Nicholas Johnston: Unspoken Speech Detection Using a Non-Invasive Brain Computer Interface

Introduction:

Electroencephalography (EEG) has been utilized by doctors and physicians for nearly a century, in the diagnosis of epilepsy and other brain related ailments (Berger, 1929). Studies have been conducted that involve detection of motor signals for the purpose of restoring control of extremities to disabled persons (Hoffman et al, 2008), or for the purpose of creating an alternative input modality for computer devices (Kim, S., 2007). Other applications of EEG related technologies include peripheral devices for commercial entertainment. The Emotiv Epoch (www.emotiv.com) is a patented wireless headset used in the video game environment in order to control in-game avatars by thought alone.

Some recent studies have attempted to devise methods for transmitting unspoken speech using implanted brain electrodes. In one such study, patients were able to accurately communicate basic speech signals in the form of syllables through the use of an implanted microarray, coupled with several training sessions (Leuthardt, 2011). Researchers have been able to detect cadences of simple syllables with limited success using a non-invasive EEG in laboratory conditions (D'Zmura et al., 2009). However, detection of imagined, unspoken speech with non-invasive EEG technologies has not yet been achieved, despite recent advancements in noise reduction technology.

Currently, the only non-invasive EEG-based Brain Computer Interface (BCI) which allows communication by locked-in patients is the P-300 speller technique (Krusienski et al, 2006). This technique requires the user to focus on a specific letter among a matrix of different letters on a computer monitor until a brain pattern occurs (event-related potential). The system then flashes rows and columns that are connected with the last detected event-related potential until the exact row and column of the character is discovered. The process is repeated until the user has spelled out complete words. This method of communication currently produces words at an average of one per 105 seconds for a 5 letter word (Usakli, 2009). The P300 spelling process has been found to be 100% accurate in 72.8% of subjects that use it (Guger et al, 2009).

Purpose

To enable unspoken communication through direct brain input by using various signal preprocessing, feature extraction and post-processing techniques while using a novel pattern matching system in order to allow classification of English language letters and short words. The hope is that the speed with which words can be accurately detected will be improved when compared with the existing P-300 speller system. A quicker speller system would enable locked-in patients (non-mobile people such as Dr. Stephen Hawking) the ability to communicate more thoroughly.

Hypotheses

It is hypothesized that the letters of the alphabet can be accurately classified using non-invasive EEG by a combination of signal enhancement, feature extraction and post-processing algorithms and the use of a novel classification method. Letters within words imagined by the user as a series of musical notes at a specific cadence will be detected more accurately and more quickly than by the current non-invasive EEG P-300 speller technique.

Materials

An 8-channel EEG was built in order to test subjects:

- Eight channel EEG system (utilizing the Texas Instruments ADS1298 FE*) Evaluation board. The MMBO portion is designed for ECG configuration and not used in this experiment.
- Silver EEG electrodes (www.rochestermed.com) and headband to hold EEG leads in place.
- Breadboards and connectors (www.nkelectronics.com)
- EEG paste for electrode conductivity (www.rochestermed.com)
- Mobile PC (battery powered) for data recording and analysis (x86, Windows XP SP3)
- OpenViBE software platform (openvibe.inria.fr)
- Arduino development environment and UNO R3 controller board (www.arduino.cc)
- C++ development environment (Microsoft Visual C++ Express 2010 required for OpenVibe)

* The Texas Instruments 24- bit ADS1298 Analogue to Digital Converter introduced in 2009 is being utilized as it contains a complete integrated front end, including Input Buffer, Low Pass DC filter, High Pass filter and the Analog to Digital Converter (ADC) itself. This low power device removes the complexity and potential noise caused by soldering these components separately.

Laboratory Setup

A custom-made laboratory setup was created to conduct the experiment.

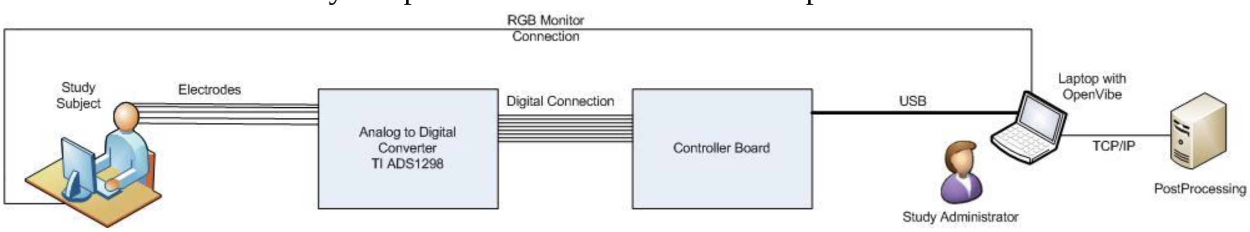


Figure 1: Laboratory setup to record participant EEG readings.

Methods

In order to conduct the experiment, an 8-channel EEG device was constructed as per the materials listed above. The TI ADS1298 AD chip was programmed in Arduino-C using an Arduino Uno R3 controller, which communicated via a programmed SPI bus. The OpenVibe development environment was programmed with four components: an Acquisition Step for acquiring subject data from the device, an LDA Training Step for the off-line programming of the LDA classifiers, a Bayes Training step for the off-line programming of the Bayes Classifier and the Online Step for the real-time testing of subjects. The experimental flow matching audible sounds heard by the test subjects and the recording of brainwave data was controlled by a script written in LUA for this purpose (www.lua.org). A Windows Driver was built in C++ for OpenVibe to communicate with the purpose-built EEG device. A Bayes classifier was created in C in order to complete the classification system. Each of the custom scripts was called by the OpenVibe C++ environment.

Study subjects had 6 electrodes attached at points Fz, Cz, Pz, Oz, P3, P4 as per the International 10-20 system of electrode placement (Thompson, 2003). The electrode locations were chosen in order to reduce electromyographic artifact commonly found on the lowest positions on the head, low on the temple, close to the eyes, at or below the ear or below the external occipital protuberance (D'Zmura, 2009). They sat in a quiet minivan, free from mains electrical interference with no distractions and were asked not to move their eyes during the 2 and 3 second recording periods.

The subjects were instructed to silently verbalize specific musical cadences representing different letters. Subjects were asked to silently verbalize "doe" for the lowest note, "ray" for the medium note and "me" for the high note. This was done to ensure a limited length reproducible tone, rather than just hum the note. Musical cadences were audibly played to subjects via shielded speakers immediately before the recording.

Testing Procedure

1. Created a connection between the purpose-built EEG system and OpenViBE software platform with the study subject attached with electrodes.
2. Recorded readings from human subject and recorded results:
   - A training set of 6 letters recorded 10 times each
   - After creating classifiers based on the training set, the user's brainwaves were recorded as they attempted to spell the word "AFTER." The subjects were asked to think of humming the pattern for each letter without actually verbalizing it. The time required to detect the intended letter was then timed. The total time required to correctly detect all five letters was used as the recording time for the word "AFTER".
   - The total time taken recording the five letter word "AFTER" was recorded 4 times for each test subject.
3. Each letter was encoded with a specific cadence and rhythm. Two sets of rhythms were initially tested on a single subject to see which creates the highest level of recognition by test subjects.
   a) 2 second Rhythm patterns using three tones per letter:

Letter	Tones	Explanation
A	1_1_1	> Each number represents a musical note
F	2_2_2	> 3 = musical note F
T	3_3_3	> 2 = musical note B
E	1_2_3	> 1 = musical note E
R	3_2_1	
backsp	2_3_1	

Figure 2: 2-second rhythm for each of 5 tested letters and backspace (unused)

   b) 3 second Rhythm patterns using four tones per letter:

Letter	Tones	Explanation
A	1_X_1_X	> Each number represents a musical note
F	2_X_3_1	> 3 = musical note F
T	3_X_X_3	> 2 = musical note B
E	1_2_3_X	> 1 = musical note E
R	3_2_1_X	X = space (no sound)
backsp	X_2_2_X	

Figure 3: 3-second rhythm for each of 5 tested letters and backspace (unused)

4. Recorded brainwaves were uniquely identified as follows:
    a. Each signal was decomposed into 6 different frequency ranges: 8- 11 Hz, 12- 15 Hz, 16- 19 Hz, 20- 23 Hz, 24- 27 Hz and 28- 31Hz.
    b. Each brainwave recorded was subjected to a Linear Discriminant Analysis (LDA) to classify the signal in each range. The LDA yielded a 0 or 1 for each sample.
    c. Each LDA classification result (binary) was used by a Bayes classifier to calculate the highest probability class based on the presented binary features.
5. With data created using the new pattern matching system, the time taken to correctly identify all letters in the 5 letter word “AFTER” was recorded. This time was then compared to the known spelling time of the P-300 speller for a word of the same length. Theoretically, with a 3 second cadence, the quickest expected time was 5 letters X 3 seconds per letter = 15 seconds.

Results:

The three second rhythm more quickly and accurately determined the intended letter than the two second rhythm by a considerable margin (figure 4). This was the reason why the three second rhythm was used in the multi-subject determination of time to detect the word “AFTER”.

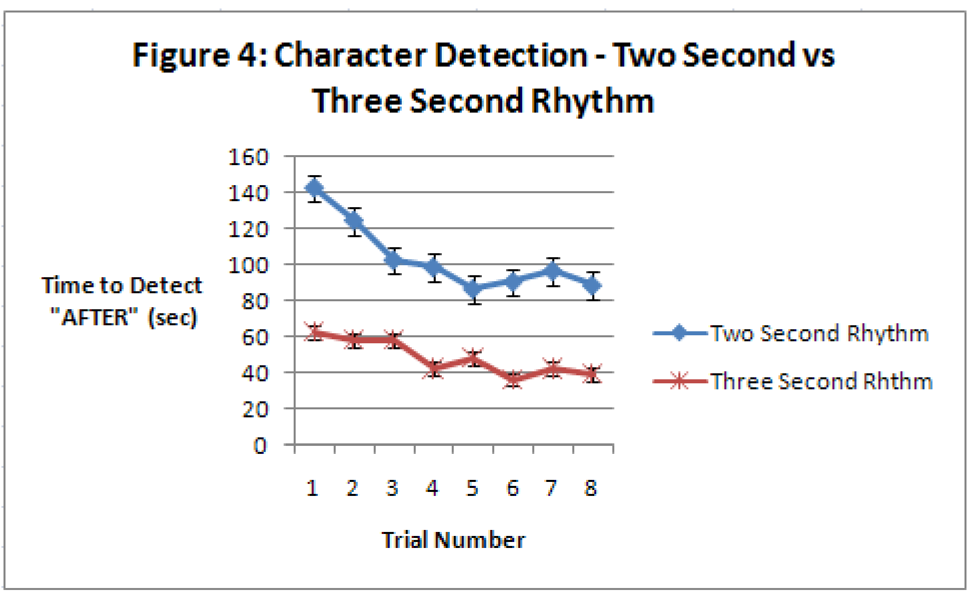
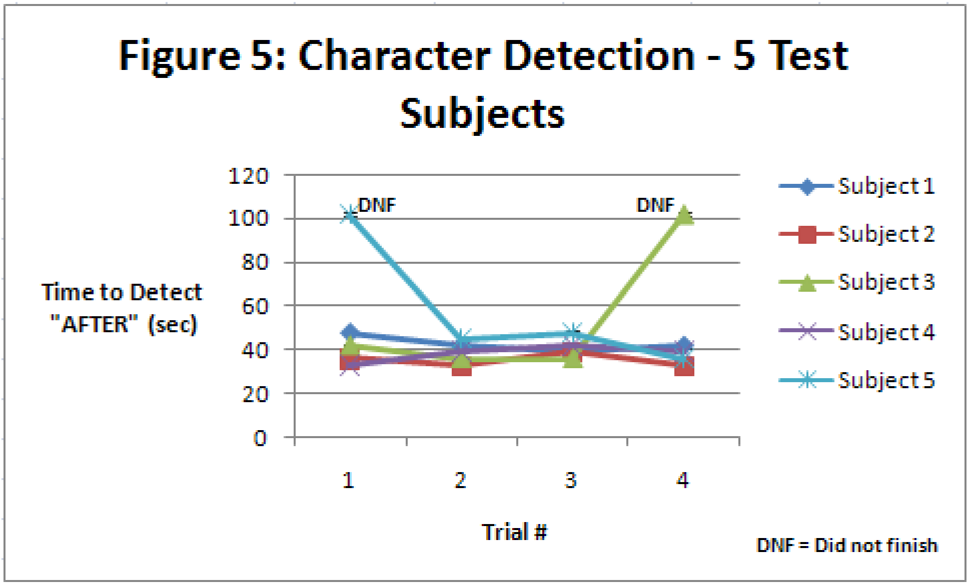


Figure 4: Character Detection - Two Second vs Three Second Rhythm

The three second rhythm created more significantly unique waveforms for each letter depicted and was used as the rhythm for the detection of letters with 5 test subjects (Figure 5).



The rate of spelling a five letter work using the non-invasive ADS1298/Arduino EEG system described here was found to be an average of 39.3 seconds with a standard deviation of 2.0 seconds for all of the test results (18 data points). Note that two subjects were each unable to finish spelling the test word in one of their trials. This is a significant improvement over the P-300 Speller system which has been tested many times, but is known to have a speed of 105 seconds for a 5 letter word (Usakli, 2009).

Conclusions:

The ADS1298/Arduino system of determining words from letters encoded using musical notes over a 3 second/4 note rhythm was found to be significantly quicker than the current P-300 Speller system. The use of the LDA Classifiers and Bayes Classification for analysing multiple states (letters) was found to be effective.

Acknowledgements:

This work was accomplished with the assistance of Dr. Raymond Lam, who reviewed the project for safety and practicality. The review of the project by the Institutional Review Board comprised of Ms. Bea Hadikan, Ms. Nancy Arsenault and Ms. Lynn Porpaczy is also greatly appreciated. Mr. Tom Hendrick from Texas Instruments was helpful in debugging the ADS1298/controller board connectivity. Lastly, thanks to Mr. Randy Johnston who volunteered to be the project supervisor.

References

Bashashati, A., Fatourechi, M,. Ward, R.K., and G.E. Birch. A Survey of Signal Processing Algorithms in Brain-Computer Interfaces Based on Electrical Brain Signals. *Journal of Neural Engineering,* Volume 3, Number 4, pages R32- R57.

Brigham, K. and Kumar, V. (2010). Subject identification From Electroencephalogram (EEG) Signals During Imagined Speech. *Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on Biometrics*. Pages 1- 8.

Berger, Hans (1929). *Über das Elektroenkephalogramm des Menschen.* Archiv für Psychiatrie und Nervenkrankheiten, 1929, 87: 527-570.

D'Zmura, M., Deng, S., Lappas, T., Thorpe, S. and R. Srinivasan. (2009). Toward EEG Sensing of Imagined Speech, In J.A. Jacko (Ed.), *Human Computer Interaction, Part 1, Learning Notes in Computer Science,* 5610, pages 40- 48.

Farell, L.A. and E. Donchin (1988). Taking the Top Off Your Head: Toward a Mental Prosthesis Utilizing Event-Related Brain Potentials. *Electroencephalography and Clinical Neurophysiology,* 70 (6), pages 510- 523.

Guger, Christoph, Daban, S, Sellers, E., Holzner, C., Krausz, G., Carabolona, R., Gramatica, F. and E. Guenter (2009). How Many People Are Able to Control a P300-based Brain-Computer Interface (BCI)? *Neuroscience Letters* 462 (2009), pages 94- 98.

Hoffmann, Ulrich, Vesin, J-M., Ebrahimi, T., and K. Diserens. (2008). An Efficient P300-Based Brain-Computer Interface for Disabled Subjects. *Journal of Neuroscience Methods.* Volume 167, Issue 1, pages 115- 125.

Kim, S-P., Simeral, J.D., Hochberg, L.R., Donoghue, J.P., Friehs, G.M., and M.J. Black. (2007). Multi-state Decoding of Point-and-click Control Signals from Motor Cortical Activity in a Human with Tetraplegia. In *Proceedings of the 3rd International IEEE EMBS Conference on Neural Engineering,* May 2- 5, 2007, pages 486- 489.

Krusienski, D.J, Sellers, E.W., Cabestaing, S.B., McFarland, D.J., Vaughn, T.M. and J.R. Wolpaw. (2006). A Comparison of Classification Techniques for the P300 Speller. *Journal of Neural Engineering,* Volume 3, Pages 299- 305.

Leuthardt, E.C., Gaona, C., Sharma, M., Szramal, N., Roland, J., Freundberg, J.S., Breshears, J., and Gerwin Shalk. (2011). Using the Electrocorticographic Speech Network to Control a Brain-Computer Interface in Humans. *Journal of Neural Engineering,* Volume 8, Number 3, pages 36- 39.

Usakli, A.B., Gurkan, S., Aloise, F., Vecchiato, G., and F. Babolini. (2009). On the Use of Electrooculogram for Efficient Human Computer Interfaces. *Computational Intelligence and Neuroscience,* Volume 2010, pages 1- 5.

Bibliography

Acharya, V. Improving Common-Mode Rejection Using the Right-Leg Drive Amplifier. Retrieved January 16, 2012 from http://www.ti.com/lit/an/sbaa188/sbaa188.pdf.

Arduino Forum. Accessed January – March 2012 at http://arduino.cc/forum/.

Bashashati, A., Fatourechi, M,. Ward, R.K., and G.E. Birch (Oct. 2007). A Survey of Signal Processing Algorithms in Brain-Computer Interfaces Based on Electrical Brain Signals. *Journal of Neural Engineering*, Volume 3, Number 4, pages R32- R57.

Brigham, K. and Kumar, V. (2010). Subject identification From Electroencephalogram (EEG) Signals During Imagined Speech. *Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on Biometrics*. Pages 1- 8.

Datasheet ADS1298, Low-Power, 8-Channel, 24-Bit Analog Front-End for Biopotential Meaurements (REV.C). Retrieved January 10, 2012 from http://www.ti.com/lit/ds/sbas459i/sbas459i.pdf.

Gray, R.M. and L.D. Davisson (2010). An Introduction to Statistical Signal Processing. Cambridge University Press. Boston MA. 478 pages.

Hendrick, Tom. Personal Communication February 29- March 6, 2012 via Forum at http://e2e.ti.com/support/data_converters/precision_data_converters/f/73/t/168648.aspx

Lotte, F., Congedo, M., Lecuyer, A., Lamarche, F. and B. Arnaldi. A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces (2007). *J. Neural Engineering* 4: R1- R13.

Margolis, M. (2011). Arduino Cookbook – Recipes to Begin, Expand and Enhance Your Projects. O-Reilly Publishing. Sebastopol, CA. 631 pages.

Mueller, J.P. and J. Cogswell (2009). C++ All-in-one Desk Reference for Dummies – Second Edition. Wiley Publishing. Hoboken NJ. 865 pages.

Nathan, M. Fixing Arduino's Serial Latency Issues. Retrieved March 3, 2012 from http://hackaday.com/2011/05/18/fixing-arduinos-serial-latency-issues/.

Romero, Renato. (n.d.). *FFT For Dummies.* Retrieved Feb 10, 2011 from VLF website: http://www.vlf.it/fft_beginners/fft_beginners.html.

Schertz, P.(2006). Practical Electronics for Inventors – Second Edition. McGraw-Hill Publishing. Toronto. 952 pages.

Wolpaw J.R., Birbaumer N., Heetderks W.J., McFarland D. J., Peckham P.H., Shalk G., Donchin E., Quatrano L.A., Robinson J.C. and Vaughan T.M. (2000). *Brain-Computer Interface Technology: A Review of the First International Meeting,* IEEE Trans. Rehabilitation Engineering, 8, pages 161-163.