**Introduction**

Speaking is the most natural way of communicating information however we have always dreamt of transmitting information to others through extra-sensorial channels (also known as telepathy). A large spectrum of different methods of communication exists between speaking and communicating telepathically. In this paper we explore Electromyography as other medium such as to see if it possible to reduce the gap between speaking and telepathy.

**Electromyography principles**

Muscle tissues conducts electrical potentials, also known as muscle action potential. These muscle action potentials are generated with movement. Electromyography (EMG) consist in detecting and measuring those action potentials.

There are two types of EMG; intra-muscular EMG and surface EMG. As intra-muscular EMG uses intrusive electrodes we decided to use the most common, non-intrusive surface EMG.

A set of EMG surface electrodes is formed of two detection electrodes and a reference electrode. The muscle action potential is measured between the detection electrodes, it is then amplified. The positioning of the detection electrodes is important, they should not be placed too far apart. Furthermore it is important to place them such as to limit cross-talk between different muscles. This is particularly true on smaller muscles. The reference electrode provides a common reference point to the differential input, it should be placed on an electrically neutral tissue (with bony prominence).

The EMG electrodes have electrically conductive gels permitting stability and a good electrical contact with the skin.



To limit noise acquired by the EMG signals such as electrical noise and maximise the quality of the signal, signal processing is applied.

**Subvocalization**

Subvocalization involves saying words in your head while reading. This inner speech controls the larynx and other muscles involved in speech. Our original objective for this study was to capture electrical signals from subvalization using electromyography. However after preliminary tests we concluded that it was impossible to detect those electrical signals.

To reduce the gap between subvocalization and speaking we decided to use electromyography such as to expose whether or not it is possible to detect speech using electromyography and if there is a difference between speaking and mumblings.

**Facial anatomy**

Speech production can be divided in two sections: the larynx and the facial muscles.

The larynx (the voice box) is responsible for voice production, it controls the vocal folds producing a set of sounds which are formed into speech by the mouth and the tongue.

The mouth is controlled by many different facial muscles:

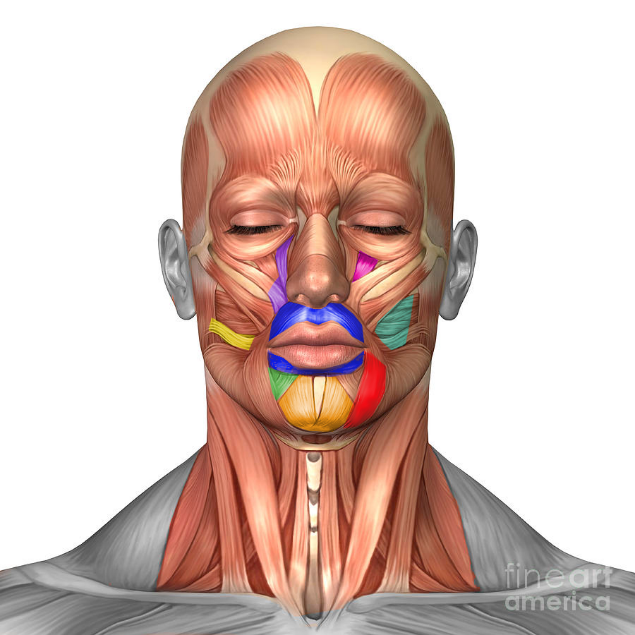


Image credit: fine art America

: Orbicularis oris: purses the lips

: Levator labii: raises upper lip

: Risoris: draws the lips in a smile

: Buccinator: pulls the lips wide and tight

: Depressor labii: lowers the lips

: Mentallis: pulls chin down

: Depressor anguli oris: lowers the bottom corners of the lips

: Levator anguli oris: raises the upper corner of the lips

*//add mylohoid*

**Study design**

The objective of our study is to analyse the difference, if any, of EMG signals between speaking and mumbling words. This method is used for word recognition.

There were different variables to consider in the study, the main ones are locations and word set.

**Location**

There are different parameters to consider when deciding where to place the surface electrodes. For example, the electrodes used have a skin contact size that can be too large for some muscles hence some locations were disregarded as the signals detected would have information from neighbouring muscles, this is known as cross-talk. Hence Levator anguli oris and depressor labii were not chosen.

Furthermore to determine were the surface electrodes would be placed, different combinations of locations were tested. The chosen combination obtained the highest accuracy ratio when the machine learning algorithm was applied.

The chosen set of locations is: Depressor anguli oris, orbicularis oris upper and lower, and mylihoid.

*Chosen positions: (graph or image )*

**Word set chosen**

For the study 10 words were chosen. These words are directionally related, this choice was made such as to link the study to future uses.

*List words chosen*

**System**

*Hardware*

To measure the muscle action potentials at the different locations we used 4 sets of EMG surface electrodes. Each set of electrodes are connected to a muscle sensor v3, the sensor is designed to be used directly and hence outputs an amplified, rectified and smoothed signal removing all noises. Each muscle sensor requires two 9V batteries, we used rechargeable ones.

Furthermore an Arduino was used as a bridge between the software and hardware parts of the study.

*Software*

When speaking, a word is formed using a combination of muscles. For example when pronouncing ‘o’ the main muscle used will be Orbicularis Oris. This implicates that each word has a unique pattern that combines different muscle action potentials.

Extracting a subset of features obtained from the signals of a word we are able to use them to train a machine learning algorithm. Later this trained machine will be able to determine words from a sample set of features.

For determining the best machine learning algorithm to be used we tested Support Vector Machine (SVM), Neural Networks, Linear Discriminant Analysis and Gaussian classifier.

In the end we chose the LDA since it was proven that it gave the best result.

We used Scikit for machine learning algorithms.

(*graph of different results with the different classifiers*)

**User study**

10 participants were recruited to take part in the user study; 2 were female and 4 men. The participants were asked to sit in front of a computer. After the electrodes were placed they were asked to follow the instructions displayed on the computer screen. The video showed the user which word to say and an indicator gave visual cues for the timing of the speech.

The conditions for the experiment were mumbling and speaking. Each of the 6 participants performed both conditions.

For both conditions the participant had to say a total of 10 words repeating them 10 times each (words changing every 5 pronunciation).

Therefore the study involved: 6 participants X 2 conditions X 10 words X 10 repetitions = 1200 words. The average study lasted 40 min, including the time to place the electrodes and explain the study to the participant.

The measures captured were then used to train the machine learning algorithm. The accuracy of the classifier is calculated across a 10-fold cross validation classification results.

(*display them*)

The difference condition accuracy is shown in: … *(graph of some sort)*

**Discussion**

The accuracy in word recognition is quite high, between 74% and 97% with an average of 84%. The lowest accuracy obtained, 74%, could have suffered from external factors. This assumption can be confirmed when looking at the results from the two different conditions. Whilst the results of the mumbling condition obtained for the participants are similar and mostly better than the speaking condition the participant who obtained the lowest accuracy result in the speaking condition has a ratio 14% lower to the speaking condition. This ratio is significantly higher than any other results. As this condition was performed last we can assume that some of the electrodes must have lost contact with the skin during the study.

Furthermore we can hypothesis that differences between certain English accents have effects on how one forms words with their facial muscles, this could be a factor that influences the study as the participants displayed different accents.

In summary the results from the study suggests that there is no major difference between the speaking and mumbling conditions in classification. In fact we can say that the mumbling condition tends to be slightly better.

**Limitations**

**Future uses**