A SELECTIVE ATTENTION AUDITORY BCI

Hilary A. Sandberg
Advisor: Dean Krusienski, Ph.D.
Old Dominion University

Abstract

A brain-computer interface (BCI) uses hardware and software to acquire and then translate brain activity to control devices. Device control is achieved through neural signals alone, and does not use muscle contractions or vocalizations. This study sought to develop a BCI that relied on a subject’s response to auditory stimuli, without the need for visual attention, by monitoring a subject’s auditory attention during 16 different conditions. Eight healthy subjects with normal hearing participated in the study. While neural signal classification accuracies for users across all conditions averaged 75%, a specific condition did not emerge as an overall choice for use as a selective auditory attention BCI. However, by first determining the condition with the greatest classification for each user, this selective attention auditory BCI demonstrates a 75% accuracy.

Introduction

A Brain-Computer Interface (BCI) enables a person to control devices or to communicate with others without using vocalizations or muscular contractions. A BCI relies solely on neural signals to control devices and send messages. A BCI uses an electroencephalogram (EEG) to measure the electric signals from the brain. These signals are recorded with the EEG from electrodes on the scalp. The signals are then amplified and sent to a computer. The computer extracts the signals and translates this information in real-time into a device command, such as selecting a letter on a computer screen. Thus, communication and device control are accomplished through brain signals alone. Applications of BCIs include allowing a means to communicate and to control devices, and a method to monitor attention and neural activity.

Several types of BCIs exist, using different measures of brain activity to control devices and communicate. One of the most common types of BCIs measures the P300 evoked potential, generated by a user's selective attention, to select letters from a computer screen, allowing the user to spell words [2]. Another type of BCI uses steady-state visual evoked potentials (SSVEP) that occur in the brain in response to a flashing visual stimulus to control devices such as prosthetic limbs [9]. However, BCIs that use these designs rely on a response to a visual stimulus, and require the user's eye gaze and visual attention to remain on a screen or flashing stimulus [2]. This requirement presents drawbacks. A user cannot look freely about the environment, or concentrate visual attention on a task.

A type of BCI that can overcome these drawbacks is an auditory BCI. Auditory BCIs have been developed that rely on the P300 response. Furdea (2009) developed an auditory P300-based BCI with 65% averaged accuracy, and Klubassa (2009) developed a P300 BCI with 66%
averaged accuracy. Potential alternatives to P300 classification in auditory BCIs are the auditory steady-state response and selective auditory attention.

Selective attention to an auditory stimulus and ASSRs can be monitored and classified. Rimmele (2010) showed that attention to auditory stimuli increased target detection. Looney (2010) used empirical mode decomposition to classify selective auditory attention when music was played in one ear and speech in the other, with 71% accuracy. Deng (2010) determined the greatest difference in ASSRs occurred when listening to meaningful speech compared to listening to speech without linguistic content.

Based upon these previous studies, this research sought to develop an auditory BCI through monitoring selective auditory attention and auditory steady-state responses (ASSRs). This BCI would translate brain activity generated by attention to auditory stimuli, thus no visual attention or eye gaze would be required. We evaluated the auditory attention focused on the left or right ear using randomized amplitude-modulated stimuli.

**Methods**

**Participants**
Eight healthy subjects participated in this study. Four subjects were native speakers of English and four subjects were non-native speakers of English. Seven subjects were right handed and one subject was left handed. Five males and three females took part in the study, with ages ranging from 22-35. The experiment was approved by the IRB and each subject gave informed consent.

**Procedure**
Auditory stimuli were presented to each ear using Etymotic research-grade insert earphones. Subjects were comfortably seated in a chair in a quiet room. Subjects were directed to keep their eyes closed and limit movements, and did not receive feedback during the experiment.

The auditory stimuli were 10 seconds in length. The stimuli consisted of female normal speech, male normal speech, female reversed speech, male reversed speech, guitar music, piano music, a 440 Hz tone, and a 250 Hz tone. The stimuli were amplitude modulated at 40 Hz and 50 Hz, for a total of 16 conditions. The 40 Hz segments were presented to the left ear, and the 50 Hz segments were presented to the right ear.

The auditory stimuli were presented 2 at a time simultaneously for binary classification. Subjects were directed to focus first on the left ear for the first set of stimuli, then on the right ear for the next set of stimuli. The auditory stimuli were randomized for presentation. Each subject completed five runs of the experiment during the same session, taking short breaks in between each run.

**Data Acquisition**
BCI2000 was used for data recording and auditory stimuli presentation. EEG data was collected using an electrode cap with 64 channels, which were grounded to the left mastoid. The EEG data was amplified and sampled at 1200 Hz.

**Data Analysis**
EEG data was downsampled to 140 Hz and decimated. Alpha rhythm artifacts were removed. A large Laplacian reference was obtained by referencing an electrode to the mean of its four nearest neighboring electrodes. A common average reference spatial filter was applied, thereby enhancing the signal-to-noise ratio of the EEG signals. A common average reference uses the mean of all electrodes as reference. The data was segmented, with five second windows with a
.25 second overlap. The stepwise linear discriminant analysis method (SWLDA) was used for classification.

**Results**

**Classification Accuracies**
All subjects were able to achieve a greater than chance classification in at least one condition. The lowest accuracy achieved for each individual subject was 60%, and the highest accuracy was 100%. The average accuracy was for each subject’s best performance on a condition was 75% across subjects.

**Discussion**

In this study, we attempted to develop an auditory BCI based on selective auditory attention. We sought to determine one condition that provided the best classification for all subjects from the 16 presented conditions.

Through offline analysis of the data from eight subjects we averaged a classification accuracy of 75% for each subject’s best condition performance from the 16 total conditions.

The classification accuracies varied greatly among subjects, with one subject achieving only 60% classification and subject five achieving 100% classification. Subject five outperformed the other subjects by a large margin for all conditions. Excluding classification from muscle movements or faulty processing, this anomaly may be due to the fact that subject five is left-handed. While approximately 5% of those who are right handed have language lateralized to the left hemisphere, 30% of those who are left handed are right-hemisphere dominant for language [7]. The right hemisphere is thought to have a role in both auditory attention and in the syllable coding of speech signals that contribute to the processing of speech in noise [1]. Due to this, a person who is right-hemisphere dominant for language may have an advantage when focusing selective auditory attention in the presence of a competing auditory signal. Subject five’s high classification accuracy suggests that an auditory BCI might perform best when used by a person who has language lateralized to the right hemisphere.

The overall classification accuracy across all subjects of 75% compares favorably to the Furdea (2009) 65% averaged accuracy and the Klubassa (2009) 66% averaged accuracy. However, this accuracy does not exceed the averages for P300 visual-based spellers.

The average classification accuracy in our study originated from each subject’s best performing condition. No one condition from the 16 conditions averaged a higher classification than the others, as performance was highly individualized across subjects. Based on Looney (2010), we hypothesized that the condition in which a subject focused attention on speech that contained linguistic meaning in competition with speech that contained no meaning would achieve the highest accuracy. The fact that this did not occur could be due to several reasons. Our experiment used significantly more conditions than the eight conditions previous studies used, and stimuli duration was 10 seconds in comparison to the 30 second duration of previous studies. Subject fatigue from continuously switching auditory attention could contribute to the variations in performance. Half of our subjects were not native speakers of English, which may also have contributed to the processing of stimuli containing linguistic information. Four of the stimuli consisted of the voices of two of the subjects, which also may have affected classification accuracies.

Because our study did achieve 75% averaged accuracy on an individual basis,
further research in developing an auditory BCI based on selective auditory attention is warranted. In future studies we plan to reduce the number of conditions, remove amplitude modulation from the stimuli, and lengthen the auditory segments. We also will use auditory stimuli that are not familiar to any of the subjects, and provide a feedback component to facilitate online classification.

Acknowledgements

The author gratefully acknowledges the guidance and support of Dr. Dean Krusienski and Dr. Anastasia Raymer from Old Dominion University.

The author gratefully acknowledges the Virginia Space Grant Consortium for sponsoring this research.

The author is thankful for the support of undergraduate research from Dr. Ivan Ash of Old Dominion University.

The author is thankful for the assistance of Garret Johnson and Nicholas Waytowich of Old Dominion University.

References


