sentences for dyslexia detection from SVM

Sentence	Fixation	Fixation	Saccade Number	Saccade
2	0.909 %	0.818 %	0.727 %	0.909 %
3	0.545 %	0.636 %	0.727 %	0.545 %
4	0.727 %	0.727 %	0.454 %	0.636 %
5	0.363 %	0.545 %	0.636 %	0.636 %
6	0.636 %	0.727 %	0.727 %	0.727 %
7	0.545 %	0.818 %	0.727 %	0.818 %
8	0.727 %	0.636 %	0.727 %	0.545 %
9	0.636 %	0.727 %	0.636 %	0.727 %
10	0.454 %	0.636 %	0.727 %	0.636 %
11	0.727 %	0.636 %	0.636 %	0.818 %
12	0.545 %	0.272 %	0.545 %	0.545 %
13	0.454 %	0.454 %	0.272 %	0.272 %
14	0.636 %	0.545 %	0.545 %	0.363 %
15	0.545 %	0.545 %	0.545 %	0.545 %

5.2 Dyslexic Readability Analysis

The correlation between Gunduz Index scores in Table 9 and the results of the reading experiments mentioned in Table 4.6 and Table 4.7 was examined. The values obtained as a result of the Pearson correlation analysis for this purpose are given in Table 5.2.

Table 5.2 Correlation between scores and experiment results

	Fixation Number	Fixation Duration	First Fixation Duration	Saccade Number	Saccade Duration
Gunduz Index Dyslexia	R = -0.03 P = 0.90	R = 0.60 P = 0.018	R = 0.83 P = 0.0001	R = 0.08 P = 0.75	R = 0.43 P = 0.10
Atesman Dyslexia	R = 0.18 P = 0.51	R = -0.19 P = 0.51	R = -0.43 P = 0.10	R = 0.25 P = 0.35	R = 0.03 P = 0.89
Gunduz Index Non- Dyslexia	R = -0.24 P = 0.38	R = 0.36 P = 0.18	R = 0.76 P = 0.001	R = -0.19 P = 0.49	R = -0.0007 P = 1
Atesman Non-Dyslexia	R = 0.04 P = 0,86	R = -0.47 P = 0.074	R = -0.61 P = 0.015	R = 0.16 P = 0.55	R = -0.17 P = 0.54

5.3 Shape Sound Connection

The SVM algorithm was used only to determine whether the participant was dyslexic or non-dyslexic over the test completion times. The success rate was 72.7% according to the results of SVM algorithm which was used to detect dyslexia. Apart

from this, the rates of spending of the participants are shown in Table 5.3 for each stage.

Table 5.3 Percentage of time students spend on which stage

Difficulty Stage	Dyslexia	Non-Dyslexia
1	10.201 %	18.451 %
2	22.465 %	32.597 %
3	67.331 %	47.725 %

T-test was applied to examine the difficulty hierarchy of the geometric shapes defended in the hypothesis. Whether the time spent in each difficulty stage had a significant effect on dyslexia was investigated. The results obtained in this context are as in Table 5.4 and Table 5.5.

Table 5.4 Group Statistics

	Label	N	Mean	Std. Deviation	Std. Error Mean
Total Time	0	29	247363,828	45788,8165	8502,7698
	1	22	219511,568	37798,4735	8058,6616
Total True Number	0	29	6,66	2,979	,553
	1	22	6,73	2,434	,519
Stage_1	0	29	,1845	,23522	,04368
	1	22	,1036	,08926	,01903
Stage_2	0	29	,3260	,17045	,03165
	1	22	,2237	,11660	,02486
Stage_3	0	29	,4773	,22326	,04146
	1	22	,6727	,16724	,03566

Table 5.5 Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Total Time	1,129	,293	2,315	49	,025	27852,2594	12029,8225	3677,4259	52027,0929
			2,378	48,605	,021	27852,2594	11714,9102	4305,4213	51399,0975
Total True Number	1,380	,246	-,092	49	,927	-,072	,780	-1,640	1,495
			-,095	48,692	,925	-,072	,758	-1,597	1,452
Stage_1	4,963	,031	1,529	49	,133	,08090	,05292	-,02545	,18724
			1,698	37,822	,098	,08090	,04765	-,01557	,17737
Stage_2	1,346	,252	2,415	49	,020	,10225	,04234	,01716	,18734
			2,540	48,561	,014	,10225	,04025	,02135	,18314
Stage_3	2,378	,129	-,3435	49	,001	-,19540	,05688	-,30970	-,08110
			-,3573	48,998	,001	-,19540	,05468	-,30529	-,08551

CHAPTER 6

CONCLUSION AND DISCUSSION

As mentioned in the introduction, there are 3 hypotheses defended in this thesis. Initially, we did not have a data set to defend these hypotheses. For this reason, hypotheses were created by conducting literature research in different disciplines.

To be able to comment on the hypotheses created, a data set was needed first. A set of experiments was created with reading and matching sections. A desktop application was coded so that this set of experiments could be implemented on the computer screen. Experiments were made using this desktop application and an eye tracker integrated with this application.

In the other parts of the thesis, the students who participated in these experiments were defined as dyslexic reader or non-dyslexic reader. But there is a detail that should not be forgotten at this point. The group we have defined as dyslexic reader may have difficulty in reading not only because of dyslexia but also for other reasons. For this reason, it would be more accurate to say this group to students with reading difficulties. But this detail does not produce any contradictions with the hypotheses that the thesis advocates.

The data obtained from the experiments were not sufficient for analysis. For this reason, these data were passed through many pre-processing stages. Afterwards, separate data sets were created for each of the defended hypotheses. Evaluating the results by remembering these hypotheses respectively will make the thesis understandable.

Our first hypothesis, as you remember, it is possible to detect dyslexia in children by testing the time spent on the learnability of the connection between geometric shapes and corresponding phonetics (H1). The data obtained from the experiment for this part of the thesis are shown in Table 4.9.

We can approach these data in two ways. The first one is based on the values shown in Table 5.3. These results show how much of each group students spend their time at which stage. The results of this table are some of the results that can be examined:

The time spent by the dyslexic readers between the difficulty stages varies greatly. However, the results of non-dyslexic readers are more uniformly distributed. Even if it is possible to defend the H1 hypothesis based on this information, a second analysis has been made. In this analysis, the relationship between each difficulty level and the group of participants (dyslexic or non-dyslexic reader) was investigated. In addition, in order to determine whether there is a differentiation between the two groups in terms of correct response numbers and total completion times, they are also included in the data. For this, independent T-test is applied, and the results are shown in Table 5.4 and Table 5.5.

The first parameter to be considered for the T-test is P value (Sig.). If this value is greater than 0.05 there is a normal distribution between the variances. The next parameter to be considered is the Sig (2-tailed) values. If this parameter is less than 0.05 there is a significant difference between the two groups. According to these criteria, the most significant difference between the two groups is seen in stage 3. There was no significant difference in difficulty level 1 and this hypothesis (H1) was supported. In addition, it can be seen that the number of correct answers does not indicate a significant difference.

The last analysis supporting this hypothesis is that we can distinguish these two groups with a success rate of 72%, even by looking at the students' time to finish the experiment (with SVM). We can reach the same result by looking at the value of Sig (2-tailed) which corresponds to the Total Time line in Table 5.5.

In another hypothesis (H2), we proposed that it is possible to detect dyslexia from eye movements in children. In this context, we applied pre-processing steps to convert the raw experiment outputs (Table 4.1) into meaningful data. As a result, the data from Table 4.3, which consists of more meaningful data, has been reached. After this stage, we obtained the answer to our question, which turned into a classical

classification problem, by applying SVM. As a result of this process, the results obtained are shown in Table 5.1.

The first thing that can be said based on the results in Table 5.1; it is possible to detect dyslexia in children from their eye movements with the highest 90% accuracy. This result confirms the second hypothesis (H2) mentioned in the Introduction.

The other thing that can be extracted from this result is which gaze characteristic is more successful in separating the two groups. However, there was no significant difference between these 4 characteristics. The exact result is that there is a big difference between these two groups for the 4 characteristics (Table 4.6 and Table 4.7).

Of course, there is another reason behind this high success rate. As it is seen in the Table 5.1, every sentence is not a perfect separator. For example, the last 5 sentences are good examples of this situation. It is wise to remember the last hypothesis (H3) we have been defending to understand what is causing it.

According to the latest hypothesis (H3), we can develop a formula/algorithm that is more successful than the current readability formulas that computes the dyslexic reading difficulty for Turkish texts. For this purpose, a scoring algorithm was developed which took into account the dyslexic parameters on the sentences planned to be taught in the experiment. The results of this algorithm and the Gunduz Index, which is the new dyslexic readability index obtained from these results, are shown in Table 4.5.

Of course, this new index should be verified. For this purpose, recorded eye movements were used while reading scoring sentences (Table 4.6 and Table 4.7). Because it was the most realistic solution to understand that a sentence was harder to read than the other sentence. For this purpose, we performed a correlation analysis between Gunduz Index score and eye characteristics data of a sentence and used the Pearson approach. Correlation scores obtained from this result are shown in Table 5.2. The first parameter to be considered in these results is P value. If P value is less than 0.05 there is a significant relationship between the two groups and the amount

of this relationship is R value. If P value is greater than 0.05, R value is not required. Indicates that there is no relationship between the two groups.

When the correlation results in Table 5.2 are analysed in accordance with these criteria, we can say that Gunduz Index has a positive correlation with Fixation Duration and First Fixation Duration information. The other thing that makes these correlation values valuable is that Atesman scores, which is the readability formula for Turkish, cannot establish a meaningful correlation with any eye characteristic. The other thing that makes these correlation values valuable is that Atesman scores, which is the readability formula for Turkish, cannot establish a meaningful correlation with any eye characteristic. In fact, this is not a situation to be surprised. Because Gunduz Index is based on only dyslexic parameters, Atesman is a general readability score. When these two readability indexes are evaluated with the results of non-dyslexic readers, we can see that Atesman makes more significant correlations. All these results show that the current general readability formulas are not sufficient for dyslexic readers. However, Gunduz Index developed in this thesis can successfully score texts for dyslexic readers (H3).

When talking about H2, we said that the last 5 sentences failed to separate two groups. This is because the last 5 sentences are difficult for both dyslexic readers and non-dyslexic readers. We can see this by looking at the score from the Gunduz Index (Table 4.5) and the score (Table 4.8) from Atesman.

In summary, we have already thought that H2 will provide high success in the hypothesis defended in this thesis. Because there are similar studies in other languages. But we share our results with H1 and H2 with great enthusiasm. The first reason for this is that there are no other studies defending these hypotheses for Turkish or other languages. The second reason is the success of these two hypotheses, opening the door to a large research topic. Successful results from H1, for example, promise that the diagnosis of dyslexia can be established in the preschool period. We believe that the success results obtained from H2 will make a great contribution to the choice of books for children, their families and teachers who have difficulty reading.

REFERENCES

- [1] G. R. Klare, Readability. Handbook of reading research, ed. P.D Pearson. New York: Longman, 681-744, 1984.
- [2] R. Biedert, J. Hees, A. Dengel, and G. Buscher. A robust realtime reading-skimming classier. In Proc. ETRA 2012, pages 123130, 2012.
- [3] Kollmorgen, Sepp and Kenneth Holmqvist. Automatically Detecting Reading in Eye Tracking Data. 2007.
- [4] K. Kunze, Y. Utsumi, Y. Shiga, K. Kise and A. Bulling, “I know what you are reading”, Proceedings of the 17th annual international symposium on International symposium on wearable computers ISWC ’13, 2013.
- [5] L. Rello and M. Ballesteros, “Detecting readers with dyslexia using machine learning with eye tracking measures”, Proceedings of the 12th Web for All Conference on - W4A ’15, 2015.
- [6] A. Gonzalez-Garduo and A. Sgaard, “Using Gaze to Predict Text Readability”, Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications, 2017.
- [7] M. J. Barrett, Z. Agic & A. Sgaard, The Dundee Treebank. in Proceedings of the Fourteenth International Workshop on Treebanks and Linguistic Theories: TLT14. Association for Computational Linguistics, Warsaw, Poland, pp. 242-248, 2015.
- [8] T. Saikh, S. Bangalore, M. Carl and S. Bandyopadhyay, Predicting source gaze fixation duration: A machine learning approach, 2015 International Conference on Cognitive Computing and Information Processing (CCIP), 2015.
- [9] M. Carl, The CRITT TPR-DB 1.0: A Database for Empirical Human Translation Process Research. AMTA 2012 Workshop on PostEditing Technology and Practice (WPTP-2012), 2012

- [10] A. S. Najar, A. Mitrovic, and K. Neshatian, Utilizing Eye Tracking to Improve Learning from Examples, *Lecture Notes in Computer Science Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge*, pp. 410418, 2014.
- [11] Eivazi, Shahram and Roman Bednarik. Predicting Problem-Solving Behavior and Performance Levels from Visual Attention Data. 2011.
- [12] Y. Lou, Y. Liu, J. Kaakinen and X. Li, “Using support vector machines to identify literacy skills: Evidence from eye movements”, *Behavior Research Methods*, vol. 49, no. 3, pp. 887-895, 2016.
- [13] J. Henderson, S. Shinkareva, J. Wang, S. Luke and J. Olejarczyk, “Predicting Cognitive State from Eye Movements”, *PLoS ONE*, vol. 8, no. 5, p. e64937, 2013.
- [14] C. Hulme, M. Snowling, and P. Quinlan, ‘Connectionism and learning to read: Steps towards a psychologically plausible model’, *Reading and Writing*, vol. 3, no. 2. pp. 159–168, 1991.
- [15] M. J. Snowling, ‘Reading Development and Dyslexia’, *Blackwell Handbook of Childhood Cognitive Development*. pp. 394–411.
- [16] M. J. Snowling, “Language skills and learning to read: The dyslexia spectrum”, Editörler: Snowling, M. J., Stackhouse, J., *Dyslexia—Speech and Language* (pp. 1–14). Chichester, West Sussex: Whurr Publishers, 2006.
- [17] J. Nijakowska, *Dyslexia in the foreign language classroom*. London: Multilingual Matters, 2010.
- [18] E. Erkan, I. Kızılaslan, S. Yıldırım, “A Case Study of a Turkish Dyslexic Student Learning English as a Foreign Language”, *International Online Journal of Educational Sciences*, 4 (3), 529-535, 2012.
- [19] M. Yildiz, K. Yildirim, S. Ates, T. Rasinski, “Perceptions of Turkish Parents with Children Identified as Dyslexic About the Problems that They and Their

Children Experience”, Reading Psychology, 33:5, 399-422, DOI: 10.1080/02702711.2010.515907, 2012.

[20] Ministry of National Education. “İlkogretim Turkce dersi ogretim programi ve kilavuzu [The elementary school Turkish class teaching program and guidance]”, Ankara, Turkey: Government Books Press, 2005.

[21] B. Baydik, “Okuma guclugu olan cocukların sozcuk okuma becerileri [Word recognition skills of children with reading difficulties]”, Ankara Universitesi Egitim Bilimleri Dergisi, 7, 29–36, 2006.

[22] A. Bingol, “Ankara’daki ilkokul 2. ve 4. sinif ogrencilerinde gelismisel disleksi oranı [Rate of developmental dyslexia in second and fourth graders in Ankara]”, Ankara Universitesi Tip Fakultesi Mecmuasi, 56, 67–82, 2003.

[23] G. Erden, F. Kurdoglu, R. Uslu, “İlkogretim okullarına devam eden Turk cocuklarının sinif duzeyine gore okuma hizi ve hatalari normlarının gelistirilmesi [Development of grade level norms for reading speed and writing errors of Turkish elementary school children]”, Turk Psikiyatri Dergisi, 13, 5–13, 2002.

[24] S. Ates, K. Yildirim, K., M. Yildiz, “Opinions of classroom teachers and prospective classroom teachers about the learning difficulties encountered in the teaching process of reading and writing”, Elementary Education Online, 9, 44–51, 2010.

[25] K. Yildirim, S. Ates, “Learning disabilities faced during the period of learning to read and write and the necessary measures to be taken”, Paper presented at the Annual Meeting of Affective Education in Action, Adana, Turkey, 2007.

[26] K. Rayner, “Eye movements and the perceptual span in beginning and skilled readers”, Journal of experimental child psychology, 41(2), 211–236, 1986.

[27] G. W. McConkie, K. Rayner, “Asymmetry of the perceptual span in reading”, Bulletin of the Psychonomic Society, 8, 365–368, 1976.

- [28] K. Rayner, A. D. Well, A. Pollatsek, “Asymmetry of the effective visual field in reading”, *Perception and Psychophysics*, 27, 537–544, 1980.
- [29] A. Kennedy, J. Pynte, “Parafoveal-on-foveal effects in normal reading”, *Vision Research*, 45, 153–168, 2005.
- [30] L. Blomert, “The neural signature of orthographic-phonological binding in successful and failing reading development”, *Neuroimage*, 57, 695–703, 2011.
- [31] C. Prado, M. Dubois, S. Valdois, “The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span”, *Vision Research*, 47, 2521–2530, 2007.
- [32] M. P. Bucci, D. Bremond-Gignac, Z. Kapoula, “Latency of saccades and vergence eye movements in dyslexic children”, *Experimental Brain Research*, 188(1), 1–12, 2008.
- [33] S. Hawelka, B. Gagl, H. Wimmer, “A dual-route perspective on eye movements of dyslexic readers”, *Cognition*, 115, 367–379, 2010.
- [34] F. Hutzler, H. Wimmer, “Eye movements of dyslexic children when reading in a regular orthography”, *Brain and Language*, 89(1), 235–242, 2004.
- [35] S. Yiğiter, “Sınıf öğretmenlerinin özel öğrenme güçlüğüne ilişkin bilgi düzeyleri ile özel öğrenme güçlüğü olan çocukların kaynaştırılmasına yönelik tutumları arasındaki ilişkinin incelenmesi”, *Yayımlanmamış Yüksek Lisans Tezi*, Marmara Üniversitesi İstanbul, 2005.
- [36] G. T. Pavlidis, “Eye Movement in Dyslexia: Their Diagnostic Significance”, *Journal of Learning Disabilities*, 18(1), 42-50, 1985.
- [37] J. Hyönä, “The use of eye movements in multimedia learning”, *Learning and Instruction*, 20, 172-176, 2010.

- [38] M. J. Yap, D. A. Balota, “Visual Word Recognition”. Editörler: Pollatsek, A., Treiman, R., The Oxford Handbook of Reading, (pp. 26-43), Oxford: Oxford University Press, 2015.
- [39] C. Medland, H. Walter, J. M. Woodhouse, “Eye moverments and poor reading: does the developmental eye movement test measure cause or effect?”, *Ophthalmic and Physiological Optics*, 30, 740-747, 2010.
- [40] S. Bayram, M. Camnalbur, E. Esgin, “Analysis of dyslexic students’ reading disorder with eye movement tracking”, *Cypriot Journal of Educational Sciences*, 7(2), 129-148, 2012.
- [41] K. Rayner, “Eye movements in reading and information processing: 20 years of research”, *Psychological Bulletin*, 124, 372–422, 1998.
- [42] J. Hyona, R. K. Olson, “Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency”, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6), 14-30, 1995.
- [43] P. Zoccolotti, M. De Luca, E. Di Pace, A. Judica, M. Orlandi, D. Spinelli, “Markers of developmental surface dyslexia in a language (Italian) with high grapheme-phoneme correspondence”, *Applied Psycholinguistics*, 20, 191–216, 1999.
- [44] M. De Luca, M. Borrelli, A. Judica, D. Spinelli, P. Zoccolotti, “Reading words and pseudowords: An eye movement study of developmental dyslexia”, *Brain and Language*, 80(3), 617–626, 2002.
- [45] M. MacKeben, S. Trauzettel-Klosinski, J. Reinhard, U. Durrwachter, M. Adler, G. Klosinski, “Eye movement control during single-word reading in dyslexics”, *Journal of Vision*, 4(5), 388–402, 2004.
- [46] S. Trauzettel-Klosinski, M. MacKeben, J. Reinhard, A. Feucht, U. Dürrwachter, G. Klosinski “Pictogram naming in dyslexic and normal children assessed by SLO”, *Vision Research*, 42, 789–799, 2002.

- [47] M. Lobier, R. Zoubrinetzky, S. Valdois, “The visual attention span deficit in dyslexia is visual and not verbal”, *Cortex*, 48(6), 768–773, 2012.
- [48] L. Rello and R. Baeza-Yates, “Good fonts for dyslexia,” *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS 13*, 2013.
- [49] M. W. Jones, M. Obregón, M. L. Kelly, H. P. Branigan, “Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: An eye-tracking study”, *Cognition*, 109, 389–407, 2008.
- [50] M. De Luca, E. Di Pace, A. Judica, D. Spinelli, P. Zoccolotti, “Eye movement patterns in linguistic a nonlinguistic task in developmental surface dyslexia”, *Neuropsychologia*, 37, 1407–1420, 1999.
- [51] C. Cappa, S. Giulivi, *Dyslang Module 3 Dyslexia across Europe*. Project Report http://dyscovery.research.southwales.ac.uk/media/files/documents/20140116/Module_3.pdf. [Accessed: 1-December-2016].
- [52] F. N. Babür, B. Haznedar, E. Erdat-Çekerek, G. Erçetin, D. Özerman, “Çocuklarda okuma güçlüğü: kelime okuma testlerinin geliştirilmesi”, 19. Ulusal Özel Eğitim Kongresi, Marmaris, Aydın, 2009.
- [53] J. L. McClelland, D. E. Rumelhart, “An interactive activation model of context effects in letter perception: Pt. 1. An account of basic findings”, *Psychological Review*, 88, 375-407, 1981.
- [54] K. R. Paap, S. L. Newsome, J. E. McDonald, R. W. Schvaneveldt, “An activation–verification model for letter and word recognition: The word superiority effect”, *Psychological Review*, 89, 573-594, 1982.
- [55] D. Norris, “A quantitative multiple-levels model of reading aloud”, *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1212-1232, 1994.

- [56] J. Grainger, A. M. Jacobs, “Orthographic processing in visual word recognition: A multiple read-out model”, *Psychological Review*, 103, 518-565, 1996.
- [57] B. Ans, S. Carbonnel, S. Valdois, “A connectionist multiple-trace model for polysyllabic word reading”, *Psychological Review*, 105, 678-723, 1998.
- [58] M. Zorzi, G. Houghton, B. Butterworth, “Two routes or one in reading aloud? A connectionist dual-process model”, *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1131- 1161, 1998.
- [59] D. Norris, “The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process”, *Psychological Review*, 113(2), 327-357, 2006.
- [60] M. Coltheart, K. Rastle, C. Perry, R. Langdon, J. C. Ziegler, “DRC: A dual route model of visual word recognition and reading aloud”, *Psychological Review*, 108, 204- 256, 2001.
- [61] E. D. Reichle, K. Rayner, A. Pollatsek, “The E-Z Reader model of eye movement control in reading: Comparison to other models”, *Brain and Behavioral Sciences*, 26, 445–476, 2003.
- [62] D. C. Plaut, J. L. McClelland, M. S. Seidenberg, K. E. Patterson, “Understanding normal and impaired word reading: Computational principles in quasi-regular domains”, *Psychological Review*, 103, 56-115, 1996.
- [63] I. K. Bjaalid, T. Høien, I. Lundberg, “Dual-route and connectionist models: A step towards a combined model”, *Scandinavian Journal of Psychology*, 38(1), 73-82, 1997.
- [64] “Tobii Pro SDK,” Develop eye tracking applications for research, 25-Aug-2015. [Online]. Available: <https://www.tobiipro.com/product-listing/tobii-pro-sdk/>. [Accessed: 18-May-2019].
- [65] “International Dyslexia Association – ...until everyone can read!” International Dyslexia Association. [Online]. Available: <https://dyslexiaida.org/>. [Accessed: 18-May-2019].

- [67] Rye, J. (1982). Cloze Procedures and the Teaching of Reading. London: Heineman Educational.
- [68] R. Flesch, A New Readability Yardstick, Journal of Applied Psychology, 32, pp. 221-233, 1948.
- [69] M. Coleman and T.L. Liao, A Computer Readability Formula Designed for Machine Scoring, Journal of Applied Psychology, Vol. 60, pp. 283-284, 1975.
- [70] R. Gunning, The Technique of Clear Writing, New York: McGrawHill, 1952.
- [71] G.H. McLaughlin, SMOG Grading - A New Readability Formula, Journal of Reading 22, pp. 639-646, 1969.
- [72] E. Ateşman, Türkçede Okunabilirliğin Ölçülmesi.A.Ü. TÖMER Dil Dergisi, sayı:58, s.171-174, 1997.
- [73] S. Dehaene, Reading in the brain. New York: Penguin Books, 2014.

APPENDICES

Appendix A - Sentences of the different difficulty levels

Table A.1 Sentences of the different difficulty levels (English)

Difficulty Groups for Gunduz Index	Sentences
1- Easy	özge played a yellow ball
2- Easy	kemal was convinced calmly
3- Easy	incoming letter with short answer
4- Easy	we're sitting previous condition
5- Easy	dog head in another direction
6- Medium	demir fishing in the creek
7- Medium	bade powder drove to me
8- Medium	edip bring wheat from store
9- Medium	chef opened the pasture menu
10- Medium	the subcutaneous device was discarded.
11- Hard	I finally cleaned the acetone with acid
12- Hard	equip epic ended in friendship
13- Hard	doğuş Tumbled Down
14- Hard	those who went to Eminönü left satisfied.
15- Hard	we made from curved bamboo pump

CURRICULUM VITAE

PERSONAL INFORMATION

Name Surname : Akın GÜNDÜZ

E-mail : akngndz93@gmail.com



EDUCATION

High School : Ayten Kemal Akınal Anatolian High School

Bachelor : Electrical and Electronics Engineering in Gaziantep University,
06.06.2017

Master Degree : Computer Engineering in Yıldırım Beyazıt University, 28.06.2019

WORK EXPERIENCE

Data Scientist: AVEO Bilişim Mühendislik Proje ve Denetim Danışmanlık LTD.
(2018- Now)

TOPICS OF INTEREST

- Cognitive Science
- Artificial Intelligence
- Computational Linguistics
- Computer Vision