COGNITIVE ASPECTS OF BRAIN-COMPUTER COMMUNICATION: AN IMPLEMENTATION AND EXTENSION OF THE P300 SPELLER PARADIGM

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

BY

UĞUR ACAR

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

MARCH 2018

**COGNITIVE ASPECTS OF BRAIN-COMPUTER COMMUNICATION:
AN IMPLEMENTATION AND EXTENSION OF THE P300 SPELLER PARADIGM**

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

Prof. Dr. Deniz Zeyrek Bozşahin _____________________
Dean, **Graduate School of Informatics**

Prof. Dr. Cem Bozşahin _____________________
Head of Department, **Cognitive Science**

Assoc. Prof. Dr. Annette Hohenberger _____________________
Supervisor, **Cognitive Science Dept., METU**

Assoc. Prof. Dr. Tolga Esat Özkurt _____________________
Co-Supervisor, **Health Informatics Dept., METU**

Examining Committee Members:

Assist. Prof. Dr. Murat Perit Çakır _____________________
Cognitive Science Dept., METU

Assoc. Prof. Dr. Annette Hohenberger _____________________
Cognitive Science Dept., METU

Assoc. Prof. Dr. Tolga Esat Özkurt _____________________
Health Informatics Dept., METU

Assist. Prof. Dr. Umut Özge _____________________
Cognitive Science Dept., METU

Assoc. Prof. Dr. Ali Albert Salah _____________________
Cognitive Science Dept., Boğaziçi University

Date: 15/03/2018

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

Name, Last name: **Uğur Acar**

Signature : ____________

ABSTRACT

COGNITIVE ASPECTS OF BRAIN-COMPUTER COMMUNICATION:
AN IMPLEMENTATION AND EXTENSION OF THE P300 SPELLER PARADIGM

Acar, Uğur
Master of Science, Department of Cognitive Science
Supervisor: Assoc. Prof. Dr. Annette Hohenberger
Co-Supervisor: Assoc. Prof. Dr. Tolga Esat Özkurt

March 2018, 64 Pages

The present thesis aims to contribute to the understanding of the effects of stimulus types and also cognitive aspects on Brain Computer Interfaces(BCIs) to promote the integration between BCIs and applications. In the model of the present thesis, three types of stimuli are presented. The first type are alphabet letters, the second type are symbols for basic needs, and the last type are words for creation of sentences. Technically the present model has two parts, a training phase and a testing (prediction) phase. In the training phase, the user concentrates on one of the stimuli among the entire set of alternative stimuli and the application detects a P300 ERP for that target stimulus which captures the selection of the subject. The training phase is used for understanding the neural patterns of each user and thus, it generates a user-specific training model. After the completion of the training phase, the subject is able to use the application for communication purposes. In the subsequent test phase, the subject carries out two sets of tasks. First, subjects use the three sets of items (letters, icons, and words), and create pre-specified target items, e.g., words (from letters), icon and word sequences. In case of incongruence of the intended and the produced item (by the application), a specific neural ERP signal, N400, is expected, which is statistically analyzed. Second, the cognitive aspects of subjects' performance are investigated with respect to their motivation and mood. Thus, the effect of subjects' current motivation and mood is measured before the experiment and after the trials of each modality and correlated with the success rate of the P300 speller performance. In sum, the proposed thesis aims at combining technical and cognitive aspects of the usage of BCI devices for communication. The analysis on predicting the success rate for the different stimuli types revealed that the P300 speller is performing better for characters and words than icons. Furthermore, The ERP data analysis within the time window of one second after onset of the predicted stimulus showed that there is no statistically significant N400 effect. Furthermore, the obtained results revealed that the current motivation of the participants might have a prominent role during the trials. The results are discussed with respect to the literature.

Keywords: Brain-Computer Interface (BCI), Cognitive Aspects of Brain Computer Communication,
P300 Speller Paradigm, Electroencephalography (EEG), N400 Event Related Potential (ERP)

ÖZ

BEYİN BİLGİSAYAR İLETİŞİMİNİN BİLİŞSEL YÖNLERİ:
P300 HECELEYECİ PARADİGMASININ UYGULANMASI VE GENİŞLETİLMESİ

Acar, Uğur
Master, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Doç. Dr. Annette Hohenberger
Ortak Tez Yöneticisi: Doç. Dr. Tolga Esat Özkurt

Mart 2018, 64 Sayfa

Bu çalışma beyin bilgisayar arayüzlerine(BBA) uyaran tipleri ve aynı zamanda bilişsel faktörlerin etkilerinin BBA ile uygulamaları arasındaki entegrasyonu artırmak amacıyla incelenmesini amaçlar. Çalışmada kullanılan model içerisinde üç farklı tip uyaran bulunmaktadır. Bu uyaran tipleri alfabenin harflerinden, temel ihtiyaçlar için sembollerden, ya da cümle oluşturmak için kelimelerden oluşmaktadır. Teknik olarak, sunulan model öğrenme (sistemin kullanıcıyı tanıması) ve test (sistemin seçilen uyaranı tahmin etmesi) olmak üzere iki kısımdan oluşmaktadır. Öğrenme safhasında kullanıcı alternatif uyaranlar arasından herhangi bir uyarana odaklanır ve sistem P300 uyaran bazlı elektrik potansiyelini algılayarak kullanıcının hangi uyarana odaklandığını tespit eder. Öğrenme safhası herbir kullanıcının sinirsel örüntülerinin anlaşılması için kullanılır. Bu sayede sistem kullanıcı bazlı öğrenme modelleri oluşturur. Öğrenme safhası bittikten sonra, kullanıcı sistemi iletişim amaçlı kullanabilir. Takip eden test safhalarında ise, kullanıcı iki farklı görevi yerini getirir. Bu görevlerden ilki, kelimeler, ikonlar veya harfleri kullanarak daha önceden belirlenmiş hedefleri gerçekleştirmesidir. Bu görevleri gerçekleştirirken istenilen ile seçilen uyaranın farklı olması durumunda tutarsızlık oluşması ve bu durumda N400 olay bazlı potansiyelinin oluşması beklenmektedir. N400 olay bazlı potansiyeli de istatiksel olarak analiz edilir. İkinci olarak, kullanıcının performansı kişinin motivasyonu ve modu bilişsel bilimler açısından değerlendirilir. Sonuçta, kullanıcının o anki motivasyonu deney öncesinde ve herbir uyaran tipi ile ilgili oturum başlamadan önce ölçülecek ve deney sonrasındaki başarı oranıyla ilişkilendirilecektir. Özetle, önerilen tez beyin bilgisayar iletişiminin teknik ve bilişsel yönlerini biraraya getirmeyi hedefler. Farklı uyarı tipleri için tahmin başarı oranındaki analiz sonucu, P300 heceleyicisinin simgelere göre karakter ve kelimeler için daha iyi performans sergilediğini göstermiştir. Bu sonuca ek olarak, tahmin edilen uyaranın gösteriminden bir saniye sonraki olay bazlı potansiyel verisi incelendiğinde istatistiksel olarak anlamlı bir N400 etkisi gözlemlenmemiştir. Son olarak, elde edilen sonuçlar katılımcıların mevcut motivasyonlarının denemeler sırasında belirgin bir rol oynayabileceğini ortaya koymuştur. Tüm sonuçlar literatüre uygun bir şekilde incelenmiştir.

Anahtar Kelimeler: Beyin Bilgisayar Arayüzü (BBA), Beyin Bilgisayar İletişiminin Bilişsel Yönleri, P300 Heceleyici Paradigması, Elektroensefalografi (EEG), N400 Olay Bazlı Potansiyeli

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Assoc. Prof. Dr. Annette Hohenberger and co-supervisor Assoc. Prof. Dr. Tolga Esat Özkurt for their invaluable help in my academic career. Studying with them for the last three years was a great privilege for me.

I also express my deep sense of gratitude to my friend Dr. Balkar Erdoğan for supporting me about P300 speller application design and implementation.

My thanks also go to Dr. Tuna Çakar for sharing his critical ideas about experimental design and experience in EEG analysis.

Last, but not least, I would like to thank my wife Hatice Acar for her understanding and love during the past few years. Her support and encouragement was in the end what made this dissertation possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background and Purpose of the Study

Direct communication between brain and computer has been a matter of interest and investigated since the discovery of the electrical activity of the human brain and the advances of EEG (Allison, 2011). Brain Computer Interface (BCI) is a direct communication path between a brain and a device that enables signals from the brain to direct some external activity, such as control of a virtual keyboard, a wheelchair or a prosthetic limb. The generated signals are directly affected by environmental factors and users' current state of minds with respect to mood, motivation, cognitive workload etc. The understanding of the effects of these cognitive aspects on BCI can help boosting the performance of integration between BCIs and applications. For instance, a wheelchair application can be adapted to detect the subject's mood and respond accordingly.

Recently, advances in BCI devices have made them available to a wider population of end users – not just clinical populations. Such devices are usually based on neurophysiological techniques such as EEG, which allows researchers to develop brainwave related experiments and BCI applications. Currently, there are many mobile EEG devices on the market such as MindWave (Neurosky Inc., San Jose, CA, www.neurosky.com), Avatar EEG Solutions (Avatar EEG solutions Inc., Calgary, Canada, www.avatareeg.com), the system Enobio (Neuroelectrics, Barcelona, Spain, http://neuroelectrics.com/enobio), the EPOC headset developed by Emotiv (Emotiv, Hong Kong, http://www.emotiv.com) and several others. These devices allow researchers to conduct experiments for studying brain function in everyday life. In this study, the Emotiv EPOC headset with 14 saline sensors and 2 reference electrodes is used as a multi-channel mobile EEG data acquisition device. Although Emotiv has some disadvantages in terms of usability (e.g., placing headset correctly) compared to medical EEG systems (Duvinage et al, 2013), it has also many advantages (e.g., wireless, mobile, low cost) in order to conduct BCI experiments. Taylor and Schmidt (2012) evaluated the detection accuracy of Emotiv EPOC by conducting an experiment in which university students were supposed to command a computer with the shown mental action on the computer screen by their thought only. According to their findings, Emotiv EPOC has an acceptable level of accuracy (87.5%) which is increasing over time by more training. Similarly, Badcock et al. (2013) also support this view by comparing Emotiv EPOC and a laboratory-based EEG system. Their study suggests that Emotiv EPOC is a technically valid alternative to measure auditory event-related potentials (ERPs) with respect to other high-cost, laboratory-based EEG systems.

In the experimental paradigm proposed by Farwell and Donchin (1988), a six by six matrix of alphanumeric characters is presented to the subject on a computer screen. The user selects and concentrates on one of the characters while rows and columns are flashed on and off repeatedly. When the selected character is flashed on, a P300 ERP is elicited. At the same time, it is detected through a brain-computer interface device (BCID). The correlation between the intended stimulus and its evoked P300 potential allows researchers to develop applications for direct communication between brain and computers.

P300 speller applications are developed in order to express users' thoughts with mostly three alternative building blocks consisting of letters, words, and icons. The matrix size is usually six by six, which means the number of different alternative inputs is thirty-six since it is the optimum capacity in order to be efficiently predicted through the P300 speller application (Farwell and Donchin, 1988). In order to enhance P300 potential-based applications, researchers focused on several aspects of the paradigm such as flash rate, stimulus type and shape, matrix size, classification and signal enhancement methods. Previous studies demonstrated that rotating and zooming of flashed characters could provide higher accuracies in P300 speller applications than ordinary whitened (flashed) characters with greyed out background (Liu et al., 2010). Similarly, Kennedy et al. (2014) found that emotional pictures elicited higher magnitude P300 potentials as compared to words. Furthermore, in order to obtain successful performance, different machine learning algorithms are analysed and performed on P300 speller applications. For example, Kaper et al. (2004) suggested using Support Vector Machines (SVMs) with Gaussian kernel transformation as classification algorithm.

Furthermore, cognitive aspects of the paradigm such as subjects' motivation, mood, self-satisfaction and cognitive workload have also been investigated. Kleih and Kübler (2013) defined motivation as motivation to help paralyzed people. They investigated the P300 BCI with a high and low motivated healthy group of subjects. They debate that the motivation for helping paralyzed patients does not influence either BCI performance or the P300 amplitude. As a result, they discuss the necessity of further investigation of the motivation effect on P300 amplitude and speller performance. Later, in a recent study, motivation was found as a significant factor that influences speller performance (Baykara et al., 2015).

The present empirical study aims to contribute to the development of communication applications based on the P300 speller paradigm by making the developed application an open source that can be used by other researchers as well and investigating different stimulus sets (characters, words and icons), N400 ERP indicating an incongruent situation (Nieuwland et al, 2006) and cognitive aspects such as motivation or mood level.

1.2 Motivation and Aim of the Thesis

The thesis focuses on the below issues and aims at promoting the progress in this application.

- *Investigation of the P300 speller performance with different stimulus types such as letters, words and icons:*

The first alternative is letters which are selected by users in order to construct words letter by letter. The second alternative is words which enable users to build up sentences word by word. In this case, words are selected from among the most frequently used and most needed ones for (clinical) users. The final alternative is icons that are representations of words or holistic requests. The icons and words are offered in order to allow users to express their thoughts more rapidly. Through the decrease of time for expressing their thoughts, the usability of the system is thus being improved. All three modalities are investigated with respect to speller performance for efficient communication according to the results of the experiments.

- *Investigation of the correlation between incongruity between intended and produced items and N400 potentials:*

While subjects are expressing their thoughts through the BCI system, they try to select a letter, a word or an icon among different options. The produced words, icons or characters have some form and semantics. These forms and semantics elicit several different brainwave patterns. In some cases, it is hard to correct mistakenly selected input or users do not want to recover although the system allows correcting an unintended letter, word or icon. For instance, the user intends to write ‘a’ but he/she eventually writes ‘b’ or he/she intends to write ‘TV’ but the system produces ‘Radyo’. In such kinds of cases, false prediction of targeted stimuli is a candidate for an N400 since it is an unexpected result. Since N400 is a short-lived ERP, it is only expected immediately after the production of an unexpected unit, however, not when a whole sequence has been completed. According to Nieuwland et al. (2006), if the items produced by the device are incongruent with the intended meaning, this is expected to elicit N400 potentials. The correlation between incongruity and the detection of the elicited brainwave pattern makes several contributions to the communication applications developed with the P300 speller paradigm. The results may affect the order and the layout of the words in the six by six matrix. In addition, further syntactic and semantic research based on the P300 speller paradigm may be informed. For instance, one might try adding suffixes and prefixes to the words in the layout.

- *Investigation of the cognitive aspects of P300 speller paradigm:*

In a recent study, motivation was found to be influencing speller performance (Baykara et al., 2015) so the effect of the subject’s current motivation is measured before the experiment and between each blocks. The measurements of subject’s motivation are correlated with the success rates. This means the present study will contribute to the understanding of the subject’s role in the success rate.

To sum up, the following cognitive aspects of the P300 speller paradigm represent three novel themes of the present study:

- Stimulus modality (letters, words, icons)
- N400 as an ERP signature of wrong prediction
- Relation between subjective factors (motivation, mood) on success rate of prediction

1.3 Outline of the Thesis

The remainder of this thesis is organized as follows: Chapter 2 contains a general literature review on cognitive aspects of BCI, EEG, BCI applications, ERPs and the P300 speller paradigm. Chapter 3 gives information on the design and methodology of the conducted experiment. In Chapter 4, an analysis of the data obtained from the experiments is presented and in Chapter 5, the analyzed data is discussed in detail. Chapter 6 sums up the results of the experiments, discusses the limitations of the study, and suggests directions for further studies.

CHAPTER 2

LITERATURE REVIEW

2.1 What are Brain Computer Interfaces?

Nicolas-Alonso and Gomez-Gil (2012) define Brain Computer Interfaces (BCI) as a hardware and software system that enables commanding computers or any other external devices by using the brain's neural activity. According to them, the main goal of BCI is to allow disabled people suffering from neuromuscular diseases to communicate with their environment. In addition, by using computers, these people can perform several tasks such as playing games, reading books, surfing on the internet, communicating with their friend and doing research (Fouad et al., 2015).

BCIs are divided into two categories according to their relative position to the skull such as invasive and non-invasive. Invasive systems require surgery to implant sensors to the brain in order to measure electrical activity. On the other hand, non-invasive systems do not need surgery and measure electrical activity of mental actions from outside the skull (Allison, 2011).

Usually, every BCI system is designed as a machine learning system for detecting certain patterns in the electrical activity of the brain, as measured by the EEG data. Standard BCI software has several consecutive steps to transform electrical activity to a digital command such as signal acquisition to capture EEG data, preprocessing to shape EEG data in a better form to do further processing, feature extraction, classification based on extracted discriminative features and connecting the resulting command to a device or an application (Fouad et al., 2015). In other words, a BCI system allows giving pre-defined commands to an application or a device such as an air conditioner, a robotic device, television, radio, wheelchair, virtual keyboard etc. by interpreting the brain activity (Mark and Wolpaw, 2009).

2.2 EEG

EEG, discovered by Hans Berger in 1929, is the recording of electrical activity of the neurons in the brain using either invasive or non-invasive forms of electrodes (Niedermeyer and Da Silva, 2005). It is used in most of the BCI studies since it is relatively cheap, non-invasive and portable (Hill et al., 2006). On the other hand, EEG signals are highly sensitive to movement of sensors and physiological artifacts such as head movements, eye blinks etc. In order to decrease the noise in the signal and enhance the contact, usually conductive gel is applied between scalp and EEG sensors. Current EEG research focuses on using dry electrodes instead of using conductive gels. In a recent study with dry electrodes, Tautan et al. (2014) investigated the signal quality of dry electrodes and compared them with the gel electrode recordings in different sessions such as closed eyes, open eyes and steady-state visual evoked potentials (SSVEP). Their experiment results imply that there is no direct relationship between signal quality and contact type of electrodes in either dry or gel sensors although lower signal quality is observed in dry sensors. On the other hand, Lopez-Gordo et al. (2014) described that their performance is as good as wet sensors by benchmarking different dry electrode types. They also state that this technology will enhance the usability of BCI devices in several novel cases of usage. Furthermore, Mathewson et al. (2017) compared dry EEG electrodes to wet electrodes in terms of single-trial noise, average ERPs, scalp topographies, ERP noise, and ERP statistical power. Although the dry electrode system could measure ERP components reliably comparing to wet electrode system, they find that dry systems show very high interelectrode impedance resulting in increases in average noise levels. This implies that a dry electrode system needs more trials in order to capture significant effects.

2.3 Cognitive Aspects of BCI

In recent studies, cognitive aspects such as the effect of users' motivation, mood, self-satisfaction and cognitive workload on the success rate are also investigated by researchers besides of hardware and software design of BCIs. For instance, Kleih and Kübler (2013) discuss the necessity of further investigation of the motivation effect on the amplitude of P300 event related potentials (ERP's) and P300 speller performance. They grouped subjects with respect to perspective taking that is an ability to understand other people's mental states. Before the experiment, they made presentations about BCI and the possibility to help disabled patients to subjects. According to Kleih and Kübler, lower ability for perspective taking implies lower empathy, and vice versa. Their experiment results showed significantly higher P300 potentials on parietal electrodes in participants with lower empathy. Therefore, they are speculating on the possibility that subjects with higher empathy are paying less attention to the BCI task because of their emotional involment to the need of disabled people to BCI.

Later, in a recent study, motivation was found to be influencing speller performance (Baykara et al., 2015). In this study, users' motivation was measured during each experiment session. Results show that there is a direct relationship between users' motivation and P300 amplitude but it is not specific to BCI. It is the result of users' attention to a given task. In another study, Kathner et al. (2014) found that P300 amplitude was lower in high cognitive workload or fatigue conditions. According to their results, further researching of online detection of fatigue or high workload conditions by detecting specific signal patterns is encouraging. Furthermore, Daly et al. (2015) measured subjects' satisfaction level rated from zero to ten on the Visual Analogue Scale (VAS) after each experiment session. Their results indicate that subjects are satisfied when the BCI responses are expected according to their mental commands. Otherwise, they feel frustration every time the system predicts incorrect output. Thus, further studying cognitive aspects of BCI is a necessity for advancing the P300 BCI with respect to performance and functionality.

2.4 Brain Computer Interface Flow

In Figure 4, the typical building blocks of a BCI system and the information flow between them are visualized.

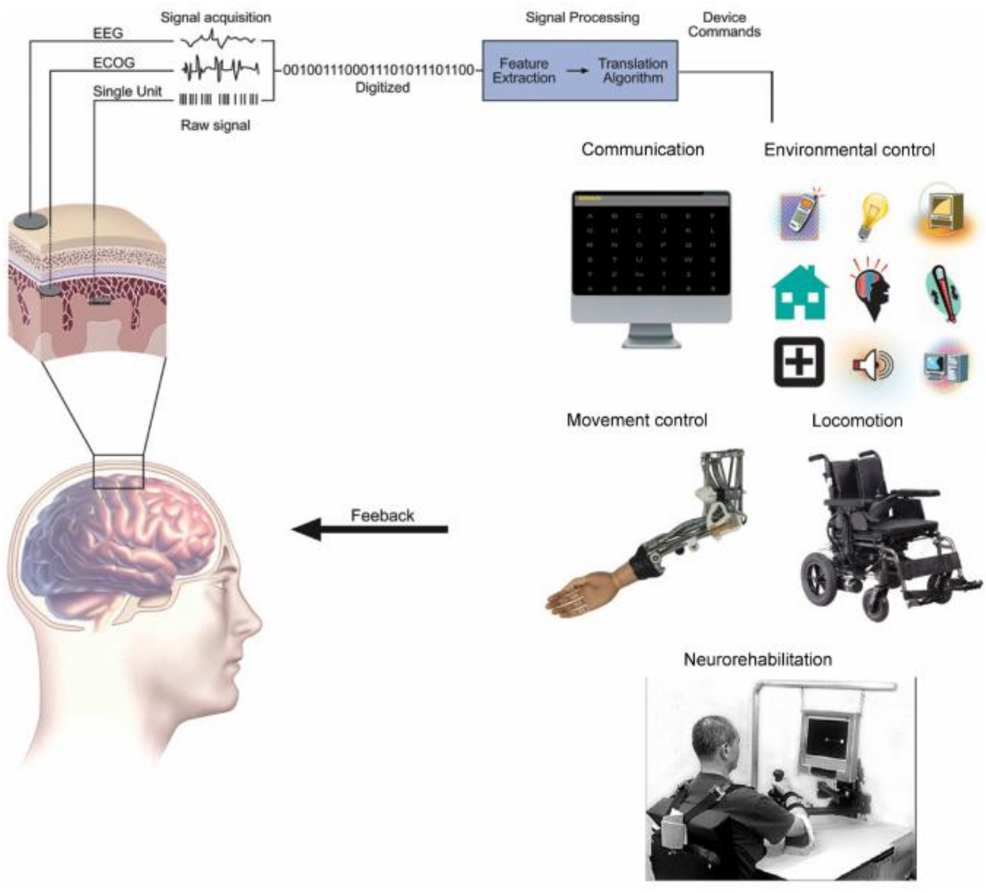


Figure 1 The brain computer interface flow (Mark and Wolpaw, 2009, p.191)

The first step is the signal acquisition phase, which is the measurement of electrical activity of the mental processes/actions through sensors. The sensors can be placed inside the skull by surgery or outside of the skull. According to the sensors' position, measurement methods are divided into invasive and non-invasive. EEG and magnetoencephalography (MEG) are non-invasive measurement methods. Electrocorticography (ECoG), single micro-electrode (ME), micro-electrode array (MEA), and local field potentials (LFPs) are invasive methods. The second step is signal processing which contains two sub-steps such as feature extraction and translation algorithm. In the feature extraction phase, firstly unnecessary and redundant information (such as muscular activity) is removed (by reducing noise in the signal, selecting channel, etc.), which is called as preprocessing step. Afterwards, the most important features implying the intent of the user are extracted in order to maximize the classification accuracy. Finally, the extracted features, which are amplitudes or latencies of event related potentials such as P300, N400, are used in a translation algorithm to map onto a specific device command. In this case, a wide variety of devices can be commanded such as a wheelchair, robotic arm, environmental controls, computer etc. (Mark and Wolpaw, 2009).

2.5 Event Related Potentials (ERPs)

Event-related potentials (ERP) are the measured brain response that is the direct result of a specific sensory, cognitive, or motor event (Luck, 2014). ERPs include specific signatures such as the P300, N400 potentials and Steady State Visual Evoked Potentials (SSVEP), among many others.

2.5.1 P300 Signals

The P300 is a wave pattern elicited in the process of decision making which occurs after the presentation of a rare audio or visual event (Wolpaw et al., 2002) (figure 2). It is observed nearly 300ms after the stimulus onset which names the signal pattern. It is usually detected when the subject is focusing attention on a specific stimulus by using the oddball paradigm, in which rare target items are encountered with respect to frequent but irrelevant (non-target) items (Farwell and Donchin, 1988).

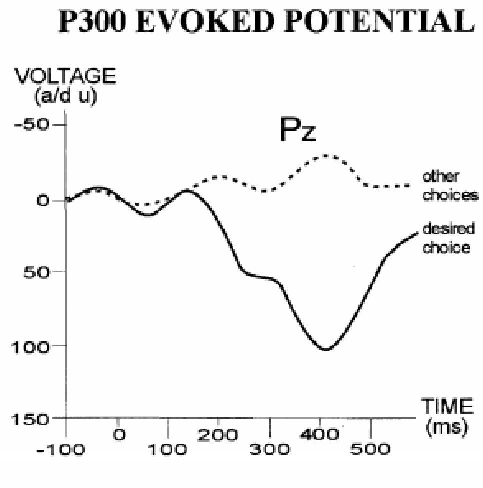


Figure 2 Positive potential 300 milliseconds after the flashed stimulus (Wolpaw et al., 2002, p.774)

2.5.2 N400 Signals

The N400 is a negative wave that peaks around 400 milliseconds post-stimulus onset related to processing meaningful stimuli including visual and auditory words, sign language signs, pictures, faces, environmental sounds, and smells (Kutas and Federmeier, 2000). On the other hand, when an unexpected stimulus or some semantic anomaly is presented, but with a suitable discourse context, this elicits an N400 potential. “The girl comforted the clock” and “The peanut was in love” are examples of such anomalies (Nieuwland and Berkum, 2006). Similarly, Holcomb (1993) states that semantically inappropriate words elicit N400 signals (figure 3).Hagoort et al. (2004) also observed that interpreting the meaning of a semantically incongruous expression elicits N400 effect in the brain.

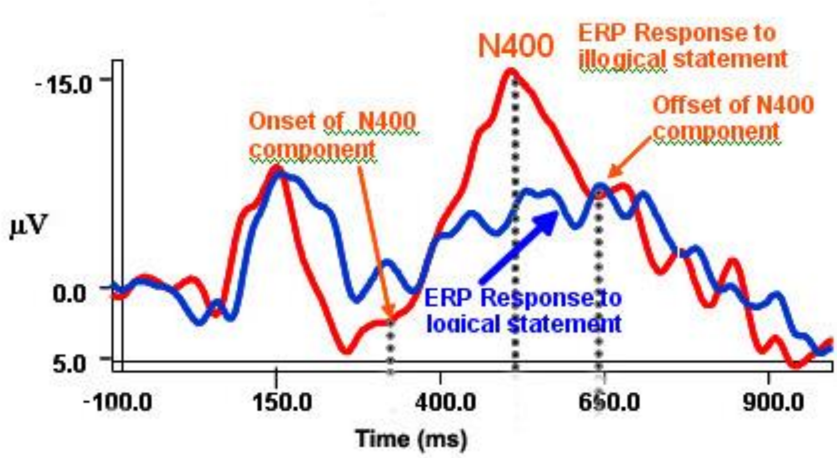


Figure 3 N400 event related potential (National Research Council of Canada, 2005, p.1)

2.5.3 SSVEP

Steady State Visually Evoked Potentials (SSVEP) are signals elicited from visual areas of the brain after presenting visual stimuli at specific frequencies. SSVEP is used in BCI applications by the presentation of several flickering light sources or flickering images on the computer screen with different frequencies (Vialatte et al., 2010). Therefore, an SSVEP-based BCI application can be developed by the detection of the focused light sources from these signal patterns. For example, Muller et al. (2015) analyzed SSVEP-based BCI applications in order to command a robotic wheelchair through low/high frequency SSVEP. Their preliminary results show that disabled people can use robotic wheelchairs more comfortably with high frequency stimuli. On the other hand, Koo and Choi (2015) state that low stimulation frequencies show higher SSVEP detection on virtual reality head-mounted displays. In another study, Wang et al. (2015) propose a hybrid BCI paradigm based on P300 and SSVEP in order to enhance P300 classification. Their results indicate almost 20% increase in SSVEP classification by using the hybrid BCI paradigm. Thus, SSVEP-based BCIs are advancing by ongoing research on different cases of usage.

2.6 Emotiv EPOC BCI Device for ERP Detection

The Emotiv EPOC is a low-cost BCI device to measure EEG signals. It has 14 channels (Figure 7), and a 128-Hertz (Hz) sampling frequency, which is the number of oscillations of the perpendicular electric and magnetic fields per second (Figure 4).

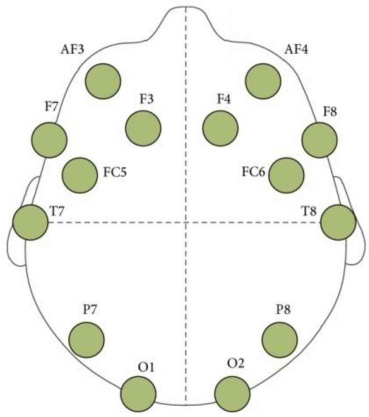


Figure 4 Localization of electrodes in the Emotiv BCI Headset (http://emotiv.com/epoc.php)

The Emotiv EPOC headset can be used for transfering EEG signals to a computer (Lievesley et al., 2011) and is a valid alternative to laboratory ERP systems (Badcock et al., 2013). It is a low-cost headset relative to medical EEG systems. According to Duvinage et al. (2013), it should only be chosen for non-critical applications such as games, communication systems, not for medical purposes. In addition, SSVEP-based BCI through Emotiv EPOC is also implemented and tested by Liu et al. (2012) in order to control environmental devices for disabled people. It is found that the online experiments have acceptable levels of accuracy greater than 90 percent. Similarly, Debener et al. (2012) experimented with the auditory oddball paradigm by using Emotiv EPOC. In their experiments, they classified P300 potentials while their subjects were walking both indoors and outdoors. P300 classification accuracies were sufficiently high in both indoor (77%) and outdoor (69%) environments. Thus, Emotiv EPOC has an acceptable level of accuracy for EEG recordings. However, Emotiv EPOC has only 14 channels implying low spatial accuracy for using it for research purposes.

2.7 BCI applications

In this section BCI applications in the literature will be reviewed. According to Bruner et al. (2011), BCI application types are divided into four main categories such as basic research, clinical/translational research, consumer products, and emerging applications. Their results imply that BCI technology is moving from research to commercial development. Emerging applications can be classified as communication, environmental control, rehabilition and games. P300 ERP-based applications are one of the most popular research areas in which researchers are striving for providing consistent and reliable communication (Gavett et al., 2012). In this application, the user focuses attention on characters in order to construct words. The P300 speller predicts a focused character since the target stimulus is expected to evoke a P300 potential in the brain while other flashes of rows and columns constitute the non-target stimuli and do not evoke a P300. After this process, prediction results of the P300 speller are displayed on the screen. In another study, Rupp et al. (2014) describe that wheelchairs and telepresence robots can also be commanded by BCIs. For instance, Toyota implemented a BCI-controlled wheelchair in 2009. The wheelchair is directed by just imagining the directions such as left, right, foward and backward (Hippe and Kulikowski, 2012). BCI systems are also widely used in games. The first-person shooter (FPS) game implemented by Cho and Lee (2014) is a good example: it that can sense players' emotional state through BCI and change the environment (graphics, sounds, interfaces etc.) accordingly. Another BCI application category is rehabilitation. Neurofeedback and biofeedback-based BCI applications have been advancing as treatment for people with autism spectrum disorder (ASD). Friedrich et al. (2014) designed a game providing neural and pyhsical feedback to the user according to measured signals for ASD. Thus, BCI applications and fields of usage are still evolving.

2.7.1 Potential Users

BCI is described as a translation of human intentions as represented by a neural activation pattern into a control signal without using the muscles, as detailed in previous sections (Section 2.1). In fact, the main objective of BCI is to provide communication and control for people with severe motor disabilities. Research in this field has been rapidly growing in cognitive neuroscience, bioengineering and machine learning. Specifically, this technology is promising for users with motor neuron diseases. Severely disabled people, especially amyotrophic lateral sclerosis (ALS) patients can use a visual P300-based BCI application for communication with an acceptable level of accuracy (Kübler et al., 2001, McCane et al., 2014).

2.8 P300 Speller Paradigm

2.8.1 Different Stimulus Types

Several studies with the aim of enhancing P300 BCI classification performance have been conducted by researchers with respect to stimulus type. Kennedy et al. (2014) analyzed differences between emotional pictures and words with respect to P300 potentials (Figure 5). They applied Repeated Measures ANOVA to P300 amplitude from specific channels. They found a significant main effect of stimuli in which pictures had greater amplitudes than words ($F(1,20)=72.466$, $p<0.001$). Pictures, as a result of visual and emotional processing yield higher magnitude P300 and late positive potentials as compared to words for positive, neutral and negative emotional states.

Positive stimulus	Neutral stimulus	Negative stimulus	Task stimulus
Panda	Water	Snake	House

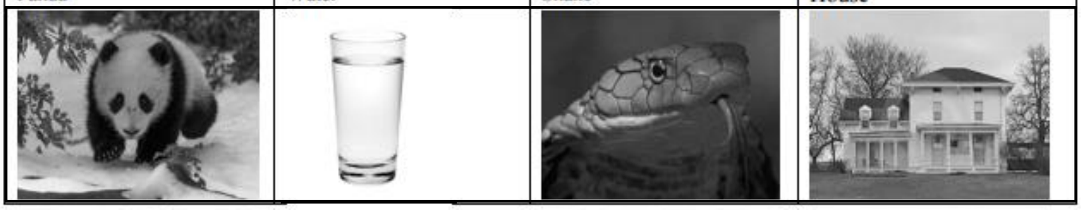


Figure 5 Example words and corresponding images for each emotional category and task stimuli used by Kennedy et al. (2014, p.5)

Another study describes the effect of motion, zoom in/out, rotation and sharpening of stimulus types (Liu et al., 2010). While they presented to subject specific stimulus types, results indicate that rotation and zooming types yield better performance than the other ones for most of the subjects (Figure 6).

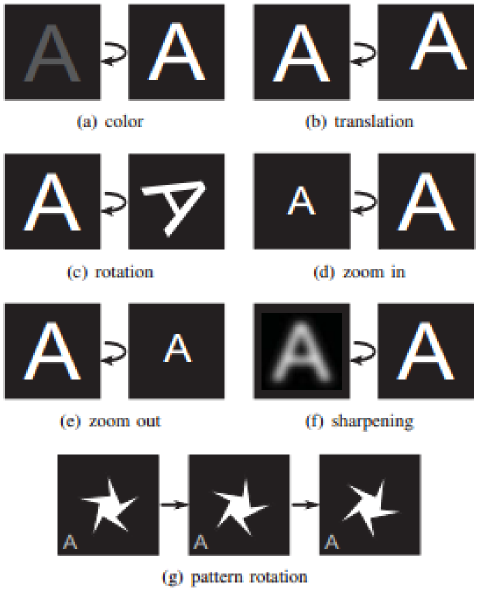


Figure 6 Example stimulus types used by Liu et al. (2010, p.274)

Chang et al. (2013) analysed auditory, visual and audiovisual modalities with respect to P300 responses. They used five Hiragana characters as stimuli in their experimental setups. The result of their experiments showed that auditory, audiovisual and visual paradigms have 52.5%, 91.3% and 95.0% mean accuracies respectively. They state that the auditory modality needs to be further investigated in order to be a possible alternative to the visual modality.

Jin et al. (2012) proposed a face paradigm within the P300 speller. They compared alternative ways to change items in a P300 BCI, which is changing characters to faces instead of flashing them as in P300 BCI (Figure 7). Experimental results show that faces produce significantly better performance in classification accuracy than flashing. They found that the face conditions (mean accuracy >91%) resulted in significantly better performance than the flash condition (mean accuracy = 75%).

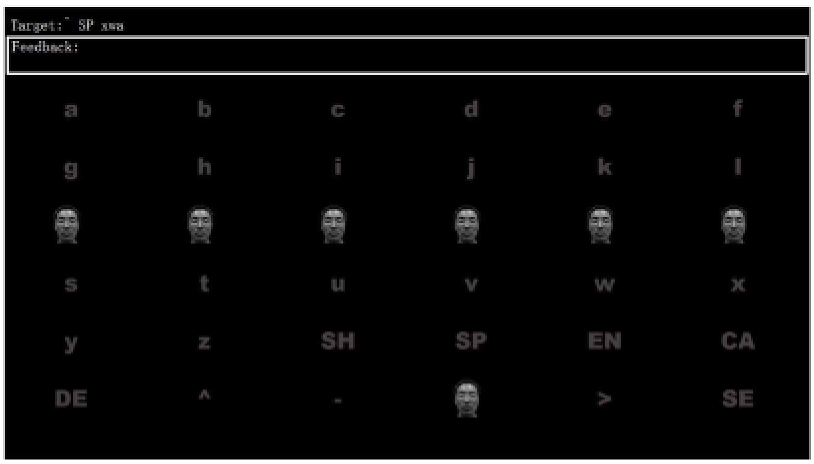


Figure 7 Example changing characters to faces (Jin et al., 2012)

Another study carried out by Takano et al. (2009) investigated a green/blue flashing matrix instead of a white/gray flashing matrix as visual stimuli. In their experiment, the speller matrix was presented in three types of intensification options such as a luminance, a chromatic and both luminance and chromatic condition. The results of their experiments indicate that the green/blue matrix under both chromatic and luminance condition yields better results than other types of intensifications.

In addition, Jirayucharoensak et al. (2011) compared P300 speller accuracy and spending time with respect to picture and character-based modalities. Although subjects spend less time to communicate in the picture-based modality, the average spelling accuracy is slightly higher than in the character-based speller modality.

To sum up, one can expect classification accuracy in the following order (least to most): words, pictures, characters with respect to stimulus type according to these studies.

2.8.2 Experiment Parameters

A six by six matrix is selected for the experiments in the present study, in order to gain best performance for the P300 speller paradigm. This choice is based on Sellers et al. (2006) finding that a six by six matrix with 175 ms inter-stimulus interval (ISI) yields higher P300 magnitude with respect to a three by three matrix. Another study carried out by McFarland et al. (2011) shows that stimulus on and off-times do not essentially influence classification performance. On the other hand, inter-stimulus interval influences the performance significantly. Although the ideal value depends on the subject, longer inter-stimulus intervals are preferred to increase the magnitude of the P300.

CHAPTER 3

METHODOLOGY

3.1 Participants

Participants in the present study are volunteering university students without any disabilities and within the age group of nineteen to twenty-two. Twelve participants from Boğaziçi University attended the main experiment, two of them were female, and the rest was male. Participants with short hair were preferred in order to maintain the headset sensors in contact with the skin. The participants were right-handed and had no psychiatric treatment for the last six months. No previous experience with the experiment or domain knowledge was required.

Prior to the experiments, the participants were informed orally and with a written guideline about the content of the experiment. They were asked to sign the informed consent form after reading it carefully. Demographic information (age, job, education degree etc.) was collected from the participants at the end of the experiment. The study has been approved by Middle East Technical University (METU)'s Ethics Committee [Appendix D].

3.2 Selection of Words and Icons

The selection of words and icons is crucial for this study. According to Glennen and Decoste (1997), selection of appropriate vocabulary requires a comprehensive evaluation of the Alternative and Augmentative Communication (AAC) user with respect to the following domains:

- Profile: age, gender, personal interests
- Communication Environment and Partners: where and with whom he/she will be communicating
- Cognition and Language: developmental skill levels with regard to his/her cognition, language
- Literacy

In the present study, one of our motivations has been to extend the usability of BCI devices by first testing them with normal subjects before attempting to test them with disabled people who lack communication abilities. Thus, disabled people are not part of any experiment in this study. In the experiments, 36 cells containing a stimulus set with different modalities are presented to the subjects. We chose a relatively small number of cells so only core vocabulary are selected as stimuli, comprising typical entries beginning with items related to basic functional needs, brief social exchanges, and other information necessary across environments. Since the participants are university students, who do not have any disabilities, we have the initial assumption that they do not have any limitations in the number and type of vocabulary words they are able to comprehend and express.

Three different sets of stimuli are introduced in the present study. In the first set, participants look at presented letters (Figure 8). In the second set, symbols for basic needs (WC, yes, no, hungry etc.) are used (Figure 9). The third set contains words (subjects, verbs and objects) (Figure 10).

Figure 8 Stimulus Set 1 (Letters)

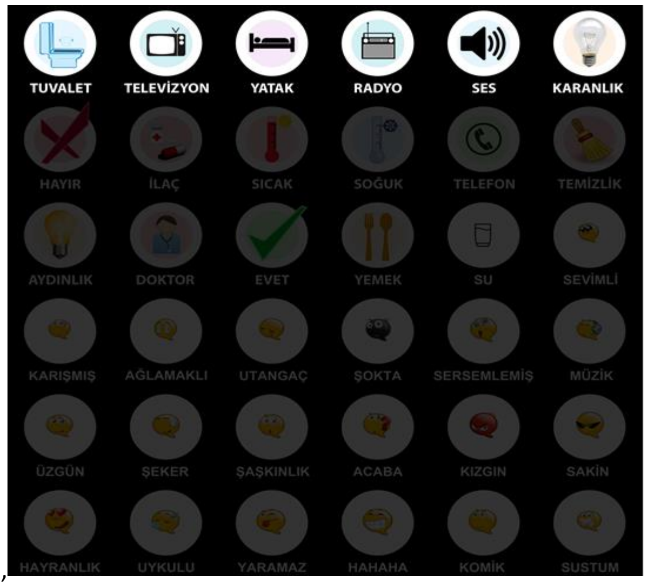


Figure 9 Stimulus Set 2 (Icons)

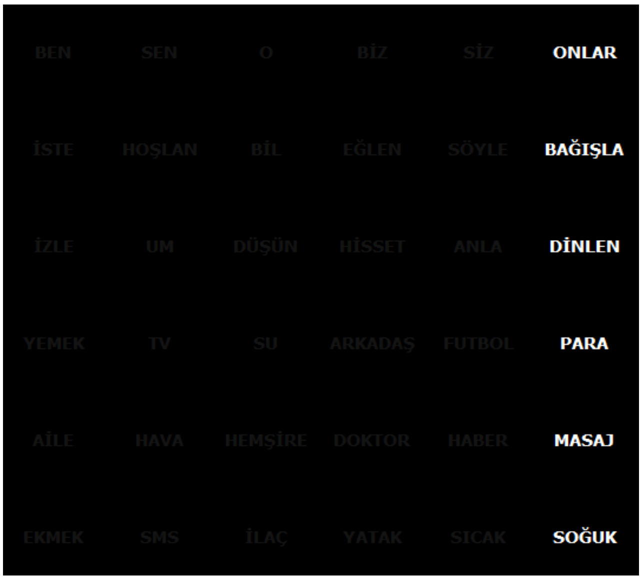


Figure 10 Stimulus Set 3 (Words)

3.3 Questionnaires

In a recent study carried out by Baykara et al. (2015), the effect of motivation, mood and cognitive workload on an auditory P300 speller paradigm was analyzed through questionnaires. Similarly, the present study uses the same questionnaires for measuring the impact of current motivation on the visual P300 speller paradigm.

Before starting each trial in one session (characters, words, and icons), motivation and mood are measured through visual analogue scales (VAS). Participants mark their level of motivation and mood on the VAS, a 10 cm line ranging from 0 (not motivated at all) to 10 (extremely motivated) (Appendix A). Another measurement tool for subject’s current motivation is the Questionnaire for Current Motivation (QCM) designed by Rheinberg et al. (2001) and modified by Nijboer et al. (2008). The QCM includes 18 items rated on a 7-point Likert scale and categorized as mastery confidence, incompetence fear, interest and challenge, (Appendix B). Participants fill the QCM once before starting the session. Both the motivation questionnaire and the items in the visual analogue scales were translated to Turkish by the author.

3.4 Materials and Apparatus

The experiments are carried out in an isolated room in order to eliminate environment noise and distraction of the participant.

The following hardware, software and setup components were used.

Hardware

- ASUS Laptop with 15 "4 inch screen size
- Emotiv EPOC BCI headset for measuring and transmitting EEG data to computer (see Section 2.6).

Software

- Windows XP/7 32bit Operating System
- Custom-developed machine learning software for testing and training the subject while showing varying stimuli on the screen

Setup (Figure 12)

- Distance between user eye and laptop is 50 centimeters (cm).
- Distance between user body and laptop d is 40 cm.

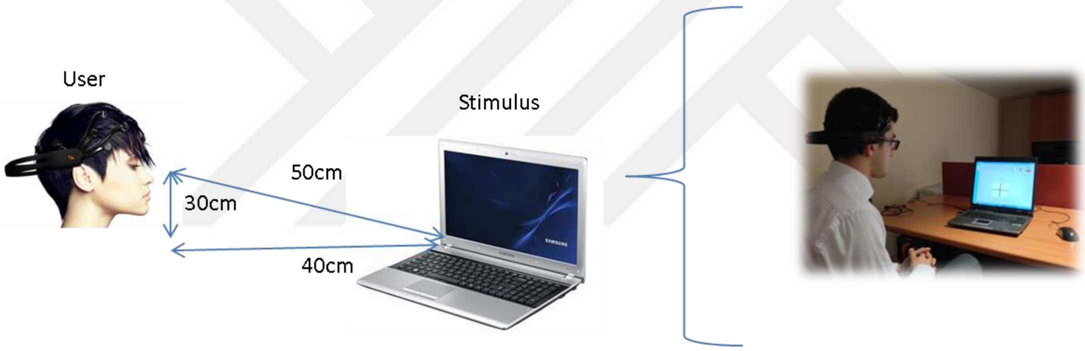


Figure 11 Experiment Setup (The headset picture is available online, https://www.wired.com/2013/03/10-companies-chasing-innovations-that-really-matter/, accessed at 01/13/2018)

3.5 Procedure

The aim of this application is to predict the focused character, icon or words such that the subject expresses his/her thoughts through his/her brain waves by spelling the characters, words or icons on the computer screen. In order to achieve this goal, a six by six matrix of each modality (characters, words and icons) is presented to the subject (see Figure 1-3 for spelling matrix examples) (Farwell and Donchin, 1988). The subject is asked to concentrate on a specific predefined cell in the matrix while the rows and columns of the matrix are flashed on and off sequentially in a random order. The cell concentrated on is called the target cell, the intersection of one row and column. Similarly, the other five rows and columns are called non-target rows/columns. The target cells are expected to evoke P300 ERPs by stimulating subjects visually. On the other hand, non-target cells are supposed to produce different patterns than the target cells. By considering this underlying principle of the P300 speller paradigm, the prediction procedure has two phases such as training and testing. In the training phase, subjects are expected to focus on the predefined cells selected by the application randomly. While subjects are focusing on the cells, their EEG-based brain wave data are recorded. Usually, it is difficult to construct the prediction model in one trial that corresponds to the duration when all the rows and columns of the matrix are flashed on only once. Several trials are carried out in order to decrease any random effect of the user or environmental noise in the signals. At the end of this phase, the subject's prediction model is constructed. In the testing phase, the subject concentrates on the cell he/she wants to type, and the application is supposed to predict the cell (character, icon or word) intended by the participant through the detection of a P300 potential in the brain wave signals by providing the detected signals as an input to the constructed model in the training phase. Again, several trials are performed for the detection of the target character in order to decrease the error in the prediction. At the same time, incongruities are possible in the produced target items, e.g., letters, words or icons. When presented with an anomaly for a given expression, it is expected to evoke a N400 potential. The N400 potential is analysed offline in the collected data at the time when the subject sees the result of the prediction of the system before moving to the next prediction trial. For each modality, training and testing phase is conducted, respectively. The order is selected randomly within the modalities given below (Figure 12):

- Characters modality training and testing
- Words modality training and testing
- Icons modality training and testing

Each subject performs training for each modality (15 repetitions for each character/word/icon). Then the system uses that training data in the testing phase to predict the participant's choice (10 trials for each character/word/icon). In other words, participants have seen a single stimulus 15 times during the training session. The participant is asked to count the number of flashes for the target stimuli in order to understand that the next stimulus will be shown on the screen after the last flash. For instance, if the target character is L, the participant is required to increase the number of counts by one when the character L flashes with the other characters on the same row or column. During the test phase, the participants see each stimulus 10 times because of time limitation.

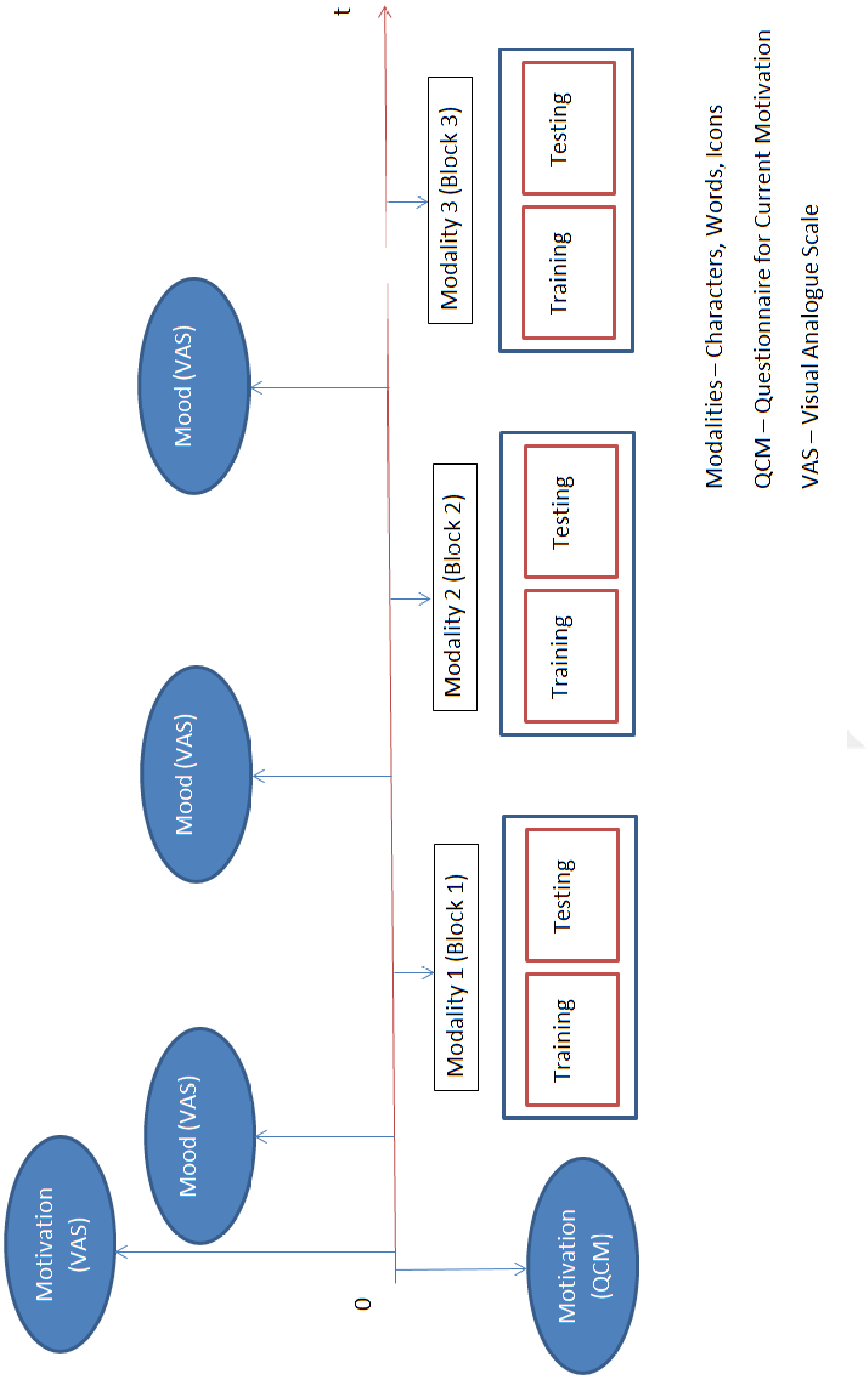


Figure 12 Session Timeline for One Participant

The same sample stimuli of letters, words, and icons are presented to every subject in order to avoid unnecessary inter-individual variation. Thus, from among the 3 stimulus sets 7 letters, 7 words, and 7 icons are selected from the 6 by 6 input matrix. Each modality is presented as one block within which items are randomized, and the order of the 3 blocks is also randomized. At the end of each block, there is an assessment of motivation and mood. Although this assessment may be considered off-line, it is yet in close proximity to the trials (Figure 12). Two criteria such as reducing eye movement effect between each trials and fixing targeted cell positions between each stimuli types are taken into account for selecting the sample stimuli. Thus, they are chosen from the center of the input matrix, which have the same cell positions for all stimuli types and are next to each other.

The sample stimuli of letters, words, and icons are as follows:

- Characters: N, O, P, Q, T, U, V
- Words: UM, DÜŞÜN, HİSSET, ANLA, TV, SU,ARKADAŞ
- Icons: DOKTOR, EVET, YEMEK, SU, AĞLAMAKLI, UTANGAÇ, ŞOKTA

3.6 Data Collection

The developed software (Section 3.8) first collects participants' EEG data, obtains the ERPs and then constructs participants' prediction model using machine-learning techniques that will be discussed in the following section according to the P300 speller paradigm in the training phase. Afterwards, the EEG data set is trained by the constructed model and the software makes online prediction during the testing phase. At the same time, during the prediction (testing) phase, EEG data is recorded for the offline analysis of N400 detection.

EEG data are transferred from subjects to the computer through the Emotiv EPOC BCI device. Data acquisition is done for the following 14 channels: AF3, F7, F3, FC5, P7, O1, O2, P8, FC6, F4, F8, AF4 (Figure 4). Finally, the sampling rate for the data collection is 128 Hz.

3.7 Data Analysis

The present study approaches the P300 speller application software by using similar classification algorithms proposed by Erdogan (2009). One of the purposes of this study is enhancing the efficiency of communication applications based on P300 brainwave potentials. In these applications, an EEG data set is translated into basic actions through machine learning algorithms. In addition, the P300 speller paradigm is commonly developed using supervised machine learning techniques containing training and testing phase. Firstly, a training phase is performed for the purpose of constructing a classification model by instructing the users. Afterwards, the constructed model is used for predicting the target stimulus. Support Vector Machine (SVM) is selected as the main classification method due to its success in BCI competition datasets (Chang and Lin, 2011) and its common use in recent BCI studies (Aguilar et al., 2015; Raju et al., 2015).

SVMs are supervised machine learning models that are used for classification and regression analysis (Cortes and Vapnik, 1995). In order to separate a two-class classification data set as in the present study, infinitely many hyperplanes can be found (Figure 13).

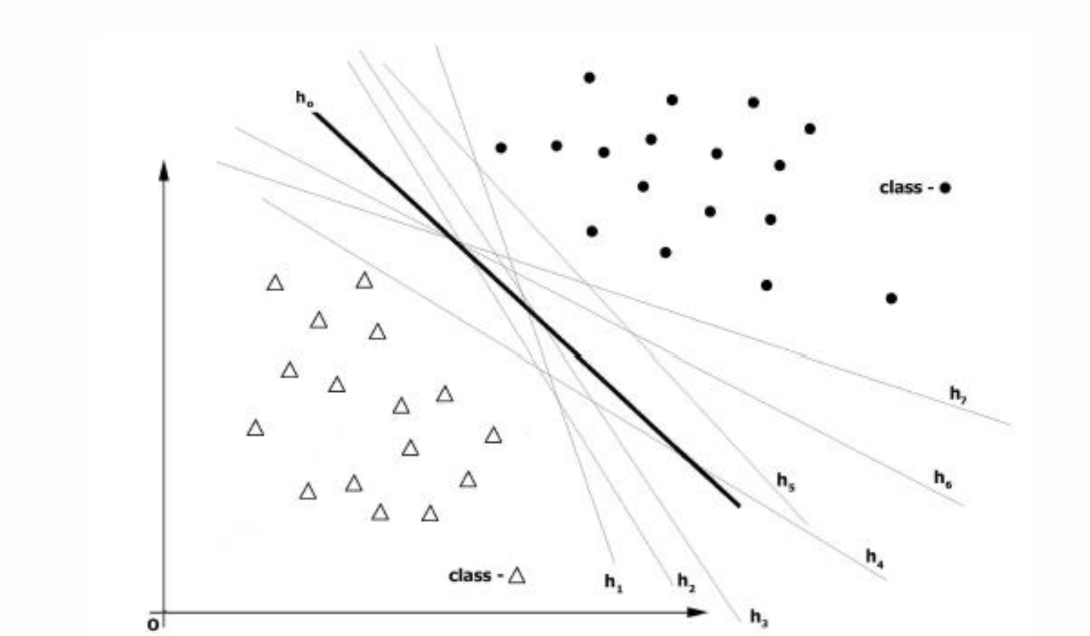


Figure 13 Several possible solutions for a two class classification task (h0, h1, h2, ...). h0 is the result of SVM that maximizes the distance between two classes (Erdogan, 2009, p.64)

Burges (1998) suggests that a good classification is performed by maximizing the distance between the hyperplane and the nearest training data points of any class (Figure 14). This is called as optimum separating hyperplane in SVMs. Intuitively, the larger the margin, the better performance in classification and the lower the generalization error of the classifier results (Figure 14).

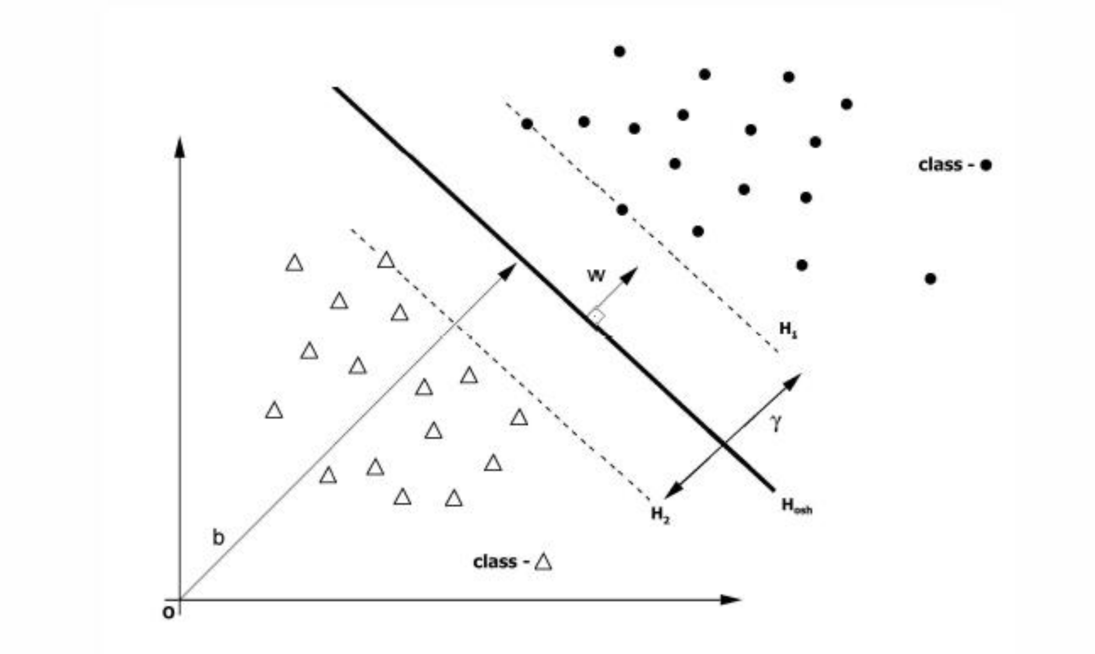


Figure 14 H1 and H2 are support vectors that are the nearest training samples to the optimum separating hyperplane (Hosh) (Erdogan, 2009, p.65)

H1 and H2 in Figure 14 are called support vectors that pass through the nearest training samples to the optimum separating hyperplane (OSH). Once H1 and H2 are calculated, they are used in the testing phase during the prediction. The OSH can be represented as (w.x+b=0) where b is the bias vector and w is the weight vector for the specific data point which is the maximum distance between H1 and H2 for the linearly separable classification of two classes. This means the linear classifier can be characterized by the set of pairs (w,b) that satisfies the following inequalities for any pattern x_i in the training data:

- Class 1 (Circle) if $w.x_i + b > 0$
- Class 2 (Triangle) if $w.x_i + b < 0$

In order to carry out experiments, the P300 speller software was adapted for the purpose of the present study. This application has three components: a signal processor, SVM classifier and a user interface. The signal processor enhances raw data coming from the Emotiv EPOC headset by applying signal filtering and normalization techniques (Section 3.8.2). This unit enhances the brainwave signals by applying averaging and band-pass filters. After enhancing signals, they are accumulated from F3, F4, FC5, FC6, P7, P8, O1, O2, F7, F8, AF3, AF4 channels for either target or non-target stimulus types. These feature vectors are used as observation matrix. In the training phase, each row in the observation matrix is also labelled with two classes as either target or non-target depending on the target intensification of the row/column. The observation matrix with class labels is pushed into libSVM with default parameters from an open source library developed by Chang and Lin (2011) to construct a SVM classification model. Two separate classification models are prepared for rows and columns respectively. In the testing phase, the flow is similar to the training phase except the labelling of the class. The class of the observation data is predicted through the constructed SVM model in the training phase. Additionally, raw data is collected from specified channels for offline analysis in order to analyse N400 potentials while showing the outputs of the prediction to the participants online during the testing phase.

3.7.1 Evaluation of the Outputs

In order to evaluate the outputs of the system, the outputs are coded due to their success of the match between the predicted item and the intended item. However, even though the output may not be a perfect match, it might be matched on the correct row or column. These kinds of matches are called Half-Match (HM). When both of the column and row is predicted correctly then it is a full-match (FM). Finally, when both of these features are missed, then it is a mismatch (MM).

- Full Match: FM (Row and Column are correct)
- Half Match: HM (Row or Column is correct)
- Mismatch: MM (No Row and No Column is correct)

3.7.2 Analysis of Success Rate

The P300 speller application allows users to express their thoughts in terms of patterns of brain activity with three alternative building blocks, namely letters, words, and icons. All three modalities are compared with respect to speller performance according to the number of FMs, HMs and MMs. In this study, the predictive success rate of the developed application is calculated using the following formula:

$$\textbf{Success Rate} = \frac{\boldsymbol{4FM + 2HM + 0MM}}{\boldsymbol{4(FM + HM + MM)}}$$

In the above equation, FM, HM an MM counts are weighted by four, two and zero respectively and summed together. Then this sum is divided by the total counts multiplied by four that are weighted as if all of the predictions are FM's. In addition, HM counts contribute to the success rate since they represent correct row or column classification but the contribution is half of the FM weight. In addition, the effect of the subject's current motivation to the success rate is also analyzed statistically.

3.7.3 Analysis of N400 ERPs

False prediction of targeted stimulus is a candidate for a N400 potential since it is an unexpected result for the participant. The correlation between elicited N400 potentials and the total FMs, MMs conditions are analysed statistically over collected raw data during each prediction. MM conditions include HM conditions since it is indistinguishable for participants whether any predicted item is HM or MM. They only see whether an outcome is correct or incorrect but not whether it is half correct/incorrect.

3.8 Developed P300 Application for Experiments

3.8.1 User Interfaces

The application has two main interfaces. The first interface is settings that users can configure and do tweaks on data acquisition, training, testing and graphical user interface. The second interface is application user interface for testing and training purposes of the P300 speller paradigm.

<u>Settings</u>

A. *Acquisition Settings* is used for storing the user name and configuring one of the three EEG data recording options such as “Record None”, “Record Train Only” and “Record Both Train and Test Data”.

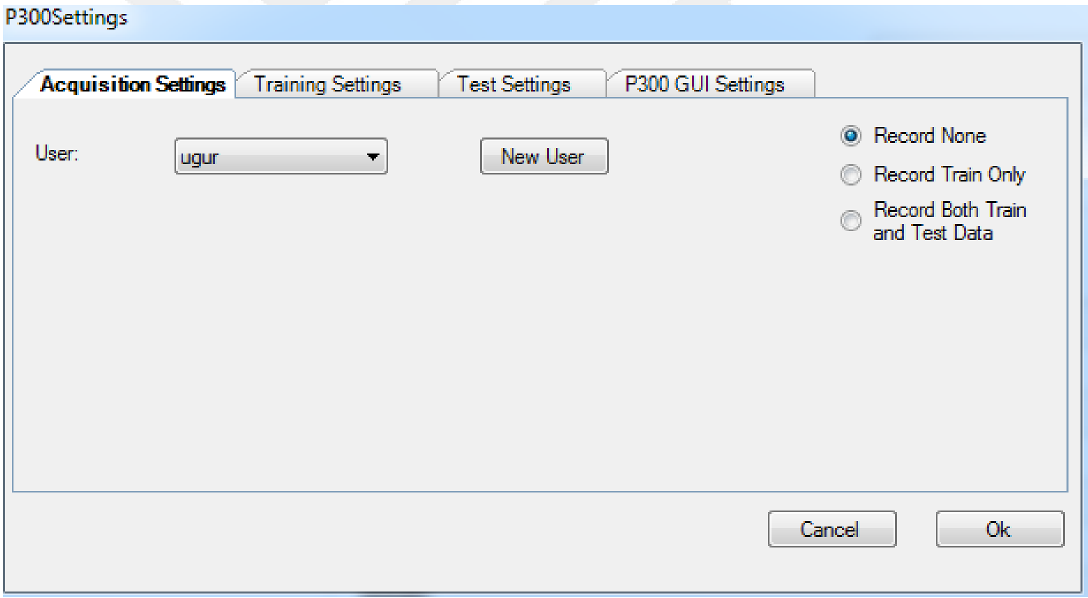


Figure 15 Acquisition Settings

B. *Training Settings* is used for configuring the number of repetitions in each trial and the number of training items. This section of the settings also allows users to choose row and column classifiers that is generated by the system after training and will be used in the testing phase. In order words, one can choose any stored classifier to use in the testing phase.

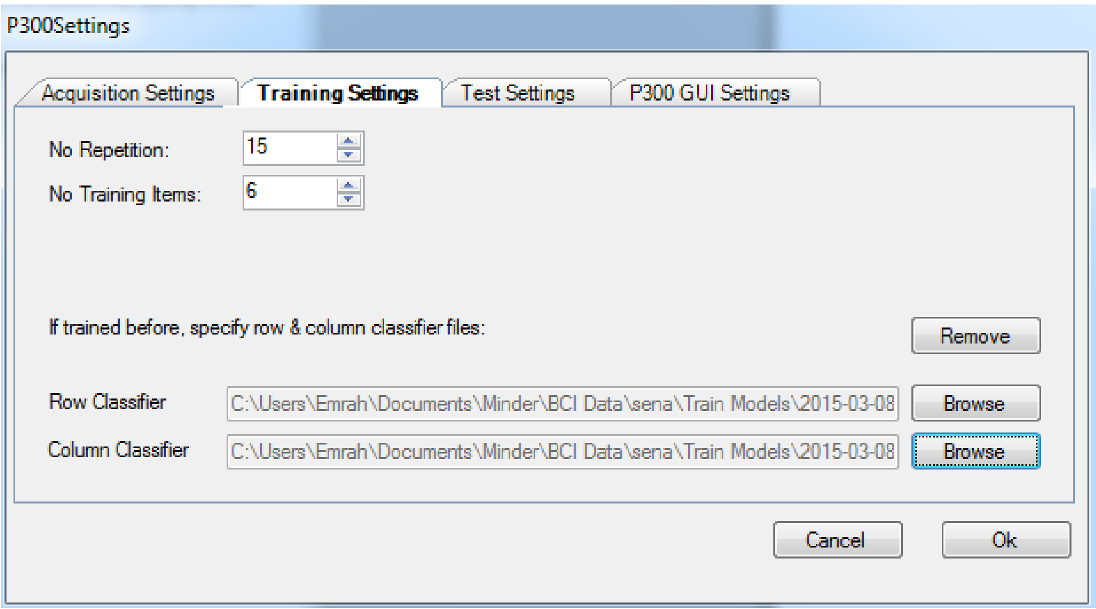


Figure 16 Training Settings

C. *Test Settings* provides configuration options for the number of repetitions of each trial, number of items to be predicted, number of tasks and finally targeted stimuli and their indices in the six by six matrix to be able to compare actual and targeted stimuli.

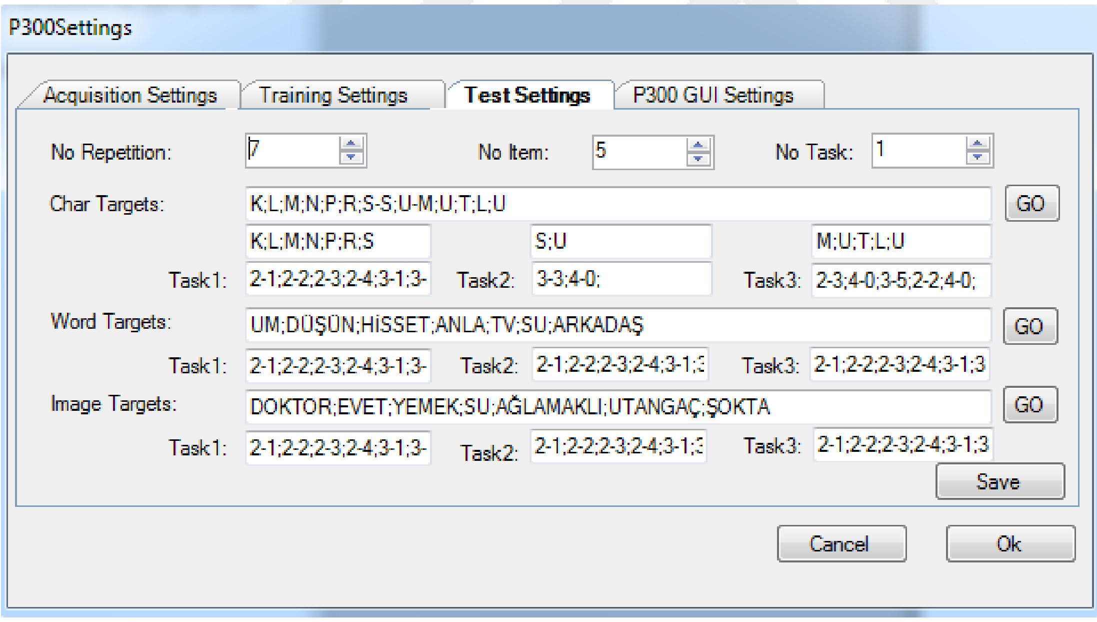


Figure 17 Test Settings

D. *P300 GUI Settings* provides options to change the background color of the stimuli, the selection of modality, flashing on and off duration, delay between each trial, and stimulus (one cell) size in pixels.

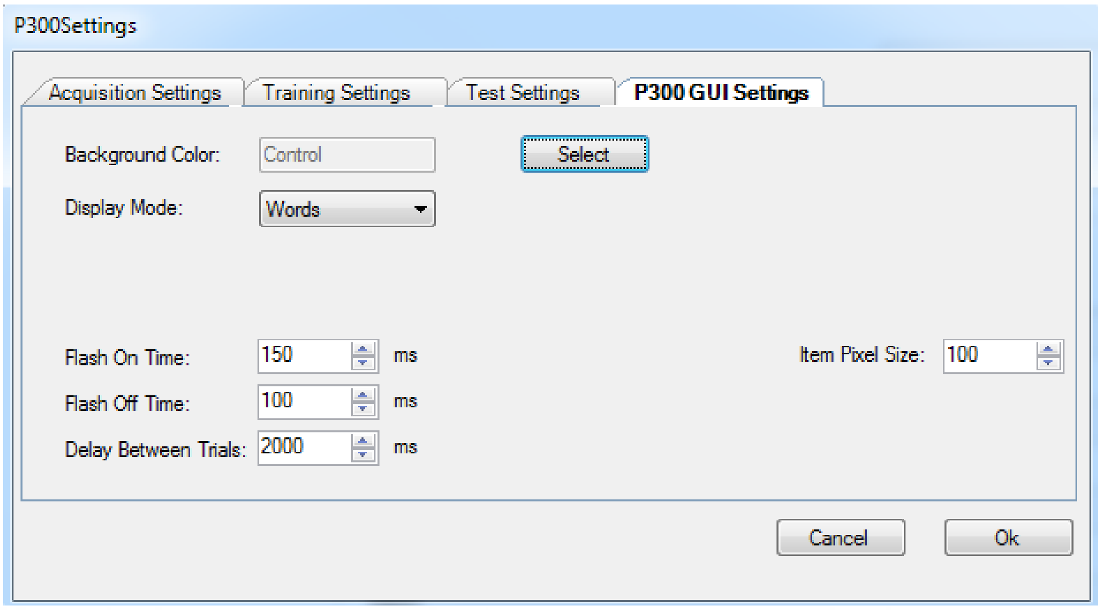


Figure 18 P300 GUI Settings

<u>Application User Interface</u>

A. The *Main Screen* displays connectivity and battery level indicator, Turkish and English language options, realtime EEG monitor, training and testing options.

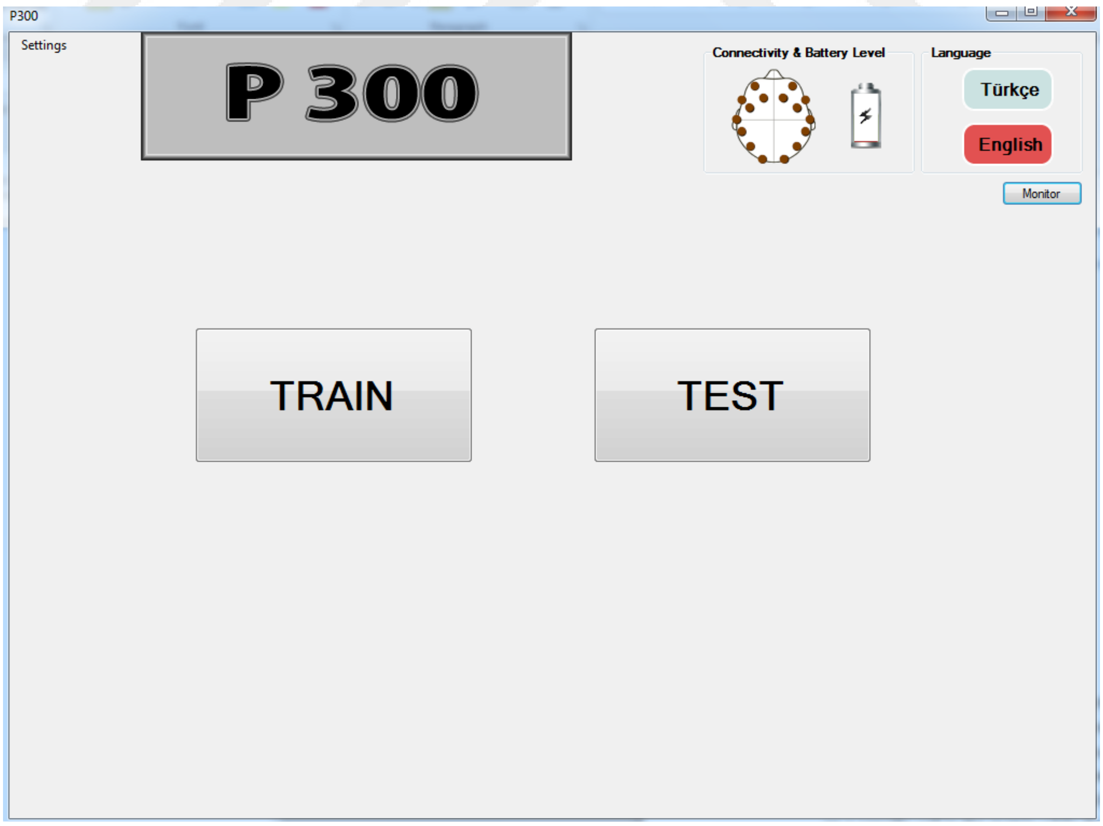


Figure 19 Main Screen

B. The *Monitor Screen* shows near realtime EEG signals measured from all sensors and all other data that Emotiv EPOC Software Development Kit provides.

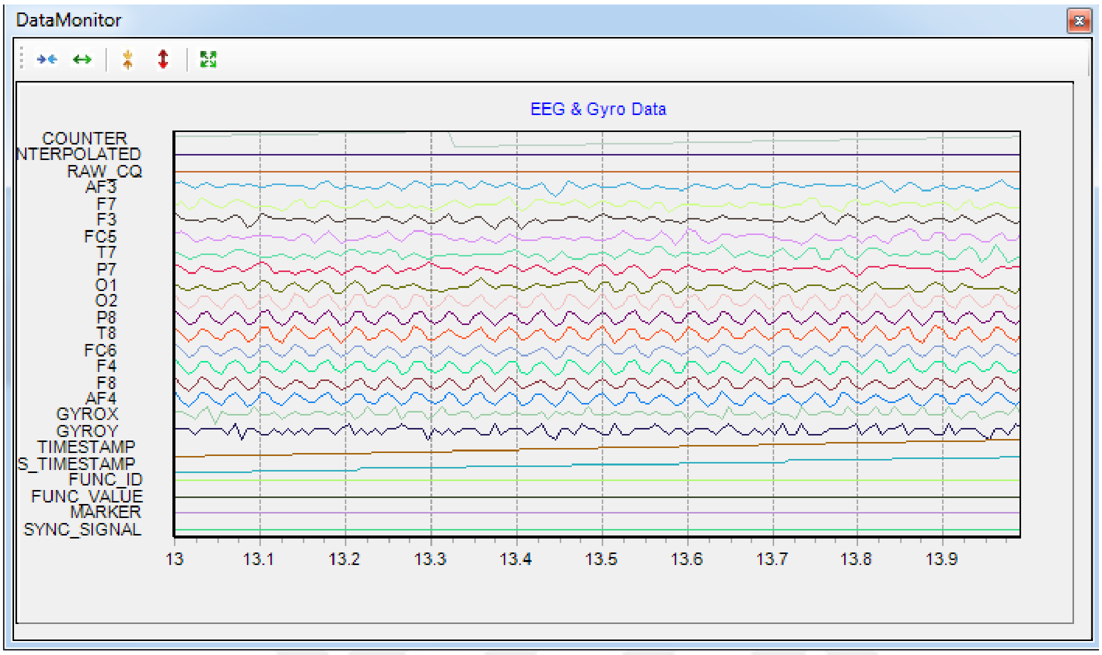


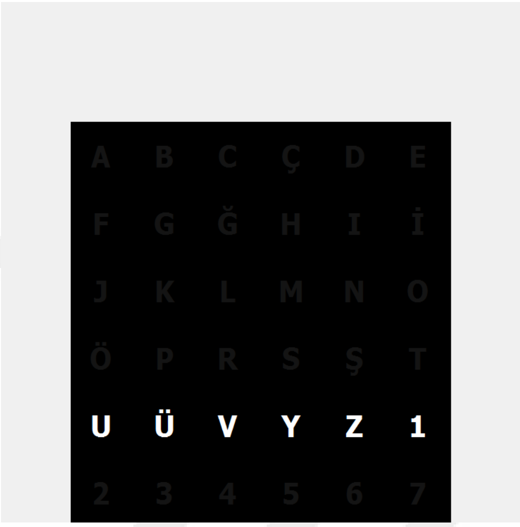
Figure 20 EEG & Gyro Data Monitor

C. The *Training Screen* shows the target item that needs to be focused on the top of the screen and the six by six stimuli matrix. The application stores gathered data on each flash on and off to construct the user row and column classifiers.

Figure 21 Training Screen

D. The *Test Screen* shows only the six by six stimuli matrix to the user. The user focuses on one item from the given list in each trial. At the end of each trial, the application predicts user's focused cell and displays it for one second on the top of screen. When the predicted stimulus is displayed for one second on the top of the screen. During this duration, the captured EEG data is stored in the filesystem for the offline analysis of N400 signals.

Table 1 Testing Screen

	
During the Trial	**At the End of Trial-Predicted Letter**

3.8.2 Application Business Logic & Algorithms

The developed P300 speller application software is using a similar methodology and classification algorithms as proposed by Erdogan (2009).

<u>Classification Problem Definition</u>

The prediction of the target cell in the P300 Speller application requires the determination of the target row and column. Although there are six columns and rows in the presented stimuli, the problem is to distinguish target and non-target rows and columns, so it is a binary classification task.

<u>Collected Data</u>

The data given for training and testing the classifier contain raw data information of all available channels of the device. This raw data consists of EEG time segments of 800 ms duration after the stimulus onset from the sensor locations AF3, F7, F3, FC5, P7, O1, O2, P8, FC6, F4, F8, and AF4. The dimension of the feature vector depends to the sampling rate of the EEG recording device so the sampling rate is 128Hz in this study. The collected data set contains observation matrix (X), stimulus type (Y) as the target or nontarget stimulus and intensified stimulus class (Y_StimClass) (Figure 22). Y_StimClass is a variable to identify different row/column intensifications.

Data Set

X = [Samples X Channels]

Y = [StimulusType X 1] (1: NonTarget Stimulus, 2: Target Stimulus)

Y_StimClass = [StimulusClass X 1] (Intensified Stimulus Classes)

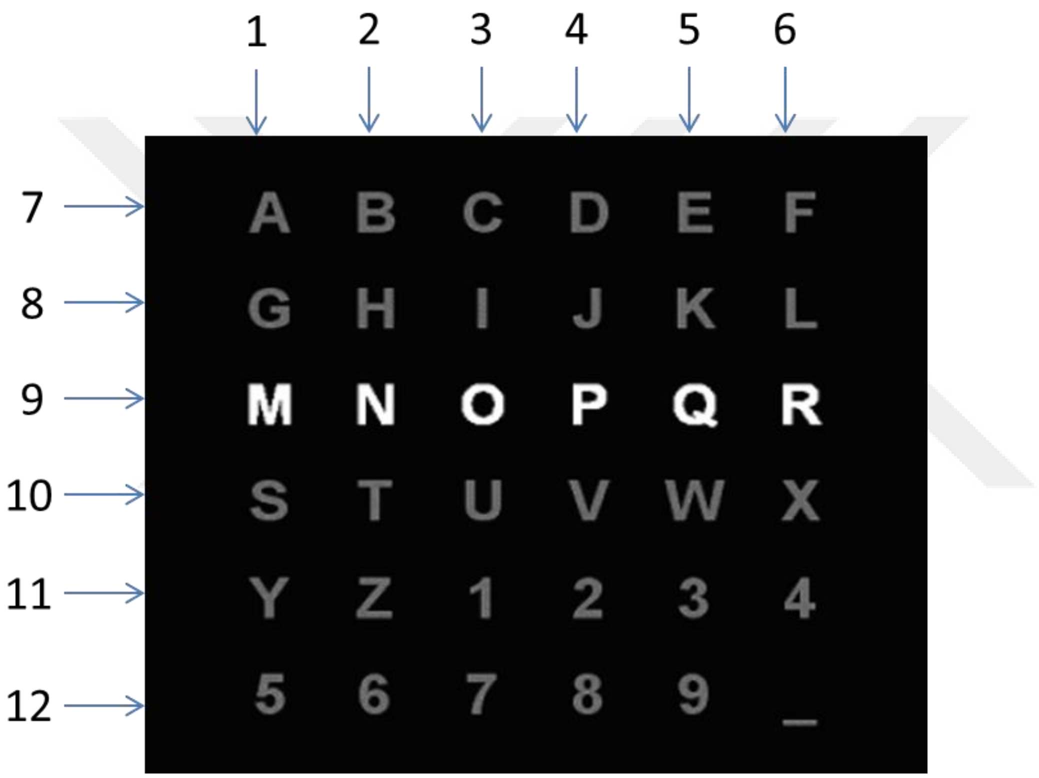


Figure 22 Illustration of the assignment of the stimulus class (Y_StimClass) for different row or column intensifications

Preprocessing: Filtering, Feature Vector and Normalization

Transferred signals are averaged between each trial. Averaging is a common signal filtering approach in classification problems. By applying averaging between trials, the noise is reduced to a level that allows obtaining the main target response or P300 response. Afterwards, the feature vector is formed by concatenating filtered EEG signals by channel.

Normalization & Classification

During each on or off flash, the accumulation of feature vectors constructs the target and non-target observation data set. Although each feature vector is filtered, the samples in a training set usually exhibit a large variance. In order to remove the extremity of the samples and increase the correlations between the samples from the same class, Gaussian normalization is applied at the end of the training and prediction phase. It is a well known type of normalization technique which is commonly used for decreasing the variance of samples in the set of interest.

Finally, SVMs construct a classification model by using the normalized data set. The constructed model predicts the targeted cell during the testing phase. (Section 3.7)

Cross Validation:

After the classifier is generated, the training data is divided into small parts; one part is used to train the classifier. The remaining part is used for validation to test the classifier performance. This method is called cross-validation, which is a testing method in supervised learning used for assessing the generalization performance of a classification model. Application of this technique provides a priori information about the classifier prediction performance and allows the modification of the parameters that are used to construct the classifier. In this study, a certain cross-validation threshold is applied to subjects' training model in order to move on to the testing phase which is 75% for both row and column classifiers. This value is found by averaging all cross-validation results of the subjects in the pilot experiments.

3.8.3 N400 Recording and Offline Analysis

A one-second EEG raw data window was recorded after the predicted item had been shown on the P300 speller application during the testing phase. In addition, the application stores the prediction result as MM, HM and FM. For offline analysis, all participants' data for all stimuli types were pooled together and analyzed. All analyses are performed by using Fieldtrip toolbox – a Matlab software toolbox for MEG and EEG analysis developed by Donders Institute for Brain, Cognition and Behavior. After preprocessing the raw EEG data, high-pass and low-pass filters are applied in order to get rid of the noise. Then Independent Component Analysis (ICA) is performed for excluding undesired components (artefacts) that are mostly related to eye-blinks and muscle-movements (The source Code is available at https://gitlab.com/uguracar/P300/blob/master/N400%20Matlab/n400_last.m).

3.9 Hypotheses of the Thesis

To sum up, the thesis focuses on the below hypotheses:

- *Stimulus modality (letters, words, icons)*

The success rate of the P300 speller application is expected to be in the following increasing order: words, pictures, characters with respect to stimulus type.

- Relation between subjective factors (motivation, mood) and success rate of the prediction

Cognitive aspects such as the effect of users' motivation and mood on the success rate of the prediction have been investigated by BCI researchers as well – beyond mere interest in hardware and software design. It is expected that the motivation measured with QCM or VAS and mood affects the success rate proportionately.

- N400 as an ERP signature of wrong prediction

While subjects are expressing their thoughts in terms of brain activation patterns through the BCI system, they try to select a letter, a word or an icon among different options. The produced words, icons or characters have some form and semantics which elicit several different brainwave patterns. A wrong and unintended prediction by the system is expected to elicit an N400 potential.

CHAPTER 4

RESULTS

4.1 Pilot Studies

Several pilot sessions were conducted before the main experiment in an effort to calibrate the experimental design. We used participants' feedback and the overall success rate to develop our task instructions and the way the problems are presented to the subjects.

The first pilot study involved 10 participants. The software included a training phase only for letters. The testing phase included letters, words, and icons, respectively. The Character Set contained 36 letters and numbers as well as symbols. The Word Set contained 36 words. The Icon Set contained 16 icons. Every subject performed training with the character modality (15 repetitions for 5 characters). Then the system used that training data in the testing phase to predict the participants' intented items (7 repetitions for each character/word/icon). In the results of the overall predictive success rates for all modalities, there were more mismatch outputs than initially expected (MM=118, HM=79 and FM=28). The first reason was related to the application of a single training session (for characters only) which was not sufficient in predicting also icons and words. The other reason was related to the motivational differences of the participants. Some of the participants did not seem to care about the task but seemed to be bored towards the end of the experiment. One final factor was related to the problems that we experienced with Emotiv EPOC. The battery of the device generally lasted for an hour but there appeared a recharging time issue until the next experiment.

The second pilot study involved five participants. The software was modified to include training sessions for each modality. The icon set also increased to 36 icons in all other modalities. In order to increase the success rate, the number of trained characters was increased from five to ten. Additionally, a questionnaire was added to measure the motivation of the subjects quantitatively. After finishing each block of experiments, subjects marked their motivation level with a value between one (least) and six (most). Similarly, based on the results, there was a significant positive relationship between motivation and success rate. Pearson correlation coefficients were as follows:

- characters: r (30) = 0.35 (p = 0.05)
- words: r (30) = 0.43 (p = 0.01)
- Icons: r (30) = 0.46 (p = 0.01).

The results implied that the predictive success rate for all modalities was increasing with the current motivation of the subjects.

In addition, increasing the number of trained characters affected the predictive success rate of all modalities positively. The order of training and testing was first characters, then words and finally icons. The success rate was also decreasing in parallel with the order (**Table 2**). This might feflect a sequence effect.

Table 2 Second Pilot Study Predictive Success Rates

	MM Count	HM Count	FM Count	Total	Success Rate(SR)	SR SD (%)	SR SE(%)
Characters	17	39	42	98	62.76	19.26	3.52
Words	61	31	24	116	34.05	19.10	3.49
Icons	66	38	16	120	29.17	17.06	3.12

In addition, the EEG recording during the prediction was added to the software for offline analysis of N400 ERPs. The offline analysis of EEG recordings for the N400 potential implied that there were significant differences between 300 and 500 milliseconds for the F3, F4, F8, FC6 and O1 channels with respect to elicited N400s by mismatch (MM) conditions (Figure 24, Figure 25, Figure 28). All participants' data for all stimuli types in this pilot experiment were pooled together across all channels and analyzed. The green vertical dashed bars in the figures indicate the time window in which the N400 effect is observed. The blue lines and red lines represent overall event related potentials for FM conditions and MM conditions respectively. Besides the N400, there are also significant differences between 700 and 800 milliseconds for the F3, F4, F8, FC6 and O1 channels with respect to negative ERPs elicited by mismatch (MM) conditions (Figure 24, Figure 25, Figure 28). Also, there are significant differences between 300 and 600 milliseconds for the FC5, F7, P7, P8 and O1 channels with respect to positive ERPs elicited by mismatch (MM) conditions (Figure 25, Figure 26, Figure 27, Figure 28).

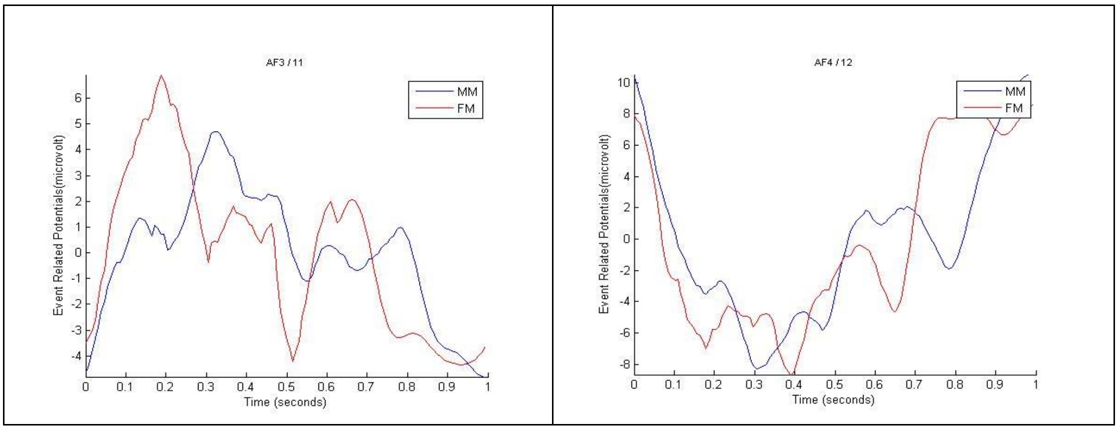


Figure 23 **ERP vs Time Graph of AF3 and AF4 channels for MM and FM conditions**

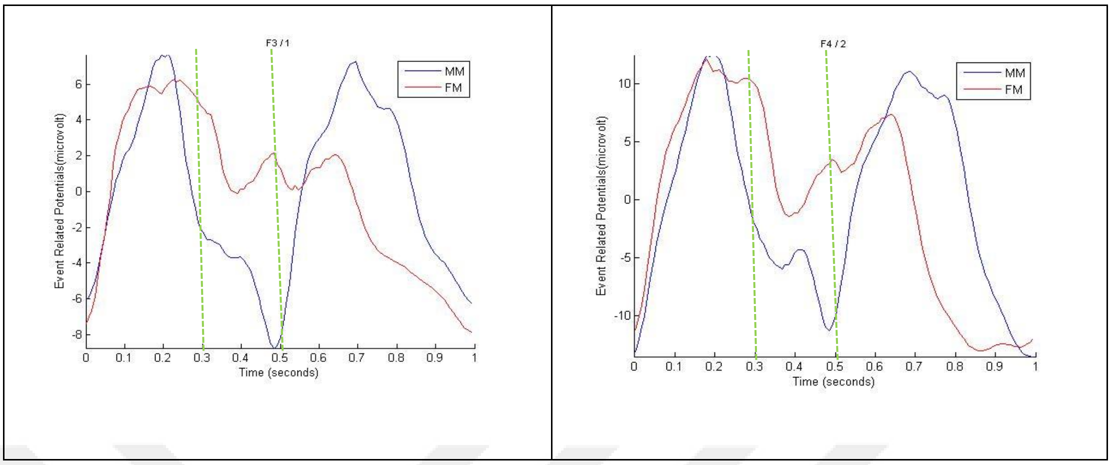


Figure 24 ERP vs Time Graph of F3 and F4 channels for MM and FM conditions

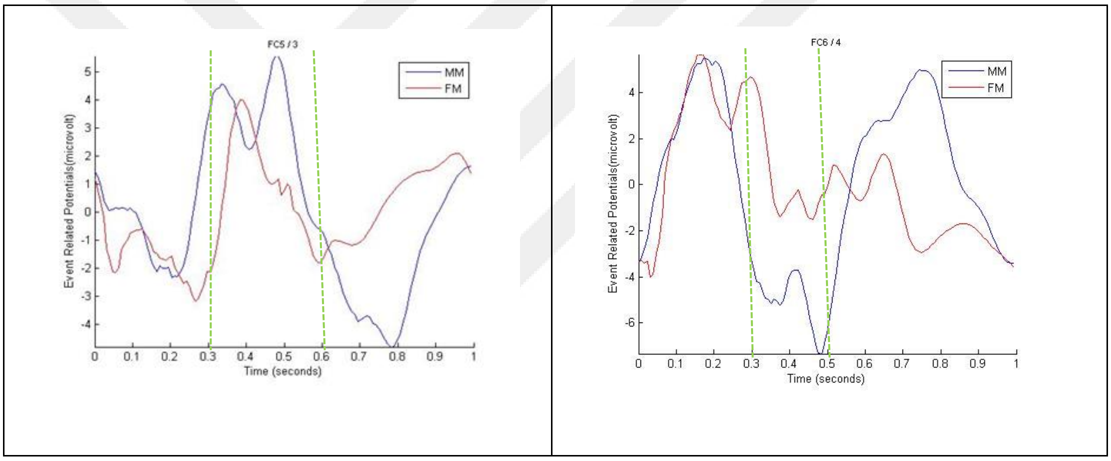


Figure 25 ERP vs Time Graph of FC5 and FC6 channels for MM and FM conditions

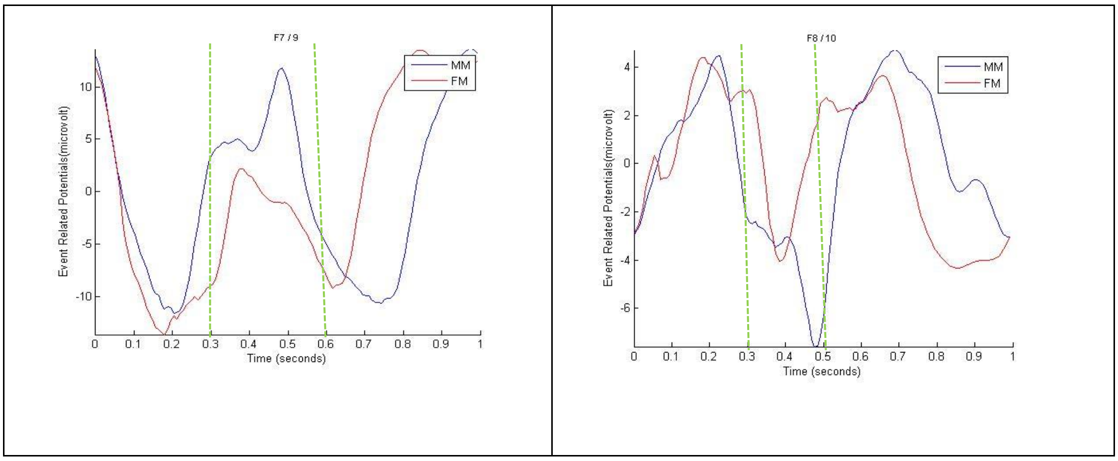


Figure 26 ERP vs Time Graph of F7 and F8 channels for MM and FM conditions

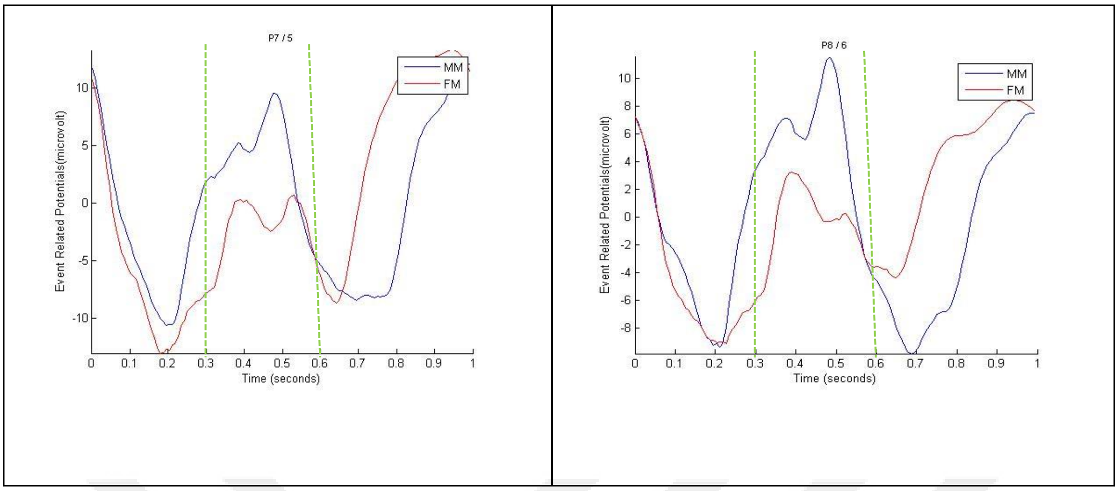


Figure 27 **ERP vs Time Graph of P7 and P8 channels for MM and FM conditions**

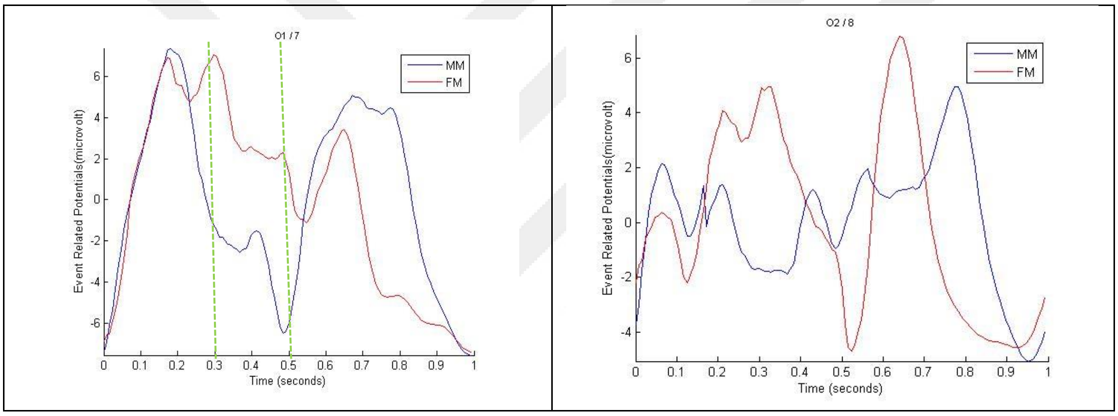


Figure 28 **ERP vs Time Graph of 01 and O2 channels for MM and FM conditions**

In the final experimental design, the training and testing parameters were kept the same as in the second pilot study. The order of performing each modality was randomized, in order to avoid the observed sequence effect. The pre-session and in-session questionnaires (Baykara et al., 2015) were added in order to measure motivation. One block contains the training and testing phase for one modality. Pre-session refers to the beginning of the first modality training phase. In-session refers to the end of the any modalities' testing phase (Figure 12). In addition, in order to eliminate participant-dependent negative effects on the success rate such as muscle movements or lack of motivation, certain thresholds were defined for the cross-validation results of row and column classifiers. If any participant's cross validation results was below the defined thresholds (75%), these participants were not allowed to move to the testing phase.

.

4.2 Main Experiment

4.2.1 Analysis of Predictive Success Rate

All three modalities are compared with respect to predictive success rates according to the number of FMs, HMs and MMs. The collected data is analyzed between different stimuli types in order to see the effect of condition in every stimuli type. Characters have the highest full-match counts and icons the lowest. In addition, characters and words' full match rates are close (Table 3).

Table 3 Main Experiment Predictive Success Rate

	MM Count	HM Count	FM Count	Total	Success Rate(SR)	SR-SD	SR-SE
Characters	5	40	39	84	70.24	13.90	4.01
Words	19	33	32	84	57.74	23.68	6.84
Icons	38	41	5	84	30.36	12.97	3.74

Table 3 shows full match, mismatch and half match counts and the calculated success rates. Similar to the results in full-match counts, success rates of characters and words are 70.24% and 57.74% respectively. On the other hand, icons success rate is the lowest, 30.36%.

Three repeated measures ANOVAs were conducted to compare the effect of modalities (characters, words and icons) on the accuracies (FM, HM and MM counts):

1) The effect of modality on predictive success counts (**FM**) was analysed. The result of the repeated measures ANOVA was statistically significant, $F(2, 22) = 12.88$, $p < 0.01$, $p_\eta^2=0.84$. Furthermore, there was a statistically significant difference between characters and icons, $F(1,11)=25.43$, $p<0.01$. Words and icons were also significantly different, $F(1,11)=20.72$, $p<0.01$. On the other hand, there were no significant relationship between characters and words. Table 4 shows the mean values and standard errors of each modality's predictive success counts (FM).

Table 4 Marginal Means and Standard Errors of each Modality's FM Counts

Modality	Mean	SE
Characters	3.25	0.50
Words	2.66	0.56
Icons	0.41	0.19

2) The effect of modality on **HM** counts was analysed. The result of the repeated measures ANOVA was not statistically significant, $F(2, 22) = 0.635$, $p = 0.54$, $p_\eta^2=0.08$. Table 5 shows the mean values and standard errors of each modality's predictive success rate (HM).

Table 5 Marginal Means and Standard Errors of each Modality's HM Counts

Modality	Mean	Standard Error
Characters	3.33	0.52
Words	2.75	0.56
Icons	3.41	0.19

3) The effect of modality on success rates, which is a weighed measure between FM, HM and MM was analysed. This analysis was performed to give justice to the success of the training model/algorithm implemented in this study and make it comparable to the findings in the literature. Note that for the model it is a partial success to predict the column or the row correctly – whereas for the participant a partial success is not visible. The repeated measures ANOVA results showed that the effect of modality on predictive success rate (FM) was statistically significant, $F(2, 22) = 18.62$, $p < 0.01$, $p_\eta^2=0.87$. Furthermore, there was a statistically significant difference between characters' and icons' predictive success rates, $F(1,11)=41.60$, $p<0.01$. Words and icons were also significantly different, $F(1,11)=28.66$, $p<0.01$ but not characters and words. Table 6 shows the mean values and standard errors of each modality's predictive success rates.

Table 6 Marginal Means and Standard Errors of each Modality's Success Rates

Modality	Mean	SE
Characters	70.24	4.01
Words	57.74	6.84
Icons	30.36	3.74

4.2.2 Analysis of Success Rate with respect to Motivation

Motivation was measured with two different questionnaires, the Questionnaire for Current Motivation (QCM) and a visual analogue scale (VAS). The QCM includes 18 items, rated on a 7-point Likert scale, corresponding to four different components of motivation: mastery confidence, incompetence fear, interest and challenge. Participants also indicated their level of motivation on the VAS, a 10 cm line ranging from 0 (not motivated at all) to 10 (extremely motivated).

The QCM was conducted once before starting the session. On the other hand, VAS measurements for motivation were collected three times before starting each block of different modalities (see section 3.5, figure 13). All participants' data for all stimuli types were pooled and tested for correlations to see the effects of the psychological factors on the predictive success rate.

Firstly, the motivation (VAS) effect on each block (characters, words, icons) in one session was analyzed. There are two variables, namely the order of the blocks and the motivation (VAS). Three different statistical analyses were performed to investigate the sequential effects:

1) We ran a repeated measures ANOVA for FM counts over the three consecutive blocks. The order of modality in each block was random for all participants so this analysis only looks at how the FM developes over time irrespective of modality. The effect of order on predictive success counts (FM) over 3 consecutive blocks was not statistically significant, $F(2, 22) = 2.86$, $p = 0.62$. In addition, there was no statistically significant difference between first and second block, $F(1,11)=0.04$,

p=0.85 or second and third block, $F(1,11)$=0.86, p=0.37. Table 7 shows the mean values and standard errors of each modality's predictive success counts (FM) in each block.

Table 7 Marginal Means and Standard Errors of Success Rates' of each Block

Modality	Mean	SE
FM-Order1	1.75	0.62
FM-Order2	1.92	0.48
FM-Order3	2.67	0.61

1) We ran a repeated measures ANOVA for motivation (VAS) over the three consecutive blocks. This analysis only looks at how the motivation of participants develops over time irrespective of modality since the order of modality in each block is random for all participants. The effect of motivation of participants over 3 consecutive blocks was statistically significant, $F(2, 22) = 27.42$, $p < 0.01$. In addition, there was no statistically significant difference between first and second block, $F(1,11)$=1.37, p=0.27 but second and third block was significantly different, $F(1,11)$=37.11, p<0.01. Table 8 shows the mean values and standard errors of three consecutive block's motivation (VAS) scores.

Table 8 Marginal Means and Standard Errors of Motivation Scores over three Consecutive Blocks

Modality	Mean	SE
Motivation(VAS-Block1)	6.92	0.54
Motivation(VAS-Block2)	6.17	0.72
Motivation(VAS-Block3)	3.00	0.39

2) We ran three simple linear regressions for analyzing the effect of motivation (VAS) scores on predicted success counts in the subsequent blocks.
   i. Motivation (VAS) scores before the first block were regressed against FM counts in the first block. The result of this analysis indicated that motivation (VAS) was significantly contributing to the success count $F(1,10) = 9.36$, p=0.01 with an R of 0.70, R^2 of 0.48.
   ii. Motivation (VAS) scores before the second block were regressed against FM counts in the second block. The result of this analysis indicated that motivation was not significantly contributing to the success count $F(1,10) = 1.24$, p=0.29 with an R of 0.33, R^2 of 0.11.
   iii. Motivation (VAS) scores before the third block were regressed against FM counts in the third block. The result of this analysis indicated that motivation was not significantly contributing to the success count $F(1,10) = 1.15$, p=0.31 with an R of 0.32, R^2 of 0.10.

Table 9 shows that average FM counts increase over consecutive blocks but there is no statistically significant change in them. On the other hand, the motivation is decreasing slightly from block 1 to block 2 but significantly from block 2 to block 3. In other words, there is an increase of FM counts from block 1 to block 3 (not significant) but motivation decreases significantly from block 1 to block 3.

Table 9 Average motivation (VAS) before starting each session across all participants vs average full match rate over the three blocks

	FM-Order			Motivation (VAS-Order)		
	1	2	3	1	2	3
Mean	1.75	1.92	2.67	6.92	6.17	3.00
SD	2.14	1.68	2.10	1.88	2.48	1.35
SE	0.62	0.48	0.61	0.54	0.72	0.39

Secondly, the overall motivation data including the QCM and VAS was analyzed with two different regression models comprising all participants' FM data, which is the sum of all FM scores of each stimulus type:

1) We first ran a multiple linear regression for the overall success rate (FM) by using all predictors (mood (VAS), motivation (VAS) and QCM (interest, mastery confidence, challenge, incompetence fear)) by using the "backward" elimination method (see Table 10). This method first considers all predictors and then step-wise eliminates those that do not contribute significantly such that it retains all those predictors that are significantly related to the criterion variable. In the first model that "backward" multiple regression produced, the results indicated that all predictors (except Challenge (QCM) which measure incompetence fear were significantly contributing to the overall success count (FM) $F(6,5) = 4.215$, $p=0.06$ with an R of 0.914, R^2 of 0.835. In the second model, Challenge (QCM) variable was excluded and all remaining predictors remained the same. In this model, the linear combination of the remaining predictors was significantly predicting the overall success count (FM), $F(5,6)=5.972$, $p = 0.025$ with an R of 0.913, R^2 of 0.833. The change effect of removing Challenge (QCM) variable from the predictors did not significantly improve the model (R^2 *Change*=-0.002, F *Change*=0.068, p=0.804), confirming that it was not a significant predictor. Although this regression model and most of the predictors' coefficients (mastery confidence (QCM), interest (QCM), Mood (VAS) and first order motivation (VAS)) were significant but the model explained a suspiciously high amount of variance. Tests for collinearity indicated that the explained variance was inflated since all predictors' VIF values were greater than 1 and some of them (i.e interest (QCM) and mood (VAS)) were even above 10 (Table 10). The collinearity diagnostics showed that a very high level of multicollinearity was present between predictors (Figure 29) as well, i.e., the variance proportions of the six variables overlapped considerably. In order to avoid collinearity of variables, inflation of variance and "over-fitting" of the model, we used a forward technique in the following second analysis (Field, 2013, pp. 324-326).

Table 10 The Coefficients, *p*-values and Collinearity Statistics of the Regression Model in the 1st Analysis (Backward)

			Collinearity Statistics	
Predictor	*Coefficient*	*P-value*	*Tolerance*	*VIF*
(Constant)	-18.77	0.011		
Mastery Confidence(QCM)	3.31	0.006	0.154	6.486
Challenge(QCM)	-1.298	0.069	0.198	5.039
Interest(QCM)	-7.561	0.014	0.015	64.952
Mood(VAS)	4.705	0.024	0.022	46.090
Motivation(VAS)	1.935	0.011	0.205	4.869

Collinearity Diagnostics[a]

				Variance Proportions						
Model	Dimension	Eigenvalue	Condition Index	(Constant)	QCM_Mastery Confidence	QCM_Challenge	VAS_Mood	QCM_Interest	VAS_MOTIV_1	QCM_IncompetenceFear
1	1	6.652	1.000	.00	.00	.00	.00	.00	.00	.00
	2	.235	5.320	.00	.00	.00	.00	.00	.00	.08
	3	.071	9.651	.00	.02	.10	.00	.00	.02	.00
	4	.022	17.291	.04	.02	.16	.00	.01	.18	.01
	5	.012	23.924	.28	.18	.08	.01	.00	.00	.21
	6	.007	30.443	.12	.25	.07	.05	.01	.19	.55
	7	.001	103.197	.57	.53	.60	.94	.98	.61	.13
2	1	5.829	1.000	.00	.00	.00	.00	.00	.00	
	2	.076	8.744	.01	.01	.05	.00	.01	.03	
	3	.063	9.635	.06	.01	.12	.00	.00	.00	
	4	.022	16.463	.01	.03	.30	.00	.01	.33	
	5	.010	24.349	.07	.49	.00	.04	.00	.08	
	6	.001	90.384	.86	.45	.53	.95	.98	.55	

a. Dependent Variable: all_FM

Figure 29 the Collinearity Diagnostics of the Regression Model

2) Second, we ran a multiple linear regression with the forward entering method of variables as a result of the overfitting of the first regression analysis (backward). By using this method, the strongest predictor can be found when the model selects them by magnitude. The result of this analysis indicated that only mastery confidence (QCM) was significantly contributing to the overall success count (FM) $F(1,10) = 6.218$, p=0.03 with an R of 0.61, R^2 of 0.38. In addition, there was only one significant coefficient, Mastery Confidence (QCM) (Table 11). This analysis thus provided a leaner and more valid explanation To summarize, this second model is the most meaningful and parsimonious model that explains overall predictive success counts after eliminating multicollinearity between predictors.

Table 11 The Coefficients, *p*-values and Collinearity Statistics of the Regression Model in the 2nd Analysis (Forward)

			Collinearity Statistics	
Predictor	*Coefficient*	*P-value*	*Tolerance*	*VIF*
(Constant)	-.123	0.964		
Mastery Confidence(QCM)	1.174	0.032	1.000	1.000

4.2.3 Analysis of N400 Potential

All main experiment participants' data for all stimuli types were pooled together across all channels and analyzed. The green vertical dashed bars in the figures below indicate the time frame in which the grand average of the positive or negative ERPs were visually analyzed. The blue lines represent overall event related potentials for FM conditions. Similarly, the red lines represent MM conditions. Similar to pilot experiments, MM conditions include HM conditions as well.

One of the hypotheses of this study was that an N400 neural pattern would be observable for the mismatch cases that are in line with the literature on the N400 (Section 2.5). A permutation cluster test was performed to observe the expected N400 effect. Contrary to our initial hypothesis, there was no significant cluster or ERP effect difference between FM and MM conditions in the data set. However, the mean amplitude for all these channels was calculated within a larger time window of one second after stimulus onset. Figure 30, 31, 32, 33 and 34 show the grand average wavelengths of FM and MM stimuli for F3, F7, P7 and P8 channels. Within this larger time window, there are indications of neural patterns of late positivity occurring between 600-700 milliseconds after stimulus onset on F3 channel. Interestingly enough, these neural patterns are almost completely consistent for character and word conditions on one of the frontal channels (F3) (Figure 30 and Figure 31).

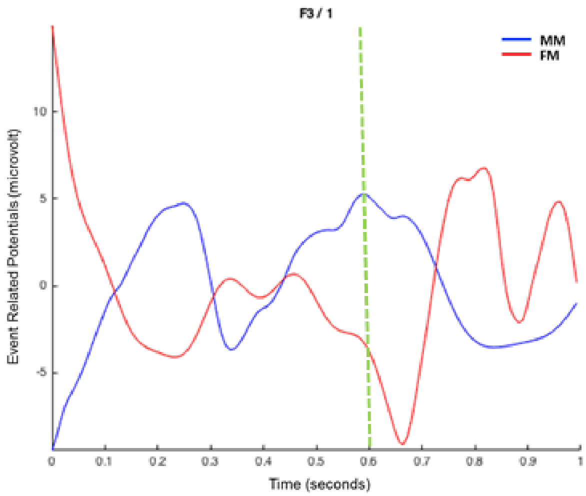


Figure 30 ERP vs Time Graph of F3 Channel for MM and FM conditionsfor Characters. The green vertical dashed line indicates the period of the potential P600 effect.

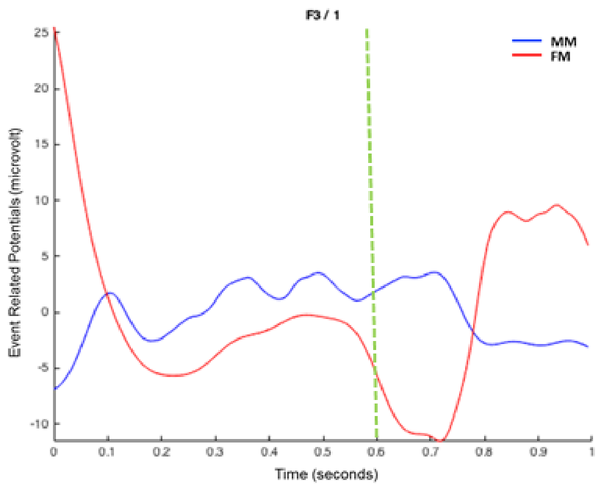


Figure 31 ERP vs Time Graph of F3 Channel for MM and FM conditions for Words. The green vertical dashed line indicates the period of the potential P600 effect.

Another observation is that there are indications of negative ERPs that occurs around 500 milliseconds over some of the parietal channels (Figure 32, Figure 33 and Figure 34). Although the N400 peaks around 400 milliseconds post-stimulus onset, it can be observed between 250 and 500 milliseconds as well and it has maximum potentials over centro-parietal electrodes (Kutas and Federmeier, 2000). Similarly, the event related potential occurring around 500 milliseconds in Figure 35 and Figure 36 appears to be similar to a N400 response. In addition, P7 and P8 are parietal channels.

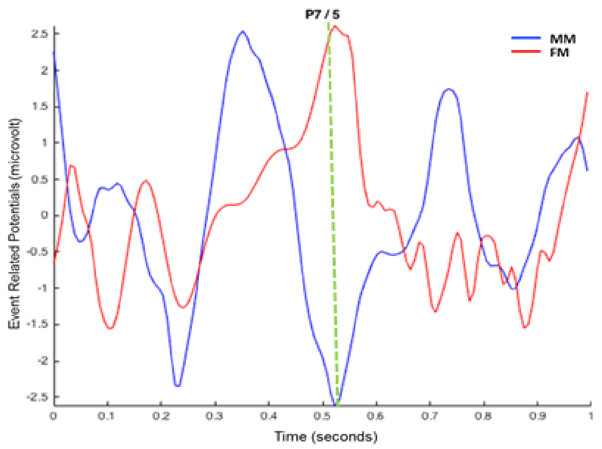


Figure 32 ERP vs Time Graph of P7 Channel for MM and FM conditions. The green vertical dashed line indicates the period of the potential N500 effect.

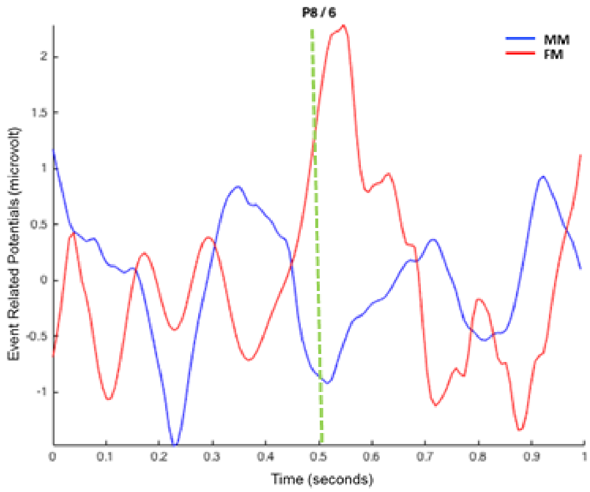


Figure 33 ERP vs Time Graph of P8 Channel for MM and FM conditions. The green vertical dashed line indicates the period of the potential N500 effect.

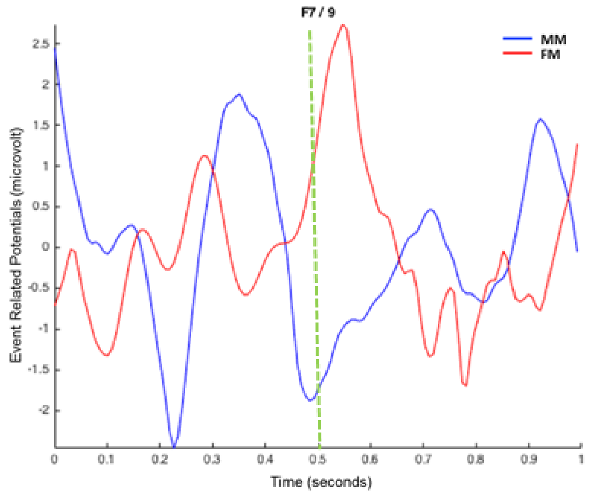


Figure 34 ERP vs Time Graph of F7 Channels for MM and FM conditions. The green vertical dashed line indicates the period of the potential N500 effect.

CHAPTER 5

DISCUSSION

This thesis has focused on understanding cognitive aspects of the P300 speller paradigm with respect to three main aspects;

1. stimuli were presented to subjects in three modalities (letters, words, icons),
2. the prediction success rate of the system in the light of motivation measurements including the QCM and VAS questionnaire, and
3. the N400 as an ERP signature of wrong prediction.

In order to achieve the aims of the thesis, a P300 speller application was designed and implemented. The speller application has configurable training and testing options, real time EEG raw data processing and a prediction engine and EEG data recording capabilities for offline analysis. In addition, in order to measure subject's motivation, questionnaires (QCM and VAS) were conducted for each participant at various points of the experiment. The application of supervised learning algorithms (SVM) in the P300 speller, offline EEG data analysis methods and motivation-related questionnaires were all in line with the literature.

5.1 Results According to the Three Modalities

To investigate the P300 speller performance, words and icons in addition to characters were offered in order to allow users to express their thoughts more rapidly so that speller performance for communication efficiency might be improved. The results indicate that participants' predictive success rate for characters and words significantly differed from icons. In addition, it is important to note that there were very few cases of full match for icons. This could be due to the fact that the main algorithm for the predictive model had been developed for characters or letters but not icons. Success rates of characters and words are 70.24% and 57.74% respectively. On the other hand, icons success rate is the lowest, 30.36%. In a recent study, Speier et al. (2017) compared stimulus types in online classification of the P300 speller using language models. The first method is the same as this study's approach, consisting of highlighting flashed characters by "intensifying" the font color to white. The second method is changing the background to white and the character to black. The third method overlays the character with an image of a famous face as proposed by Kaufmann and colleagues (2012). They used g.tec amplifiers, active EEG electrodes that consisted of 32 channels, and an electrode cap (Guger Technologies, Graz, Austria) for data acquisition. They applied two different classifiers to acquired data such as stepwise linear discriminant analysis (SLDA) and particle filtering (PF) with dynamic stopping. In their preliminary experiment, comparing traditional and inverted stimuli showed accuracy rates of 93.39% and 92.13% respectively. Due to these results, they focused on inverted and famous face stimuli in their main experiment. The average accuracy by using inverted stimuli and particle filtering classifier was 91.67%. Famous faces stimuli had an average accuracy of 96.00% with the same classification method. Comparing Speier and colleagues' findings on accuracy of character prediction clearly shows the superiority of their results (70.24% (including HMs) vs (93.39% or 91.67%)). However, our findings comply with Duvinage et al.'s experiment results and confirm that the Emotive Epoc headset is significantly worse than a 32 channel research/medical grade cap - although the Emotive device's performance is far above the chance level. Furthermore, Elsawy et al. (2014) analyzed the performance of a Principal Component Analysis (PCA) ensemble classifier for P300-based spellers by using Emotiv Epoc headset. Another debatable aspect of this result could be the differences in semantic and visual cognitive processing between characters, words and icons. Characters might be primitive and therefore easy to process compared to words and icons. In a recent study, Huang et al. (2015) analysed how icons are processed in the brain by using neuroimaging techniques (via functional magnetic resonance imaging). Their experiment results indicated that icons, images or pictures are not cognitively processed like words or letters although both words and icons stimulate the semantic system that is needed for language processing in the brain. The brain spent more effort to process icons than characters and words. This could be an important factor for the P300 speller performance which would mean that the probability of detecting P300 ERP could be decreasing depending on visual and semantic processing of the stimulus types. Similarly, characters and words predictive success rates are close (characters: 70.24% words: 57.74% icons: 30.36%) compared to icons but characters still have a higher success rate than words. This result could be an indication of cognitive processing difference between words and letters that affect event related potentials in the brain. Words may stimulate semantic system of the brain differently since they have a corresponding lexical meaning in the participants. Another differentiator factor could be the structural difference between words and letters since a word is composed of many letters and has different lengths (e.g "su", "televizyon" etc.) whereas a letter is just one character.

5.2 Results of the Motivation Questionnaires and the Correlation between Success Rate and Motivation

The second goal of this study was to understand of the subject's role in P300 potential-based applications since brain activity was depending on user's mental or emotional state. Kübler and Kleih (2015) listed some main psychological factors that may affect BCI performance such as attention, concentration, motivation, or visuo-motor coordination. Baykara et al. (2015) also supported Kübler with their study in which they showed that participants' current motivation was an influencing factor for speller performance. Heckhausen (1977) defines motivation as an "energizer" to our behaviour and a "trigger" to our actions for the achievement of our goal and divides motivation into two kinds: intrinsic and extrinsic motivation. On the one hand, intrinsic motivation is described motivation due to the act of performing the task itself. On the other hand, extrinsic motivation appears when your goal is to win a reward, such as respect, status or money as a result of performing an action. In this study, the effect of intrinsic motivation on P300 speller performance was analyzed. The QCM was found to be able to predict participants' performance in the P300 speller task – however not all of its scales. "Mastery confidence" indicates a participant's belief of being competent to successfully perform the P300 speller task. Similarly, "interest" indicates how attractive the P300 speller task is for a participant. On the other hand, "challenge" assesses the performance aspect of the P300 speller application experienced by the participant while "incompetence fear" measures participants' judgement about him- or herself being unable to perform the P300 speller task. Although the initial expectation was that all of the QCM sub-components were supposed to influence success rate either negatively or positively, only mastery confidence had a significant effect on the success rate according to the correlation analysis between success rate and motivation (VAS), mood (VAS) and motivation sub components (QCM). These results suggest that the mastery confidence with which a participant entered the session might be an influencing factor for the adaptation speed of the participant to the P300 speller task. The anticipated mastery may yield a higher magnitude of P300 amplitude which may have eventually resulted in better performance. However, further research needs to be conducted regarding mastery confidence and its possible influence on the P300 speller task performance to validate our findings. However, participants' overall motivation (VAS) was decreasing throughout the experiment although their success rate increases over three consecutive blocks. This result might be the outcome of the overall low predictive success rate or usability of the BCI device. On the other hand, Baykara and colleagues (2015) found that participants' overall motivation was not changing throughout the experimental period. What appears to be a contradictory result is in fact none. While three consecutive blocks in random order (characters, words, icons) were shown in one session in the present study, one session consisted of only one block during which participants run twelve trials of copy spelling plus two optional free spellings in Baykara et al.'s study. Furthermore, the success rate is increasing towards the end of the session. This result supports the co-adaptation hypothesis that is explained by Mattout et al. (2015) as a closed-loop interaction between the participant's brain and the P300 speller application. They related the closed-loop interaction in terms of three main aspects: increasing the success rate of the P300 speller system by detecting and correcting errors automatically, providing adaptive behaviour to optimize the computer's speed-accuracy trade-off and simplifying the setup time of the environment such as wearing the headset, calibrating the speller application. These improvements support that subjects' getting more comfortable with the hardware and software setup towards the end of the session.

5.3 Results of N400 Analysis

Finally, the P300 speller application can make wrong predictions to the subject's intended words, icons or characters. In this case, a response of the subject after the wrong prediction would be detected and corrected automatically by the application. In order to achieve this aim, an analysis of the N400 ERP that would be elicited by unexpected or wrongly predicted inputs by the system was one of the novel contributions of this study to the autonomous feedback mechanisms in the P300 speller application research area. One of the challenges for the N400 ERP analysis is the generated noise by eye movements of the participants. Sometimes such noise can mask the small N400 potentials although participants are instructed not to move their muscles. Another factor that might affect the N400 potential is the priming effect since the words and icons were selected according to the basic needs of the ALS patients and they belong to the same semantic context. Kutas and Federmeier (2000) report that the N400 amplitude is reduced when a target word is preceded by a word that is semantically, morphologically, or orthographically related to it. However, the mismatch in this study which is supposed to elicit an N400 potential is not a semantic mismatch but a mismatch of the intended and predicted word. However, there are indications of neural patterns of late positivity occurring between 600-700 milliseconds after stimulus onset on the F3 channel. Late positive potentials (LPP) are known to be correlated with emotional processing (frustration) and syntactic aberrations. The positive event-related brain potentials in the 600 milliseconds range (P600) is known as a language-relevant ERP which is thought to be elicited by reading grammatical errors and other syntactic anomalies. P600 ERPs are also known to occur when a sentence is parsed in a different way than the reader originally expected (Gouvea et al., 2010). In this context, Kutas et al. (1998) states that grammatical violations show large positive potentials around 500-1000 ms after stimulus onset, namely a P600, and these violations do not elicit a large N400 response. Following this line of argumentation, in the present case, wrong predictions of the P300 speller application for words and characters appear to reflect the subject's surprise upon encountering an unexpected stimulus. This incongruency then elicited a P600 response. This may not be a linguistic response since the P300 speller only produces single characters, words or icons not sentences. Another debatable area is that some of the participants were also familiar to the experiment setup since they both attended the pilot and the final experiment sessions. Although, this condition appears to be an advantage at first sight, this could be a disadvantage or a factor that is negatively affecting the N400 potential due to the semantic priming and knowing the context before starting the session. Finally, the orders of the stimuli to be focused by the participants were selected in sequential order in the final experiment. In the half-match conditions, if either the row or the column is predicted correctly by the system so that despite the predicted word was wrong, a character or icon in close proximity to the targeted stimulus was selected. This could be another factor reducing the N400 amplitude due to semantic priming.

Another interesting finding is that, although the N400 ERP could be seen in some of the channels in the analysis of the pilot experiments, it was not commonly visible in the main experiment. Rather than an N400, an indication of negativity between 500-600 milliseconds was observed on P7, P8 and F7. Featherstone et al. (2013) state that N500 ERP is typically elicited in response to semantic incongruities in language similar to N400. Furthermore, Xiao et al. (2011) analysed the size incongruity effect which is an increase in response time due to the difference between the physical size of the letters that represents an object and the actual size of the object name. Their analysis provided that conflict detection and conflict resolution in the size incongruity effect were eliciting two different negative potentials, N200-N400 and N500-700 respectively. This might explain the N500 effect in the mismatch condition. The positive pattern is called P600 in the literature. In this study, we were envisioning to detect N400 ERPs during processing of wrong predictions or incongruities. On the other hand, indications of P600 ERPs were observed on some of the channels (F3) for mismatch conditions that are known to be elicited by reading grammatical errors and other syntactic anomalies (Kaan et al., 2000). However, Brouwer et al. (2017) state that some of the recent studies revealed that semantically anomalous, syntactically correct sentences failed to elicit the expected N400 effect; on the other hand they produced a P600 effect. They also propose a single stream model saying that the N400 amplitude is the result of the activation of the conceptual knowledge associated with the incoming word in the memory, whereas the P600 is the reflection of the processing of the word-by-word construction, reorganization or updating of an utterance interpretation in a sentence. These findings might imply that the P600 effect can be the result of the interpretation of consecutive predictions of characters, sentences or icons to conclude a meaning.

CHAPTER 6

CONCLUSION

This thesis work is an exploratory study to investigate the role of motivation and mood, stimulus type and wrong prediction on participants' performance during usage of the P300 speller application. The obtained results revealed that the current motivation of the participants might have a prominent role during the trials. The results of prediction success rate for the different stimuli types revealed that P300 speller is performing better for characters and words than icons, probably since they need more visual and semantic processing. Participants' success rate for characters and words were significantly greater than icons but there was no significant difference between them. That implies that words can be used efficiently as input stimulus layout in the P300 speller applications instead of characters, which may increase the speed of communication through brainwaves. However, the overall success rate of the Emotiv EPOC device is still too low and needs to be improved before the application may be used by patients. Furthermore, while our initial hypothesis predicted an N400 ERP during wrong prediction, the ERP analysis of the main experiment during prediction suggests that there is no significant ERP effect within the time window of one second after stimulus onset but the grand average wavelengths of some specific channels (F3, F7, P7, P8) showed symptoms of late positive potentials 600-700 ms after stimulus-onset and a negativity between 500-600 milliseconds. They might be caused by the surprise upon encountering an unexpected stimulus and the interpretation of consecutive predictions of characters, sentences or icons to conclude a meaning.

6.1 Limitations and Future Studies

This study has some limitations which might be addressed in future studies. First, one session including one training and one testing phase for each modality takes more than 1 hour for one participant. This time limitation is reducing the number of samples that feeds into the SVM classifier. The amount of the sample is an influencing factor on classification performance. This may be a drawback for the present study with students who have limited time but might not be so detrimental in the case of patients who can afford much more training time. This short-coming might thus not carry over to the target population - patients. Another drawback of this constraint may be the boredom effect on the subjects towards the end of the experiment. One can devise an experimental design such that subjects do the same task overall but attend to shorter sessions on different days and times. Motivation or mood of subjects could also be part of this longitudinal design. For instance, the same level of mood or motivation on different days could select the training samples for classifiers respectively to increase the classifiers predicting performance. Another limitation of the study is the usability of the Emotiv EPOC headset. Wearing and calibrating the connectivity of all sensors takes up to 10 minutes depending on the hair volume of the subjects. Furthermore, the Emotiv headset's sensitivity to eye blinks and muscle movements produces noisy signals. In the future, advances in sensor technology, especially in dry electrodes may bring BCI devices to the market that are easy to use and have low signal-to-noise ratios. Finally, icons showed the lowest success rate compared to words and characters. This unexpected result implies that the visual and semantic processing of the icons elicits different forms of ERPs from words and characters. In addition, the developed P300 speller application uses the same SVM classifier parameter sets such as kernel function, the degree of the kernel function, class weight etc. for all stimulus types although these parameters are fine-tuned for characters only. Therefore, to reach better performance with icons, they could be simplified to basic, characteristic forms without color and also a SVM classifier with a customized parameter sets specific to icons could be designed, in future studies.

REFERENCES

Allison, B. (2011). Trends in BCI research: progress today, backlash tomorrow?. XRDS: Crossroads, *The ACM Magazine for Students, 18(1)*, 18-22.

Aguilar, J. M., Castillo, J., & Elias, D. (2015). EEG Signals Processing Based on Fractal Dimension Features and Classified by Neural Network and Support Vector Machine in Motor Imagery for a BCI. *In VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014, Springer International Publishing*, 615-618.

Badcock, N. A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., & McArthur, G. (2013). Validation of the Emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs. *PeerJ,* 1:e38 https://doi.org/10.7717/peerj.38.

Baykara, E., Ruf, C. A., Fioravanti, C., Käthner, I., Simon, N., Kleih, S. C. Kleih, Kübler A. & Halder, S. (2015). Effects of training and motivation on auditory P300 brain-computer interface performance. *Clinical Neurophysiology*, *127* (1), 379-387.

Brouwer, H., Crocker, M. W., Venhuizen, N. J. and Hoeks, J. C. J. (2017), A Neurocomputational Model of the N400 and the P600 in Language Processing. *Cognitive Science, 41*(S6), 1318-1352.

Brunner, P., Bianchi, L., Guger, C., Cincotti, F., & Schalk, G. (2011). Current trends in hardware and software for brain–computer interfaces (BCIs). *Journal of Neural Engineering, 8*(2), 025001.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*(2), 121-167.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST), 2*(3), 27.

Chang, M., Nishikawa, N., Struzik, Z. R., Mori, K., Makino, S., Mandic, D., & Rutkowski, T. M. (2013). Comparison of P300 responses in auditory, visual and audiovisual spatial speller BCI paradigms. *arXiv preprint arXiv:1301.6360.*

Cho, O. H., & Lee, W. H. (2014). BCI sensor based environment changing system for immersion of 3D game. *International Journal of Distributed Sensor Networks*, *10*(5), 620391.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273-297.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*(1), 21-58.

Daly, J., Armstrong, E., Thomson, E., Andreas, P., & Martin, S. (2015). P300 brain computer interface control after an acquired brain injury. *International Journal on Recent and Innovation Trends in Computing and Communication, 3*(1), 318-325.

Debener, S., Minow, F., Emkes, R., Gandras, K., & Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk?. *Psychophysiology, 49*(11), 1617-1621.

Donders Institute for Brain, Cognition and Behavior, "FieldTrip: a Matlab software toolbox for MEG and EEG analysis," 2010, http://www.ru.nl/neuroimaging/fieldtrip.

Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., & Dutoit, T. (2013). Performance of the Emotiv Epoc headset for P300-based applications. *Biomedical Engineering Online, 12*(1), 56.

Elsawy, A. S., Eldawlatly, S., Taher, M., & Aly, G. M. (2014). Performance analysis of a Principal Component Analysis ensemble classifier for Emotiv headset P300 spellers. *In Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 5032-5035.

Erdogan, H. B. (2009). A Design and Implementation of P300 Based Brain Computer Interface, *MSc. Thesis*, Department of Electrical and Electronics Engineering, METU.

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. *Advances in Psychophysiology, 2*(S 1), 78.

Farwell L.A., Donchin E. (1988). Talking off the top of your head: Toward a Mental Prosthesis Utilizing Event-Related Brain Potentials, *Electroencephalography and Clinical Neurophysiology, 70*, 510-523.

Featherstone, C. R., Morrison, C. M., Waterman, M. G., & MacGregor, L. J. (2013). Semantics, syntax or neither? A case for resolution in the interpretation of N500 and P600 responses to harmonic incongruities. *PloS one, 8*(11), e76600.

Field, Andy (2013). *Discovering statistics using IBM SPSS statistics*. 4th edition. Los Angeles: Sage..

Fouad, M. M., Amin, K. M., El-Bendary, N., & Hassanien, A. E. (2015). Brain Computer Interface: A Review. *In Brain-Computer Interfaces*, Springer International Publishing, 3-30.

Gavett, S., Wygant, Z., Amiri, S., & Fazel-Rezai, R. (2012). Reducing human error in P300 speller paradigm for brain-computer interface. *In Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2869-2872.

Glennen, S., & DeCoste, D. C. (1997). *The Handbook of Augmentative and Alternative Communication*. Cengage Learning.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes, 25*(2), 149-188.

Hagoort, Peter, Hald, Lea, Bastiaansen, Marcel, & Petersson, Karl Magnus (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438-441.

Hajcak, G., Dunning, J. P., & Foti, D. (2009). Motivated and controlled attention to emotion: time-course of the late positive potential. *Clinical Neurophysiology, 120*(3), 505-510.

Heckhausen, H. (1977). Achievement motivation and its constructs: A cognitive model. *Motivation and Emotion, 1*(4), 283-329.

Hill, N. J., Lal, T. N., Schröder, M., Hinterberger, T., Widman, G., Elger, C. E. & Birbaumer, N. (2006). Classifying event-related desynchronization in EEG, ECoG and MEG signals. *Lecture Notes in Computer Science, 4174*, 404-413.

Hippe, Z. S., & Kulikowski, J. L. (2012). *Human-Computer Systems Interaction: Backgrounds and Applications 2*. J. Wtorek (Ed.). Rzeszów: Springer

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30*(1), 47-61.

Huang, S. C., Bias, R. G., & Schnyer, D. (2015). How are icons processed by the brain? Neuroimaging measures of four types of visual stimuli used in information systems. *Journal of the Association for Information Science and Technology, 66*(4), 702-720.

Jin, J., Allison, B. Z., Kaufmann, T., Kübler, A., Zhang, Y., Wang, X., & Cichocki, A. (2012). The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement. *PloS one, 7*(11), e49688.

Jirayucharoensak, S., Hemakom, A., Chonnaparamutt, W., & Israsena, P. (2011, July). Design and evaluation of a picture-based P300 AAC system. In *Proceedings of the 5th International Conference on Rehabilitation Engineering & Assistive Technology,* p.30. Singapore Therapeutic, Assistive & Rehabilitative Technologies (START) Centre.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159-201.

Kaper, M., Meinicke, P., Grossekathoefer, U., Lingner, T., & Ritter, H. (2004). BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm. *Biomedical Engineering, IEEE Transactions on, 51*(6), 1073-1076.

Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biological Psychology, 102*, 118-129.

Kennedy, L., Dorrance, S., Stoneham, T., Bryant, M., Boyd, K., Flippen, K., & Nichols, D. F (2014). Event-Related Potentials in Humans for Emotional Words versus Pictures, *IMPULSE: The Premier Undergraduate Neuroscience Journal*, *Roanoke College*, 1-16.

Kleih, S. C., & Kübler, A. (2013). Empathy, motivation, and P300 BCI performance. *Frontiers in Human Neuroscience, 7*, 642.

Kleih, S. C., & Kübler, A. (2015). Psychological Factors Influencing Brain-Computer Interface (BCI) Performance. *In Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference,* 3192-3196.

Koo, B., & Choi, S. (2015). SSVEP response on Oculus Rift. In *Brain-Computer Interface (BCI), 2015 3rd IEEE International Winter Conference on*, 1-4.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463-470.

Kübler, Andrea, Kotchoubey, Boris, Wolpaw, Jonathan, R. & Birbaumer, Niels (2001). Brain-Computer Communication: Unlocking the Locked In. *Psychological Bulletin, 127*(3), 358-375.

Lievesley, R., Wozencroft, M., & Ewins, D. (2011). The Emotiv EPOC neuroheadset: an inexpensive method of controlling assistive technologies using facial expressions and thoughts? *Journal of Assistive Technologies, 5*(2), 67-82.

Liu, Y., Zhou, Z., & Hu, D. (2010). Comparison of stimulus types in visual P300 speller of brain-computer interfaces. In *Cognitive Informatics (ICCI), 2010 9th IEEE International Conference*, 273-279.

Liu, Y., Jiang, X., Cao, T., Wan, F., Mak, P. U., Mak, P. I., & Vai, M. (2012, July). Implementation of SSVEP based BCI with Emotiv EPOC. *In Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2012 IEEE International Conference*, 34-37.

Lopez-Gordo, M. A., Sanchez-Morillo, D., & Valle, F. P. (2014). Dry EEG electrodes. *Sensors, 14*(7), 12847-12870.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Mark, J. N., & Wolpaw, J. R. (2009). Clinical applications of brain-computer interfaces: current state and future prospects. *Biomedical Engineering, IEEE Reviews in, 2*, 187-199.

Mathewson, K. E., Harrison, T. J., & Kizuk, S. A. (2017). High and dry? Comparing active dry EEG electrodes to active and passive wet electrodes. *Psychophysiology, 54*(1), 74-82.

Mattout, J., Perrin, M., Bertrand, O., & Maby, E. (2015). Improving BCI performance through co-adaptation: applications to the P300-speller. *Annals of physical and rehabilitation medicine, 58*(1), 23-28.

McCane, L. M., Sellers, E. W., McFarland, D. J., Mak, J. N., Carmack, C. S., Zeitlin, D., Wolpawa, J.R. & Vaughan, T. M. (2014). Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 15*(3-4), 207-215.

McFarland, D. J., Sarnacki, W. A., Townsend, G., Vaughan, T., & Wolpaw, J. R. (2011). The P300-based brain–computer interface (BCI): effects of stimulus rate. *Clinical Neurophysiology, 122*(4), 731-737.

Müller, S. M. T., Diez, P. F., Bastos-Filho, T. F., Sarcinelli-Filho, M., Mut, V., Laciar, E., & Avila, E. (2015). Robotic Wheelchair Commanded by People with Disabilities Using Low/High-Frequency SSVEP-based BCI. *In World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada,* 1177-1180. Springer International Publishing.

National Research Council of Canada (2005), *N400 Event Related Potentials*, p.1 [online] http://www.nrc-cnrc.gc.ca/eng/achievements/highlights/2005/brain_activity.html

Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors, 12*(2), 1211-1279.

Niedermeyer, E., & Da Silva, F. L. (Eds.). (2005). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins.

Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for 2the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098-1111.

Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., & Kübler, A. (2008). An auditory brain–computer interface (BCI). *Journal of Neuroscience Methods, 167*(1), 43-50.

Raju, V. N., Ra, I., & Sankar, R. (2015). A P300-Based BCI Classification Algorithm Using Least Square Support Vector Machine. *International Journal of Software Engineering and Its Applications, 9*(5), 247-254.

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern-und Leistungssituationen (Langversion, 2001). *Diagnostica, 2*, 57-66.

Rupp, R., Kleih, S. C., Leeb, R., Millan, J. D. R., Kübler, A., & Müller-Putz, G. R. (2014). Brain–Computer Interfaces and Assistive Technology. *In Brain-Computer-Interfaces in their Ethical, Social and Cultural Contexts,* 7-38. Springer Netherlands.

Sellers, E. W., Krusienski, D. J., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. *Biological psychology, 73*(3), 242-252.

Speier, W., Deshpande, A., Cui, L., Chandravadia, N., Roberts, D., & Pouratian, N. (2017). A comparison of stimulus types in online classification of the P300 speller using language models. *PloS one, 12*(4).

Support Vector Machines, [online] http://www.support-vector-machines.org/SVM_osh.html

Takano, K., Komatsu, T., Hata, N., Nakajima, Y., & Kansaku, K. (2009). Visual stimuli for the P300 brain–computer interface: a comparison of white/gray and green/blue flicker matrices. *Clinical neurophysiology, 120*(8), 1562-1566.

Tautan, A. M., Mihajlovic, V., Chen, Y. H., Grundlehner, B., Penders, J., & Serdijn, W. (2014, March). Signal Quality in Dry Electrode EEG and the Relation to Skin-Electrode Contact Impedance Magnitude. *In Proceedings of the 7th International Conference on Biomedical Electronics and Devices, Angers, Loire Valley, France,* 3-6.

Taylor, G. S., & Schmidt, C. (2012). Empirical evaluation of the Emotiv EPOC BCI headset for the detection of mental actions. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 193-197. SAGE Publications.

Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in Neurobiology, 90*(4), 418-438.

Wang, M., Daly, I., Allison, B. Z., Jin, J., Zhang, Y., Chen, L., & Wang, X. (2015). A new hybrid BCI paradigm based on P300 and SSVEP. *Journal of Neuroscience Methods, 244*, 16-25.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology, 113*(6), 767-791.

Xiao, X., Luo, J., Lei, M., Cui, Q., Jia, L., & Zhang, Q. (2011). The electrophysiological correlates of the size incongruity effect. *Neuroreport, 22*(4), 157-160.

Zhang, D., Maye, A., Gao, X., Hong, B., Engel, A. K., & Gao, S. (2010). An independent brain–computer interface using covert non-spatial visual selective attention. *Journal of Neural Engineering, 7*(1), 016010.

APPENDICES

APPENDIX A: QUESTIONNAIRE 1

Motivasyon Seviyesi Anketi

Lütfen aşağıdaki çizgiye şu anki motivasyonunuzu (**I**) işareti ile işaretleyiniz.

Hiç Motive Değil   Kesinlikle Motive

APPENDIX B: QUESTIONNAIRE 2

Mod Seviyesi Anketi

Lütfen aşağıdaki çizgiye şu anki modunuzu (**I**) işareti ile işaretleyiniz.

Hiç Modunda Değil — Kesinlikle Modunda

APPENDIX C: QUESTIONNAIRE 3

The QCM includes 18 items, rated on a 7-point likert scale, corresponding to four different components of motivation: mastery confidence, incompetence fear, interest and challenge.

Item	Statement	Factor
1	Bugün Beyin Bilgisayar Arayüzü ile çalışmak için sabırsızlanıyorum	I
2	Bu görevin zorluklarıyla başa çıkabileceğimi düşünüyorum	M
3	Muhtemelen bu çalışma bugün iyi gitmeyecek	M
4	Stratejilerimi geliştirmeyi ve yeni stratejiler denemeyi seviyorum	I
5	Kendimi iyi performans baskısı altında hissediyorum	F
6	Bu çalışma benim için büyük bir meydan okuma(challenge)	C
7	Bugünkü çalışma için sabırsızlanıyorum	I
8	Bugünkü performansımın nasıl olacağını çok merak ediyorum	C
9	Burada kendimi utandırmaktan biraz korkuyorum	F
10	Çalışmada elimden gelenin en iyisini yapmak için çok kararlıyım	C
11	Çalışmaya katılım için bir ödüle ihtiyacım yok; çalışmayı yaparken zaten eğlenirim	I
12	Çalışmada başarısız olmak beni utandırır	F
13	Herkesin beyin aktivitesini kontrol edebileceğini düşünüyorum	M
14	Çalışmayı bugün tamamlayamayacağımı düşünüyorum	M
15	Bugün çalışmada başarılı olduğum zaman, kendimle gurur duyacağım	C
16	Çalışmayı düşündüğüm zaman endişeleniyorum	F
17	Çalışma saatleri dışında da çalışmayı denemek isterdim	I
18	Çalışmayı düşünmek bile beni felç ediyor	F

The third column contains information which motivational factor the item measures: I, interest; M, mastery confidence; F, incompetence fear; C, challenge. Item 3 and 14 have to be reversed.

APPENDIX D: METU's Ethics Committee Approval

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
ueam@metu.edu.tr
www.ueam.metu.edu.tr

Sayı: 28620816/06

13 OCAK 2016

Gönderilen: Yrd.Doç.Dr.Annette HOHENBERGER

Bilişsel Bilimler

Gönderen: Prof. Dr. Canan SÜMER

İnsan Araştırmaları Komisyonu Başkanı

İlgi: Etik Onayı

Sayın Yrd.Doç.Dr.Annette HOHENBERGER danışmanlığını yaptığınız yüksek lisans öğrencisi Uğur ACAR'ın "Beyin Bilgisayar İletişiminin Bilişsel Yönleri: P300 Heceleyici Paradigmasının Uygulaması ve Genişletilmesi" başlıklı araştırması İnsan Araştırmaları Komisyonu tarafından uygun görülerek gerekli onay **2015-FEN-074** protokol numarası ile **01.01.2016-31.12.2016** tarihleri arasında geçerli olmak üzere verilmiştir.

Prof. Dr. Canan SÜMER

Uygulamalı Etik Araştırma Merkezi

İnsan Araştırmaları Komisyonu Başkanı

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü	x
Sosyal Bilimler Enstitüsü	X
Uygulamalı Matematik Enstitüsü	X
Enformatik Enstitüsü	X
Deniz Bilimleri Enstitüsü	X

YAZARIN

Soyadı : ACAR
Adı : UĞUR
Bölümü : Bilişsel Bilimler

TEZİN ADI (İngilizce) : COGNITIVE ASPECTS OF BRAIN-COMPUTER COMMUNICATION: AN IMPLEMENTATION AND EXTENSION OF THE P300 SPELLER PARADIGM

TEZİN TÜRÜ : Yüksek Lisans X Doktora

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın. X

2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullancılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)

3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)

Yazarın imzası: Tarih: